Minireview: Epigenetics of Obesity and Diabetes in Humans

Howard Slomko,* Hye J. Heo,* and Francine H. Einstein

Department of Pediatrics (H.S.), Division of Nephrology, and Obstetrics and Gynecology and Women’s Health (H.J.H., F.H.E.), Diabetes Research Center, Albert Einstein College of Medicine, Bronx, New York 10461

Understanding the determinants of human health and disease is overwhelmingly complex, particularly for common, late-onset, chronic disorders, such as obesity and diabetes. Elucidating the genetic and environmental factors that influence susceptibility to disruptions in energy homeostasis and metabolic regulation remain a challenge, and progress will entail the integration of multiple assessments of temporally dynamic environmental exposures in the context of each individual’s genotype. To meet this challenge, researchers are increasingly exploring the epigenome, which is the malleable interface of gene-environment interactions. Epigenetic variation, whether innate or induced, contributes to variation in gene expression, the range of potential individual responses to internal and external cues, and risk for metabolic disease. Ultimately, advancement in our understanding of chronic disease susceptibility in humans will depend on refinement of exposure assessment tools and systems biology approaches to interpretation. In this review, we present recent progress in epigenetics of human obesity and diabetes, existing challenges, and the potential for new approaches to unravel the complex biology of metabolic dysregulation. (Endocrinology 153: 1025–1030, 2012)

The obesity has reached epidemic proportions throughout the world, and the rapid rise in prevalence rates has made this a major focus of public health concern. The environment has inarguable impact on normal development and health throughout the lifespan. Some have suggested that the environment plays a role in nearly 85% of all diseases (1). Our modern living environment may even play a dominant role in the current epidemic of obesity and diabetes (2). Increased accessibility to low-cost food, the end of obligated daily physical activity for survival, and a growing reliance on technology are some of the relevant components of modern living (2–4). In addition to the importance of diet and activity, growing concern surrounds the unavoidable exposure to a wide range of man-made chemicals in industrialized countries. Other environmental exposures of interest can occur through ambient particles, water, food, and use of consumer or personal care products (5). Endocrine-disrupting chemicals, often used in the production of plastics and resins, are ubiquitous, and may interfere with insulin action, growth, and metabolic rate, among other physiological functions (6). Some chemical substances may have low-dose effects, meaning that the typical exposure levels, despite being below the Environmental Protection Agency’s standard toxicity testing, may have relevant biological effects (5).

Advances in sequencing techniques have revealed a new level of complexity with the use of metagenomic analysis to study the complex ecosystems of the human gut (7–9). Our internal ecosystem is adaptable, and continual shifts in phyla occur in response to changes in the host diet (10). Recent studies have shown profound changes in the composition and metabolic function of the gut microbiota in obese individuals (11–15). In turn, each host’s unique biological relationship with its gut microbiota may influence an individual’s risk of disease (16, 17).

* H.S. and H.J.H. contributed equally to this work.

Abbreviations: GWAS, Genome-wide association studies; VMR, variably methylated regions.
The scope of what constitutes our environment and complexity of understanding its biological impact on human health and disease is rapidly emerging. Despite the substantial advances that have been made in the ability to sequence and map the human genome, the other half of the gene-environment equation has been much more difficult to define, much less to quantify (18). Our inability to accurately assess relevant exposures prevents the clear delineation of each exposure’s specific contribution to disease or condition. The study of the epigenome using new approaches offers great hope in facing the daunting task of understanding complex gene-environment interactions.

**Epigenetics**

Epigenetics is the study of heritable changes in gene expression that occur without changes in DNA sequence (19). The genome is organized in a layered framework, the context of which influences accessibility and function. Each of the over 200 different human cell types has essentially the same genomic sequence in a given individual. The epigenome varies among different cell types and is more dynamic compared with the largely static genome (20). Cellular phenotype and responsiveness to external cues are governed through variations in DNA methylation, histone tail modifications, and chromatin binding. DNA methylation and histone modifications are two major epigenetic regulators in mammalian cells, which are functionally linked in transcription and may provide a mechanism for the stable propagation of gene activity from one generation of cells to the next (21). Gene and protein expression are also posttranscriptionally regulated by micro-RNA and other small RNA, which have greater temporal flux, making interpretation variability of these more formidable (22). Because exogenous influences can induce epigenetic modifications, epigenetic variation among individuals may be genetically or environmentally determined (23).

The term epigenetics was first coined by Conrad Waddington in 1942 to mean the study of causal mechanisms of development, which bring phenotype into being (24). In addition, he emphasized the importance of developmental processes, interrelatedness, and the dynamic system of genes and gene expression. Not surprisingly, the potential for adverse environmental impact on normal developmental is well recognized and perhaps best exemplified by the relationship between teratogens and congenital anomalies. More recently, a significant mass of epidemiological evidence has linked early life conditions and poor fetal growth with adult-onset diseases, such as cardiovascular disease and metabolic disorders. This hypothesis, described by David Barker and often referred to as the Barker hypothesis, arose from epidemiological observations linking low birth weight and risk for death from cardiovascular disease later in life (25). Since then, many others using different populations have made similar observations that early adverse conditions are associated with diabetes and metabolic dysfunction later in life (26). In addition, a large number of rodent and nonprimate animal models have been used to explore the phenomenon of developmental origins of adult disease [see review by Seki and colleagues (27) in this issue], many of which demonstrate phenotypic changes in offspring and tissue-specific changes in gene expression. Not only do these provide significant insight into the molecular changes associated with adverse intrauterine conditions, but they also implicate an epigenetic basis for common chronic disease.

Direct translation of findings from experimental animal studies is not assured, and confirmation in human subjects is not an effortless endeavor. To date, only a limited number of studies have documented direct evidence of epigenetic changes in association with suboptimal early life conditions in humans. Heijmans et al. (28) in a follow-up study of the Dutch Famine Cohort demonstrated subtle differences (−5% decrease) in methylation at the IGF2 differentially methylated region in individuals that had been exposed prenatally to maternal famine and their unexposed siblings. Changes in DNA methylation in umbilical cord blood cells have been documented, but these are either limited to global methylation changes or include only a limited number of human subjects (29–31). Using a restriction enzyme-based assay, our group has examined genome-wide changes in DNA methylation associated with intrauterine growth restriction compared with controls in a single population of multipotent hematopoietic stem cells also from a small group of neonates (32). Changes in DNA methylation were found in a restricted number of loci, including the hepatocyte nuclear factor 4α (HNF4A) gene, a well-known diabetes-associated gene. Using a large prospective cohort, Godfrey et al. (33) assessed DNA methylation in a set of candidate gene promoters using umbilical cords (tissue, not blood) and found positive associations between hypermethylation of RXRα and NOS3 with childhood adiposity at 9 yr of age. The interpretation of these findings in samples of mixed cell types and whether such changes have functional significance has come into question (34). Furthermore, in a fairly large cohort of mono- and dizygotic adolescent and middle-aged twin pairs, methylation levels of H19 and IGF2 differentially methylated regions in whole blood samples were more attributable to heritable factors and single-nucleotide polymorphisms, rather than environmental or stochastic events (35).
Potential hindrance to identification of direct links between epigenetic modification and environment in previous human studies can be partially attributed to use of tools that limit the evaluation of DNA methylation alterations to specifically cytosine that precede a guanine (CpG methylation). Previously, DNA methylation in mammalian cells was thought to occur only in the CpG regions. However, genome-wide, single-base-resolution maps of methylated cytosine from both human embryonic stem cells and fetal fibroblasts have identified non-CpG methylation in human cell lines (36). In embryonic stem cells, nearly one quarter of all methylation identified were in non-CG context, which disappeared upon induction of differentiation and was restored only in pluripotent stem cells. In skeletal muscle biopsies from both type 2 diabetes differentiation and in pluripotental stem cells makes them propensity to persist in embryonic stem cells before differentiation and is understood and requires further investigation. However, the functional role of non-CpG methylation is not well understood. Non-CpG methylation may provide greater clues to the direct linkage between genes, environment, and functional expression.

Other human studies have also identified personalized epigenomic signatures characterized by dynamic and stable variably methylated regions (VMR) that can be used as potential strategies for identifying patients at risk of common disease as well as for the identification of potential genomic regions of environmental vulnerability (38). In an attempt to identify VMR with covariation with body mass index, 74 random samples from the Age, Gene/Environment Susceptibility (AGES) study underwent comprehensive high-throughput array-based relative methylation analyses to compare 4.5 million CpG sites genome wide. Individuals included in the study were between 69 and 96 yr of age who had DNA samples obtained from two time points, about 11 yr apart. The study identified four VMR that showed covariation with body mass index consistently over a decade near genes previously implicated in regulation of body weight or diabetes. VMR are regions of extreme variability across individuals defined by 10 or more consecutive probes with an average SD (median absolute deviation) of more than 0.125. VMR are classified as stable when they remain static over time within individuals and as dynamic when they have high intra-individual differences. Stable VMR may represent the actual epigenetic changes associated with the disease process, whereas dynamic VMR may represent the potentially vulnerable regions that are more prone to environmental effects over a time course. Applying these findings in the setting of other disease processes may provide a better understanding of the epigenetic basis for developmental origins of obesity and metabolic disease.

**Existing Challenges to Studying the Epigenomics of Human Disease**

The assessment of environmental determinants of health and disease is clearly complex. Numerous endocrine-disrupting chemicals associated with increasing prevalence of obesity have been identified. These obesogens act by dysregulating lipid metabolism, basal metabolic rate, and regulation of appetite to promote obesity (6). These ubiquitous substances are found in plastics, personal care products, and food packaging and may be ingested, inhaled, or absorbed through the skin. For instance, epidemiological studies have shown a direct correlation between the increased presence of mono-benzyl and mono-ethylhexyl phthalates, a family of man-made compounds used in the manufacture of plastics, in urine and waist circumference in men (39, 40). Bisphenol A, another component of polycarbonate plastics and epoxy resins, has been shown to leach from the lining of food cans, baby bottles, dental sealants, and deposits, such that humans are routinely being exposed to these chemicals (6). Evidence of human exposure has been reported in urine, serum, breast milk, and maternal and fetal tissues (6). Animal studies have linked bisphenol A to numerous adverse health effects including impaired fertility and insulin resistance, and human studies are underway. Maternal exposure to several other chemical substances in pregnancy has been associated with increased body mass index in offspring (40–42). The mechanisms by which these chemicals influence the health of the offspring is not clear, but disruption of normal epigenetic regulation is likely to be involved.
Environmental health science is dedicated to the study of the impact of many environmental factors and their capacity for causation of disease. Since the first half of the 20th century, exposure-response relationships have been investigated for their potential role in occupation-related disease (43). Unfortunately, legal sanctions emerging from regulatory noncompliance with standards set by Occupational Safety and Health Administration and Environmental Protection Agency provides a great disincentive for companies to measure personal exposures. As a result, exposure science has shifted increasingly to predict exposure from models based on observational data, spatiotemporal determinants, or sampling of ambient air and water (43). Although the challenge of assessment and accurate measurement of environmental exposures is daunting, the nearly continuous development of innovative technologies offers great promise. The successful application of high-throughput technologies can be seen in application of DNA sequence analysis for genome-wide association studies (GWAS). GWAS have identified hundreds of single-nucleotide polymorphisms associated with many common diseases and traits (44). Although many of the loci identified are low-penetrant genes with low relative risk that may not be clinically relevant, results from GWAS are generally more highly regarded than environmental exposure studies revealing similar or even slightly higher relative risks. This has been attributed to the lower error rate of genotyping techniques and the low reproducibility of environmental exposure assessment in human populations (45).

Christopher Wild, a molecular epidemiologist, recently coined the term exposome to help envisage an equitable representation of both sides of the gene-environment interaction and to counterbalance the prevailing dominance of the genome (18). The exposome encompasses the totality of a lifetime of environmental exposures beginning at conception, which influence the internal cellular and chemical milieu of an individual. Low-level exposures, which may fluctuate of time, have the potential to exact their impact over long periods or have cumulative effects when combined with other exposures. Not only does the exposome vary among individuals, but also the influence of each exposure is construed within the context of an untold number of potential responses to that exposure as dictated by the individual’s genotype. However, unlike the high precision and reproducibility of genomic technologies, currently available methods for exposure assessment are rudimentary. Indeed, tools that enable accurate assessment of the exposome will have great implication in biomedical science and for the prediction of human disease.

For assessment of the cumulative effects of environmental exposures and prediction of individual disease susceptibility, epigenetic-based assays offer significant advantages as biomarkers if the challenges of study design, validation, and interpretation can be overcome. The evidence of an epigenetic basis for chronic adult disease, like obesity and diabetes, is limited not only by a number of practical issues related to study design (46) but also by difficulty in the interpretation of epigenetic variations. In 2005, Fraga et al. (47) published a landmark paper demonstrating increasing epigenetic discordance between monozygotic twins with advancing age. The epigenetic drift associated with aging is postulated to be a result of differences in environmental exposures. As a result, variation in epigenetic marks may be due to genotype, a countless number of environmental exposures, or stochastic events (20, 47, 48).

Biomarkers are used in many clinical settings to identify points between exposure and disease. Meaningful biomarkers are specific and sensitive and reliably mark a particular biological endpoint (49). Epigenetic assays that include measures of cellular toxicity, chromosomal alterations, and changes in expression can be used to measure an interval between low-dose exposure and disease onset (45). -Ome is a term that generally refers to the molecular techniques that generate a complete, or near-complete, set of biological molecules with high-throughput techniques (50, 51). These powerful tools provide comprehensive analysis of the cellular complement of specific constituents, such as DNA, RNA, proteins, intermediary metabolites, etc. (52). In the near future, integration of layered -omic technologies could allow for quantification of the effects of multiple, cumulative exposures and an individual’s biological responsiveness to those exposures. If these techniques could be leveraged together, they have the potential to provide tools that quantify an individual’s susceptibility to disease as well as predict their inherent protection against disease.

For epigenetic biomarkers to be a useful tool to distinguish points along the continuum of exposure to disease, normal or healthy will need to be defined. Several national and international consortia have begun this process. Encyclopedia of DNA Elements (ENCODE), the International HapMap Project, the 1000 Genomes Project, and the NIH Roadmap Epigenomics Mapping Consortium have been established to further the understanding of epigenetic features and decipher how they interact with genomic sequence to contribute to human health and disease (53). In particular, the NIH Roadmap Epigenomics Project has established several initiatives that will provide a public resource for epigenomic maps of normal human stem cells and primary tissues in addition to supporting technology development and funding research in epigenetic changes associated with specific disease.
(www.roadmapepigenomics.org). The creation of these groups represent a significant investment of resources and may provide the critical mass of epigenomic investigators needed to move the field forward.

-Omics-based research is frequently criticized for producing large amounts of uninterpretable data. The future of research involving obesity, diabetes, and gene-environment interaction will undoubtedly involve systems biology approaches. Systems biology is the comprehensive, quantitative analysis that integrates the manner in which all of the components of a biological system interact over time (54). A systems approach usually incorporates -omics-based assays with iterative measures and may include layering of multiple global sets of biological data. The large datasets generated are then used to construct new predictive models, which can be refined until they will allow for the prediction of the behavior of the system given any perturbation (55). Such models would enable a researcher or clinician to predict disease susceptibility or response to treatment or provide prognosis in a specific individual.

Conclusion

In conclusion, the environment has great biological impact on human health and disease, particularly in common complex disorders, like obesity and diabetes. Direct evidence linking specific environmental exposures and metabolic disease in humans is limited. Progress in this area has been hampered by the availability of rudimentary tools for exposure assessment. The challenge arises from the need to integrate the impact of fluidly changing environmental influences over time, which govern within the context of potential responses that are determined by the individual genotype. At the interface of gene-environment interactions lies the epigenome, which may provide an accessible recording of the exposome and provide insight into the origins of specific disease. Use of technologies that create complete biological datasets and the development of systems approaches to interpretation of complex iterations offer the hope for significant, paradigm changing discoveries in our understanding of common human disease.

Acknowledgments

Address all correspondence and requests for reprints to: Francine H. Einstein, M.D., Department of Obstetrics, Gynecology, and Women’s Health, Division of Maternal Fetal Medicine, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Belfer Building, Suite 512, Bronx, New York 10461. E-mail: francine.einstein@einstein.yu.edu.

The authors receive funding support from the National Institutes of Health (HD063791 to F.H.E.), the Reproductive Scientist Development Program (K12HD00849 to H.J.H.), the March of Dimes (to H.J.H.), and the Diabetes Research Center at Einstein (DK20541). The funders had no role in the decision to publish or preparation of the manuscript.

Disclosure Summary: The authors have no conflicts of interest or financial sponsorship to report.

References

17. Sartor RB 2004 Therapeutic manipulation of the enteric microflora
in inflammatory bowel diseases: antibiotics, probiotics, and prebiotics. Gastroenterology 126:1620–1633