VIEWPOINT

SCIENTIFIC DISCOVERY AND THE FUTURE OF MEDICINE

Epigenetics at the Crossroads of Genes and the Environment

Andrew P. Feinberg, MD, MPH
Center for Epigenetics, Johns Hopkins University School of Medicine, Baltimore, Maryland.

M. Daniele Fallin, PhD
Department of Mental Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland.

Corresponding Author: Andrew P. Feinberg, MD, MPH, Center for Epigenetics, Department of Medicine, Johns Hopkins University School of Medicine, 720 Rutland Ave, Ross 1064, Baltimore, MD 21205 (afeinberg@jhu.edu).

Epigenetics refers to information transmitted during cell division other than the DNA sequence per se, and it is the language that distinguishes stem cells from somatic cells, one organ from another, and even identical twins from each other. Examples include (1) DNA methylation, a covalent modification of the nucleotide cytosine, that is copied during cell division at CpG dinucleotides by the maintenance enzyme DNA methyltransferase I; (2) posttranslational modifications of nucleosome proteins about which the DNA double helix is wrapped; and (3) the density of nucleosomes and higher-order packaging of chromatin within the nucleus, including its relationship to the nuclear lamina.

In contrast to the DNA sequence, the epigenome is relatively susceptible to modification by the environment as well as stochastic perturbations over time, adding to phenotypic diversity in the population. A classic example is the environmentally modified phenotype of the *Agouti* gene, which regulates coat color and weight in mice and is manifest through epigenetic changes. A repetitive DNA variant within this gene can activate the gene in a nonregulated way, causing a yellow coat and obesity in the mice. This activation can be suppressed by DNA methylation, which can be modulated by supplying more or less dietary methionine, the essential amino acid that is the precursor of all DNA methylation in the cell.[1] Similar epigenetically mediated changes can be evoked via chemical exposure and even maternal behavior.[2] In humans, there has long been suspicion of the importance of nutrition in early life in modulating the epigenome. The classic epidemiological example of Överkalix, Norway, showed lower life expectancy for boys whose grandfathers had experienced famine prepubertally.[3] During the Dutch Hunger Winter and Great Leap Forward, which involved starvation of huge numbers of people, children exposed in utero to the famine during their first trimester show DNA methylation changes in genes associated with birth weight and low-density lipoprotein cholesterol.[4] A recent study of in utero nutritional deprivation in mice showed epigenetic changes continuing via the germline to the next generation but not beyond.[5]

For nearly all common diseases, both genetic predisposition and environmental influences shape risk, which typically increases with age. Even though both genetic and environmental exposures can be measured, there are currently limitations to how well they inform disease risk prediction or the underlying biological mechanisms leading to disease. Epigenetic marks on the genome may provide critical data to inform both prediction, in the age of precision medicine, and etiologic insight. This is because epigenetic marks are biologically related to both environmental exposure experience and to genes, and thus may be a measurable gauge of both genetic and environmental influence on disease risk. The interplay between genes, environment, epigenetics, and disease is complicated and still poorly understood. However, it is clear from both animal models and human studies that epigenetic marks such as DNA methylation can be modified by multiple types of environmental change, may be partially controlled by genetic variation, and certainly regulate gene expression, and thus how and when the genetic code is translated into biological action.

In the era of precision medicine, genetic risk scores are emerging as a potentially useful metric of risk. However, a risk metric based on inherited genes alone is static, because it does not incorporate age and environmental experience. Thus, most sophisticated approaches to risk prediction can include age, other demographic features, and specific environmental risk factors if known. This is challenging because many disease-specific environmental risk factors are not yet known, or the cumulative individual exposure (over time or across exposures) is not easily measured. Epigenetic marks can be used as biological measures of age,[6] and have shown promise as biomarkers of cumulative or specific exposure such as prenatal exposure to maternal cigarette smoking.[7] DNA methylation is stable in stored blood samples, and thus could have clinically feasible function for adding a metric of age and exposure risk to estimated inherited genetic risk to inform individual risk prediction.

The fields of genetic medicine and environmental health are not only concerned with risk prediction, but perhaps more importantly, with identifying and understanding causes of disease that can be acted upon for prevention and treatment. This is an important distinction because risk prediction does not necessarily need a specified mechanism. If a set of genetic markers can accurately predict which individuals will get (or already have) disease, and prediction is the goal, the mechanism of such genes, or whether they are simply proxy markers of some unmeasured mechanism, is irrelevant. However, if the goal is biological understanding of the disease, to inform prevention or treatment strategies, this information is critical. Epigenetic epidemiology is beginning to show utility in this regard. Some genetic associations with autoimmune disease have been shown to be mediated through epigenetic mechanisms. For example, a vari-

ant within a noncoding RNA that is thought to regulate genes in the major histocompatibility complex regulates the distribution of DNA methylation at that locus, but the susceptibility to rheumatoid arthritis is related to the methylation levels, an example of epigenetic mediation of a genetic variant.[8] Moreover, the location of epigenetic variants may be closer to the true mechanistic target and at considerable distance from the location of the disease-associated genetic variant.

More complicated scenarios connecting genes, environment, and epigenetic marks are possible and likely. For example, because epigenetic regulation controls when and where genes are expressed, a mutated gene that could cause disease is irrelevant in the context of normally low expression. However, that mutated gene may have adverse biological consequences if that regulatory context changes due to epigenetic modifications resulting from age or exposure. This would appear as an age-by-gene or exposure-by-gene interaction that could be mechanistically explained by epigenetics. It is also plausible that genetic variation contributes to inherent differences in epigenetic state between individuals, making some people more susceptible to environmental insults that would destabilize epigenetic regulation resulting in pathogenic biology.

A direct link between genetics, epigenetics, and the environment in disease risk can be seen in recent studies of diet-induced obesity and diabetes. Bl/6 mice fed a high-fat diet develop diet-induced obesity and diabetes very similar to human disease. These mice show widespread changes in DNA methylation in adipocytes and liver, and these same regulatory regions are highly conserved in humans, with similar changes in DNA methylation that are reversible by bariatric surgery.[9] Moreover, many of these genomic regions point to single nucleotide polymorphisms (SNPs) that show association with human diabetes but were opaque to previous genome-wide association study analysis because of multiple-testing correction for the whole genome when the environmental influence was not included in the analysis. In the smaller universe of genes epigenetically altered by diet, these SNPs show significant association. Previously identified relatively large-effect SNPs from traditional genetic association analysis are often members of the same biochemical or signaling pathways as the newly uncovered genetic/epigenetic diabetes genes.[9] Exercise is also strongly related to changes in DNA methylation, suggesting that the epigenome is highly dynamic in a way relevant to cardiovascular disease risk.[10]

The field of epigenetics and epigenetic epidemiology have much to do to improve measurement of epigenetic marks, inform natural variation in such marks, and the biological and population-level relationships between genes, environment, and epigenetics. This is an important emerging area as it holds promise for better risk prediction in precision medicine as well as for clarification of disease mechanisms among the existing opaque landscape only partially informed by traditional genetic and environmental studies to date.

ARTICLE INFORMATION

Conflict of Interest Disclosures: Both authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Dr Feinberg reports receipt of grants from the National Institutes of Health, licensed technology royalties and consulting fees from Orion Genomics outside the submitted work, and 3 issued patents: methods and kits for diagnosing and determination of the predisposition for diseases, methods for identifying cancer risk, and an epigenetic test for colorectal cancer risk. Dr Fallin reports no disclosures.

REFERENCES

1. Wolff GL, Kodell RL, Moore SR, Cooney CA. Maternal epigenetics and methyl supplements affect agouti gene expression in Avy/a mice. *FASEB J*. 1998;12(11):949-957.

2. Jirtle RL, Skinner MK. Environmental epigenomics and disease susceptibility. *Nat Rev Genet*. 2007;8(4):253-262.

3. Pembrey ME, Bygren LO, Kaati G, et al; ALSPAC Study Team. Sex-specific, male-line transgenerational responses in humans. *Eur J Hum Genet*. 2006;14(2):159-166.

4. Tobi EW, Goeman JJ, Monajemi R, et al. DNA methylation signatures link prenatal famine exposure to growth and metabolism. *Nat Commun*. 2014;5:5592.

5. Radford EJ, Ito M, Shi H, et al. In utero effects: in utero undernourishment perturbs the adult sperm methylome and intergenerational metabolism. *Science*. 2014;345(6198):1255903.

6. Horvath S. DNA methylation age of human tissues and cell types. *Genome Biol*. 2013;14(10):R115.

7. Joubert BR, Håberg SE, Nilsen RM, et al. 450K Epigenome-wide scan identifies differential DNA methylation in newborns related to maternal smoking during pregnancy. *Environ Health Perspect*. 2012;120(10):1425-1431.

8. Liu Y, Aryee MJ, Padyukov L, et al. Epigenome-wide association data implicate DNA methylation as an intermediary of genetic risk in rheumatoid arthritis. *Nat Biotechnol*. 2013;31(2):142-147.

9. Multhaup ML, Seldin MM, Jaffe AE, et al. Mouse-human experimental epigenetic analysis unmasks dietary targets and genetic liability for diabetic phenotypes. *Cell Metab*. 2015;21(1):138-149.

10. Barrès R, Yan J, Egan B, et al. Acute exercise remodels promoter methylation in human skeletal muscle. *Cell Metab*. 2012;15(3):405-411.