PRENATAL NUTRITION AND THE RISK OF ADULT OBESITY: LONG-TERM EFFECTS OF NUTRITION ON EPIGENETIC MECHANISMS REGULATING GENE EXPRESSION

Running title: Epigenetic memory of gestational nutritional stress

Estanislau Navarro1,*, Anna N. Funtikova2,3,4, Montserrat Fito2,5, Helmut Schröder2,3

1Molecular Oncology Laboratory, Bellvitge Biomedical Research Institute (IDIBELL)
2Cardiovascular Risk and Nutrition Research Group (CARIN), IMIM (Hospital del Mar Medical Research Institute), Barcelona, Spain.
3CIBER Epidemiology and Public Health (CIBERESP), Instituto de Salud Carlos III, Spain
4PhD candidate, Food and Nutrition program, University of Barcelona, Spain
5CIBER Physiopathology of Obesity and Nutrition (CIBEROBN), Instituto de Salud Carlos III, Spain

*Author to whom correspondence should be addressed at:
Molecular Oncology Laboratory, Bellvitge Biomedical Research Institute (IDIBELL)
Gran Via 199
L’Hospitalet de Llobregat
08908 Barcelona, Spain
Emails: estanis.navarro@gmail.com
Phone: +34932607339
### ABBREVIATIONS:

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AS-RNA</td>
<td>antisense RNA</td>
</tr>
<tr>
<td>BAT</td>
<td>brown adipose tissue</td>
</tr>
<tr>
<td>BET</td>
<td>bromodomain and extraterminal</td>
</tr>
<tr>
<td>BMI</td>
<td>body mass index</td>
</tr>
<tr>
<td>CAD</td>
<td>coronary artery disease</td>
</tr>
<tr>
<td>CVD</td>
<td>cardiovascular diseases</td>
</tr>
<tr>
<td>DHA</td>
<td>docosahexaenoic acid</td>
</tr>
<tr>
<td>DMR</td>
<td>differentially-methylated region</td>
</tr>
<tr>
<td>DNMTs</td>
<td>DNA methyltransferases</td>
</tr>
<tr>
<td>DOAD</td>
<td>developmentally methylated region of adult disease</td>
</tr>
<tr>
<td>DWF</td>
<td>Dutch Winter Famine</td>
</tr>
<tr>
<td>EGCG</td>
<td>(-)-epigallocatechin gallate</td>
</tr>
<tr>
<td>eRNAs</td>
<td>enhancer RNAs</td>
</tr>
<tr>
<td>ERV</td>
<td>endogenous retrovirus</td>
</tr>
<tr>
<td>GEN</td>
<td>genistein</td>
</tr>
<tr>
<td>HAT</td>
<td>histone acetyl transferase</td>
</tr>
<tr>
<td>HDAC</td>
<td>histone deacetylase</td>
</tr>
<tr>
<td>IAP</td>
<td>intracisternal A particle</td>
</tr>
<tr>
<td>IUGR</td>
<td>intrauterine growth retardation</td>
</tr>
<tr>
<td>I3C</td>
<td>indol-3-carbinol</td>
</tr>
<tr>
<td>LINE</td>
<td>long interspersed nuclear element</td>
</tr>
<tr>
<td>IncRNA</td>
<td>long non-coding RNA</td>
</tr>
<tr>
<td>LSS</td>
<td>Leningrad siege study</td>
</tr>
<tr>
<td>LTR</td>
<td>long terminal repeat</td>
</tr>
<tr>
<td>miRNAs</td>
<td>microRNAs</td>
</tr>
<tr>
<td>ME</td>
<td>metastable epiallele</td>
</tr>
<tr>
<td>NCD</td>
<td>non-communicable disease</td>
</tr>
<tr>
<td>NuRD</td>
<td>nucleosome remodeling and deacetylating complex</td>
</tr>
<tr>
<td>PASR/PALR</td>
<td>promoter-associated short RNAs/promoter-associated long RNAs</td>
</tr>
<tr>
<td>PcG</td>
<td>polycomb group protein</td>
</tr>
<tr>
<td>PRE</td>
<td>polycomb responsive element</td>
</tr>
<tr>
<td>PRC1/2</td>
<td>polycomb repressive complex</td>
</tr>
<tr>
<td>PTM</td>
<td>post-translational modification</td>
</tr>
<tr>
<td>SFN</td>
<td>sulforaphane</td>
</tr>
<tr>
<td>SGA</td>
<td>small for gestational age</td>
</tr>
<tr>
<td>T2D</td>
<td>type 2 diabetes</td>
</tr>
<tr>
<td>WAT</td>
<td>white adipose tissue</td>
</tr>
</tbody>
</table>
LIST OF GENES

α-MSH: α-melanocyte-stimulating hormone
ABCA1: ATP-binding cassette subfamily A member 1
BRD2: bromodomain-containing 2 protein
CEBP-α/β/γ: CCAAT/enhancer binding protein-α/β/γ
CDKN2A\textsuperscript{p16}: cyclin-dependent kinase inhibitor 2A; CDKN2A
DNMTs: DNA methyltransferases
GLUT2: glucose transporter type 2
GNAS-AS: GNAS-antisense lncRNA gene
GR1-C: glucose transporter 3, group C, member 1 (glucocorticoid receptor)
IGF2: insulin-like growth factor II
IGF2R: insulin-like growth factor II receptor
INS-IGF2: INS-IGF2 readthrough IncRNA
IL-10: interleukin-10
JHDM1: lysine (K)-specific demethylase 2A
LEP: leptin
LDLRAP1: low-density lipoprotein receptor adapter protein 1
LSA1: lysine (K)-specific demethylase 1A (KDM1A)
HNF4A: hepatocyte nuclear factor 4α
MEG3: maternally expressed 3 lncRNA
MLL1: mixed-lineage leukemia 1
MTA1: metastasis-associated protein 1
OGA: O-GlcNAcase
OGT: O-linked GlcNAc transferase
PDX1: pancreatic and duodenal homeobox 1
PI3K-p85: phosphatidylinositol 3 kinase-p85
PGC1α: peroxisome proliferator-activated receptor gamma, coactivator 1 alpha
PLRL: prolactin receptor
PLZF: promyelocytic leukemia zinc finger protein
PPAR-γ: peroxisome proliferator-activated receptor-γ
p15\textsuperscript{INK4B}: cyclin-dependent kinase inhibitor 2B, CDKN2B
p16\textsuperscript{INK4A}: cyclin-dependent kinase inhibitor 2A; CDKN2A
p21\textsuperscript{WAF1/Cip1}: cyclin dependent kinase inhibitor 1A; p21/WAF1
RUNX2: runt-related transcription factor 2
SIRT1: sirtuin 1
TET2/3: (ten-eleven translocation 2/3) methylcytosine dioxygenase
UCP1: uncoupling protein 1
WIF1: Wnt inhibitory factor 1
ZEB1: zinc finger E-box binding homeobox 1
ABSTRACT

Solid epidemiological evidence indicates that part of the risk of obesity in adulthood could be programmed during prenatal development by the quality of maternal nutrition. Nevertheless, the molecular mechanisms involved are mostly unknown, which hinders our capacity to develop effective intervention policies. Here, we discuss the hypothesis that mechanisms underlying prenatal programming of adult risk are epigenetic and sensitive to environmental cues such as nutrition. While the information encoded in DNA is essentially stable, regulatory epigenetic mechanisms include reversible, covalent modifications of DNA and chromatin, such as methylation, acetylation etc. It is known that dietary availability of methyl donors has an impact on the patterns of gene expression by affecting DNA methylation at regulatory regions, a likely basis for reprogramming developmental plasticity. The Agouti and Axin-fused genes, as well as the embryonic growth factor IGF2/H19 locus are examples of diet-induced modulation of phenotypic traits by affecting methylation of gene-regulatory regions. Furthermore, recent work has evidenced an unsuspected role for chromatin as metabolic sensor. Chromatin is susceptible to a number of post-translational modifications that modulate gene expression, among them the GlcNAcylation of histone proteins and other epigenetic regulators. Intracellular levels of the precursor molecule UDP-GlcNAc, and hence the degree of global chromatin GlcNAcylation, depend on the energetic state of the cell, making GlcNAcylation a functional link between nutrition and regulation of gene expression. Dietary interference with these regulatory mechanisms could effectively counteract the early-life programming of adult risk.

KEYWORDS:

nutrient sensors, gestational programming, nutrition, epigenetic memory, O-GlcNAcylation, regulatory exhaustion.
1. INTRODUCTION

Obesity and related conditions are currently explained by a linear view of disease progression, in which expansion of white adipose tissue (WAT) generates a proinflammatory microenvironment that spreads through organs and tissues, hindering their function and causing the devastating comorbidities associated with obesity. This model can be refined by adding genetic predispositions and factors linked to lifestyle, such as diet or exercise, that influence the onset and progression of obesity and obesity-related conditions (see [1, 2] for recent reviews). In this way, obesity, although modulated by genetic predispositions, could be considered as an “acquired” environmental disease to be treated by targeting causative elements such as defective or excessive nutrition or sedentary lifestyle.

In recent years, however, the concept of obesity as an “epigenetic disease” has begun to be discussed. According to this perspective, environmental impacts on maternal nutrition during early embryo life leave a “nutritional imprint” with long-term effects on the promotion of obesity and related conditions in adulthood [3]. The present review addresses the hypothesis that adult obesity and related conditions may be programmed, at least in part, very early in life, focusing special attention on the recent explosive increase in the availability of data on epigenetic regulation and modification (DNA methylation, histone modifications, and non-coding RNAs). Here, we review recent literature on the epigenetic mechanisms putatively involved in programming the long-term effects of early-life nutrition, and how these could be modulated by environmental/nutritional cues.

Nutrition can affect the expression of a number of genes by its impact, in the form of methyl donors such as folate, on the degree of methylation of their regulatory regions, although folate supplementation could also have some undesired consequences on the stability of the epigenome. In the present review, we especially highlight recent reports that underscore a role for chromatin as a metabolic/nutrient sensor through the post-translational modification of histone proteins, which links regulation of gene expression to availability of nutrients and metabolites from the intermediary metabolism (see [4] for a recent review).
2. LONG-TERM EFFECTS OF EARLY-LIFE NUTRITION ON OBESITY AND OBESITY-RELATED RISK FACTORS

2.1 Long-term effects of fetal/neonatal malnutrition on adult risk factors

Sound epidemiological evidence has shown that malnutrition in early life not only affects fetal/neonatal growth but also has long-term consequences for adult health. Almost a century ago, Kermack, McKendrick and McKinlay hypothesized that the first 15 years of life influence the lifelong health of the individual [5], while Forsdahl demonstrated a significant association between infant mortality in the early years and adult mortality (ages 40-69 years) by coronary artery disease (CAD), and concluded that “great poverty in childhood and adolescence, followed by prosperity, is a risk factor for arteriosclerotic heart disease” [6]. Similarly, Barker and Osmond evidenced a consistent, positive correlation between infant mortality in 1921-1925 and ischemic heart disease mortality rates in 1968-1978 in England and Wales [7], results confirmed by their further work showing a strong positive correlation between low weight at birth and at one year of age and rates of CAD death in adulthood [8]. Several additional hypotheses have been proposed. The idea that growth impairment at conception or during fetal development could be linked to a number of chronic adult diseases, such as CAD [7], glucose intolerance [9], type 2 diabetes (T2D) [10], or obesity and hypertension [11] is known as the “developmental origins of adult disease” (DOAD) hypothesis [3]. This challenges the consideration of the womb as a highly protective milieu that buffers the embryo from environmental stressors, and acknowledges that the embryo is sensitive to outside stimuli and that the prenatal period is critical for modulating long-term disease predisposition by reprogramming the developmental plasticity of the embryo, likely by selectively impairing cell differentiation and organ growth [12, 13]. The thrifty phenotype hypothesis states that fetal undernutrition will cause the embryo to prepare for a life of scarcity, but subsequent accelerated overgrowth (catch-up) in conditions of affluence in the early post-natal period could overcome these protective barriers, facilitating fat deposition, obesity, and obesity-related conditions [10, 14]. Fetal/neonatal overnutrition could have similar long-term effects on health ([15] and references therein), as reported for infants who are large for their gestational age when born to mothers with diabetes [16], indicating that postnatal fast weight gain could be the critical factor explaining their increased risk for cardiovascular diseases (CVD) [17]. This idea has been supported by the results of two
prospective trials in which randomly assigned cohorts of small-born infants were fed either a nutrient-enriched formula or a control formula, which showed that the nutrient-enriched diet increased fat mass later in childhood [18].

2.2 Historical cohorts to study the association of early nutrition and later disease.

Well-documented historical episodes of hunger have been used to study the association between prenatal nutritional stress and the risk for adult chronic diseases[19], with the first demonstration coming from the cohort of people affected by the Dutch Winter Famine of 1944-1945 (DWF) [20]. Beginning in October 1944, civilian transport in the western part of Holland was totally blocked by the German Army, causing a dramatic shortage in food supply and a subsequent reduction in the daily rations, which reached their lowest value (400-800 Kcal/per person/day) between December 1944 and April 1945. Upon the end of the war, the blockade was raised and daily rations soon reached normal pre-war values [21]. The DWF cohort includes the original inhabitants of the blocked areas as well as their descendants born around the time of hunger, who have been stratified by famine exposure at the first, second, or third trimester of fetal development [21]. In the current DWF population older than 50 years, a number of studies have linked prenatal famine exposure with numerous characteristics: decreased fetal growth and reduced tolerance to glucose in adulthood (especially when exposure occurred at late gestation [22]); higher BMI and waist circumference, in women but not in men [23]; a more atherogenic lipid profile [24], and obesity, but only when exposure was during the first trimester of pregnancy [20, 21], an effect that affected up to the third generation [25]. Furthermore, prenatal famine exposure has been related with increased risk of T2D [26], overall adult mortality in women [27], and CAD [28, 29], although other authors have failed to find this association between prenatal famine exposure and adult CAD [30, 31] (see Table 1 for a summary of the long-term effects of fetal nutritional stress on obesity-related risk factors).

The Leningrad Siege Study (LSS) also investigated the relationship between maternal malnutrition and risk factors in adult offspring. The siege of Leningrad (now St. Petersburg) by the German Army for almost three years (September 1941-January 1944) resulted in more than one million deaths, mainly of starvation due to daily rations of only 300 calories during the worst of the siege [32]. The LSS results contradicted but complemented those of the DWF study, showing no associations between
intrauterine/infancy starvation and glucose intolerance, dyslipidemia or CVD risk later in adulthood [32, 33], even 70 years after the end of the siege [34], probably because the LSS population had no catch-up period of accelerated growth in early childhood, but rather lived in harsh conditions until the end of the siege [35]. This suggests that a period of postnatal nourishment must be relevant in determining the impact of maternal malnutrition on future health [36].

Striking evidences indicate that disease predisposition could be transferred to successive generations. The cohort of people born in Överkalix parish in northern Sweden at the turn of the 20th century showed a direct association between food availability up to the puberty of grandparents and parents and mortality due to CVD and T2D in the grandsons [37]; cardiovascular mortality was linked to the father’s nutritional status and diabetes to the paternal grandfathers’ [38]. Furthermore, sharp changes in food supply up to the puberty of paternal grandmothers increased CVD mortality of their sons’ daughters [39]. The DWF cohort also has been studied in the context of the transgenerational effects of prenatal exposure to famine, assessing the health of the children (F2) of parents born during the famine (F1) to undernourished women (F0). In one such study, an increase in body mass index (BMI) was found in F2 of F1 fathers, but not F1 mothers, when compared with offspring of control (not famine-exposed) fathers [25]. The authors did not find significant differences in the prevalence of CAD, T2D, high cholesterol, or hypertension in the F2 of exposed parents compared to the F2 of unexposed parents [25]. Similarly, Painter et al. detected an increase in neonatal adiposity and poor adult health in the F2 of F1 women exposed “in utero” to F0 famine, but did not find transgenerational impacts of famine exposure on the F2 rates of CAD or T2D in offspring of the exposed F1 [40]. Nevertheless, this F2 cohort is still young (mean age 37 years) and rates of chronic diseases could increase in the future in this group [25].

3. EPIGENETICS: THE NEXT FRONTIER IN THE REGULATION OF GENE EXPRESSION

3.1 Challenging dogmas on genome structure and function

Recent years have seen a paradigm change in ideas about the flow of genetic information. The original linear model, “DNA makes RNA makes protein”, for many years considered the central dogma of molecular biology [41, 42], has been replaced by
a multilayered process characterized by the pervasive expression of many different regulatory RNAs from multiple genomic loci in addition to the protein-coding mRNAs (Figure 1). International sequencing efforts have shown that the human genome is composed of just 20,000 to 25,000 protein-coding genes [43-45], with a best-guess estimate of 22,333 [46], a figure that would account for a mere 2% of the total genome length [47] and far fewer than the 100,000-120,000 protein-coding genes previously estimated [48-50]. Nevertheless, the number of transcriptional units could exceed 60,000 [51] by including a plethora of previously unclassified non-protein-coding RNAs, such as microRNAs (miRNAs), antisense-RNAs (AS-RNAs), promoter-associated RNAs (PALRs, PASRs), other long non-coding RNAs (lncRNAs), etc. [52-54]. Recent data suggest that approximately 80% of the genome could be considered biochemically active [55], most of it in the form of DNase I-accessible loci or candidate regulatory sequences [56-58]. Taken together, all these data draw a new image of the eukaryotic nucleus in which the genome would be pervasively transcribed, even in intronic and intergenic sites [59], with protein-coding genes encoding additional transcriptional units for a complex population of short and long non-coding RNAs with putative regulatory functions [60].

3.2 microRNAs and long non-coding RNAs: new actors on the gene-regulation stage

MicroRNAs (miRNAs) are small RNAs (over 22 nucleotides long) with important roles in post-transcriptional gene regulation. MiRNAs are transcribed as pri-miRNA precursors; after two processing steps, they generate the mature miRNAs, which are subsequently exported to the cytoplasm where they associate to their target mRNAs and induce their degradation or interfere with their translation (see [61] for a review). As of March 2016, the entire human miRNAome was described as formed by 1881 precursors and 2588 mature miRNAs (www.mirbase.org, mirbase release 21 [62]). MiRNAs are highly promiscuous, so that a single miRNA can interact with a number of different mRNAs which in turn can be regulated by several different miRNAs [63]. Aberrant miRNA expression profiles have been linked to a number of human diseases, such as leukemias [64], solid tumors [65], atherosclerosis and CVD [66, 67], chronic pain [68], and renal fibrosis [69], among others. Alterations in the expression of miRNAs have been detected not only in the primary tissue but also in blood [70], urine
The family of lncRNAs constitutes an extremely heterogeneous population of functionally unrelated RNAs longer than 200 nucleotides, which have important roles in development, differentiation, and disease [74-76]. Currently estimated at more than 56,000 [77], the number of lncRNAs more than doubles the number of protein-coding genes in the human genome, thus giving a hint of their functional complexity. Most of the lncRNAs are natural antisense transcripts, transcribed from the complementary chain of target genes, which likely cis-regulate its genomic locus by recruiting the histone-modifying machinery to the regulatory regions of the target genes [78]. Another common family of lncRNAs is the enhancer RNAs (eRNAs) or lncRNAs transcribed from functional enhancer sequences and associated to specific histone marks (high H3K4me1, H3K4me2, and H3K27ac and low H3K4me3 and H3K27me3) which could have a role in the activation of genes [79, 80], probably by facilitating the long-distance interaction of enhancers and their cognate promoters by forming or stabilizing DNA loops [81].

3.3 Methylation of DNA and histones, and the regulation of gene expression

Genes are precisely regulated by controlling access of the transcriptional machinery to their regulatory regions (promoters or enhancers). This process is mediated by the covalent modification of regulatory DNA sequences, mainly by methylating cytosines from CpG dinucleotides at position 5 through the activity of the family of DNA methyltransferases (DNMTs), which transfer a methyl group from S-adenosyl-methionine (SAM) to the cytosine [82]. CpG pairs concentrate in proximal (CpG islands) or distal (CpG shores) parts of the gene promoter, and in the methylated form are usually associated to gene silencing, although they are also involved in parental imprinting (see [83] for a review). Patterns of CpG methylation are strictly regulated and stable, and their alterations have been associated to diverse diseases, including cancer.

The function of histones in chromatin is also subjected to epigenetic regulation. Histones have protruding N-terminal and C-terminal tails that expose many positively charged aminoacids (K/R), which modulate the interaction with other histones and with the negatively charged DNA [84]. Histone modifications (known as tags or marks) contribute to the regulation of gene expression by changing the net charge of histone
proteins in order to regulate the strength of their interaction with DNA (and hence its accessibility) and by providing docking sites for the interaction of components of the transcriptional machinery [85] (see the description of the histone-tag nomenclature at the legend of Table 2). Aminoacids in the protruding histone tails are susceptible to many different covalent modifications “in vivo”, such as methylation, acetylation, ribosylation, phosphorylation, ubiquitination, sumolation, biotynilation, and more than 50 different tags [86, 87]. Furthermore, the existence of numerous histone genes, which generate up to 57 histone variants of differing lengths and sequences in humans, greatly increases the number of aminoacids susceptible to modification ([88]).

The stability of epigenetic marks is critical to control the long-term effects of events that alter the epigenetic landscape. While the genetic information encoded in DNA can be regarded as essentially stable, epigenetic marks are reversible and respond to endogenous, as well as environmental (nutritional, social, hormonal, cultural) stimuli. These dynamic modifications are “interpreted” by a number of families of epigenetic regulators [89] that have been classified as “tag-writers” (histone methyltransferases and acetyltransferases) [90], “erasers” (histone demethylases and deacetylases), or “readers”, proteins that include domains of the plant, chromo, Tudor, or MBT families, among others, and recognize different tagged histones (see [91] for a recent review). Elucidating the mechanisms that promote, maintain, read, and erase epigenetic marks is currently an area of intense research aiming to develop small molecules able to modulate expression of disease-associated genes at the epigenomic level [92].

4. REGULATION OF GENE EXPRESSION BY FOLATE AVAILABILITY: THE agouti, axin-fused AND IGF2/H19 LOCI

As reported above, the DWF cohort evidenced a number of risks in adulthood that could be traced to specific periods of embryonic development during which the mother experienced nutritional stress. A likely explanation for this finding was that the stress burden could affect (reprogram) the expression of genetic loci critical for embryo development. In this sense, recent work has evidenced that the folate-dependent methylation of DNA regulatory sites constitutes a direct link between nutrition and regulation of gene expression. In this section we review data on the effects of folate availability on the regulation of gene expression in two animal models, the agouti and
axin-fused loci, and in the human IGF2/H19 locus, a complex genetic structure involved in the control of early embryo development.

4.1 Genomic structure of the agouti\(^{A^y}\)-axin-fused (Axin\(^{Fu}\)) and IGF2/H19 loci

Agouti mice show a characteristic yellow coat because the protein encoded by the agouti gene (ASP, for agouti signaling peptide) blocks eumelanin synthesis in hair follicles [93]. Interestingly, the “viable yellow” (\(A^y\)) allele, a mutation in the agouti gene due to the insertion of a partial long terminal repeat (LTR) of an intracisternal A particle (IAP) retrotransposon, prompts the ubiquitous expression of the agouti gene because this LTR functions as a non-tissue-specific cryptic promoter [94, 95]. The \(A^y\) allele is epigenetically unstable and can lose methylation at the LTR so that heterozygous (\(A^y/a\)) mice, which also carry the recessive null (a) allele for a black coat, display continuously variegated agouti fur depending on the methylation status of the cryptic promoter, ranging from yellow in mice with a demethylated promoter to obscure-agouti/pseudoagouti fur in mice with an overmethylated promoter [96]. These methylation patterns are randomly fixed early in embryonic life and are maintained throughout life [97].

Similarly to the agouti gene, the axin\(^{Fu}\) locus derives from the insertion of an intracisternal A particle (IAP) retrotransposon, in this case in the intron 6 of the axin gene. This insertion generates a number of aberrant axin transcripts, some of them using the 3’ LTR of the IAP as a cryptic promoter [98]. The wild-type AXIN protein controls the formation of embryonic axis by repressing Wnt signaling [99]. Work by Rakyan and cols. showed that hypomethylation of the cryptic promoter (IAP) of the Axin\(^{Fu}\) allele caused the upregulation of AXIN expression, with the subsequent downregulation of Wnt activity that resulted in kinking of the distal tail [100]. On the other hand, the parentally imprinted IGF2/H19 human locus encodes the insulin-like growth factor II (IGF2) gene, which promotes embryo and placental growth [101] and is expressed almost exclusively from the paternal chromosome [102], as well as the maternally expressed lncRNA-H19, which limits the growth of the placenta through the regulated processing of miR-675 that is embedded in the first exon of the H19 gene [103]. The IGF2/H19 locus includes an intergenic, CpG-rich, differential methylation region (DMR) as well as a distal enhancer region, immediately downstream of the H19 gene, which together drive the tissue-specific, allele-restricted, and developmentally confined expression of the IGF2 and H19 genes [104] (Figure 2). The intergenic DMR
regulates IGF2/H19 gene expression; its demethylated form (maternal allele) recruits the regulatory factor CTCF, which functionally blocks the downstream enhancer, repressing the growth-promoting effect of the IGF2 gene and activating the growth-limiting H19 gene. In contrast, methylation of the intergenic DMR (paternal allele) allows the enhancer to activate the IGF2 promoter while repressing the H19 gene. In this way, the degree of intergenic DMR methylation acts as a balance, promoting (through IGF2 expression) or limiting (through H19 expression) early embryonic growth [105].

4.2 Folate availability regulates expression of the viable yellow agouti (A\textsuperscript{vy}) and axin-fused (Axin\textsuperscript{Fu}) loci.

Chromatin-methylating enzymes, i.e., DNA and histone methyltransferases, use S-adenosylmethionine (SAM) as universal methyl donor and folate to regenerate SAM through the one carbon metabolism pathway [106], suggesting that folate is an important component of the gene regulatory machinery. Folate deficiencies have been acknowledged as the basis of many nutrition-related diseases [107, 108], which supports this hypothesis and highlights folate as a critical link between nutrition and gene expression. This relationship has been evidenced by work with agouti and axin-fused model mice. Thus, in agouti mice, supplementation of (a/a) dams with a methyl-rich diet shifted the coat color of the A\textsuperscript{vy}/a offspring from yellow to the brown pseudoagouti [109], a result also seen after supplementation of (a/a) dams with genistein, a major soy isoflavone [110]. In both cases, the phenotypic shift coincided with an increased methylation of the IAP insertion region at the agouti locus in the pups [111].

Change in coat color is not the only phenotypic trait associated to the agouti locus. Heterozygous mice harboring the lethal yellow (A\textsuperscript{y}) or viable yellow (A\textsuperscript{vy}) mutations showed a number of pleiotropic effects, including obesity, increased tumor susceptibility, and embryonic lethality [112]. The obesogenic role of the A\textsuperscript{vy} allele was due to the agouti gene product, the agouti-signaling peptide (ASP), since a transgene overexpressing ASP\textsuperscript{wt} induced obesity in the recipient mice [113] and incubation of mature 3T3-L1 adipocytes with recombinant ASP increased expression of adipogenic transcription factors [114]. Interestingly, this obesogenic effect proved to be sensitive to dietary supplementation with methyl donors or with genistein, which reversed the effect of the A\textsuperscript{vy} allele by overmethylating its cryptic promoter [110, 115]; this constitutes experimental evidence of a dietary modification in pregnant mothers that directly affects the obese phenotype of the offspring.
On the other hand, in the case of the Agouti locus, supplementing Axin\textsuperscript{Fu} dams with methyl donors before and during pregnancy significantly increased the methylation of the Axin\textsuperscript{Fu} locus and reduced the incidence of tail kinking in their offspring [116]. Nevertheless, the mechanisms involved in this nutrient-induced methylation of regulatory regions are complex and must include numerous (epi)genetic factors to accurately control the expression of Agouti or Axin-fused genes [117]; Dolinoy et al. showed that yellow agouti mice had less repressive histone tag H4K20me3 in the LTR of the IAP, compared to their pseudoagouti littermates [118].

4.3 The Dutch Winter Famine caused life-long alterations in the methylation pattern of the IGF2/H19 locus in the offspring of affected pregnant mothers

In an experimental study, Heijmans et al. studied the methylation pattern of the IGF2-DMR in the DWF cohort of individuals conceived during the Dutch famine. Six decades after being exposed to the famine, the IGF2-DMR was hypomethylated (see Figure 2), compared with their non-exposed siblings, when the exposure was periconceptional but not when individuals were exposed late in gestation, suggesting that early-life nutritional stress had an impact on the methylation pattern of the IGF2/H19 locus, and that this altered pattern could be maintained for decades until adulthood [119]. Furthermore, the degree of methylation of the IGF2-H19 locus was directly associated with fetal and infant growth, with children born small-for-gestational age (SGA) showing a decreased level of IGF2-DMR methylation in white blood cells obtained from umbilical cord, compared with control children [120]. On the other hand, Steegers-Theunissen et al. showed that periconceptional supplementation of pregnant mothers with 400 μg of folic acid significantly increased methylation at the IGF2-DMR locus of the child, compared with the offspring of non-supplemented mothers; however, the authors also found an inverse association between IGF2-DMR methylation and birth-weight that is difficult to explain by the current models of regulation of the IGF2-H19 locus but that could indicate involvement of other genomic loci [121].

Other loci related to growth and metabolism also have been studied. Those encoding interleukin-10 (IL-10), leptin (LEP), ATP-binding cassette subfamily A member 1 (ABCA1) or the lncRNAs maternally-expressed 3 (MEG3) and GNAS-antisense (GNAS-AS) were found to be overmethylated in the famine-exposed individuals vs. non-exposed siblings, while the INS-IGF2 lncRNA, also at the IGF2 locus, was found to be undermethylated. This result further supports the hypothesis that
early-life nutritional stress could cause widespread alterations in DNA methylation [122]. A genome-wide analysis of DMRs in the DWF cohort detected a number of differentially methylated CpG elements—such as enhancers, open chromatin regions, and developmental enhancers—associated to the prenatal famine, usually at regulatory regions, suggesting that they could influence the expression of many different genes and especially of genes linked to metabolism and growth [123]. On the other hand, the promoters of the glucocorticoid receptor GRI-C, peroxisome proliferator-activated receptor gamma (PPAR-γ) lipoprotein lipase, or phosphatidylinositol 3 kinase p85 (PI3K-p85) did not differ in their degree of methylation in the blood of individuals exposed “in utero” to famine, compared to non-exposed controls [124], suggesting that the nutrition-dependent changes in methylation could be gene- or tissue-restricted.

5. EPGENETIC REPROGRAMMING OF DEVELOPMENTAL PLASTICITY: ADIPOSE TISSUE AND PANCREAS

The interventional or longitudinal studies described above have clearly established that nutrition can modulate the global levels of methylation at gene-regulatory regions. Nevertheless, this top-down approach (from nutrition to gene expression) has produced very scanty data on the specific genes or gene networks de/methylated, so that it is difficult to delineate the mechanisms linking overall nutrition-dependent epigenetic alterations with increased adult risk of non-communicable diseases (NCDs). An alternative to the top-down approach appeals to the phenotypic plasticity of tissues and organs at critical periods of early embryo development (mainly to control stem cell proliferation and/or cell differentiation) to determine whether the expression of the genes controlling these processes could be modulated by external stimuli. In this chapter, we discuss recent work on genes important to the development of organs and systems that maintain metabolic homeostasis, such as the pancreas and the adipose tissue, and whose expression has been shown to be modulated by nutritional cues.

5.1 Maternal malnutrition affects the epigenetic regulation of pancreas development.

Pancreatic and duodenal homeobox 1 (PDX1) is a transcription factor that is essential for the survival and function of the mature pancreatic β-cells, to the extent that disruption of the Pdx1 gene (Pdx1<sup>−/−</sup> mice) results in the arrest of pancreatic
development at an early embryonic stage [125]. Adult expression of PDX1 has been shown to be modulated by early-life nutrition. Thus, in a mice model of intrauterine growth retardation (IUGR), which predisposes to T2D development in adulthood, Park et al. showed that the *Pdx1* gene underwent epigenetic modifications that downregulated *Pdx1* expression in β-cells from over 50% (vs. control) in IUGR fetuses to null values in adults [126]. Furthermore, the authors were able to map this regulatory behavior to a region of the *Pdx1* proximal promoter that included a CpG dinucleotide in a binding site for the USF-1 transcription factor, which was progressively methylated in IUGR rats but not in control animals. Methylation of the *Pdx1* proximal promoter was accompanied by histone deacetylation, since inhibition of HDAC partially restored *Pdx1* expression [126]. Similar results have been recently described by Abuzgaia et al., who used a rat model of gestational maternal protein restriction, which increases susceptibility to T2D. In a low-protein group, they detected downregulation of *Pdx1* mRNA and protein expression, as well as of its downstream target genes insulin and glucose transporter Glut2, compared with controls [127]. Furthermore, a genome-wide analysis of CpG-methylated sites made in pancreatic islands from IUGR rats showed that changes in cytosine methylation in over 1400 loci preceded the development of T2D, giving a hint about the number of loci potentially involved in the pathogenesis of this disease [128], while a high-throughput analysis of gene expression led to the detection of 253 differentially expressed genes in the livers of exposed vs. non-exposed offspring, among them many encoding enzymes related to fat metabolism and potentially involved in the development of metabolic diseases [129].

Diet-induced changes in the levels of histone tags also have been described in a number of developmentally important genetic loci. Maternal low protein diets have been associated to downregulation of mRNA and protein expression of the tumor suppressors CDKN2A\(^{P16}\) and p21\(^{WAF1/Cip1}\) in the mammary glands of offspring, concomitant with a reduction in acetylated H3 and in H3K4me2 in their promoter regions [130, 131]. In a rat model of gestational maternal protein restriction, Sandovici et al. showed the downregulation of *Hnf4a*, a transcription factor gene critical to pancreatic function, in the islets of the exposed pups, compared with control pups, as well as an increase in the repressive tag H3K9me2 and a decrease in the active mark H3K4me1 in the *Hnf4a* intronic enhancer, with a subsequent reduction in its interaction with the *Hnf4a* promoter [132].
5.2 Maternal malnutrition affects the regulation of adipogenic genes.

The adipose tissue of an adult organism is mainly composed of white adipose tissue (WAT) plus some remnants of the thermogenic brown adipose tissue (BAT). WAT has a double physiological role as regulator of the body’s energy balance and as an endocrine organ that controls a number of physiological and pathophysiological mechanisms through the release of hormones and adipokines. This highly dynamic tissue adapts to external stimuli by increasing the number or size of adipocytes (see [133] for a recent review). Preadipocyte differentiation is controlled by the early transcription factors C/EBP-β and C/EBP-γ, which regulate PPAR and C/EBP-α, adipogenic master genes that promote adipocyte terminal differentiation through their ability to activate the expression of adipocyte-specific genes (see [134] for a recent review). Among the genes activated by PPAR-γ are many involved in lipid metabolism and insulin/IGF signaling [135], as well as the uncoupling protein-1 (UCP1), a functional marker for BAT whose mRNA peaks soon after birth in large mammals and then is rapidly lost, together with its transcription activating factors PPARα and coactivator PGC-1α, as part of a developmental switch from BAT to WAT [136].

Prenatal undernutrition has been shown to have an impact on the fetal fat mass as well as in the expression of UCP1 mRNA. In this regard, ovine maternal nutrient restriction during the early gestation resulted in offspring with more adipose tissue and higher levels of UCP2 and PPARα mRNAs than unrestricted controls, while no changes were detected in the levels of UCP1, PPAR-γ, or the long and short forms of the prolactin receptor (PRLR) [137]. Furthermore, in mice, protein restriction of pregnant mothers resulted in lower levels of leptin mRNA and protein in the offspring [138]. In contrast, pups born to mothers submitted to a high-fat diet showed worse glucose tolerance and higher leptin levels and lower adiponectin levels in the adipose tissue than their control littermates, which was associated with higher H3K9Ac and lower H3K9me2 levels at the adiponectin promoter and lower levels of H4K20me at the leptin promoter [139].
6. CHROMATIN AS A NUTRIENT SENSOR: NUTRITION-INDUCED POST-TRANSLATIONAL MODIFICATIONS OF HISTONE PROTEINS AND THEIR EFFECTS ON CHROMATIN FUNCTION.

Nutrition has long been considered a major environmental factor responsible for changes in the transcriptome, mainly by the binding of small ligands to nuclear receptors, transcription factors, or cofactors [140]. Nevertheless, recent years have seen a number of reports in which small metabolites or nutrients controlled patterns of gene expression not by binding to single, specific targets but by modulating the global level of post-translational modifications (PTMs) of DNA, histones, and other epigenetic regulators (see [141] for a recent review). These “epigenetic metabolites” constitute a functional link between nutrition and regulation of gene expression, with poor nutrition and metabolite shortage affecting the appropriate deployment of epigenetic tags. This highlights chromatin architecture as a marker of the energetic and metabolic state of the cell. According to this model, chromatin would act as a metabolic sensor to adjust global patterns of gene expression to the dynamic changes in the concentration of specific metabolites, while ensuring its epigenetic inheritance [142]. This novel perception of metabolite-dependent chromatin function will surely provide researchers with conceptual tools to unravel the epigenetic components underlying reprogramming of gene expression by early-life nutritional stress, as well as to study the mechanisms that translate these regulatory alterations in risk increases to adult disease. In this section, we review the regulatory function of the best characterized of these epigenetic metabolites: O-linked-N-acetylglucosamine (O-GlcNAc), NAD+, and acetyl-CoA, as well as their functional relationship with chromatin and Sir2uin1, two of the better characterized metabolic sensors.

6.1 Intracellular O-GlcNAcylation depends on the extracellular levels of glucose through the nutrient sensor UDP-GlcNAc

Among the different intermediate metabolites with a confirmed role in the establishment or maintenance of histone PTMs, O-linked-N-acetylglucosamine constitutes a class of its own. Addition of an O-linked-N-acetylglucosamine group (O-GlcNAc) to Ser/Thr residues of nuclear proteins [143] targets hundreds of proteins, among them epigenetic regulators and chromatin-associated proteins such as histones, members of the polycomb repressive complex, and a number of HDACs directly involved in the regulation of gene expression (see [144] for a recent review).
GlcNAc is reversibly added to its target proteins as a single residue by the O-linked GlcNAc transferase (OGT), which uses the UDP-GlcNAc sugar as substrate, and is removed by the O-GlcNAcase (OGA) enzyme. Both reactions constitute the “O-GlcNAc cycle”. O-GlcNAc and O-GlcNAcylated products can be found not only in the nucleus or cytoplasm, but also in the endoplasmic reticulum-Golgi axis where they are used in the biosynthesis of glycoproteins, glycosaminoglycans, or glycolipids [145].

The precursor form UDP-GlcNAc, which is synthesized by the hexosamine biosynthetic pathway, can be considered a sensor of the metabolic state of the cell [146], since its synthesis requires four metabolites, glucose (carbohydrates), glutamine (aminoacids), acetyl CoA (lipids), UTP (nucleotides), and ATP as energy donor; their availability directly determines UDP-GlcNAc levels and subsequently those of intracellular O-GlcNAcylation [146]. Thus, the levels of UDP-GlcNAc reflect the nutrient status of the cell, with extracellular levels of glucose modulating the extent of intracellular O-GlcNAcylation [147]. Moreover, the O-GlcNAc cycle is functionally integrated with other nutrient sensing pathways such as mTOR, AMPK (which is activated by a low ATP/AMP ratio), MAPK, and insulin-AKT and the direct interaction of OGT with tyrosine kinase receptors, PIP3 sites, or p38 MAPK has been described, suggesting that these signaling pathways could also be modulated by O-GlcNAcylation (see Figure 3 and [148]).

6.2 The activity of a number of epigenetic regulators of gene expression is controlled by O-GlcNAcylation

In a gene expression analysis in C. elegans, Love et al. found over 800 promoters modified by O-GlcNAcylation in which disruption of the O-GlcNAc cycle led to alterations in the pattern of gene expression, many of them linked to lipid/carbohydrate metabolism or microRNA expression; this constituted a nutrient-responsive program of gene expression [149]. In a mice model, male levels of OGT and of O-GlcNAc were seen to be reduced by early prenatal maternal stress, which affected hypothalamic gene expression as well as the landscape of brain microRNA expression [150].

On the other hand, Medford et al. demonstrated that the chronic ingestion of a sugar and saturated fat-rich diet (Western diet) increased protein O-GlcNAcylation in the hearts of rats, without modifying the levels of the OGT/OGA enzymes, which suggested that accumulation of O-GlcNAcylated products was due to the increased
availability of nutrients [151]. This result has the utmost importance because it links the degree of protein O-GlcNAcylation to the historical record of nutrient ingestion. Among the proteins that were post-translationally modified by O-GlcNAcylation, the group of gene expression regulators is well represented. Several excellent reviews on the O-GlcNAcylation of transcription factors have been recently published (see [152] and references therein); here we will focus on the epigenetic mechanisms affected by O-GlcNAcylation.

On the functional side, O-GlcNAcylation has been shown to regulate chromatin dynamics through two different mechanisms: i- by tagging members of the polycomb group of proteins (PcG) [153] and other epigenetic regulators, such as the Ten-Eleven Translocation (TET) family of DNA hydroxylases/demethylases [154], and ii- by directly targeting the four core histones in chromatin [155, 156], although it seems that there is also some extent of crosstalk between the two mechanisms. The polycomb group proteins are transcriptional regulators that mediate the repression of numerous genes along the development to adulthood. Polycomb group proteins form two polycomb-repressive complexes (PRC1 and PRC2), which recognize and bind polycomb-responsive elements (PREs) at the regulatory regions of their target genes. PRC2s are H3K27 trimethyltransferases while PRC1 binds to H3K27me3 and monoubiquitinates H2AK119, with the overall effect of repressing the expression of the target genes [157].

In two thought-provoking studies, the polyhomeotic protein (Ph), one of the components of PRC1, was shown to be O-GlcNAcylated [158], while the O-linked GlcNAc transferase (OGT) enzyme was identified as the PcG member previously known as super sex combs (sxc) [158, 159]; O-GlcNAcylation-defective null-mutants (sxc/Ogt[−]) were unable to maintain polycomb-mediated transcriptional repression [158]. This last result strongly supports a role for the nutrient sensor O-GlcNAc in the repression of polycomb target genes and provides mechanistic insight on the functional relationship among nutrients and the transcriptional machinery. On the other hand, O-GlcNAc sites have been detected in all the core histones, cross-talking with other histone PTMs. Thus H3S10GlcNAc impaired the phosphorylation of the same residue [155], H2BS112GlcNAc facilitated H2BK120 monoubiquitination [160], and H3T32GlcNAc reduced phosphorylation at H3S10, H3S28, and H3T32 [161].
The O-GlcNAcylation of target histones was facilitated by the binding of OGT to the TET2 and TET3 (Ten-Eleven Translocation 2/3) proteins at active transcription start sites enriched in the activation mark H3K4me3, while loss of OGT or TET2/3 reduced the target gene expression [154, 162]. In this way, O-GlcNAcylation would have a dual role in the control of gene expression, reinforcing the repressive role of polycomb on its targets and facilitating transcription by binding to TET2/3 proteins. How these two opposite roles are regulated is currently unknown.

6.3 Alterations in the balance of O-GlcNAcylation contribute to human disease

Recent years have seen evidences that an altered flow of O-GlcNAc could contribute to obesity, CVD, and T2D [163], to the extent that determination of O-GlcNAc in blood has been proposed as a biomarker for early metabolic dysfunction in youngsters [164]. As stated above, levels of O-GlcNAc depend on the balance between the opposite activities of OGT (GlcNAcylation) and OGA (O-GlcNAc removal), and changes in their activity have been associated to human pathology, although the molecular mechanisms involved have not yet been clarified. The chemical inhibition of OGA resulted in increased levels of GlcNAc and impeded insulin-dependent phosphorylation of Akt and GSK3β, leading to insulin resistance in 3T3-L1 adipocytes [165]. In contrast, heterozygous OGA-deficient mice had reduced insulin sensitivity, hyperleptinemia, and showed a de-regulated expression of genes associated with growth, innate immunity, and the metabolism of glucose, lipids, sterols, and calcium [166]. On the other hand, increasing GlcNAc levels by the transgenic overexpression of OGT induced insulin resistance and hyperleptinemia when OGT was expressed in muscle and adipose tissue [167], and was associated with obesity and fatty liver when it was expressed in liver, likely by interfering with the metabolism of Apo B and Apo A-I [168].

6.4 Other metabolic sensors: SIRT1, Acetyl-CoA, and FAD

In addition to O-GlcNAc, other metabolic sensors link DNA and chromatin PTMs (and hence regulated gene expression) to the levels of different metabolites in a way that could be epigenetically inherited, potentially contributing to the long-term effects of nutrition on adult disease risk. One of the best characterized is Sirtuin1 (SIRT1), a nuclear NAD⁺-dependent protein deacetylase which links cellular metabolic status (such as NAD⁺ levels, modulated by dietary niacin) to the dynamics of chromatin architecture and regulated gene expression through the deacetylation of histones and
other nuclear regulatory proteins [169]. In this sense, SIRT1 has been considered as a master regulator of the transcriptional networks that control lipid metabolism, gluconeogenesis, insulin secretion, and inflammation [170]. A recent analysis of the liver acetylome from Sirt1-deficient (Sirt1/−) mice allowed the detection and identification of a group of nuclear proteins directly deacetylated by SIRT1 “in vitro”, among them a subunit of the SWI/SNF chromatin remodeling complex and five proteins involved in RNA processing [171]. Other chromatin-associated proteins have been characterized as substrates for the SIRT1 deacetylase activity, including the transcriptional repressor promyelocytic leukemia zinc finger protein (PLZF) [172], the histone methyltransferase mixed-lineage leukemia 1 (MLL1) associated to H3K4 trimethylation at certain promoters [173], and SATB1 homeobox 1, the regulator of high-order chromatin structure [174]. Also related to NAD⁺ sensing, the globular macro domain constitutes a highly conserved protein fold present in histone variants (macroH2A), modifiers (macro-PARPs), and chromatin remodelers (Alc1) (see [175] for a recent review), which functions as an ADP-ribose binding module [176], likely mediating transcriptional regulation in response to nutrient availability.

On the other hand, acetyl-CoA is the acetyl donor for histone lysine acetylation by histone acetyl transferases (HATs). Independently of its role in the synthesis of O-GlcNAc, acetyl-CoA has been considered a metabolic sensor by itself. Wellen et al. showed that suppression of the ATP-citrate lyase activity (the enzyme that converts glucose-derived citrate to acetyl-CoA) resulted in a global decline in histone acetylation, and more specifically in the hypoacetylation of the promoter region of the glucose transporter GLUT4 gene in adipocytes, thus linking glucose availability to the activation of a glucoalytic program of gene expression [177]. Finally, other metabolites whose levels change in response to the nutrient status and are involved in the epigenetic regulation of gene expression have been described ([178], and references therein). These include FAD, a riboflavin (vitamin B12) derivative that works as a cofactor of LSD1, a H3K4 histone demethylase [179].

6.5 Effect of plant bioactive compounds on epigenetic regulators

In recent years, a number of dietary bioactive compounds of plant origin have been characterized as epigenetic modulators of gene expression because of their ability to reshape histone post-translational modification marks or DNA methylation patterns, although their exact role in maintaining the altered patterns required to establish lifelong
effects remains poorly understood. In this section, we review recent reports on the direct
effect of small secondary plant metabolites, mainly polyphenols and isothiocyanates, on
the activity of enzymes directly involved in the structural/regulatory modifications of
chromatin. We will not discuss the effects of caloric intake or of changes in
micro/macronutrient consumption because there are excellent recent reviews on this
topic [180].

Polyphenols such as curcumin, resveratrol, genistein, or ECGG constitute a
complex class of plant secondary metabolites with an accepted role as epigenetic
modulators. Curcumin (diferuloylmethane) is a bioactive compound with anti-oxidative,
anti-inflammatory, and anti-lipidemic properties and contributes to epigenetic regulation
by downregulating expression of DNA methyltransferases (DNMT1, 3A and 3B) [181]
and inhibiting HAT activity [182]. Furthermore, curcumin treatment was seen to alter
expression of a large set of miRNAs, and consequently of their target genes [183], to
upregulate the tumor suppressor IncRNA MEG3 [184], as well as to suppress
expression from the H19 locus, but not from IGF2, in cancer cell lines [185].
Resveratrol is another phytochemical with an established role as epigenetic regulator.
Resveratrol activates SIRT1 by increasing its binding to laminA [186] and inhibits
mono-ubiquitination of histone H2B at K120 [187], which has an impact on the
landscape of histone marks. Furthermore, resveratrol has been shown to regulate
expression of the chromatin modifier metastasis-associated protein 1 (MTA1), a
component of the silencing nucleosome remodeling and deacetylating complex (NuRD)
[188]. On the other hand, the soybean isoflavone genistein has been observed to
promote hypermethylation of CpG islands in specific mouse genes [189], as well as to
modulate H3K19 methylation or deacetylation of specific promoters through a
reduction of SIRT1 activity [190], while gärchinol has been characterized as a potent
inhibitor of histone acetyltransferases [191], cambinol as an inhibitor of SIRT1 and
SIRT2 [192], and chaetocin as an inhibitor of histone H3 lysine 9 (H3K9), a
methyltransferase suppressor of the variegation 3-9 homolog 1 (Suv39 h1) [193]
The list of phytochemicals with epigenetic regulatory potential is not restricted
to polyphenols, but should also include sulforaphane (SFN), a dietary isothiocyanate,
and allyl sulphides [194, 195], which function as HDAC inhibitors “in vitro”, and green
tea catechins, especially epigallocatechin gallate (EGCG), which have been shown to
downmodulate DNA methylation by attenuating the effect of DNA methyltransferase 1
(DNMT1) through a yet unknown mechanism, although various authors have shown the reexpression of tumor suppressors p16\textsuperscript{INK4a} and p15\textsuperscript{INK4b} and WIF1 upon treatment of tumor cell lines with EGCG ([196] and references therein).

7. XENOMIRS AND METASTABLE EPIALLELES: DOUBLE-EDGED SWORDS FOR NUTRIGENOMIC INTERVENTION.

The safety of nutrigenomic intervention is widely accepted, mainly because the accidental ingestion of physiologically toxic amounts of polyphenols and other plant metabolites seems to be highly unlikely. Nevertheless, physiologically safe levels of ingested nutrients and metabolites could have undesired effects on epigenetic regulators [197, 198]. Of special concern are genes that harbor metastable epialleles (MEs), i.e. alleles which generate phenotypical variation through differential methylation, in their regulatory regions [199]. Metastable epialleles have been shown to promote the spreading of methylation to neighboring genes, thus contributing to their untimely epigenetic silencing [200]. What makes regulatory MEs interesting for nutraceutical research is the potential modulation of expression of their associated genes by environmental factors, with the degree of DNA methylation at their promoters reflecting the maternal intake of folic acid, as in the agouti and axin-fused genes [111, 116]. This environmental sensitivity makes ME-harboring genes likely targets for nutraceutical intervention, as well as interesting candidates to explain the long-term susceptibility to chronic diseases by linking prenatal exposures (e.g. to maternal folate) to the developmental plasticity of the epigenome [201]. In this context, Waterland et al. described a number of human MEs and studied their dynamics in a model of seasonal food shortage in Gambian subsistence farmers. They compared levels of methylation in MEs from individuals conceived in scarce vs. affluent conditions, and found them to be significantly increased in individuals conceived in the challenging season [202], probably due to seasonal variations in the intake of methyl donors by the pregnant mothers, although no data were provided on the expression variations of the ME-associated genes [203, 204]. To make things more interesting, a recent report showed that genomic regions flanking human MEs were enriched in repetitive sequences of the long interspersed nuclear element (LINE) or endogenous retrovirus (ERV) families [205], a relevant result because murine MEs have also been localized next to retroviruses and other transposable elements that generate epigenetic variability or
instability [206, 207] or affect the transcriptional pattern of the locus [208]. Other authors have detected hypermethylated, leaky elements of the Alu [200] or B1 [209, 210] families that spread methylation to flanking regions and consequently contributed to the silencing of contiguous genes. Clearly, more research is needed on the function and dynamics of MEs, on their associated genes, and on their role as spreaders of methylation and gene-silencers; nonetheless, MEs could provide the degree of phenotypic variability needed to explain some of the long-term effects of early-life nutrition.

Another risky association between nutrition and the epigenetic machinery deals with the recently described xenomiRs or xenomiRNAs, i.e., exogenous plant or animal-derived miRNAs [211]. In a highly controversial report, Zhang et al. described that the plant-specific microRNA miR-168a could be acquired by animals through food ingestion, survive the acidic conditions of stomach, target the low-density lipoprotein receptor adapter protein 1 (LDLRAP1) mRNA, and reduce the liver expression of LDLRAP1 protein [212]. Similarly, Baier et al. reported that miR-29b and miR-200c could be acquired by humans through milk consumption, and that miR-29b could target the runt-related transcription factor-2 (RUNX2) mRNA and miR-200c the ZEB1 mRNA in “in vitro” luciferase assays [213]. Furthermore, these authors showed that depletion of these miRNAs from milk was not compensated by intracellular synthesis, suggesting that humans relied on continuous exogenous supply of miR-29b and miR-200c [213]. Although these results have been openly questioned [214-217], with some laboratories being unable to replicate them [218] and others considering them as laboratory artifacts [219], the idea that xenomiRs could regulate the expression of target genes in humans or animals is fascinating (see [220] for a recent review), and could open a conceptual gate for the direct use of miRNA-characterized diets in nutritional interventions aimed to stop disease progression [221, 222]. It is likely that the coming years will see enormous progress in this highly controversial research area.

8. CONCLUSIONS AND FUTURE TRENDS

The epidemic of obesity that currently affects the western world has generated great interest in identifying the molecular mechanisms involved in the onset and progression of obesity and its related conditions. Of special interest are the effects of prenatal nutrition on the long-term risk of chronic diseases, as well as the genetic and
epigenetic mechanisms that regulate the expression of genes critical for the appropriate
development of adipogenic tissues. The present review focused on recently published
work on the interface between nutrition and epigenetic mechanisms, i.e., how nutrients
have an impact on the covalent marks of DNA and histones that have a direct role in the
regulation of gene expression. Although this field of research is quite young, enough
data are available to draw a number of conclusions that could serve as a theoretical basis
for the development of further investigations: i.- early-life malnutrition results in
increased risk of obesity and related conditions in adulthood, according to sound
epidemiological evidence from studies in historical cohorts; ii.- nutrition has a direct
impact on the expression of a number of responsive genes, and the availability of
specific nutrients, such as methyl donors or folate, can modulate expression of certain
genes (e.g., the Agouti gene) by changing the degree of methylation of their regulatory
regions; iii.- chromatin could be considered a metabolic/nutrient sensor that responds to
nutritional cues by changing the patterns of gene expression. The post-translational
addition of O-GlcNAc, whose intracellular levels reflect the metabolic status of the cell,
modulates gene expression by tagging histone proteins and members of the polycomb
group of proteins (PcG), among others. Thus poor nutrition, leading to a metabolite
shortage, could affect the appropriate deployment of epigenetic tags and, subsequently,
the growth and development of critical organs and tissues.

A general overview of this topic would highlight evidence that early
environmental exposures (such as nutrition) can affect (i.e. reprogram) the
developmental plasticity of key tissues and organs by affecting epigenetic mechanisms
that control gene expression. Nevertheless, this explanation remains a theoretical
structure because we still lack a mechanistic view of the processes involved. The many
unknowns include the genes affected, the number and nature of their nutrient-sensitive
regulatory regions, the epigenetic regulators involved, the gene-modulating role of
specific nutrients and metabolites, aa well as the developmental alterations generated in
the critical tissues and organs. On the other hand, the promising insights into the
regulating roles of xenomiRNAs or of the metastable epialleles in nutrition-dependent
changes make them worthy of study because of their ground-breaking potential as
regulatory mechanisms, as well as their yet unforeseen physiological impacts.

These will definitely be future fields of active research because they correspond
to the primary questions raised by the data here reviewed. Nevertheless, we would pose
other, perhaps more philosophical, questions. It would be very important to describe the molecular mechanisms by which the cellular memory of the initial nutritional insult is established in early life and propagated to adulthood, i.e., to understand why the predisposition “awakens” 40-50 years later. We could hypothesize a requirement for a second signal, or the interaction of nutritionally primed epigenetic mediators with the aging machinery of the cell. In addition, it is tempting to speculate that nutrition-derived reductions in prenatal nutritional/metabolic sensors (such as the methyl donors or UDP-GlcNAc, among others) could somehow affect the activity of critical, as yet unknown, regulatory regions or epigenetic regulators, leading to a gradual loss of epigenetic tags and/or other signals important for the maintenance of gene expression homeostasis. The affected cells would thus enter a state of gradual “regulatory exhaustion” that would hinder their physiological stability.

It is clear that a deep knowledge of the genetic and epigenetic mechanisms involved in nutrition-dependent reprogramming would likely provide us with tools for primary prevention as well as to reverse the process and, hopefully, its effects on long-term disease susceptibility. Of course, much more research is needed to fulfill this objective, and it is likely that the coming years will see a dramatic advance in research on the nutrition-dependent alterations in gene expression and the associated increases in disease risk.
ACKNOWLEDGMENTS

The authors appreciate the English revision by Elaine Lilly, PhD (Writers First Aid). This research was supported by a grant (2FD097-0297-CO2-01) from Fondo Europeo de Desarrollo Regional (FEDER); by a scholarship for PhD training from a national program to prepare university professors (FPU), Ministry of Education of Spain (AP2010-3198); and also partially supported by FEDER [CB06/02/0029] and Red Investigación Cardiovascular, HERACLES [RD12/0042]), and AGAUR (2014 SGR 240) and (2014 SGR 0334). The CIBERESP and CIBEROBN are initiatives of the Instituto de Salud Carlos III, Madrid, Spain. We apologize to all our colleagues whose work has not been cited for reasons of space limitations.

CONFLICTS OF INTEREST

None

ROLE OF THE FUNDING SOURCE

The funding sources had no involvement in the collection, analysis, and interpretation of data; in the writing of the report; or in the decision to submit the article for publication.
TABLE LEGENDS

Table 1: Consequences of maternal malnutrition during the Dutch famine (1944-1945) on the development of chronic diseases at adulthood of their descendants

AGE: Age at which the analysis was performed in descendants of mothers exposed to famine.

(*) When the age of the subjects is not stated, the year at which the study was published is indicated.

MATERNAL EXPOSURE TO HUNGER: gestation period in which mothers were exposed to famine.

(**) Late gestation: babies born between 7 January 1945 and 28 April 1945
Mid gestation: babies born between 29 April 1945 and 18 August 1945
Early gestation: babies born between 19 August 1945 and 8 December 1945

For the control groups not exposed to famine see each reference.

Table 2: Summary of selected histone post-translational modifications (acetylation or methylation) and their association to functional states of chromatin

Only individual labels are shown, not their combinations.

Bivalent domain: histone tag that represses developmentally expressed genes while maintaining them poised for activation upon differentiation.

TSS: Transcription Start Site

The code for histone labels is as follows: (H2B, 3, 4)-Histone 2B, 3 or 4, (K)-lysine, (ac)-acetylation, (me)-methylation. The number after the lysine code indicates the position modified, and me1, 2, 3 indicates mono-dio-tri methylations. This is a selection of significant histone labels and by no means an exhaustive inventory of all known histone post-translational modifications (PTMs).
FIGURE LEGENDS

Figure 1: New families of RNAs which control the flow of genetic information, and their functional interactions

Boxes show components of the original “Central Dogma of Molecular Biology”, which described the flow of genetic information from DNA to protein and considered mRNA essentially as an information-encoding, unstable intermediate [41, 42, 223]. This perspective has been updated by adding relevant families of non-coding RNAs, discovered recently, which are also transcribed from genomic DNA and have a functional role in the control of gene expression. These include the following: piRNAs, Piwi-interacting RNAs; snRNAs, small non-coding RNAs; lncRNAs, long non-coding RNAs; and miRNAs, micro RNAs. Solid arrows indicate transcription, the dotted arrow stands for translation, and the dashed arrows represent a regulatory relationship between two species of nucleic acids.

These relationships are described as follows:

A- microRNAs that regulate stability or translability of mRNAs [61];
B- lncRNAs that regulate mRNA processing, stability and degradation [224];
C- lncRNAs that can function as “miRNA sponges” to reduce their effective concentration [225];
D- miRNAs that target and downregulate lncRNAs [226];
E- antisense, AS-lncRNAs that regulate expression of neighboring genes, lncRNAs (as XIST) involved in imprinting and silencing of gene expression, or enhancer RNAs (eRNAs) that activate gene expression [79, 227, 228], among others;
F- generation of new regulatory regions by retrotransposition near other functional genes [229];
G- small RNAs involved in retroelement silencing in the germline [230].

This figure reviews only some of the most relevant non-coding RNA families, and is neither a systematic catalog of these elements nor of their functions. We have focused on the functional interactions between coding and non-coding RNAs and with DNA, not considering their mutual impacts on proteins.

Figure 2: Transcriptional regulation of the imprinted IgF2/H19 locus

Shown are the IgF2 and H19 genes as well as the most relevant controlling elements which regulate their parentally specific expression: the differentially methylated region...
(DMR) and the distal 3’ transcriptional enhancer. In this parentally imprinted locus the maternal chromosome keeps the DMR unmethylated so that it can bind the insulator zinc-finger protein CTCF, which blocks the communication between the IgF2 promoter and the distal enhancer [231], resulting in the transcriptional activation of the H19 gene. On the other hand, CTCF fails to bind to the heavily methylated DMR from the paternal chromosome, which facilitates the interaction between the distal enhancer and the IgF2 promoter to induce its expression.

**Figure 3: Chromatin as a nutrient sensor: GlcNAcylation of chromatin and epigenetic regulators links the nutritional status of the cell with mechanisms regulating gene expression**

Nutrient availability regulates levels of UDP-GlcNAc through the hexosamine biosynthetic pathway; this mechanism is controlled by a number of nutrient-sensing pathways (mTOR, AMPK, etc). GlcNAcylation of chromatin and other epigenetic regulators such as histones, polycomb group proteins (PcG), HDACs, TET (Ten-Eleven Translocation) proteins, and associated transcription factors (see text for details) have a global impact on the regulation of gene expression.
REFERENCES


[169] Li X. SIRT1 and energy metabolism. Acta Biochim Biophys Sin (Shanghai). 2013;45:51-60.


<table>
<thead>
<tr>
<th>CONSEQUENCES ON PRENATALLY EXPOSED DESCENDANTS</th>
<th>AGE*</th>
<th>MATERNAL EXPOSITION TO HUNGER**</th>
<th>REFERENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Significantly higher obesity rates</td>
<td>19</td>
<td>First half of gestation</td>
<td>[20]</td>
</tr>
<tr>
<td>Significantly lower obesity rates</td>
<td>19</td>
<td>Last trimester of gestation</td>
<td>[20]</td>
</tr>
<tr>
<td>Higher BMI and waist circumference in women but not in men</td>
<td>50</td>
<td>Early gestation</td>
<td>[23]</td>
</tr>
<tr>
<td>No effect of prenatal famine on systolic or diastolic blood pressure</td>
<td>50</td>
<td>Early or late gestation</td>
<td>[250]</td>
</tr>
<tr>
<td>Reduced concentrations of plasma factor VII</td>
<td>50</td>
<td>Early gestation</td>
<td>[251]</td>
</tr>
<tr>
<td>Significantly higher LDL-HDL cholesterol ratios</td>
<td>50</td>
<td>Early gestation</td>
<td>[252]</td>
</tr>
<tr>
<td>Higher prevalence of coronary heart disease</td>
<td>50</td>
<td>Early gestation</td>
<td>[253]</td>
</tr>
<tr>
<td>Decreased glucose tolerance</td>
<td>1998</td>
<td>Late gestation</td>
<td>[22]</td>
</tr>
<tr>
<td>Perceived poor health</td>
<td>50</td>
<td>Early gestation</td>
<td>[254]</td>
</tr>
<tr>
<td>Earlier onset of coronary artery disease</td>
<td>50</td>
<td>Early gestation</td>
<td>[28]</td>
</tr>
<tr>
<td>Higher rates of microalbuminuria</td>
<td>48-53</td>
<td>Mid gestation</td>
<td>[255]</td>
</tr>
<tr>
<td>Similarly decreased glucose tolerance at 50y and 58 y</td>
<td>50/58</td>
<td>All periods of gestation</td>
<td>[256]</td>
</tr>
<tr>
<td>No effect of exposure on adult mortality up to 57 years</td>
<td>57</td>
<td>All periods of gestation</td>
<td>[257]</td>
</tr>
<tr>
<td>Higher systolic and diastolic blood pressure after stress tests</td>
<td>58</td>
<td>Early gestation</td>
<td>[258]</td>
</tr>
<tr>
<td>Reduced insulin secretion (lower disposition index)</td>
<td>58</td>
<td>Mid gestation</td>
<td>[259]</td>
</tr>
<tr>
<td>Not associated to a greater prevalence of metabolic syndrome</td>
<td>58</td>
<td>All periods of gestation</td>
<td>[260]</td>
</tr>
<tr>
<td>Elevated total cholesterol, triglycerides and LDL-chol. In women only</td>
<td>58</td>
<td>All periods of gestation</td>
<td>[261]</td>
</tr>
<tr>
<td>No increase in CAD risk or Framingham risk</td>
<td>58</td>
<td>All periods of gestation</td>
<td>[30]</td>
</tr>
<tr>
<td>Reduced carotid and femoral arteries intima media thickness</td>
<td>58</td>
<td>All periods of gestation</td>
<td>[262]</td>
</tr>
<tr>
<td>Increased weight and adiposity in women but not in men</td>
<td>59</td>
<td>All periods of gestation</td>
<td>[263]</td>
</tr>
<tr>
<td>Less DNA methylation at the IGF2 gene compared with unexposed siblings</td>
<td>60</td>
<td>Early gestation</td>
<td>[119]</td>
</tr>
</tbody>
</table>

Table 1
<table>
<thead>
<tr>
<th>HISTONE LABEL</th>
<th>CHROMATIN FUNCTIONAL STATE</th>
<th>REFERENCES</th>
</tr>
</thead>
<tbody>
<tr>
<td>H2BK120ac</td>
<td>Active or poised TSS. Absent from core promoters.</td>
<td>[232]</td>
</tr>
<tr>
<td>H3K4ac</td>
<td>Marks promoters of actively transcribed genes</td>
<td>[233]</td>
</tr>
<tr>
<td>H3K9ac</td>
<td>Enriched at the promoters of actively transcribed genes</td>
<td>[234]</td>
</tr>
<tr>
<td>H3K14ac</td>
<td>Associated to active promoters</td>
<td>[235]</td>
</tr>
<tr>
<td>H3K27ac</td>
<td>Associated to active enhancers and promoters</td>
<td>[235, 236]</td>
</tr>
<tr>
<td>H3K36ac</td>
<td>Associated to active promoters</td>
<td>[237]</td>
</tr>
<tr>
<td>H3K4me1</td>
<td>Associate to enhancers of differentiation genes</td>
<td>[238]</td>
</tr>
<tr>
<td>H3K4me2</td>
<td>Associate to active chromatin</td>
<td>[239]</td>
</tr>
<tr>
<td>H3K4me3</td>
<td>Associated to actively transcribed genes</td>
<td>[240]</td>
</tr>
<tr>
<td>H3K9me1</td>
<td>Associate to enhancers of differentiation genes</td>
<td>[238]</td>
</tr>
<tr>
<td>H3K9me2</td>
<td>Repressive mark</td>
<td>[241]</td>
</tr>
<tr>
<td>H3K9me3</td>
<td>Repressive signal in gene-poor regions</td>
<td>[242]</td>
</tr>
<tr>
<td>H3K27me1</td>
<td>Associate to enhancers of differentiation genes</td>
<td>[238]</td>
</tr>
<tr>
<td>H3K27me2</td>
<td>Repressive signal mediated by polycomb repressive complex 2</td>
<td>[243]</td>
</tr>
<tr>
<td>H3K27me3</td>
<td>Temporary repressive signal in gene-rich regions Mediated by polycomb repressive complex 2</td>
<td>[242]</td>
</tr>
<tr>
<td>H4K8ac</td>
<td>Associate to active chromatin</td>
<td>[244]</td>
</tr>
<tr>
<td>H4K12ac</td>
<td>Associate to active chromatin</td>
<td>[245]</td>
</tr>
<tr>
<td>H4K16ac</td>
<td>Associate to active chromatin</td>
<td>[244]</td>
</tr>
<tr>
<td>H4K20me1</td>
<td>Repressive mark of facultative heterochromatin</td>
<td>[246]</td>
</tr>
<tr>
<td>H4K20me2</td>
<td>DNA damage response</td>
<td>[247]</td>
</tr>
<tr>
<td>H4K20me3</td>
<td>Constitutive heterochromatin</td>
<td>[248, 249]</td>
</tr>
</tbody>
</table>

Table 2
Figure 2

Hexosamine biosynthetic pathway:

- Glucose
- Acetyl-CoA
- UTP
- Glutamine
- ATP

UTP-GlcNAc

O-linked GlcNAc transferase (OGT)

Chromatin

GlcNAcylated chromatin

O-GlcNAcase (OGA)

GlcNAc

Nutrient sensing pathways:
- mTOR
- AMPK
- MAPK
- Insulin-AKT

Changes in gene expression
Figure(s)
Click here to download high resolution image

**FIGURE 3**

A diagram illustrating transcriptional activation of the *Igf2* gene on the paternal chromosome and silencing of the *H19* gene on the maternal chromosome. The diagram includes elements such as methylation, unmethylation, and the role of CTCF proteins in gene regulation.
piRNAs (sncRNAs) → DNA → mRNA → PROTEIN

RNA retroelements

IncRNAs

miRNAs

FIGURE 1