

A Meeting of the Nutrition Society, hosted by the Irish Section, was held at the O'Reilly Hall, University College Dublin, Dublin, Republic of Ireland on 18–20 June 2008

## Symposium on 'The challenge of translating nutrition research into public health nutrition'

### Session 2: Personalised nutrition Epigenomics: a basis for understanding individual differences?

John C. Mathers

Human Nutrition Research Centre, School of Clinical Medical Sciences, Institute for Ageing and Health, Newcastle University, Framlington Place, Newcastle upon Tyne NE2 4HH, UK

Epigenetics encompasses changes to marks on the genome that are copied from one cell generation to the next, which may alter gene expression but which do not involve changes in the primary DNA sequence. These marks include DNA methylation (methylation of cytosines within CpG dinucleotides) and post-translational modifications (acetylation, methylation, phosphorylation and ubiquitination) of the histone tails protruding from nucleosome cores. The sum of genome-wide epigenetic patterns is known as the epigenome. It is hypothesised that altered epigenetic marking is a means through which evidence of environmental exposures (including nutritional status and dietary exposure) is received and recorded by the genome. At least some of these epigenetic marks are remembered through multiple cell generations and their effects may be revealed in altered gene expression and cell function. Altered epigenetic marking allows plasticity of phenotype in a fixed genotype. Despite their identical genotypes, monozygotic twins show increasing epigenetic diversity with age and with divergent lifestyles. Differences in epigenetic markings may explain some inter-individual variation in disease risk and in response to nutritional interventions.

#### Epigenetic marks: DNA methylation: Histone modifications: Cellular memory: Inter-individual differences

The basis for each individual's opportunity for good health or risk of common complex disease, e.g. CVD or type 2 diabetes, is established at conception by the consortium of genetic variants acquired from each parent. However, many environmental factors acting throughout development *in utero* and for the duration of postnatal life influence the expression of those genetic variants as altered phenotype. For example, excellent epidemiological research and studies using animal models have established that being born small, especially if followed by faster than expected growth (early adiposity rebound), is a significant risk factor for several chronic diseases. These ideas are encapsulated in the developmental origins of health and disease hypothesis<sup>(1)</sup>. 'High-risk' lifestyles that are characterised by poor dietary choices, low levels of physical activity and exposure to environmental pollutants,

including those in cigarette smoke, when sustained over long time periods, contribute to premature ageing and increase the likelihood that an individual will succumb to CVD, type 2 diabetes or other age-related and obesity-related diseases. These influences on individual susceptibility have been conceptualised in the 'health pendulum'<sup>(2)</sup>.

These observations suggest a number of key questions including:

1. how is dietary exposure 'remembered' over long time periods. In the case of increased risk of type 2 diabetes in those poorly nourished *in utero*, the effects of the nutritional insult may not be apparent for decades so the 'memory' of the insult has to be retained for a substantial part of the lifespan;

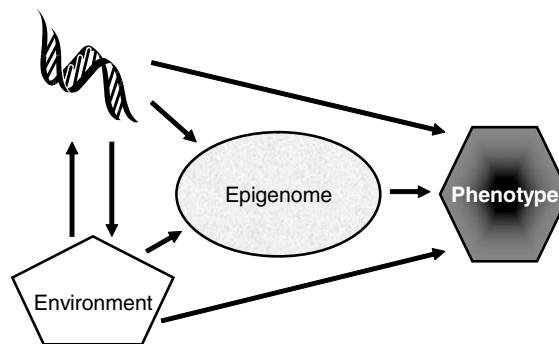
Corresponding author: Professor John C. Mathers, fax +44 191 2228943, email john.mathers@ncl.ac.uk

2. how are the effects of dietary exposure 'revealed' as altered phenotype. At its most fundamental, this process is likely to involve altered gene expression;
3. what is the mechanistic (molecular) basis for individual differences in response to diet.

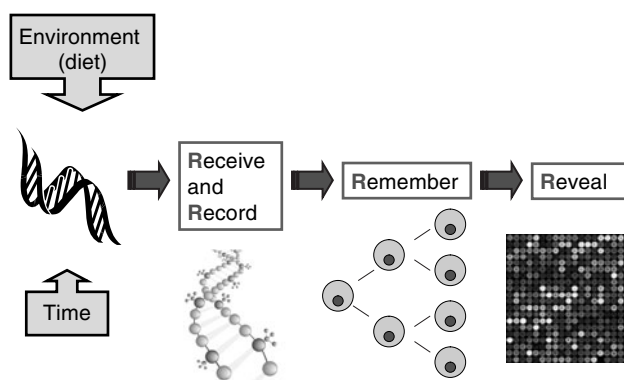
### Epigenetics

Epigenetics is the science of chromatin modifications resulting in altered regulation of gene expression without involving changes in the primary DNA sequence. Chromatin is much more than just a clever means of packaging a large amount of DNA into the cell nucleus. It seems probable that this 'smart packaging' is a highly-sophisticated information store that regulates gene expression in response to environmental signals. The best understood epigenetic marks are (a) patterns of methylation of cytosine bases in DNA and (b) post-translational modifications of the histone tails that protrude from the globular core of eight proteins around which DNA is wrapped<sup>(3)</sup>. More than 100 modifications of histones including acetylation, methylation, phosphorylation and ubiquitination have been described<sup>(4)</sup>. Although the functional consequences of most of these 'marks' remain obscure, it is possible that individually and collectively they constitute a sophisticated code (the histone code<sup>(5)</sup>) that regulates which genes are transcribed in which cells under specific circumstances. Throughout the genome most cytosine residues are methylated at position 5 in the cytosine ring. However, in certain parts of the genome, notably in the 5' region upstream of genes, there are unusually dense collections of CpG dinucleotides known as CpG islands in which the cytosines are unmethylated. Indeed, a recent detailed study of human chromosomes 6, 20 and 22 has shown that there is little cytosine methylation near transcription start sites<sup>(6)</sup>. About half human genes have CpG islands in their promoter regions and, for genes that are being transcribed, these islands are normally unmethylated. In association with chromatin modifications, promoter methylation prevents access by the transcription machinery to specific DNA domains and results in gene silencing. In effect, the epigenetic marks act as a switch turning on or off expression of specific genes. Cell differentiation, X chromosome inactivation and genetic imprinting are all consequences of epigenetic processes. The totality of epigenetic marks in a given cell is described as the epigenome.

Many plants flower only after exposure to an extended cold period (winter), which is the process of vernalisation. The suppression of sexual reproduction is caused by the protein flowering locus C, which blocks flowering by repressing those genes needed for flower development. During vernalisation, expression of the flowering locus C gene is down regulated and it is now known that this down-regulation is a result of epigenetic silencing<sup>(7)</sup>. This process is a clear example in which the environment (in this case low temperature) is a major determinant of phenotype (flowering) by regulating gene expression through an epigenetic mechanism. A second (more nutritional) example of the role of epigenetics in linking



**Fig. 1.** Role of the epigenome in orchestrating the complex interplay between the environment and the genome that determines an individual's phenotype. (Adapted from Zoghbi & Beaudet<sup>(9)</sup>.)



**Fig. 2.** The four 'Rs' of epigenomics. A conceptual model of the key processes through which altered epigenomic markings as a result of environmental (dietary) exposures are 'Received', 'Recorded', 'Remembered' and 'Revealed'. (For more details, see Mathers & McKay<sup>(11)</sup>.)

environmental exposure with altered phenotype comes from the honeybee (*Apis mellifera*). To create a few queens among a vast number of workers, female larvae are fed the poorly-characterised 'royal jelly', which results in the production of larger adult bees with a very different physiology and lifespan (2 years v. a few weeks) than their worker siblings<sup>(8)</sup>. Very recently, it has been shown that the different nutritional input to the larval queens results in a profoundly different developmental fate because of changes in epigenetic markings (altered DNA methylation<sup>(8)</sup>). More generally, evidence is accumulating to support the idea that the epigenome plays a key role in orchestrating the complex interplay between the environment and the genome that determines an individual's phenotype (Fig. 1)<sup>(9)</sup>. These ideas form the basis for the hypotheses that epigenetic markings (a) enable plasticity of phenotype in a fixed genotype and (b) connect environmental exposure with gene expression and function<sup>(10)</sup>. A simple model described as the '4 Rs of epigenomics' has been developed to articulate the key processes underlying these hypotheses (Fig. 2)<sup>(11)</sup>. This model proposes that environmental exposures (notably diet) and, of course, time impact on the genome and that at least some of these impacts are 'Received' and 'Recorded' by epigenetic

mechanisms. If such information is to be important in the medium to long term, then it must be 'Remembered' through subsequent cell divisions. For DNA methylation, this process of 'memorisation' is well understood since DNA methylation marks on the parental strand are copied to the daughter strand by the housekeeping DNA methyltransferase 1. The memorisation of histone modifications is much less well understood, but it is apparent that histones segregate randomly during cell division so that each daughter cell acquires some of the marked proteins, which then spread to the newly-deposited histones<sup>(12)</sup>. Finally, the consequences of earlier environmental exposures may be 'Revealed' as altered gene expression leading to changes in cellular function with potential to modify health.

### Nutritional influences on epigenetic marking, gene expression and phenotype in mammals

The viable yellow agouti ( $A^{vy}$ ) mouse harbours an intracisternal A-particle transposon upstream from the *Agouti* locus, which acts as an alternative promoter for this gene. When this alternative promoter is active it results in mice with yellow coats and a propensity to obesity and obesity-related diseases<sup>(13)</sup>. However, even among inborn strains, not all littermates have the same phenotype; some are born with greater or lesser amounts of brown fur and the darker-coloured mice are leaner. The latter phenotype is associated with methylation of cytosine residues in the intracisternal A-particle and provides a striking example of how phenotype can vary markedly as a result of epigenetic variability. In this mouse model altering maternal nutrition during pregnancy also affects epigenetic markings and phenotype. Higher intakes of methyl donors (folic acid, vitamin B<sub>12</sub>, choline and betaine) by the dams increase the proportion of mottled and pseudo-agouti offspring, which is accompanied by greater methylation of the cryptic promoter in the proximal end of the  $A^{vy}$  intracisternal A-particle<sup>(14)</sup>. Such an effect might be anticipated if higher methyl donor supply facilitates more methylation of this locus. However, such a mechanism does not explain the remarkably similar effects on both phenotype and DNA methylation of including genistein (250 mg/kg diet) in the diet of the pregnant dams<sup>(15)</sup>. Since genistein is not a methyl donor, an alternative mechanistic explanation is required for the epigenetic silencing of the *Agouti* cryptic promoter and the subsequent changes in coat colour and reduced adiposity<sup>(15)</sup>. Recent studies in sheep have shown that limiting folate and vitamin B<sub>12</sub> supply in the immediate peri-conceptual period only can result in fatter adult offspring with the expected higher blood pressure and greater insulin resistance<sup>(16)</sup>. Using restriction landmark genomic screening, these authors have subsequently shown that the observed phenotypic changes are associated with altered patterns of methylation at many genomic loci.

Environmental modulation of epigenetic marking producing altered phenotype is not restricted to development *in utero*. The adult offspring of rat dams that exhibit 'better' maternal behaviour (more licking and grooming of pups and arched-back nursing) during the first week of postnatal life show less fearfulness and have more modest

hypothalamic–pituitary–adrenal responses to stress<sup>(17)</sup>. Mechanistic studies have revealed that the better maternal behaviour is associated with demethylation of specific CpG sites in the glucocorticoid receptor gene within the hypothalamus affecting transcription factor binding to the glucocorticoid receptor promoter<sup>(18)</sup>. The Canadian team undertaking these studies have speculated that other environmental exposures during the early postnatal period might have profound effects on behaviour in later life through interfering with epigenetic programme driven by maternal care<sup>(19)</sup>. There are several other examples of dietary factors influencing DNA methylation but in most cases the consequences, if any, for phenotype have not been established clearly.

### Inter- and intra-individual variation in epigenetic markings

Monozygotic twins are a good model for studies of effects of environment *v.* genetics on phenotype since the members of such pairs are genetically identical. A groundbreaking study of monozygotic twins has reported that young twins are much more alike epigenetically than are older monozygotic pairs<sup>(20)</sup>. Not only are the older twin pairs epigenetically more distinct but they also have less-similar gene expression profiles<sup>(20)</sup>. In addition to the likely stochastic effects of age in modulating epigenetic marks<sup>(21)</sup>, there is *a priori* evidence of the impact of (unknown) environmental factors, since those adult twins who have lived further apart for substantial periods of their lives have more dissimilar epigenetic markings<sup>(20)</sup>. This study provides evidence of altered DNA methylation patterns associated with altered gene expression despite identical genotypes. It provides a potential mechanistic explanation for divergent phenotype (e.g. disease experience) of monozygotic pairs. Further, it provides encouragement to investigate the role of nutrition as an environmental modifier of individual epigenomic patterns.

Much of the current data on inter-individual variation in DNA methylation profiles is restricted to analyses of nucleated blood cells, but it is apparent that there are also tissue differences in methylation patterns that may be associated with cellular differentiation<sup>(22)</sup>. Indeed, studies of CpG islands in the *CYP21A2* and *TNF* genes have revealed complex inter- and intra-individual variation in patterns of methylation of cytosine residues<sup>(22)</sup>.

To a large extent, the determinants of differential DNA methylation patterns are unknown but it would be reasonable to suppose that both genetic and environmental factors will contribute, with the balance between these two types of factor differing in particular circumstances. Detailed analysis of the extent of methylation of selected domains within the maternally-imprinted gene encoding the insulin-like growth factor II has shown inter-individual variation among human twins<sup>(23)</sup>. However, the estimated heritability of methylation is high (ranging from 57% to 97% for individual CpG sites) and there is no evidence that heritability is less in middle-aged twins than in adolescent twins<sup>(23)</sup>. From these observations it appears that for methylation of the insulin-like growth factor II gene (at the

genomic sites investigated), genetic factors rather than age or environmental exposures are the dominant explanatory factors.

In seeking genetic explanations for inter-individual differences in DNA methylation patterns the genes encoding DNA methyltransferases are an obvious target. There are at least five members of this gene family<sup>(24)</sup>, with DNA methyltransferase 1 being the housekeeping enzyme responsible for maintaining DNA methylation patterns through cell division whilst DNA methyltransferase 3A and 3B play roles in *de novo* methylation. Several polymorphisms in DNA methyltransferases are known<sup>(25)</sup>, which will form the basis for genotype–epigenotype (and phenotype) association studies.

### Anticipated developments

Epigenetics has been described by the National Institutes of Health (Bethesda, MD, USA) as an emerging frontier of science<sup>(26)</sup> and there is ample evidence of burgeoning interest in the field in many biological sciences and especially research on the biological determinants of health. This aspect is best illustrated by the moves to initiate the Human Epigenome Project, which aims to identify, catalogue and interpret genome-wide methylation patterns of all human genes in all major tissues<sup>(27)</sup>. Initial work on bisulfite DNA sequencing of human chromosomes 6, 20 and 22 has provided approximately  $2 \times 10^6$  CpG methylation values from twelve tissues<sup>(6)</sup>. Over one-sixth of the 873 genes analysed show differential methylation in the 5' untranslated region and in one-third of cases, these differentially methylated 5' untranslated regions are inversely correlated with transcription<sup>(6)</sup>, which is what would be expected if methylation is acting as a transcription switch.

This growing research impetus to undertake the Human Epigenome Project has been stimulated by two major factors. The first is the realisation that, despite the explosion of genomic information emerging from the genome mapping projects, these data are of limited utility without information on what genes are expressed in given circumstances. This need drives research to understand the complexities of chromatin structure and how epigenetic information regulates gene expression. The second major factor is technological developments. Until recently, most of the techniques for studying DNA methylation have provided information on either the extent of methylation of the whole-cell DNA (without reference to genomic location) or targeted information on single genes. The earliest approach to genome-wide profiling used restriction landmark genomic scanning, which involves two-dimensional electrophoresis to separate genomic fragments after digestion with methylation-sensitive endonucleases. This approach has been successful in identifying differentially-methylated sequences in sheep exposed to a nutritional intervention *peri-conceptually*<sup>(16)</sup>. However, restriction landmark genomic scanning is labour intensive and is associated with considerable bioinformatics problems in identifying the differentially-methylated sequences. Rapid developments in oligonucleotide tiling arrays to provide

increasingly comprehensive coverage of the genome at more modest cost make this approach the current state-of-the-art. However, the dominance of array-based methodologies is being challenged by the availability of the next-generation sequencing platforms, which provide astonishingly-high throughput at lower cost. Already, massively parallel bisulfite pyrosequencing has been used to identify tumour-associated methylated DNA loci in blood samples, which may be useful as novel biomarkers of breast cancer<sup>(28)</sup>. It has been predicted that the first sequencing-based mammalian methylome (whole-genome DNA methylation profile) will be reported in 2008<sup>(29)</sup>. Looking a little further ahead (perhaps 2015), it is possible that quite different technologies such as positron emission tomography scanning will be sufficiently powerful and sensitive to provide non-invasive whole-body imaging of the methylome in specific tissues<sup>(29)</sup>.

From the nutrition researcher's perspective, these developments are exciting and will provide the tools with which to discover the impact of nutritional status and/or dietary exposures on the genome and its implications for health. In addition, if, as seems likely, epigenomic marks record a wide variety of environmental exposures there is an opportunity to explore such marks for evidence of long-term dietary choices. Such research could be the basis for the development of radically-new tools to characterise (and quantify) nutritional intake over long time periods that avoid the substantial limitations of current memory and recording-based approaches.

### Conclusions

It is now established that epigenetic processes allow plasticity of phenotype in a fixed genotype and that altered epigenetic markings are implicated in the aetiology of several diseases and may be causal for some<sup>(10)</sup>. Although there is proof of principle that dietary factors influence epigenetic markings and gene expression and through such mechanisms may modify health, research in nutritional epigenomics is just beginning and there are many unanswered questions. These questions include:

1. which genes, contributing to altered diseases susceptibility, are epigenetically deregulated by dietary and other environmental exposures;
2. which environmental (nutritional) factors, in what doses and over what time scales, affect epigenomic markings;
3. in what periods of the life course (or other circumstances) is the genome especially vulnerable to altered epigenetic markings by dietary factors and are such changes sustained over long time periods;
4. can nutritional interventions modulate the adverse effects of other environmental exposures, e.g. xenobiotics, or of ageing;
5. can epigenomic marks form the basis for novel diet-related biomarkers of disease risk<sup>(30)</sup>.

Epigenetic heterogeneity among individuals is now established as a functionally-important mechanism responsible for individual phenotypic differences<sup>(27)</sup> and looks likely to

be a very fruitful area for research aimed at understanding inter-individual differences in responses to diet.

### Acknowledgements

Nutritional epigenomics research in my laboratory is funded by the Biotechnology and Biological Sciences Research Council and the Engineering and Physical Sciences Research Council through the Centre for Integrated Systems Biology of Ageing and Nutrition (BB/C008200/1) and by 'The European Nutrigenomics Organisation; linking genomics, nutrition and health research' (NuGO; CT-2004-505944); which is a Network of Excellence funded by the EC's Research Directorate General under Priority Thematic Area 5, Food Quality and Safety Priority, of the Sixth Framework Programme for Research and Technological Development.

### References

1. Gluckman PD & Hanson MA (2004) Living with the past: Evolution, development, and patterns of disease. *Science* **305**, 1733–1736.
2. Mathers JC (2002) Pulses and carcinogenesis: Potential for the prevention of colon, breast and other cancers. *Br J Nutr* **88**, Suppl. 3, S273–S279.
3. Bernstein BE, Meissner A & Lander ES (2007) The mammalian epigenome. *Cell* **128**, 669–681.
4. Kouzarides T (2007) Chromatin modifications and their function. *Cell* **128**, 693–705.
5. Jenuwein T & Allis CD (2001) Translating the histone code. *Science* **293**, 1074–1080.
6. Eckhardt F, Lewin CD, Cortese R *et al.* (2006) DNA methylation profiling of human chromosomes 6, 20 and 22. *Nat Genet* **38**, 1378–1385.
7. Bastow R, Mylne JS, Lister C, Lippman Z, Martienssen RA & Dean C (2004) Vernalization requires epigenetic silencing of *FLC* by histone methylation. *Nature* **427**, 164–167.
8. Kucharski R, Maleszka J, Foret S & Maleszka R (2008) Nutritional control of reproductive status in honeybees via DNA methylation. *Science* **319**, 1827–1830.
9. Zoghbi H & Beaudet A (2007) Epigenetics and human disease. In *Epigenetics*, pp. 435–456 [C Allis, T Jenuwein and D Reinberg, editors]. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.
10. Feinberg AP (2007) Phenotypic plasticity and the epigenetics of human disease. *Nature* **447**, 433–440.
11. Mathers JC & McKay JA (2008) Epigenetics – Potential contribution to fetal programming. *Adv Exp Med Biol* (In the Press).
12. Hatchwell E & JM Grealley (2007) The potential role of epigenomic dysregulation in complex human disease. *Trends Genet* **23**, 588–595.
13. Morgan HD, Sutherland HGE, Martin DIK & Whitelaw E (1999) Epigenetic inheritance at the agouti locus in the mouse. *Nat Genet* **23**, 314–318.
14. Waterland RA & Jirtle RL (2003) Transposable elements: Targets for early nutritional effects on epigenetic gene regulation. *Mol Cell Biol* **23**, 5293–5300.
15. Dolinoy DC, Weidman JR, Waterland RA & Jirtle RL (2006) Maternal genistein alters coat color and protects Avy mouse offspring from obesity by modifying the fetal epigenome. *Environ Health Perspect* **114**, 567–572.
16. Sinclair KD, Allegrucci C, Singh R *et al.* (2007) DNA methylation, insulin resistance, and blood pressure in offspring determined by maternal periconceptional B vitamin and methionine status. *Proc Natl Acad Sci USA* **104**, 19351–19356.
17. Liu D, Diorio J, Tannenbaum B, Caldji C, Francis D, Freedman A, Sharma S, Pearson D, Plotsky PM & Meaney MJ (1997) Maternal care, hippocampal glucocorticoid receptors, and hypothalamic-pituitary-adrenal responses to stress. *Science* **277**, 1659–1662.
18. Weaver ICG, Cervoni N, Champagne FA, D'Alessio AC, Sharma S, Seckl JR, Dymov S, Szyf M & Meaney MJ (2004) Epigenetic programming by maternal behavior. *Nat Neurosci* **7**, 847–854.
19. Szyf M, I Weaver & M Meaney (2007) Maternal care, the epigenome and phenotypic differences in behavior. *Reprod Toxicol* **24**, 9–19.
20. Fraga MF, E Ballestar, Paz MF *et al.* (2005) Epigenetic differences arise during the lifetime of monozygotic twins. *Proc Natl Acad Sci USA* **102**, 10604–10609.
21. Martin GM (2005) Epigenetic drift in aging identical twins. *Proc Natl Acad Sci USA* **102**, 10413–10414.
22. Rakan VK, Hildmann T, Novik KL *et al.* (2004) DNA methylation profiling of the human major histocompatibility complex: A pilot study for the Human Epigenome Project. *PLoS Biol* **2**, e405.
23. Heijmans BT, Kremer D, Tobi EW, Boomsma DI & Slagboom PE (2007) Heritable rather than age-related environmental and stochastic factors dominate variation in DNA methylation of the human IGF2/H19 locus. *Hum Mol Genet* **16**, 547–554.
24. Bestor TH (2000) The DNA methyltransferases of mammals. *Hum Mol Genet* **9**, 2395–2402.
25. Cebrian A, Pharoah PD, Ahmed S *et al.* (2006) Genetic variants in epigenetic genes and breast cancer risk. *Carcinogenesis* **27**, 1661–1669.
26. National Institutes of Health (2008) NIH roadmap for medical research. <http://nihroadmap.nih.gov/epigenomics/>
27. Brena RM, Huang THM & Plass C (2006) Toward a human epigenome. *Nat Genet* **38**, 1359–1360.
28. Korshunova Y, Maloney RK, Lakey N *et al.* (2008) Massively parallel bisulphite pyrosequencing reveals the molecular complexity of breast cancer-associated cytosine-methylation patterns obtained from tissue and serum DNA. *Genome Res* **18**, 19–29.
29. Beck S & Rakan VK (2008) The methylome: approaches for global DNA methylation profiling. *Trends Genet* **24**, 231–237.
30. Jirtle RL & Skinner MK (2007) Environmental epigenomics and disease susceptibility. *Nat Rev Genet* **8**, 253–262.