Conference on ‘New technology in nutrition research and practice’
Symposium 4: Understanding molecular mechanisms

The integration of epigenetics and genetics in nutrition research for CVD risk factors

Yiyi Ma1,2 and Jose M. Ordovas2,3,4*

1Department of Medicine, Biomedical Genetics, Boston University, Boston, MA, USA
2Jean Mayer USDA Human Nutrition Research Center on Aging at Tufts University, Boston, MA, USA
3Department of Cardiovascular Epidemiology and Population Genetics, Centro Nacional de Investigaciones Cardiovasculares, Madrid, Spain
4IMDEA Alimentación, Madrid, Spain

There is increasing evidence documenting gene-by-environment (G × E) interactions for CVD related traits. However, the underlying mechanisms are still unclear. DNA methylation may represent one of such potential mechanisms. The objective of this review paper is to summarise the current evidence supporting the interplay among DNA methylation, genetic variants, and environmental factors, specifically (1) the association between SNP and DNA methylation; (2) the role that DNA methylation plays in G × E interactions. The current evidence supports the notion that genotype-dependent methylation may account, in part, for the mechanisms underlying observed G × E interactions in loci such as APOE, IL6 and ATP-binding cassette A1. However, these findings should be validated using intervention studies with high level of scientific evidence. The ultimate goal is to apply the knowledge and the technology generated by this research towards genetically based strategies for the development of personalised nutrition and medicine.

CVD is the leading cause of total global mortality. The WHO estimates that 17.3 million people died from CVD in 2008, representing 30 % of all global deaths. In the USA alone, the overall rate of death attributed to CVD was 235.5 per 100 000 based on 2010 data(1). Moreover, CVD are projected to remain the single leading cause of death worldwide.

With the goal to prevent and cure CVD, numerous risk factors have been identified, including dyslipidaemia, inflammation, obesity, hypertension, smoking, age and diabetes(2). Dyslipidaemia refers to high concentrations of TAG, total cholesterol and LDL cholesterol and low concentration of HDL cholesterol. According to the American Heart Association(3), the prevalence of adults having high TAG (>150 mg/dl), high total cholesterol (≥200 mg/dl), high LDL-cholesterol (≥130 mg/dl), and low HDL-cholesterol (≤40 mg/dl) is 33, 44·4, 31·9 and 18·9 %, respectively. Also, inflammation is part of the complex biological response to harmful stimuli, which is common to a number of chronic diseases. However, for the most part, both dyslipidaemia and inflammation are preventable or reversible by having a healthy lifestyle.

Of the factors that define a healthy lifestyle, diet is one of the most important components. Specifically, dietary fatty acids (FA) are associated with risk factors for CVD. For example, although it is still under debate, unsaturated FA tend to increase HDL-cholesterol(4), reduce TAG(5), and decrease IL-6 (6) compared with SFA. Although MUFA and PUFA differ in the magnitude of these beneficial effects(7), anti-atherosclerosis effect is...
has been demonstrated for n-3 and n-6 PUFA. In addition, individuals exhibit different physiological responses to dietary FA, reflecting, in part, the contributions of genetic variability.

The role of genetic factors in contributing to these inter-individual differences in lipid responses to dietary FA has been widely studied. Our group has found that the association between dietary intake of total fat and plasma HDL-cholesterol was modified by a genetic variant located within the hepatic lipase gene (LIPC) and that the association between dietary PUFA intake and plasma fasting TAG is modified by the genetic variants located within APOA5 gene. The effect of PUFA on HDL may differ according to different genotypes of several genes such as APOA5, APOA1IL6, NF-k-light-chain enhancer of activated B cells, TNF-α.

In addition, there is also a genetically based difference in TAG response to n-3 PUFA. As a result of wide availability of new genetic technologies such as genome-wide association studies and next generation sequencing, an enriched catalogue of common or rare SNP has been formulated. However, the variation explained by all these genetic variants only account for <20%, indicating the existence of other sources of variability, such as epigenetic mechanisms.

Epigenetics has become a research area of intense interest and growth. The definition of epigenetics underwent a series of changes as biological knowledge expanded. In 1940, epigenetics was first defined as ‘... the interactions of genes with their environment which bring the phenotype into being...’ by developmental biologists. In the 1990s, epigenetics was described as the study of changes in gene expression, which were not a result of changes in the DNA sequence. Recently, inspired by genome-wide technologies, a new term epigenomics has been coined, targeting the study of all factors contributing to changes in genome-wide chromatin structure including DNA methylation, histone modification and chromatin remodelling.

DNA methylation is the best-studied epigenetic mechanism and involves the addition of a methyl group directly onto DNA residues such as cytosine and adenine and the C5-methylcytosine modification is the most common in eukaryotes. DNA methylation can occur in different regions of the genome such as repetitive sequences, gene body, promoter-related CpG island and CpG island shore, which are located up to 2 kb upstream of the CpG island. DNA methylation patterns in different regions present different functions. For example, gene silencing is correlated with hypermethylation in promoter regions rather than in the gene body. Also, cancer and ageing are correlated with hypomethylation of repetitive elements, while this is not the case for methylation of specific genes. Considering the different functionalities of DNA methylation in different regions, studies of DNA methylation occurring in specific sites of specific genes could provide more interpretable and meaningful explanations.

Similar to all the other epigenetic mechanisms, DNA methylation may act as a biomarker of the effect of environmental factors on the genome. A wide array of factors have been identified to affect DNA methylation patterns, including aging, dietary FA, malnutrition, dietary protein, methyl-donors, chemical pollutants, sun exposure and smoking. The connection of ageing with DNA methylation was first observed in the candidate tumour suppressor genes, of which the methylation is increased with age, leading to gene silencing. Later, it was reported that the ageing effect on DNA methylation is a prevalent phenomenon across the whole genome based on studies with monozygotic and dizygotic twins, which showed that the variation in DNA methylation increases significantly with age. Also dietary FA were suggested to regulate DNA methylation patterns. The intervention of a high-fat diet was found to increase the DNA methylation of a metabolically related gene, PPARγ, coactivator 1α (PPARG1C1), however, after the intervention was withdrawn, DNA methylation of PPARG1C1 returned back to its baseline level. The methylation of this gene was also affected by palmitic acid and oleic acid. Arachidonic acid and DHA were shown to affect DNA methylation of FA desaturase 2 in mice liver. In addition, EPA was found to have a demethylation effect on the tumour suppressor gene.

Besides the effects of environmental factors, DNA methylation has been associated with different phenotypes. For instance, DNA methylation has been proposed as one mechanism of atherosclerosis. In ApoE knockout mice, DNA methylation changes were shown to precede any histological sign of atherosclerosis. In addition, the same study also found associations between global DNA hypermethylation and dyslipidaemia, characterised by the atherogenic lipoproteins. An in vitro oligonucleotide-binding assay found that a CG-rich 17-nucleotide sequence binds to inflammatory helper T cells, suggesting a relationship between lipoproteins and DNA methylation target sites, CpG dinucleotides. Besides affecting lipid concentrations, DNA methylation is also involved in inflammation. IL-6 is an acute phase protein induced during inflammation that functions as an inducer of differentiation of inflammatory helper T cells. DNA methylation has been identified as one mechanism of transcription regulation of IL6. For example, methylation of the promoter region in IL6 is negatively correlated with gene expression in peripheral blood mononuclear cells and the DNA methyltransferase inhibitor 5-aza-2'-deoxycytidime induces IL6 transcription in cancer cells. This silencing of IL6 expression may be due to the binding of methyl-CpG-binding protein 2 to the hypothetical binding sites in IL6 gene, which is close to its transcription start site.

The significance of this knowledge relates to its eventual translation into public health. The traditional concept of ‘one size fits all’ is limited, and the study of epigenetics will facilitate knowledge to further the development of personalised medical care. In this case, it is necessary to generate a more complete understanding of both genetic and epigenetic mechanisms contributing to the substantial inter-individual variations of response to environmental challenges. Moreover, we will also expand our knowledge of the molecular mechanism of gene-
Epigenetics and DNA methylation

Overview of epigenetics

Epigenetics acts as the cross-talk between the genome and environment, encompassing three major mechanisms: DNA methylation, histone modification, and chromatin remodelling. DNA methylation involves adding a methyl group onto a DNA nucleotide such as cytosine and adenine (63). With respect to histone modifications, a wide array of modifications are introduced to the histone tails, including methylation, acetylation, phosphorylation, ubiquitylation, sumoylation, ADP ribosylation, deamination, proline isomerisation, crotonylation, pimonylation, butyrylation, formylation, hydroxylation and O-GlcNAcylation (61). In terms of the chromatin remodelling mechanism, ATP-dependent enzymes remodel and control chromatin structure and assembly to make it become active or inactive to the extrinsic stimulus (62). These mechanisms play a critical role in development.

DNA methylation and DNA sequence

DNA methylation, CpG dinucleotides, CpG islands. In mammalian cells, most DNA methylation occurs on CpG dinucleotides (63). Regions enriched in CpG dinucleotides are known as CpG islands. However, the definitions of CpG islands have been evolving. In 1987, Gardiner-Garden and Frommer (64) defined CpG island as ‘a stretch of DNA sequence where moving average of % G + C was >50, and the moving average of ratio of the observed to expected CpG was greater than 0.6.’ These calculations are based on a 100 bp window and sliding across the sequence at 1 bp intervals. However, in regions rich in repetitive elements, this definition results in an overestimation of its presence. Therefore, Takai and Jones set up more stringent criteria for CpG islands, including % G + C = 55, ratio of observed to expected CpG greater than 0.65 and sequence length being ≥500 bp. With Takai and Jones’ criteria, a web page service algorithm CpGIS was developed (65). Furthermore, Ponger and Mouchiroud (66) extended their criteria to estimate the transcription start sites associated CpG islands with the algorithm CpGProD. However, both criteria are subjective and computationally inefficient for the analysis of the genome-wide DNA sequences, so a new definition, named CpGcluster (67), was proposed. This algorithm is based on the distance between two consecutive CpG and uses an integer arithmetic algorithm, which makes it fast and computationally efficient compared to previous methods. However, it has low sensitivity. Recently, a new algorithm CpG_MI (68) was developed to take into account more variability of the test such as different locations of CpG dinucleotide among different CpG islands. With the growing availability of the experimental results of DNA methylation, the prediction of DNA methylation based on machine learning approach is possible. More specifically speaking, EpiGRAPH (69) algorithm for prediction of DNA methylation was trained by the wet-lab experiments data, and then this algorithm could be used to predict the methylation probability of another stretch of DNA sequence.

DNA methylation and genetic variants. It has been shown that proximal sequence elements are both necessary and sufficient for regulating DNA methylation (70). Moreover, SNP can regulate DNA methylation (71-77). For example, the C allele of a SNP located within the promoter region of matrix metalloproteinase 1 was shown to have significantly higher DNA methylation status than the corresponding T allele (71). Also, the G allele of one SNP located within the potassium-chloride co-transporter 3 (KCC3, SLC12A6) was found to be methylated at the adjacent C nucleotide (72).

Systematic analyses of the whole human genome have identified an array of such genetic variants having regulatory effects on DNA methylation patterns, indicating that genetic regulation on DNA methylation is prevalent across the whole genome. For instance, a genomic survey using methylation-sensitive SNP analysis based on a 50 and 250 K SNP genotyping platform showed that sixteen SNP-tagged loci were confirmed to have allele-specific DNA methylation events (77). Also, in brain samples, approximately 10% of the CpG sites included in the analysis were found to be affected by the genotypes of the SNP in cis-position, while 0.1% of the analysed CpG sites were regulated by the genotypes of the SNP in the trans-position (74). Furthermore, it was suggested by studies with sixteen human pluripotent and adult cell lines that approximately one-third (23–37%) heterozygous SNP in the human genome may regulate DNA methylation patterns (73), and a big proportion of the observed loci with allele-specific DNA methylation events (38–88%) is dependent on the allele status of CpG-related SNP, a type of SNP with one allele to disrupt and the other allele to create CpG dinucleotides (73). Finally, the effect of genetic variants outweighed the influence of imprinting on DNA methylation, because it was shown that the number of methylation loci affected by genetic variants were way more than those loci influenced by the sex of parent of origin (73) and there is convincing evidence of the interesting interdependence between genetics and epigenetics underlying diversity in the human genome (74).

DNA methylation and gene function

DNA methylation has different genetic functions mostly depending on location. For example, DNA methylation within the promoter region is more likely to regulate gene transcription (70,79) while DNA methylations within the gene body tend to modify the alternative promoters and splicing events (80-84).

DNA methylation within promoter regions and gene transcription. The negative correlation between DNA methylation and gene transcription is common to most genetic regions across the whole genome with rare exceptions (79,85). The first experiment indicating the
transcription-regulatory effects of DNA methylation was conducted by McGhee and Ginder\(^\text{[86]}\). Since then, a large body of evidence has accumulated supporting an inverse correlation between DNA methylation and gene transcription for most genes, including but not limited to house-keeping genes\(^\text{[87]}\), genes located on the inactive X chromosome\(^\text{[88–90]}\), imprinted genes\(^\text{[77,91]}\), tumour suppressor genes or oncogenes\(^\text{[92–94]}\), cellular differentiation and development-related genes\(^\text{[95–98]}\), metabolic genes\(^\text{[89,92–94]}\) and inflammation-related genes\(^\text{[89,92–94]}\). However, in some instances, DNA methylation has been positively correlated with gene expression\(^\text{[106–110]}\). Most of such transcriptional regulation effects were related to DNA methylation within promoter regions\(^\text{[50]}\) by direct blocking the binding of transcriptional activators or indirect recruitment of methyl-binding proteins and co-repressor complexes to facilitate the formation of heterochromatin in a cooperative way\(^\text{[111]}\).

**DNA methylation within gene bodies and alternative promoter and splicing events.** DNA methylation is also found on CpG sites located within gene bodies\(^\text{[82,112–114]}\), suggesting a potential genetic function besides gene transcription. By comparing differential DNA methylation patterns on a genome-wide scale across different tissues (brain, heart, liver and testis) and different developmental stages of mice, approximately 16 % of the identified tissue differential methylation regions or developmental stage differential methylation regions were located within intragenic regions\(^\text{[114]}\). Also, it was found that the majority of methylated CpG sites were located within gene bodies\(^\text{[82,112,113,115]}\). According to analysis with human normal tissues (whole blood, monocyte, granulocyte, skeletal muscle, spleen and brain), 15.4 % of CpG islands located within the gene bodies were found to be methylated, which is higher than the proportion of methylated CpG islands within 5' promoter region (7.8 %) and the whole gene region (10.6 %)\(^\text{[113]}\). Using human brain tissue, Maunakea et al.\(^\text{[82]}\) generated high-resolution methylation maps with dense coverage of 24.7 million of the 28 million CpG sites across the whole genome. They found that 34 % of all intragenic CpG islands were methylated, whereas only 2 % of the CpG islands located within the 5' promoter regions were methylated, so they concluded that ‘DNA methylation may serve a broader role in intragenic compared with 5' promoter CpG islands in the human brain’. Again, the altered DNA methylation in the immune system was shown to occur predominantly at CpG islands within gene bodies based on the analysis with both mouse cells within haematopoietic lineage\(^\text{[112]}\) and human B cells\(^\text{[113]}\).

The methylation within gene bodies may be related to alternative promoters\(^\text{[82]}\) and alternative splicing events\(^\text{[80,81,82,84]}\). Based on methyleome maps of human brain tissues, differentially methylated intragenic CpG islands may act as promoters, and novel transcripts have been found to be initiated from these intragenic promoters, indicating that intragenic methylation functions to regulate cell context-specific alternative promoters in gene bodies\(^\text{[82]}\). With a computational analysis of human chromosome 6, 20 and 22 based on datasets from the Human Epigenome Project and the Human Genome Project, hypermethylated CpG sites were found to be prevalent in alternatively spliced sites, and the frequency of methylation increases in loci harbouring multiple putative exonic splicing enhancers\(^\text{[84]}\). According to the analysis of data from RNA-seq experiments and methylome data with single nucleotide resolution of human cell lines, DNA methylation was found to be enriched in included alternatively spliced exons, and inhibition of DNA methylation lead to aberrant splicing of alternatively spliced exons. Further, they found that the alternative splicing may be because of the alternative definitions of exons via recruitment of methylated CpG site-binding protein 2 to the methylated CpG sites\(^\text{[80]}\). Another potential mechanism for the regulation of DNA methylation on alternative splicing events may be the fact that DNA methylation patterns affect chromatin structure\(^\text{[83]}\). Finally, a DNA methylation related protein, CCCTC-binding factor, was shown to promote alternative splicing events on a genome-wide scale, providing potential links between DNA methylation and alternative splicing events\(^\text{[83]}\).

**DNA methylation and environmental factors**

**DNA methylation and ageing.** DNA methylation is affected by ageing partially because of its intimate relationship with development. DNA methylation patterns change during each stage of development\(^\text{[116]}\). Before implantation, almost all DNA methylation becomes erased except for those imprinting regions. During implantation, the entire genome gets methylated except for the CpG islands. After implantation, pluripotency genes are de novo methylated and tissue-specific genes are demethylated in the cell types for their expression. The correlations between ageing and DNA methylation were also suggested by in vitro studies. For example, compared to immortal cell lines, normal diploid fibroblasts were found to have a dramatic decrease in their 5-methylcytosine contents during their growth in culture\(^\text{[117]}\). Furthermore, the observation that the decrease rate in mouse primary diploid fibroblasts was faster than in hamsters and human subjects and the fact that mouse has the shortest lifespan suggested that the rate of loss of 5-methylcytosine is positively correlated with growth potential. Also, the treatment of human diploid fibroblasts with DNA methylation inhibitors, azacytidine and azadoxycytidine, were shown to inhibit the initial cellular growth\(^\text{[118]}\).

Recently, epidemiological analyses have indicated the potential relationships between ageing and DNA methylation patterns. A cross-sectional study with monozygotic twins\(^\text{[33]}\) found that younger twins have significantly lower levels of 5-methylated cytosines than older twins, and that the variance of DNA methylation of the older twins was significantly greater than that of the younger twins. The observed differences in DNA methylation were consistent with the findings with gene expression by showing that the 50-year-old twins had dramatically different expression profiles while the 3-year-old twins had almost identical ones. The observed discordance of
DNA methylation with age was consistent across different tissues within the analysis, including lymphocytes, epithelial mouth cells, intra-abdominal fat and skeletal muscle biopsies. Later, another study with thirty-four male monozygotic twins with age ranging from 21 to 55 years identified eighty-eight sites located within or near eighty genes of which DNA methylation patterns were significantly correlated with age\(^8\). Three genes from that list of eighty genes were further validated and replicated with the analysis of their correlations with age in a population-based sample of thirty-one males and twenty-nine females with age ranging from 18 to 70 years, which are Edar-associated death domain, target of myb1 (chicken)-like 1 and neuronal pentaxin II. Interestingly, all of these three genes have been reported to be associated with a wide array of age-related phenotypes, such as wound healing\(^{120}\), Parkinson disease\(^{121}\), cancer\(^{122,123}\), and loss of teeth, hair and sweat glands\(^{124}\). Also, a longitudinal study found that DNA methylation differs by age because methylation patterns of candidate genetic loci, such as the dopamine receptor 4 gene, the serotonin transporter gene, and the X-linked monoamine oxidase A gene, were shown to change during the period when these children grew from age 5 to 10 years\(^{125}\).

Finally, changes in DNA methylation patterns have been reported to be associated with a series of age-related diseases. The evidence suggests that global hypomethylation and gene-specific promoter hypermethylation were associated with different types of cancer. It was found that the number of a subpopulation of cells in human colonic mucosa increase with age, and the promoter of estrogen receptor gene in this subpopulation of cells becomes hypermethylated. This age-related hypermethylation of oestrogen receptor was found in all cells in colorectal tumours examined\(^{126}\). Also, age-dependent methylation of oestrogen receptor alpha was associated with prostate cancer\(^{127}\). The hypermethylation of several tumour suppressor genes have been suggested as biomarkers of lung cancer\(^{128}\). Alzheimer's disease was correlated with DNA methylation of CpG sites located near or within the genetic loci reported to harbour genetic susceptible risk variants for Alzheimer's disease\(^{129}\). Compared with the normal retinas, those of patients with age-related macular degeneration were found to have hypermethylation and gene repression of glutathione S-transferase isoform mu5 and glutathione S-transferase isoform mu5\(^{130}\).

**DNA methylation and fatty acids.** FA affect expressions of a wide array of genes by acting as ligands for transcription factors, such as PPAR, the liver X receptors (LXR), retinoid X receptor, hepatocyte nuclear factor 4, sterol regulatory element-binding proteins (SREBP), NF-κ-light-chain enhancer of activated B cells, cyclooxygenase and lipoxigenase\(^{131}\). PPAR and LXR are members of the nuclear hormone receptor superfamily of transcription factors, which bind to specific motifs within the promoters of genes as heterodimers with the retinoid X receptor\(^{132}\). There are three isoforms of PPAR, including PPARα, PPARβ and PPARγ. In general, PPAR bind with both saturated and unsaturated FA with a relatively more potent binding with n-6 and n-3 PUFA and their derivatives to regulate expressions of genes that control lipid and glucose homeostasis and inflammation. Regarding LXR, there are two family members, LXRα and LXRβ. As a sensor of cholesterol in the nucleus, LXR can be activated by increased intracellular cholesterol concentrations. Also, the binding of long-chain FA to LXR\(^{133}\) was shown to regulate expression of genes involved in sterol and FA metabolism\(^{134-137}\) carbohydrate metabolism\(^{138,139}\). Hepatocyte NF 4α is an orphan member of the steroid hormone receptor superfamily and functions by binding with the activated (CoA) form of FA to regulate expression of genes participating in the lipid, lipoprotein\(^{139,140}\) and glucose metabolism\(^{132}\). SREBP have three isoforms, which are SREBP-1a, SREBP-1c and SREBP-2, and all of them are transcription factors playing a critical role in controlling synthesis of FA, TAG and cholesterol\(^{144}\). PUFA were found to lower the mature form of the protein levels of SREBP by raising cellular cholesterol levels or by reducing SREBP mRNA stability and SREBP transcription or by promoting degradation of SREBP protein\(^{144-147}\). Cyclooxygenase and lipoxigenase function to convert n-6 and n-3 PUFA into pro- and anti-inflammatory signalling molecules to regulate activity of transcription factors of inflammation such as NF-κ-light-chain enhancer of activated B cells\(^{149}\).

The effect of FA on DNA methylation was also suggested by a study with mice heterozygous for disruption of cystathionine beta-synthase (Cbs\(^{-/-}\))\(^{150}\), which could be induced to have hyperhomocysteinaemia (HHcy), providing an indirect evidence because of the potential modifications on DNA methylation by homocysteine through its participation in the C1 metabolism. In that study, a dosage of HHcy (normal, mild and moderate) was developed by treating the mice (Cbs\(^{-/+}\)) with diet to induce HHcy (mild) and the mice (Cbs\(^{-/-}\)) with diet to induce HHcy (moderate). The potential relationship between homocysteine and DNA methylation was supported by the significantly inverse correlation between total homocysteine levels and liver methylation capacity, measured by the ratio of S-adenosylmethionine to S-adenosylhomocysteine. Correspondingly, mice with moderate HHcy had higher methylation of candidate CpG sites within the promoter region of FA desaturase 2 in liver, leading to lower gene expression of FA desaturase 2 and lower protein activity of 6(6)-desaturase (encoded by FA desaturase 2) in liver, compared with mice with mild and normal HHcy. Also, mice with moderate HHcy have lowest level of arachidonic acid and DHA in total liver than those mice with mild and normal HHcy.

Direct evidence for the link between FA and DNA methylation were conducted with in vitro and in vivo studies. Incubation of human skeletal muscle cells with 48 h treatment with free FA, such as palmitate and oleate, can increase DNA methylation levels of the promoter region of PPARγ coactivator-1α in primary human skeletal cells, leading to suppression of its gene
expression. Also, in vitro treatment of U937 leukemia cells with EPA was found to decrease methylation of the promoter regions of a myeloid lineage-specific transcription factor CCAAT/enhancer-binding protein, a tumour suppressor gene, resulting in an increased gene expression. One in vivo study with rats found that feeding a diet high in n-3 PUFA, mainly with EPA and DHA could significantly decrease global DNA methylation levels.

A randomised control trial with high-fat overfeeding in young adults with low or normal birth-weight supported a relationship between FA and DNA methylation. Having high-fat overfeeding (+50% energy) for 5 d increased DNA methylation in the promoter region of PPARGC1A, measured in the skeletal muscle cells extracted from healthy young men with low birth-weight. The observed induction of DNA methylation in PPARGC1A was found to be reversible because DNA methylation returned to its baseline level after the high-fat diets were withdrawn. Although DNA methylation of PPARGC1A was not found to have significant correlation with its gene expression, high-fat challenge in the subjects with low birth-weight were shown to induce peripheral insulin resistance and decrease gene expression of PPARGC1A.

DNA methylation and other environmental factors. Besides ageing and dietary FA, DNA methylation patterns are modifiable by several other environmental factors, including global nutrition status, air pollution, weather and smoking. In mice, supplementation of methyl donors during gestation was shown to have a dose–response relationship with the methylation of viable yellow agouti (Ay) locus and browness of coat colour in the offspring. Energy restriction in utero decreased the overall methylation and changes in the methylation patterns of imprinted loci in mice. Similarly in human subjects, those subjects having experienced famine prenatally because of their in utero exposure to the Dutch Hunger Winter were shown to have less DNA methylation of the imprinted gene, insulin-like growth factor 2 (Igf2) during gestation. Moreover, increased concentrations of ozone and components of fine particle mass were associated with hypomethylation of tissue factor (F3), intercellular adhesion molecule 1 and toll-like receptor 2 and hypermethylation of interferon-γ and IL6 and with decreases in global DNA methylation in whole blood. A genome-wide analysis followed by an independent replication study showed that smokers have decreased level of DNA methylation of a single CpG site, which was located at −1099 bp to the transcription start site of IL6, measured in peripheral blood mononuclear cells. In the macrophages from healthy control subjects, lower methylation of the previously identified CpG site was in line with the higher IL6 expression stimulated by lipopolysaccharide. Experiments with electrophoretic mobility shift assay provided potential mechanistic explanation for these associations by identifying the methylation-dependent affinity of protein–DNA interactions. In vitro treatment of 5-aza-2′-deoxycytidine activated IL6 expression in human pancreatic adenocarcinoma cell lines, indicating an important role of DNA methylation at the IL6 genetic locus. Also, chromatin immune-precipitation assays with the same cell lines identified a potential response element to the binding of methyl-CpG-binding protein 2, located from −666 to −426 bp to the transcription start sites, providing potentially a mechanistic explanation for the DNA methylation of IL6. A cross-sectional study with blood leucocyte found that workers living in an industrial area had the lowest, whereas rural and urban residents had the highest and intermediate methylation levels of the second intron of IL6. Another cross-sectional study with leucocytes found that a prudent diet, characterised by a high intake of vegetables and fruit, was associated with DNA methylation levels of the promoter region of IL6. According to the analysis of DNA methylation patterns of IL6 in periodontal tissues, patients with periodontitis were found to have lower methylation and higher gene expression. An in vitro study with cultured human lung cells showed that the DNA methylation levels of promoter regions of a panel of inflammation related genes (IL6, IL1β, IL8 and IL10) were higher in cancer cells than normal ones, and the higher methyllations went along with the lower gene expressions. A study with patients with paediatric obstructive sleep apnoea found that DNA methylation offorkhead box P3 had a significantly positive correlation with serum levels of high-sensitivity C-reactive protein. A cross-sectional study with blood samples from 742 community-dwelling elderly individuals found that hypomethylation of repetitive element LINE-1 was associated with increased levels of serum vascular cell adhesion molecule-1. Finally, a study with samples of leucocytes from 966 African American identified that DNA methyllations of 257 CpG sites within 240 genes contribute to serum levels of C-reactive protein.

DNA methylation and dyslipidaemia. DNA methylation patterns have been related to dyslipidaemia. After stimulation with lipoproteins, the global levels of 5-methylated cytosines within the differentiated human monocyte-macrophage cell line THP-1 were significantly increased. According
to a genome-wide DNA methylation analysis with samples of CD4+ cells from 991 individuals of the Genetics of Lipid-lowering Drugs and Diet Network (GOLDN) study, four CpG sites located within the intron 1 of carotid palmitoyltransferase 1A were found to be associated with fasting levels of VLDL-cholesterol and TAG. DNA methylation of the CpG site with top findings was further found to be associated with carotid palmitoyltransferase 1A expression. The observed association between DNA methylation, gene expression and fasting TAG was replicated in the Framingham Heart Study (160). Also, a higher methylation pattern of the promoter region of ATP-binding cassette A1 (ABCA1) in samples of whole blood was found to be associated with a lower circulating HDL-cholesterol and HDL2-phospholipid levels in ninety-seven patients with familial hypercholesterolaemia (161). Similarly, in patients with familial hypercholesterolaemia, leucocyte DNA methylation of lipoprotein lipase had positive correlations with HDL-cholesterol and HDL particle size, whereas DNA methylation of cholesteryl ester transfer protein had a negative association with LDL-cholesterol in all the participants and negative associations with HDL-cholesterol, HDL-TAG levels, and HDL particle size (162). Further, the methylation of lipoprotein lipase in visceral adipose tissue extracted from thirty men with severe obesity were found to have negative correlations with HDL-cholesterol and gene expression of lipoprotein lipase (162). The potential mechanism for the effects of lipoproteins on DNA methylation is unknown. The modifications of chromatin structure may account as one potential mechanism, because it was found that ApoA1 can physically bind to a CG-rich oligonucleotide in vitro, leading to the remodelling of chromatin structure (56).

Genetics and epigenetics integrate

The integration of genetics and epigenetics require large datasets with deep and comprehensive phenotyping. The proposed research in this subject has been facilitated by our access to such rich resources, specifically the GOLDN study and the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) consortium.

The GOLDN study was designed to identify genetic determinants of lipid response to two interventions (a high-fat meal challenge and fenofibrate treatment for 3 weeks) (163). The study ascertained and recruited families from the Family Heart Study at two centres (Minneapolis, MN and Salt Lake City, UT), who self-reported to be white. Only families with at least two siblings were recruited for a total of 1327 individuals. Volunteers were required to withhold lipid-lowering agents (pharmaceuticals or nutraceuticals) for at least 4 weeks before the initial visit to be eligible. A total of 1053 individuals met all eligibility requirements.

For methylation studies, DNA was extracted from CD4+ T cells harvested from buffy coats with the use of antibody-linked Invitrogen Dynabeads. CD4+ T cells were selected for three reasons. First, DNA methylation patterns are often tissue specific. For instance, studies of whole blood samples reflect methylation variations within each blood cell type that may act to confound epigenomic association results (164). Second, many key genes involved in lipid metabolism are expressed in lymphocytes and other immune cells (e.g. PPAR) (165). In one study, peripheral blood mononuclear cells gene expression profiles were demonstrated to reflect nutrition-related metabolic changes. Responsive genes were enriched for FA-metabolising enzymes, including carotid palmitoyltransferase 1, ACAA2 and SCL25A20 (166). Therefore, this cell type should reflect underlying epigenetic variation influencing blood lipids while minimising potential confounding. Third, blood collection is the most viable tissue collection method among healthy individuals. We used the Illumina Infinium Human Methylation450 Beadchip (Illumina Inc, San Diego, CA) to interrogate approximately 470,000 autosomal CpG sites across the genome.

The CHARGE Consortium was formed to facilitate genome-wide association studies meta-analyses and replication opportunities among multiple large population-based cohort studies, which collect data in a standardised fashion and represent the preferred method for estimating disease incidence. The initial design of CHARGE included five prospective cohort studies from the USA and Europe: the Age, Gene/Environment Susceptibility-Reykjavik study, the Atherosclerosis Risk in Communities study, the Cardiovascular Health study, the Framingham Heart study and the Rotterdam study. With genome-wide data on a total of about 38,000 individuals, these cohort studies have a large number of health-related phenotypes measured in similar ways. For each harmonised trait, within-cohort genome-wide association study analyses are combined by meta-analysis. A prospective meta-analysis of data from all five cohorts, with a properly selected level of genome-wide statistical significance, is a powerful approach to finding genuine phenotypic associations with novel genetic loci (167). Since its creation, CHARGE has incorporated many other cohorts, increasing significantly its sample size and the ability to identify new and relevant associations and interactions.

Genetic variants and methylation levels revisited

Using data from the GOLDN study, we revisited the topic of the local correlation between genetic variants and DNA methylation levels (cis-meQTL) and conducted a cis-meQTL analysis. We found that over 80% of genetic variants at CpG sites (meSNP) are meQTL loci ($P < 10^{-5}$) and meSNP account for over two-thirds of the strongest meQTL signals ($P < 10^{-200}$). Beyond direct effects on the methylation of the meSNP site, the CpG-disrupting allele of meSNP were associated with lowered methylation of CpG sites located within 45 bp. The effect of meSNP extends as far as 10 kb and can contribute to the observed meQTL signals in the surrounding region, likely through correlated methylation patterns and linkage disequilibrium. Therefore, GOLDN supports previous findings showing that
meSNP are behind a large portion of observed meQTL signals and play a crucial role in the biological process linking genetic variation to epigenetic changes.\(^{(168)}\)

**APOE gene variants, methylation and ageing**

Common APOE gene variants are associated with age-related diseases; however, the underlying mechanisms have not been entirely elucidated and DNA methylation may be a significant contributor. To test this possibility, we conducted an integrated analysis with both population (GOLDN study) and in vitro studies (Encyclopedia of DNA elements (ENCODE) consortium) to systematically explore the relationships among age, plasma lipids, DNA methylation patterns, sequence variants and gene expression of APOE\(^{(169)}\). We found that APOE methylation was correlated with gene expression, associated with age, plasma total cholesterol and sequence variants, including both promoter variant rs405509 and well-known APOE \(\varepsilon\) variants. Furthermore, the association between APOE methylation patterns within the promoter region and age were dependent on promoter variant rs405509. These associations suggest that APOE methylation may explain its ageing effects.\(^{(169)}\)

**IL6 gene variants, methylation and dietary n-3**

n-3 PUFA reduce IL6 gene expression, but their effects on transcription regulatory mechanisms are not totally elucidated. As in previous instances, we systematically explore the relationships among n-3 PUFA, DNA methylation, sequence variants, gene expression and protein concentration of IL6 by conducting an integrated analysis of data from population (GOLDN study) and in vitro studies (ENCODE consortium)\(^{(170)}\). As a result, methylation of IL6 promoter CpG site (cg01770232) was positively associated with IL6 plasma concentration. IL6 gene expression and more dosage of the A allele of rs2961298, but negatively associated with circulating total n-3 PUFA. Furthermore, there was significant interaction between rs2961298 and circulating total n-3 PUFA for cg01770232 methylation. Therefore, in GOLDN, the association between n-3 PUFA and IL6 promoter methylation was not only negative but also dependent on sequence variants.\(^{(170)}\)

**Genetic variation at lipid-related genes, methylation and dietary fatty acids**

Using data from CHARGE and ENCODE consortia, we conducted another integrated analysis to explore whether gene–diet interactions on blood lipids act through DNA methylation\(^{(171)}\). Based on predicted relations in FA, methylation, and lipids, we selected seven candidate SNP located within APOE, ABCA1, 3-hydroxy-3-methylglutaryl-CoA reductase, APOA5, proprotein convertase subtilisin/kexin-type 9 and hepatocyte NF-1 homeobox A. According to the meta-analysis of seven cohorts in the CHARGE consortium, plasma HDL-cholesterol was not only associated with genotypes of ABCA1 rs2246293, but also positively associated with circulating EPA, for which the association was further dependent on genotypes of ABCA1 rs2246293. With methylation data in GOLDN, we found that methylation level of ABCA1 promoter CpG site cg14019050 was not only associated with genotypes of rs2246293, but also negatively associated with circulating EPA, for which, again, the association was further modified by genotypes of rs2246293. We further found that the correlation between methylation level of ABCA1 cg14019050 and plasma HDL-cholesterol is negative in GOLDN. Using data from ENCODE consortium, we identified a negative correlation between methylation of cg14019050 and ABCA1 expression. In order to validate the mediation effect of cg14019050 methylation in the pathway from gene–EPA interaction to plasma HDL-cholesterol, we conducted an additional mediation analysis, which was further meta-analysed across the GOLDN study, Cardiovascular Health study and the Multi-Ethnic Study of Atherosclerosis. We did observe a mediation effect; however, the magnitude of the mediation effect did not reach statistical significance. At APOE, although we observed consistent significant interactions between promoter SNP rs405509 and circulating α-linolenic acid for both plasma TAG in CHARGE consortium and methylation level of CpG site cg04406254 in GOLDN, there is no evidence to support the mediation effect of APOE methylation. Therefore, we obtained little evidence that DNA methylation explains the gene–FA interactions on blood lipids.\(^{(171)}\)

**Conclusions**

Despite the extensive evidence for gene-environment interactions and more specifically gene–diet interactions, the underlying biological mechanisms are still unclear. The current integrated studies of genetics and epigenetics provide gene-specific preliminary evidence that DNA methylation may act as one possible mechanism for such interactions, which is consistent with the established regulatory role of DNA methylation as the interface between ‘nature’ and ‘nurture’.

DNA methylation has been demonstrated to be determined by the local nucleotide sequence and almost all of the methylation (99.98 %) in differentiated mammalian cells occurs on the CpG dinucleotides\(^{(63)}\). Furthermore, the phenomenon of allele-specific DNA methylation, suggested by observed associations between genetic variants and DNA methylation, is widespread across the human genome. For example, according to analysis of twin pairs and their parents, >35,000 CpG sites were shown to have allele-specific DNA methylation events\(^{(172)}\).

Evidence has been accumulating in support of changes to DNA methylation in response to different types of environmental factors. Studies with monozygotic and dizygotic twins suggested the potential role of environmental factors in the regulation of DNA methylation\(^{(33,34)}\).

Based on the current knowledge, there is clear genetic contribution to DNA methylation as shown by significant SNP-CpG pair associations for genes, including
Moreover, there are significant interactions between methylation-related SNP and other environmental factors of interest, such as age and circulating FA. We found significant interactions for the promoter SNP of APOE, which interacted with age and α-linolenic acid, the promoter SNP of IL6 with EPA and DHA, and the promoter SNP of ABCA1 with EPA. These interactions were not only observed for the CVD traits, but also for the DNA methylation measurements of the corresponding genes. Furthermore, the results from the correlations between methylation and CVD traits and gene expression were in the same direction of the observed genetic associations and interactions. Our integrated analysis of both genetics and epigenetics provide preliminary evidence for the potential and partial mechanistic role of DNA methylation to explain gene–environment interactions, and such role maybe loci-specific.

With respect to clinical implications, the use of common SNP in the clinical setting for primary or secondary prevention remains controversial. APOE is one example, in that the ε4 variant was demonstrated to have a dosage effect on the incidence of and on the age of onset of the late-onset Alzheimer’s disease. However, debates persist over whether the genotyping test for APOE ε4 is necessary or desirable, because there are no medications or clinical strategies to counter the deleterious effect of the ε4 isoforms. However, the finding that α4 is associated with APOE methylation and expression suggest that the deleterious effects of ε4 might be mitigated by applying appropriate lifestyle-based modifiers that reduce the difference in methylation across different APOE isoforms.

There are many gaps and limitations that need to be overcome. First, the evidence in human subjects comes primarily from observational studies and a cause–effect relationship cannot be established. Second, DNA methylation studies in human subjects are based primarily on blood cells. Overall, we need intervention studies to increase the level of evidence supporting the notion that genotype-dependent epigenetic changes are an underlying molecular mechanism for gene–environment interactions with the objective of providing reliable evidence to advance the development of more personalised approaches to nutrition recommendations and medical care.

Acknowledgements

This material is based upon work supported by the U.S. Department of Agriculture, under agreement No. 58-1950-0-014. Any opinions, findings, conclusion, or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the view of the U.S. Department of Agriculture.

Financial Support

The authors were supported by the US Department of Agriculture, agreement no. 58-1950-0-014 and by the National Institutes of Health Grant P50 HL105185.

Conflict of Interest

None.

Authorship

Y. M. and J. M. O. contributed equally to the design of the review and writing of the manuscript and both reviewed and approved the final version.

References


44. Steegers-Theunissen RP, Obermann-Borst SA, Kremer D et al. (2009) Periconceptional maternal folic acid use of 400 microg per day is related to increased methylation of the IGF2 gene in the very young child. *PLoS ONE* 4, e7845.


88. Tomoilo D, D’Urso M, Martini G et al. (1984) Specific methylation pattern at the 3’ end of the human


133. Peet DJ, Turley SD, Ma W et al. (1998) Cholesterol and bile acid metabolism are impaired in mice lacking the nuclear oxysterol receptor LXR alpha. Cell 93, 693–704.


