A Paediatrician’s guide to epigenetics
Lauren Byrne1 and Amanda J Drake1

1. University/British Heart Foundation Centre for Cardiovascular Science, University of Edinburgh, The Queen’s Medical Research Institute, 47 Little France Crescent, Edinburgh EH16 4TJ, UK

Address all correspondence to:
Amanda J Drake
University/BHF Centre for Cardiovascular Science, University of Edinburgh
The Queen’s Medical Research Institute
Edinburgh, EH16 4TJ
Telephone 44 131 2426748
Fax 44 131 2426779
Email: mandy.drake@ed.ac.uk
ORCID ID: https://orcid.org/0000-0001-8633-333X

Keywords: Epigenetics, DNA methylation, transgenerational epigenetic inheritance, Developmental Origins of Health and Disease
Abstract
Epigenetic regulation of gene expression is critical for normal development. Dysregulation of the epigenome can lead to the development and progression of a number of diseases relevant to paediatricians, including disorders of genomic imprinting and malignancies. It has long been recognised that early life events have implications for future disease risk, and epigenetic modifications may play a role in this, although further high-quality research is needed to better understand the underlying mechanisms. Research in the field of epigenetics will contribute to a greater understanding of growth, development and disease; however, paediatricians need to be able to interpret such research critically, in order to be able to utilise the potential advances brought about through epigenetic studies, whilst appreciating their limitations.

Summary
• Epigenetic regulation of gene expression is critical for normal development and disruption can result in disease

• Diseases in which epigenetic dysregulation is known to occur include those involving genomic imprinting (e.g. Prader-Willi, Angelman, Beckwith-Wiedemann and Silver Russell syndromes), Rett syndrome and many malignancies.

• Studies suggest epigenetic dysregulation is an important mediator in the Developmental Origins of Health and Disease (DOHaD) however further high-quality research is needed to better understand the underlying mechanisms.

• An understanding of epigenetics will allow paediatricians to interpret research critically, embracing the potential advances brought about through epigenetic studies, whilst appreciating their limitations.
What is meant by the term ‘epigenetics’?

‘Epigenetics’ was a term coined by Conrad Waddington as the ‘branch of biology which studies the causal interactions between genes and their products, which bring the phenotype into being’. More recent definitions include a Cold Spring Harbor meeting consensus definition of ‘stably heritable phenotype resulting from changes in a chromosome without alterations in the DNA sequence’ and an NIH Roadmap Epigenomics Project definition of ‘both heritable changes in gene activity and expression…and also stable, long-term alterations in the transcriptional potential of a cell that are not necessarily heritable’. The lack of a clear, unifying definition means the term is commonly used to describe many processes which may affect gene function, in particular DNA methylation and histone modifications.

**DNA methylation:**

In mammals, DNA methylation typically involves the addition of a methyl group to the 5’carbon of cytosine residues (5-methylcytosine; 5mC) in a cytosine-phosphate-guanine (CpG) sequence. ‘CpG islands’ are regions of the genome that contain a large number of CpG dinucleotides which are common near the transcription start site of genes and are generally unmethylated. In differentiated cells, DNA methylation patterns associate with gene silencing or activation, however there is ongoing debate over the extent to which DNA methylation drives changes in gene transcription as opposed to reflecting transcriptional changes driven by the actions of transcription factors.1 Although not fully understood, DNA methylation is thought to have an important role in development and cellular differentiation, for example, in silencing germline genes in somatic cells. Furthermore, DNA methylation plays an important role in genomic imprinting, X-inactivation and in silencing transposable elements, which make up a large fraction of the genome.

**Histone modifications:**

DNA is packaged around histone proteins in a complex known as chromatin. Post-translational histone modifications (e.g. methylation, acetylation, phosphorylation) are important in many biological processes such as gene regulation, DNA repair and chromatin condensation. In general, DNA that encodes actively transcribed genes is more loosely packaged (euchromatin) while DNA encoding inactive genes is more tightly condensed (heterochromatin).

**The epigenome and paediatric disease**

Epigenetic regulation of gene expression is critical for normal development and disruption can result in disease.2 There are several examples of paediatric diseases in which perturbations of epigenetic processes have been described.

**a) Imprinting disorders**
Most autosomal genes display bi-allelic expression and are expressed equally from maternal and paternal alleles. Imprinted genes display mono-allelic expression; at these genes, either the maternal or paternal allele is silenced during germ cell development, predominantly by DNA methylation. Imprinted genes are normally arranged in clusters and regulated by common imprinting control regions (ICRs) which are differentially methylated dependent upon which parent the gene is inherited from. Genomic imprinting is essential for normal growth and development, however any genetic/epigenetic alteration resulting in downregulation or loss of function of the active allele could have significant consequences, as there is no other active allele to compensate. The imprinting disorders primarily involve problems with growth or neurological development.

There are several different mechanisms leading to an increase or decrease in the dose of imprinted genes: (i) genetic abnormalities affecting the expressed allele e.g. mutations, deletions or duplications; (ii) chromosomal errors resulting in the re-organisation of genetic material e.g. translocations and uniparental disomy; and (iii) epigenetic errors affecting DNA methylation of differentially methylated regions controlling the expression of imprinted genes, in some cases due to an underlying genetic abnormality. Thus, although these conditions are associated with epigenetic dysregulation, the majority occur as a consequence of genetic abnormalities.

**Angelman and Prader-Willi syndromes**

Angelman (AS) and Prader-Willi (PWS) syndromes are perhaps the most well-known imprinting disorders, both caused by functional loss of genes located in the same region of chromosome 15. AS is characterised by cognitive impairment, developmental delay, speech impairment and epilepsy. 70% of cases result from deletion of a region in the maternal chromosome 15, which contains UBE3A, a gene in the ubiquitin pathway which, in the developing brain, is usually maternally-expressed and paternally-silenced. Less common causes include mutations in the maternally inherited gene, uniparental disomy (where both chromosome 15s are inherited from the father), and imprinting defects, whereby the maternal gene is also silenced. Common to all cases is the loss of expression of maternal UBE3A, with an inability to compensate for this due to the normal silencing of the paternal allele. PWS is phenotypically very different, presenting with hypotonia, delayed development, behavioural problems, hyperphagia and obesity resulting from loss of the normally expressed paternal copies of genes in the same region of chromosome 15, as a consequence of deletion, mutation or uniparental disomy. Again there is an inability to compensate for the lack of gene expression from the paternal alleles because the maternal copies are silenced.

**Beckwith-Wiedemann and Silver-Russell Syndrome**

Precise epigenetic regulation of gene expression provides fine-tuning of growth signals. Paternally-expressed genes tend to be growth-promoting, whilst maternally-expressed genes tend to growth-
restrict. Beckwith-Wiedemann syndrome (BWS) is an overgrowth disorder characterised by macrosomia, macroglossia, abdominal wall defects and an increased risk of embryonal tumours. BWS is a clinically heterogeneous condition which highlights the complexity of epigenetic regulation of growth; the condition can be caused by several different mechanisms, all affecting genes important for growth regulation in the imprinted region of chromosome 11p15.

This imprinted region is functionally divided into two domains, each of which has a different ICR (Figure 1). ICR1 regulates expression of IGF2, a growth-promoting gene, and H19, a non-coding RNA (ncRNA). ICR1 is normally methylated on the paternal allele, and unmethylated on the maternal allele leading to paternal IGF2 expression and maternal H19 expression. In contrast, ICR2 is methylated on the maternal allele and regulates expression of several genes, including CDKN1C, a negative regulator of cell proliferation. Hypomethylation of ICR2 on the maternal allele (the most common epigenetic alteration in BWS) results in reduced CDKN1C expression. BWS can also be caused by mutations in CDKN1C, or through gain of methylation at maternal ICR1 resulting in biallelic expression of IGF2 and biallelic silencing of H19. A small proportion of cases are caused by paternal uniparental disomy. This deeper understanding of the causative mechanisms of BWS can assist with predicting clinical features, as different genetic and epigenetic alterations are associated with specific phenotypes, including differing relative risks of tumour development.

Interestingly, loss of methylation of ICR1 on the paternal allele, resulting in biallelic silencing of IGF2, causes the clinical condition known as Silver-Russell syndrome (SRS). Children with SRS have impaired growth, including reduced birthweight, failure to thrive, short stature and distinct facial features. SRS can also be caused by maternal uniparental disomy of chromosome 7 although the mechanism by which this results in the same phenotype as epigenetic changes in ICR1 on chromosome 11 remains unknown. A greater understanding of these disorders and the associated epigenetic dysregulation is likely to contribute to new knowledge of the mechanisms underpinning normal growth and development.

b) Rett syndrome

Rett syndrome is a severe neurological disorder caused by a mutation in the X-linked gene Methyl-CpG-binding protein 2 (MECP2). It is inherited in an X-linked dominant pattern and is therefore typically only seen in girls, as MECP2 mutation is lethal in hemizygous males. Rett syndrome is characterised by normal growth and development for the first 6-18 months, followed by developmental slowing and loss of motor skills and speech. Affected girls also display gait abnormalities, stereotypical hand movements, seizures and intellectual disability. The role of the epigenome can be seen two-fold in this syndrome: firstly, epigenetic processes account for at least part of the considerable phenotypic variability amongst patients with Rett syndrome, through the
normal biological process of X-inactivation. In females, one X-chromosome is silenced by DNA methylation, a process which usually occurs randomly in any cell. Thus, the abnormal X-chromosome should be silenced in ~50% of cells. However, X-inactivation can sometimes display skewing, resulting in a greater or lesser proportion of cells with the abnormal X-chromosome being active, resulting in a more or less severe phenotype. Secondly, unlike the imprinting disorders, the mutation affects a protein predicted to have epigenetic properties: MECP2 encodes methyl CpG binding protein 2 which binds to methylated DNA, essentially “reading” the epigenetic code and regulating gene transcription. The protein is expressed particularly highly in the brain and is important for maintaining neuronal function. The neurological phenotype in MECP2-null mice can be reversed by re-expression of MECP2 after birth, highlighting the reversible nature of epigenetic changes and the promise of future therapeutic strategies for disorders such as Rett syndrome.

**c) Cancer**

In the disorders discussed thus far we have seen that epigenetics plays an important role in growth and development, and so it is perhaps unsurprising that epigenetic dysregulation is also seen in cancer - a disease in which normal developmental programs are disordered and there is uncontrolled cell proliferation. Altered DNA methylation is a hallmark of human cancers – including both global DNA hypomethylation and CpG island hypermethylation. DNA hypomethylation of oncogenes could result in their increased expression, whilst promoter CpG island hypermethylation often results in silencing of the associated gene, which may be of particular importance in the context of tumour suppressor genes. For example, DNA methylation changes occur in human gliomas, with some studies showing CpG island hypermethylation in a subset of gliomas. A ‘CpG island methylator phenotype’ (CIMP) resulting in cancer-specific CpG island hypermethylation has been described in a subset of cancers including gliomas. Notably, the CIMP occurs as a consequence of gene mutations, for example mutations in the genes isocitrate dehydrogenase type 1 (IDH1) and BRAF (a proto-oncogene encoding for the B-Raf protein) have been described in association with CIMP in a subset of gliomas.

Our knowledge of the role of histone modifications in controlling gene expression and contributing to oncogenesis has increased dramatically over recent years. This is clearly demonstrated in paediatric high-grade gliomas (HGGs), which account for around half of paediatric gliomas, themselves the most common type of paediatric brain tumour. Although many oncological disorders of childhood have seen significant improvements in survival rates over recent years, children with HGGs continue to have very poor outcomes with 2-year survival ranging from 10-30%. Recent analysis of paediatric HGGs has identified genetic mutations in histone 3 variants, which result in disruption of the epigenetic post-translational modification of histones. Importantly, distinct mutations affecting modifications at separate sites on the histone tail associate with the neuroanatomical site of the
tumours, effectively dividing patients into subgroups, with differing demographics, response to treatment and prognosis. Furthermore, these mutations are not seen in adult HGGs, suggesting that the mechanisms of tumorigenesis in adult and paediatric HGGs are molecularly distinct.

This increased awareness and understanding of the complex interplay between genetic and epigenetic dysregulation in paediatric cancers has implications for treatment, as it opens the door for novel therapeutic strategies, particularly targeting the epigenome. The dynamic nature and potential reversibility of epigenetic modifications makes epigenetic therapies an attractive proposition, indeed a number of drugs affecting histone modifications are already in clinical trials for paediatric gliomas, and epigenetic therapies are already successfully used in the treatment of other malignancies.

Epigenetics and the Developmental Origins of Health and Disease (DOHaD)

Numerous studies in humans and in animal models have demonstrated that exposure to adverse environmental conditions in utero or in early postnatal life increases the risk of cardio-metabolic, neurobehavioural and reproductive disorders in adulthood. Adverse early life experiences include prenatal and early postnatal factors e.g. under- or overnutrition, glucocorticoid overexposure or being born preterm. The dynamic nature of the epigenome may enable an organism to respond and adapt to environmental influences, and epigenetic modifications may provide a mechanism by which rapid and reversible phenotypic changes can be achieved. Many animal and human studies have reported that early life environmental factors associate with alterations in DNA methylation, histone modification and/or ncRNA (reviewed in ). This is distinct from the previous examples, since it represents a purely non-mutational mechanism of disease susceptibility. However, although the substantial amount of interest in the area means that it is now almost accepted that the early life environment influences long-term health through induced changes in epigenetic marks during development, it is important to note that there is a general lack of mechanistic evidence.

Examples of human studies suggesting a role for the epigenome include those showing that individuals exposed to famine in utero during the Dutch Hunger Winter have altered DNA methylation at IGF2 and a number of other metabolic genes when compared with their unexposed same-sex siblings. Maternal consumption of an unbalanced diet associates with changes in DNA methylation in the offspring in adulthood. Exposure to maternal obesity and/or gestational diabetes is associated with an increased risk of obesity and diabetes in the offspring, and changes in DNA methylation have been reported in individuals in these studies. Individuals born preterm are at higher risk of neurodevelopmental disorders and we have demonstrated alterations in DNA methylation at sites relevant to neurodevelopment in preterm infants in comparison to term born controls. In order to further explore the role of the epigenome in the programming of disease risk, animal models have been developed. These include manipulations of maternal diet (including under-
or overnutrition), prenatal glucocorticoid overexposure or exposure to environmental toxins or ‘endocrine disrupting chemicals’ in species as diverse as rodents, sheep and non-human primates. Work in such models has described effects on many organ systems in association with changes in DNA methylation and histone modifications at genes important in development and function (reviewed in 27).

Epidemiological data and animal studies have also shown that programmed effects are not limited to the first-generation offspring (i.e. those directly exposed to an insult), but can be transmitted to subsequent generations through non-genomic mechanisms. Reports describing alterations in DNA methylation, histone modifications and/or ncRNAs in the germline suggest a role for ‘transgenerational epigenetic inheritance’ in the transmission of disease risk. However, while epigenetic inheritance does occur in plants and in the nematode worm, C. elegans, the evidence for transgenerational epigenetic inheritance in mammals remains limited. Further studies are necessary to understand whether (and how) induced alterations in the germline epigenome escape the extensive epigenetic reprogramming which occurs in the germline and post-fertilisation embryo and to understand the mechanisms by which alterations in germ cells lead to the reproduction of the programmed effects in subsequent generations.

Interpreting epigenetic studies in humans
The specific disorders discussed at the start of the review (e.g. imprinting disorders, Rett syndrome and cancers) involve large changes in the epigenome (e.g. complete loss or gain of DNA methylation), generally occur as a consequence of genetic abnormalities, and are supported by detailed mechanistic studies. In contrast, most clinical studies of relevance to paediatrics, particularly those in the DOHaD field (and the associated animal studies), are largely descriptive and usually report very small changes in DNA methylation. Although these studies are often cited as suggesting a causal link between changes in DNA methylation and disease, in reality many of the findings are of uncertain significance. In order to prove causality, i.e. to show that epigenetic differences (i) are a consequence of the insult (ii) affect gene regulation in the tissue of interest and (iii) impact on disease risk, more detailed functional studies are necessary, and these are generally lacking. In this section we review some of the other common problems with population-level epigenetic research and discuss factors to consider when interpreting DNA methylation studies in particular.

Study design
In observational studies, recruitment bias can mean that even with carefully chosen control populations and correction for confounders, the influence of additional socio-economic/environmental factors is missed. Many epigenome-wide association studies (EWAS) in human populations are limited by their small sample size – indeed, given the normal variability in DNA methylation between
individuals, many (perhaps most) are underpowered. Additionally, studies are frequently confounded by multiple comparisons, making it likely that erroneous inferences are made.

**Which samples to profile?**

In human studies, particularly in children, DNA may only be available from accessible tissues e.g. placenta, buccal cells or blood. However, since DNA methylation and histone modification patterns are highly tissue-specific, methylation/histone marks in one tissue are unlikely to be a good proxy for those present in other, more inaccessible tissues. Some studies have attempted to address this e.g. by profiling DNA methylation in both blood and post mortem tissue and have shown that although there are some similarities, there are also many differences between tissues, suggesting that blood-based profiling may provide only limited information about underlying pathology elsewhere. Nevertheless, the use of accessible tissues may still be useful for biomarker identification, although this requires the development of robust algorithms, since many of the changes described are small, and replication in other cohorts is essential.

A key confounder in many studies is that tissues and blood are comprised of a mix of different cell types, and alterations in the proportions of different cells (which will have different DNA methylation profiles) will affect the percentage DNA methylation found in any tissue or in blood. In other words, measurable differences in DNA methylation may simply reflect differences in cell/tissue composition (Figure 2). Lifestyle choices or the development of disease may impact on cell-type distribution - for example, changes in blood cell composition occur with smoking or diabetes. Most population-based DNA methylation studies report small differences in DNA methylation (typically ~5%) and rarely assess whether this represents a change in DNA methylation in a single cell type, or differences in the cellular population. Further, differences in cell-subtype populations between individuals may be a major confounder in many studies.

**Genetic differences are important**

Up to 80% of inter-individual differences in DNA methylation can be attributed to DNA sequence polymorphisms, i.e. underlying genetic differences result in inter-individual differences in DNA methylation profiles. Where such genetic differences are also important in disease risk, this could result in the assumption that the disease occurred as a result of changes in DNA methylation, rather than as a result of differences in the DNA sequence itself (Figure 3). To make it more complicated, DNA methylation differences may also be a consequence of inter-individual differences in gene expression. To fully address both of these issues, additional analyses require profiling of both the genome and gene expression.

**Cause or consequence?**
It is also important to consider the phenomenon of ‘reverse causation’, where the disease itself leads to changes in DNA methylation rather than vice versa, such that the changes in the epigenome are a consequence rather than a cause of disease (Figure 3). This is particularly the case for cross-sectional studies, for example although studies report differences in DNA methylation at the HIF3A gene in obese individuals, a recent analysis suggests that rather than there being a causal effect of HIF3A methylation on body mass index, instead, obesity affects HIF3A methylation.\textsuperscript{33} Reverse causation is impossible to exclude in studies only performed in adults exposed to an adverse early life environment, including the Dutch Famine or Motherwell cohorts.\textsuperscript{19, 20} Longitudinal studies are therefore important to understand the chronology of changes and assist in the identification of biomarkers of risk.

Summary

In conclusion, it is clear that dysregulation of the epigenome occurs in a number of paediatric diseases. Expanding our knowledge of these disease processes in turn contributes to a greater understanding of normal growth and development. It has long been recognised that early life events have implications for future disease risk, and although studies suggest this may at least in part be mediated by epigenetic modifications; further high-quality research is needed to better understand the underlying mechanisms. A working knowledge of epigenetics will allow paediatricians to interpret such research critically, embracing the potential advances brought about through epigenetic studies, whilst appreciating their limitations.

Most importantly, building a strong foundation in epigenetic research will allow paediatricians to take part in the conversations that will undoubtedly become part of our future clinical practice, as we begin to apply this knowledge to risk stratify patients and move towards more personalised medicine. As an increasing number of epigenetic therapies are developed, it will be essential for paediatricians to understand their mechanisms and their role in treating disease, in order to select the most appropriate therapies for each individual patient.
Figure Legends

**Figure 1: Imprinted regions of chromosome 11p15.5.** (P=paternal allele, M=maternal allele. Red = silenced gene. Green = active gene. Filled hexagon = methylated DNA, unfilled hexagon = unmethylated DNA). **A)** Normal methylation pattern - ICR1 is methylated on the paternal allele and unmethylated on the maternal allele, resulting in paternal expression of IGF2 and maternal expression of H19. ICR2 is methylated on the maternal allele and unmethylated on the paternal allele resulting in maternal expression of CDKN1C. **B & C)** Two causes of BWS – B) Loss of methylation at ICR2 on the maternal allele resulting in reduced expression of CDKN1C. C) Gain of methylation at ICR1 on the maternal allele resulting in bi-allelic expression of IGF2. **D)** One cause of SRS – Loss of methylation at ICR1 on the paternal allele resulting in bi-allelic expression of H19.

**Figure 2: Interpreting changes in DNA methylation in mixed cell populations.** Tissues and blood are comprised of a mix of different cell types and individual cell types have different methylation profiles. **A)** Schematic diagram of a tissue comprising a mixed cell population (four cell types indicated by circles, squares, triangles and pentagon). One specific CpG site is normally methylated in the square cells (filled) whereas it is unmethylated in the other cell types (open), giving a % DNA methylation of 25% when the whole tissue is profiled. **B)** Following exposure to an ‘insult’, the CpG also becomes methylated in a proportion of the circular cells such that profiling gives a % methylation of 40% at that locus. **C)** The same result is also seen when rather than changing DNA methylation within a cell type, exposure to an insult instead results in a change in cellular population.

**Figure 3: Interpreting changes in DNA methylation cause versus consequence.** **A)** Although the results of many studies have been used to suggest that changes in DNA methylation cause disease, there have been very few mechanistic studies to support this. **B)** Spurious associations can arise when a third factor leads to both a change in DNA methylation and the disease. Up to 80% of inter-individual differences in DNA methylation can be attributed to DNA sequence polymorphisms i.e. underlying genetic differences result in inter-individual differences in DNA methylation profiles. Where such genetic differences are also important in disease risk, this could result in the assumption that the disease occurred as a result of changes in DNA methylation, rather than as a result of differences in the DNA sequence itself. **C)** Disease states may lead to changes in DNA methylation rather than vice versa - a phenomenon termed reverse causation.
References


Figure 3

Causal association: CpG methylation → Disease

Spurious association:
- CpG methylation
- Disease

Common cause

Reverse causation:
- CpG methylation
- Disease