

REVIEW ARTICLE

DNA methylation in human diseases

Zelin Jin, Yun Liu*

*The Ministry of Education Key Laboratory of Metabolism and Molecular Medicine,
Department of Biochemistry and Molecular Biology, School of Basic Medical Sciences,
Fudan University, Shanghai, China*

Received 25 December 2017; accepted 9 January 2018

Available online 31 January 2018

KEYWORDS

DNA Methylation;
Epigenetic
epidemiologic
studies;
Epigenetics;
EWAS;
Human diseases

Abstract Even though the importance of epigenetics was first recognized in light of its role in tissue development, an increasing amount of evidence has shown that it also plays an important role in the development and progression of many common diseases. We discuss some recent findings on one representative epigenetic modification, DNA methylation, in some common diseases. While many new risk factors have been identified through the population-based epigenetic epidemiologic studies on the role of epigenetics in common diseases, this relatively new field still faces many unique challenges. Here, we describe those promises and unique challenges of epigenetic epidemiological studies and propose some potential solutions. Copyright © 2018, Chongqing Medical University. Production and hosting by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

The term “epigenetics” was coined by Conrad Hal Waddington, a British developmental biologist in 1942, to describe the “whole complex of developmental processes” linking genotype and phenotype.^{1,2} Since then, this concept has changed several times,^{3,4} and recent studies on epigenetics at the molecular level mainly cover changes in DNA methylation, histone modifications, non-coding RNAs, and higher-order chromatin structure. Even though the role of epigenetics was first recognized in development, an increasing amount of evidence has shown that it is also related to the development and progression of many

common diseases. Here, we first review the role of epigenetics, and mainly DNA methylation, in common human diseases.

Introduction to DNA methylation

DNA methylation is one of the earliest epigenetic modifications found in humans.⁵ It is a type of post-replication modification that often occurs in cytosines of the CpG dinucleotide sequence with the help of DNA methyltransferases (DNMTs),⁶ which transfer a methyl group from S-adenyl methionine to the fifth carbon of a cytosine residue to form 5-methylcytosine (5mC). The process of demethylation is more complex and can be passive or active. Ten-eleven translocation (TET) enzymes oxidize 5mCs^{7,8} and promote locus-specific removal of DNA methylation.⁹

* Corresponding author. Fax: +86 021 64033738.

E-mail address: yliu39@fudan.edu.cn (Y. Liu).

Peer review under responsibility of Chongqing Medical University.

Specifically, in the presence of water, oxygen, and α -keto-glutarate, 5mC becomes 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC), and 5-carboxylcytosine (5caC) through stepwise oxidation, yielding carbon dioxide and succinate.⁹ Then both 5fC and 5caC can be replaced by an unmodified cytosine through thymine DNA glycosylase-mediated base-excision repair.^{8,10}

Genomic imprinting

The role of DNA methylation in human diseases was first explored in the context of genomic imprinting. Genetic imprinting is an epigenetic phenomenon in which the maternal and paternal alleles are expressed in a parent-of-origin-specific manner. Genomic imprinting is stable and heritable during mitosis.¹¹ A relatively well-studied imprinting locus in humans is the *insulin-like growth factor 2 (IGF2)/H19* region, located on chromosome 11p15.5, in which DNA methylation on the imprinting control region (ICR) regulates the binding of zinc-finger CCCTC-binding factor (CTCF) and subsequent gene expression.¹² Loss of imprinting (LOI) causes many disorders such as Beckwith–Wiedemann Syndrome and cancer (which will be discussed later). Similarly, LOI of paternally-inherited chromosome 15q11.2–q13 is observed in Prader–Willi syndrome,¹³ and LOI of this region is also implicated in Angelman syndrome.¹⁴

DNA methylation and cancer

Aberrant DNA methylation of imprinted loci has been widely observed in many cancer types, such as colon, breast, liver, bladder, Wilms, ovarian, esophageal, prostate, and bone cancers. In addition, recent studies on the applications of omics technologies have shown that there are many differential DNA methylation related to cancers, such as hepatocellular carcinoma, glioblastoma, breast cancer, squamous cell lung cancer, thyroid carcinoma, and leukemia.^{15–25} Furthermore, DNMT mutations, different expression levels of DNMTs or as well as dysregulation of TETs are frequently observed in cancer, all of which suggest a strong connection between DNA methylation and cancer. There are many good reviews discussing the role of epigenetic changes in cancer (see Feinberg et al Nat. Rev. Genet. 2016²⁶; Timp et al Nat. Rev. Cancer 2013²⁷). Even though epigenetic alterations have not been recognized as a “Hallmark” of cancer yet,^{28,29} recent works point out that epigenetic changes in cancer over large chromatin regions can lead to epigenetic instability and the alteration of gene expression. While epigenetic changes can regulate stem cell reprogramming and cellular plasticity in normal development, disrupted epigenetic changes in cancer may eventually lead to tumor cell heterogeneity. Moreover, difference in epigenetic reprogramming has been observed between primary tumor and distant metastases in the same individual,³⁰ while, in contrast, no driver mutations were identified among the metastases.³¹ All these suggest that epigenetic dysregulation plays an important role in tumor development and metastasis, and also indicates its potential application in the diagnosis,^{32–34} prognosis³⁵ and the treatment of cancer.^{36,37}

DNA methylation and common diseases

In addition to genomic imprinting and cancer, the important role of DNA methylation in common human diseases has recently been revealed. In the subsequent sections, we discuss some recent findings on DNA methylation in these common human diseases using autoimmune diseases, metabolic disorders, psychological disorders, and aging as examples.

DNA methylation in autoimmune diseases

Autoimmune diseases are caused by low activity or over activity of the immune system. Immune cells are the primary targets in disease and are easily accessible from the blood, making them a relatively good model for studying the link between DNA methylation and phenotypes. Moreover, the low concordance rates in monozygotic twins for many of them such as rheumatoid arthritis (RA, 12.3–21%),^{38,39} systemic lupus erythematosus (SLE, 11.1–24.4%),^{40,41} and multiple sclerosis (MS, 16.7%)⁴² indicates that the environment or epigenetics may play important roles in disease etiology.

RA is a chronic autoimmune inflammatory disease with an estimated global prevalence of 0.3–1.0%.⁴³ It can cause symmetrical polyarthritis in small and large joints.⁴⁴ Genome-wide DNA methylation analysis of peripheral blood mononuclear cells showed that altered DNA methylation of human leukocyte antigen (HLA) class II can mediate the genetic risk of developing RA.⁴⁵ In adjuvant arthritic rats (animal models of RA), a high methylation level of the patched 1 (PTCH1) gene and decreased expression of PTCH1 protein activated the Hedgehog signaling pathway, leading to the increased secretion of interleukin 6 (IL-6) and tumor necrosis factor alpha.⁴⁶ In addition, the level of DNA methylation at diagnosis is associated with the response to disease-modifying anti-rheumatic drug treatment in early RA,⁴⁷ indicating that these sites may potentially be involved in the pathway associated with RA development. Moreover, our recent data showed that DNA methylation mediates the interaction between genotype and smoking in the development of RA, indicating that it can potentially integrate both internal genetic and external environmental risk factors.⁴⁸ This highlights the potential central role of DNA methylation in the development of RA.

SLE is an autoimmune disease in which the immune system mistakenly attacks healthy tissue in many parts of the body. A genome-wide assessment of DNA methylation identified differential DNA methylation in SLE patients in genes involved in autoantibody production.⁴⁹ In peripheral blood of SLE patients, low levels of DNA methylation was observed in the promoter of IL-6 gene.⁵⁰ Furthermore, the methylation status of naïve CD4⁺ T cells is related to different cutaneous manifestations and an investigation with neutrophils from SLE patients found DNA hypomethylation at the specific insulin receptor substrate for long interspersed nuclear element-1 (LINE-1).⁵¹

MS is the most common cause of neurological disability in young adults with a lifetime risk of 1 in 400.⁵² It is a chronic inflammatory and neurodegenerative disease that is self-immune mediated.⁵² Although the etiology of MS

remains largely unknown, genetic and environmental factors are thought to play a role. Changes in DNA methylation have been observed in multiple immune cell types in MS patients including CD4⁺ T cells, CD8⁺ T cells, and CD44⁺ encephalitogenic T cells,⁵³ and it is also altered in the hippocampi of patients with MS after demyelination.⁵⁴ Additionally, smoking, a risk factor for MS, can alter the DNA methylation level in MS patients through the interaction of major genetic risk factors (female and HLA risk haplotypes).⁵⁵ This indicates that, similar to RA,⁴⁸ DNA methylation can be influenced by gene and environment and may play a role in the pathogenesis of MS.

DNA methylation in metabolic disorders

Hyperglycemia (which can lead to type I and type II diabetes), hyperlipidemia (such as obesity-related conditions), and many diseases associated with these two phenomena, such as cardiovascular diseases, are currently a huge risk of death. Understanding the epigenetic mechanisms leading towards these conditions will have huge public health benefits. The symptoms of diabetes include increased thirst, frequent urination, and unexplained weight loss.⁵⁶ These are consequences of insulin resistance or insulin insensitivity, which lead to hyperglycemia. Type 2 diabetes mellitus (T2BM) is a complex disorder caused by a series of factors including genetics, epigenetics, and environmental factors. A comparison of the methylation level of islets in patients with T2BM with those from normal donors showed that a series of sites had different DNA methylation levels which led to differential gene expression.⁵⁷ That included the decreased methylation of cyclin-dependent kinase inhibitor 1A and phosphodiesterase 7B promoters, which demonstrates impaired glucose-stimulated insulin secretion both *in vitro* and in humans.⁵⁷ Shortened leukocyte telomere length was also proposed to affect the DNA methylation level of LINE-1, which may increase the risk of T2BM in Chinese patients.⁵⁸ Methylation of thioredoxin-interacting protein is sensitive to glucose concentrations, and its hypomethylation has been observed in patients with hyperglycemia.⁵⁹ Obesity is another common disease in which the environment plays an important role. Genome-wide analysis showed that increased DNA methylation of hypoxia-inducible factor 3 alpha in both blood cells and adipose tissue is associated with an increased body mass index (BMI).⁶⁰ Offspring DNA methylation of Aryl-Hydrocarbon Receptor Repressor (AHRR) correlates with maternal BMI, gestational age, and birth weight.⁶¹ In addition, the DNA methylation level of two proteins important for fat metabolism, leptin and adiponectin, is positively correlated with BMI and waist girth,⁶² whereas the methylation level of the leptin gene in blood cells is negatively correlated with BMI.⁶² It has also been proposed that low-density lipoprotein cholesterol (LDL-C) may be a key regulator of the DNA methylation profile of leptin and adiponectin in adipose tissues. Changes in genotype-dependent DNA methylation have been observed in the promoter of the actin-related protein 2/3 complex subunit 3, a gene related to lipogenesis and lipid accumulation, and this was proposed to be a risk factor for obesity-related disorders.⁶³ Moreover, studies have shown the correlation

between DNA methylation and many metabolic traits such as high-density levels of lipoprotein cholesterol,^{64–66} LDL,⁶⁷ and triglycerides.⁶⁷

DNA methylation in neurological disorders

Neurological disorders are other types of diseases in which DNA methylation is believed to play an important role in development. This correlation was suggested when frequent mutations in the Methyl-CpG-binding protein 2 (MeCP2) gene was observed in patients with autism spectrum disorder (ASD)⁶⁸ and Rett Syndrome,^{69,70} a neurodevelopmental disorder affecting about 1 in 10,000 female babies. MeCP2 binds to DNA sequences with methylated CpG and methylated CH (H refers to A, C, or T), which is key to regulating neuronal gene expression *in vivo*.^{71,72} It has been proposed that changes in the DNA methylation profile lead to the differential expression of genes related to synaptic activity,⁷³ and that the different binding affinity of MeCP2 protein to methylated DNA can regulate the expression level of genes related to the pathophysiology of autism⁷⁴ and Rett Syndrome such as brain derived neurotrophic factor.⁷⁵ In addition to genetic mutations related to the epigenetic regulation of gene expression, studies have shown that some environmental factors (such as exposure to garden pesticide) can affect DNA methylation levels in placental tissue, which put children at high risk for ASD.⁷⁶ A genome-wide study measuring neuron-specific methylation situations in ASD identified some differentially methylated regions between the ASD and control groups.⁷⁷ Combined genetic and epigenetic measurements also revealed both known and new pathways involved in ASD that had not previously been made known by genetic findings alone.⁷⁸ The role of DNA methylation in other mental disorders has also been explored. For example, the methylation level of catechol-O-methyl transferase in peripheral blood was lower in patients with schizophrenia (SZ) than in healthy controls in a Malaysian study,⁷⁹ which seemed to be affected by the severity of the clinical symptoms as well as the pharmacological treatment.⁷⁹ Moreover, crosstalk between different epigenetic signals including DNA methylation, histone modifications, and micro RNAs was also suggested in the development of SZ.^{80,81} In addition, some genome-wide studies found that differentially methylated CpGs are strongly related to the development of Parkinson's disease (PD), either from the cerebral cortex,⁸² or from blood and saliva.⁸³ The DNA was significantly hypomethylated in the promoter regions of α -synuclein and parkin in early-onset PD patients.⁸⁴ Although relatively successful, most of these studies were performed on surrogate tissues, such as blood, or in a relatively small sample size, due to difficulties in obtaining enough samples from neuropathological examinations (especially brain tissues), and large brain biobanks are needed for validation and replication of these findings in the neural bio specimens.

DNA methylation and aging

Aging is a process of becoming old with many physical and psychological consequences. Even though it is a natural process, it can be biologically classified as a disease.⁸⁵ Many

years ago, age-related DNA methylation changes were observed in several species such as salmon,⁸⁶ rats,⁸⁷ and mice.⁸⁸ In recent years, work from Steve Horvath's group showed that measuring DNA methylation levels at 353 CpG sites can be a good estimator of the biological age of a tissue, cell type, organ, or even individuals.^{89,90} This type of DNA clock is known as an epigenetic or DNA methylation clock, and has proved to be a good age predictor in many heterogeneous tissues such as whole blood, PMBCs, cerebellar samples, occipital cortex, buccal epithelium, colon, adipose, liver, lung, saliva, and uterine cervix, as well as in individual cell types such as CD4 T cells and CD14 monocytes (myeloid lineage).^{89,90} Interestingly, in contrast with normal people, accelerated DNA methylation age is found in individuals with diseases such as PD,⁹¹ Huntington's disease,⁹² and Alzheimer's disease,⁹³ as well as some viral infections.⁹⁴ All these findings suggest that DNA methylation may not only be suitably viewed as a biomarker of aging, but may also be related the pathological process of aging itself.

Challenges in epigenetic epidemiological studies focused on DNA methylation

With the abundance of emerging evidence indicating the important role of DNA methylation in common diseases, researchers have attempted to use DNA methylation as a biomarker to identify epigenetic changes that are associated with disease status. It is believed that this population-based epidemiologic approach for studying the role of epigenetics in common diseases can be used to identify new risk factors that might be missed by conventional genetic epidemiologic approaches such as the genome-wide association study (GWAS); this epigenetic epidemiologic approach is also known as the epigenome-wide association study (EWAS). Even though epigenetic epidemiological studies hold much promise⁹⁵ and are conceptually relatively simple, there are in fact, many unique challenges to this approach. Here, we describe some of them and propose some potential solutions.

Since DNA methylation profiles differ for each tissue, it is understandable that it is more relevant to perform experiments on primary disease-affected tissues in EWAS. Even though this may not be difficult for autoimmune diseases, as immune cells are easily accessible from blood, it can be quite challenging for many other diseases such as neurological disorders in which human biopsies of target organs from healthy controls are difficult to obtain, especially when a large number of samples are needed for epidemiology studies. Many studies have proposed to use blood as a surrogate tissue under the assumption that some epigenetic changes in target tissue can also be observed in the blood; however, the suitability of using a surrogate tissue is still a subject of debate.^{96,97} Although disease-associated epigenetic changes observed in surrogate tissue such as blood may be useful as a biomarker in certain clinical settings, to prove the biological importance of these associations, it is necessary to validate these changes in disease-related tissue. This also highlights the importance of establishing biobanks that not only collect blood, as is done for genetic epidemiology studies, but also collect tissues directly involved in disease development.

The second challenge in performing many EWAS studies is that the starting materials used to generate DNA methylation profiles are usually derived from heterogeneous cell populations. For example, DNA samples readily available from most genetic epidemiology studies are usually from whole blood, which includes many different cell types with distinct biological functions. Moreover, the proportion of each cell type is variable among individuals and can be affected by disease status, making it a potential confounding factor in association analyses. Many statistical approaches have been proposed to perform cell-type deconvolution and identify epigenetic changes that are not the result of a shift in cell type. Some reference-based approaches have used unique DNA methylation signatures from each cell type to estimate their proportions and then performed subsequent adjustments with these estimate.⁴⁵ Obviously, this type of approach depends on the availability of the reference DNA methylation profiles of each cell type, and to date, has been applied in PBMC,^{45,98,99} cord blood,^{100,101} and the prefrontal cortex.¹⁰² Additionally, this approach assumes that the DNA methylation signature used to estimate cellular proportions is not affected by the disease status. Because of these limitations, several reference-free approaches such as EWASher,¹⁰³ ReFACTOR,¹⁰⁴ RefFreeEWAS,¹⁰⁵ RUVm,¹⁰⁶ and SVA^{107,108} have been proposed. However, the performance of these reference-free approaches strongly depends on the validity of model assumptions.¹⁰⁹

The third challenge is that because DNA methylation can be influenced by both genetic factors and environmental changes, its role in disease etiology can be rather complicated (Fig. 1). This is straightforward if disease phenotype is only affected by genotype (as in Mendelian disorders), or the environment (as in infectious diseases). However, in the development of common diseases, DNA methylation can be a mediator, modifier, or even consequence of the disease. For example, as a mediator, DNA methylation can be an intermediary in the disease etiology of genetic^{45,98} or environmental risk factors.^{110,111} In some cases, DNA methylation can even act as an integrator of both internal genetic and external environmental risk factors in disease development, as demonstrated in RA⁴⁸ and MS⁵⁵. As a modifier, DNA methylation can also influence the relationship between genotype (or environment) and disease phenotype. Even though this relationship has not been explored so far, the gene-environment interaction (GxE) has been widely observed in the development of many diseases, and it is not surprising that DNA methylation plays a role in this process. Additionally, because DNA methylation is relatively dynamic and can be influenced by many factors (e.g., age, gender, race, life style, living conditions), they are all potential confounders that should be considered in the epidemiology study. All these make it difficult to build one simple model to study the potential roles of DNA methylation, and so careful experimental design is vital to any successful epigenetic epidemiology study.

Another challenge for epigenetic epidemiology studies is to establish the causal role of epigenetic changes. While we are interested in the way changes in DNA methylation play a role in disease etiology, DNA methylation can also be affected by the disease itself, and this needs to be

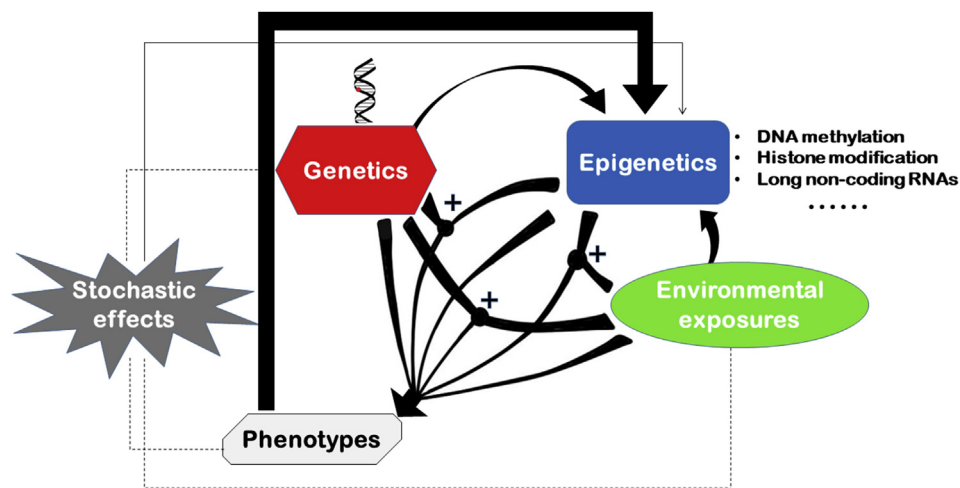


Figure 1 The relationship between genetics, epigenetics and environmental exposures on disease phenotype.

distinguished from the biology of the disease. Unlike GWAS, this is another unique challenge in performing EWAS. In addition to using molecular biochemistry experiments (e.g., *in vitro* luciferase reporter assays, manipulating methylation levels on targeted CpGs,¹¹² etc.) to directly test the biological role of methylation changes, other methods have been proposed. For example, this challenge can be overcome with a longitudinal experiment design, because epigenetic alterations that occur before or in the very early stage of the disease process are more likely to be related to disease etiology. In addition, some statistical approaches such as causal inference test (CIT)^{45,98} and Mendelian randomization (MR)¹¹³ have also been proposed and applied to investigate the causal role of epigenetic changes.^{113–115} Both approaches use genetic variants as an instrument to infer the causal effect of DNA methylation on the outcome. Because the mediation-based CIT test and MR approach use genetic instruments in different ways, each method has its own advantages as well as some disadvantages.¹¹⁶

Conflicts of interest

The authors declare no conflicts of interest.

Acknowledgement

This work was supported by from National Basic Research Program of China (973 program) (No. 2015CB910401 to Yun Liu) and the National Natural Science Foundation of China (No. 31471212 and No. 31771451 to Yun Liu).

References

1. Waddington CH. The epigenotype. *Endeavour*. 1942;1:18–20.
2. Waddington CH. *The Strategy of the Genes: A Discussion of Some Aspects of Theoretical Biology*. Ruskin House/George Allen and Unwin Ltd. London: London George Allen and Unwin Ltd.; 1957.
3. Deichmann U. Epigenetics: the origins and evolution of a fashionable topic. *Dev Biol*. 2016;416(1):249–254.
4. Pujadas E, Feinberg AP. Regulated noise in the epigenetic landscape of development and disease. *Cell*. 2012;148(6):1123–1131.
5. Jaenisch R, Bird A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat Genet*. 2003;33(suppl):245–254.
6. Moore LD, Le T, Fan G. DNA methylation and its basic function. *Neuropsychopharmacology*. 2013;38(1):23–38.
7. Kohli RM, Zhang Y. TET enzymes, TDG and the dynamics of DNA demethylation. *Nature*. 2013;502(7472):472–479.
8. Pastor WA, Aravind L, Rao A. TETonic shift: biological roles of TET proteins in DNA demethylation and transcription. *Nat Rev Mol Cell Biol*. 2013;14(6):341–356.
9. Wu H, Zhang Y. Reversing DNA methylation: mechanisms, genomics, and biological functions. *Cell*. 2014;156(1-2):45–68.
10. Zhang W, Xia W, Wang Q, et al. Isoform switch of TET1 regulates DNA demethylation and mouse development. *Mol Cell*. 2016;64(6):1062–1073.
11. Peters J. The role of genomic imprinting in biology and disease: an expanding view. *Nat Rev Genet*. 2014;15(8):517–530.
12. Yang Y, Hu J-F, Ulaner GA, et al. Epigenetic regulation of Igf2/H19 imprinting at CTCF insulator binding sites. *J Cell Biochem*. 2003;90(5):1038–1055.
13. Cassidy SB, Schwartz S, Miller JL, Driscoll DJ. Prader-Willi syndrome. *Genet Med*. 2012;14(1):10–26.
14. Williams CA, Driscoll DJ, Dagli AI. Clinical and genetic aspects of Angelman syndrome. *Genet Med*. 2010;12(7):385–395.
15. McCann AH, Miller N, O'Meara A, et al. Biallelic expression of the IGF2 gene in human breast disease. *Hum Mol Genet*. 1996;5(8):1123–1127.
16. Cui H, Onyango P, Brandenburg S, Wu Y, Hsieh C-L, Feinberg AP. Loss of imprinting in colorectal cancer linked to hypomethylation of H19 and IGF2. *Cancer Res*. 2002;62(22):6442–6446.
17. Cui H, Cruz-Correa M, Giardiello FM, et al. Loss of IGF2 imprinting: a potential marker of colorectal cancer risk. *Science*. 2003;299(5613):1753–1755.
18. Ulaner GA, Vu TH, Li T, et al. Loss of imprinting of IGF2 and H19 in osteosarcoma is accompanied by reciprocal methylation changes of a CTCF-binding site. *Hum Mol Genet*. 2003;12(5):535–549.
19. Murphy SK, Huang Z, Wen Y, et al. Frequent IGF2/H19 domain epigenetic alterations and elevated IGF2 expression in epithelial ovarian cancer. *Mol Cancer Res*. 2006;4(4):283–292.

20. Byun H-M, Wong H-L, Birnstein EA, Wolff EM, Liang G, Yang AS. Examination of IGF2 and H19 loss of imprinting in bladder cancer. *Cancer Res.* 2007;67(22):10753–10758.
21. Jelinic P, Shaw P. Loss of imprinting and cancer. *J Pathol.* 2007;211(3):261–268.
22. Zhao R, DeCoteau JF, Geyer CR, Gao M, Cui H, Casson AG. Loss of imprinting of the insulin-like growth factor II (IGF2) gene in esophageal normal and adenocarcinoma tissues. *Carcinogenesis.* 2009;30(12):2117–2122.
23. Vu TH, Nguyen AH, Hoffman AR. Loss of IGF2 imprinting is associated with abrogation of long-range intrachromosomal interactions in human cancer cells. *Hum Mol Genet.* 2010;19(5):901–919.
24. Bhusari S, Yang B, Kueck J, Huang W, Jarrard DF. Insulin-like growth factor-2 (IGF2) loss of imprinting marks a field defect within human prostates containing cancer. *Prostate.* 2011;71(15):1621–1630.
25. Leick MB, Shoff CJ, Wang EC, Congress JL, Gallicano GI. Loss of imprinting of IGF2 and the epigenetic progenitor model of cancer. *Am J Stem Cells.* 2012;1(1):59–74.
26. Feinberg AP, Koldobskiy MA, Göndör A. Epigenetic modulators, modifiers and mediators in cancer aetiology and progression. *Nat Rev Genet.* 2016;17(5):284–299.
27. Timp W, Feinberg AP. Cancer as a dysregulated epigenome allowing cellular growth advantage at the expense of the host. *Nat Rev Cancer.* 2013;13(7):497–510.
28. Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell.* 2000;100(1):57–70.
29. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell.* 2011;144(5):646–674.
30. McDonald OG, Li X, Saunders T, et al. Epigenomic reprogramming during pancreatic cancer progression links anabolic glucose metabolism to distant metastasis. *Nat Genet.* 2017;49(3):367–376.
31. Makohon-Moore AP, Zhang M, Reiter JG, et al. Limited heterogeneity of known driver gene mutations among the metastases of individual patients with pancreatic cancer. *Nat Genet.* 2017;49(3):358–366.
32. Guo S, Diep D, Plongthongkum N, Fung H-L, Zhang K, Zhang K. Identification of methylation haplotype blocks aids in deconvolution of heterogeneous tissue samples and tumor tissue-of-origin mapping from plasma DNA. *Nat Genet.* 2017;49(4):635–642.
33. Li W, Zhang X, Lu X, et al. 5-Hydroxymethylcytosine signatures in circulating cell-free DNA as diagnostic biomarkers for human cancers. *Cell Res.* 2017;27(10):1243–1257.
34. Song C-X, Yin S, Ma L, et al. 5-Hydroxymethylcytosine signatures in cell-free DNA provide information about tumor types and stages. *Cell Res.* 2017;27(10):1231–1242.
35. Lapinska K, Faria G, McGonagle S, Macumber KM, Heerboth S, Sarkar S. Cancer progenitor cells: the result of an epigenetic event? *Anticancer Res.* 2018;38(1):1–6.
36. Byler S, Goldgar S, Heerboth S, et al. Genetic and epigenetic aspects of breast cancer progression and therapy. *Anticancer Res.* 2014;34(3):1071–1077.
37. Choi SJ, Jung SW, Huh S, Chung Y-S, Cho H, Kang H. Alteration of DNA methylation in gastric cancer with chemotherapy. *J Microbiol Biotechnol.* 2017;27(8):1367–1378.
38. Aho K, Koskenvuo M, Tuominen J, Kaprio J. Occurrence of rheumatoid arthritis in a nationwide series of twins. *J Rheumatol.* 1986;13(5):899–902.
39. Bellamy N, Duffy D, Martin N, Mathews J. Rheumatoid arthritis in twins: a study of aetiopathogenesis based on the Australian Twin Registry. *Ann Rheum Dis.* 1992;51(5):588–593.
40. Deapen D, Escalante A, Weinrib L, et al. A revised estimate of twin concordance in systemic lupus erythematosus. *Arthritis Rheum.* 1992;35(3):311–318.
41. Järvinen P, Aho K. Twin studies in rheumatic diseases. *Semin Arthritis Rheum.* 1994;24(1):19–28.
42. Bammer H, Schaltenbrand G, Solcher H. Examinations of twins in multiple sclerosis. *Dtsch Z Nervenheilkd.* 1960;181:261–279.
43. Murphy D, Hutchinson D. Is male rheumatoid arthritis an occupational Disease? A review. *Open Rheumatol J.* 2017;11:88–105.
44. Majithia V, Geraci SA. Rheumatoid arthritis: diagnosis and management. *Am J Med.* 2007;120(11):936–939.
45. 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.
46. Sun Z-H, Liu Y-H, Liu J, et al. MeCP2 regulates PTCH1 expression through DNA methylation in rheumatoid arthritis. *Inflammation.* 2017;40(5):1497–1508.
47. Glossop JR, Nixon NB, Emes RD, et al. DNA methylation at diagnosis is associated with response to disease-modifying drugs in early rheumatoid arthritis. *Epigenomics.* 2017;9(4):419–428.
48. Meng W, Zhu Z, Jiang X, et al. DNA methylation mediates genotype and smoking interaction in the development of anti-citrullinated peptide antibody-positive rheumatoid arthritis. *Arthritis Res Ther.* 2017;19(1):71.
49. Chung SA, Nititham J, Elboudwarej E, et al. Genome-wide assessment of differential DNA methylation associated with autoantibody production in systemic lupus erythematosus. *PLoS One.* 2015;10(7):e0129813.
50. Cai XY, Lu Y, Tang C, et al. Effect of interleukin-6 promoter DNA methylation on the pathogenesis of systemic lupus erythematosus. *Zhonghua Yi Xue Za Zhi.* 2017;97(19):1491–1495.
51. Sukapan P, Promnarate P, Avihingsanon Y, Mutirangura A, Hirankarn N. Types of DNA methylation status of the interspersed repetitive sequences for LINE-1, Alu, HERV-E and HERV-K in the neutrophils from systemic lupus erythematosus patients and healthy controls. *J Hum Genet.* 2014;59(4):178–188.
52. Compston A, Coles A. Multiple sclerosis. *Lancet (London, England).* 2002;359(9313):1221–1231.
53. Li X, Xiao B, Chen X-S. DNA methylation: a new player in multiple sclerosis. *Mol Neurobiol.* 2017;54(6):4049–4059.
54. Chomyk AM, Volsko C, Tripathi A, et al. DNA methylation in demyelinated multiple sclerosis hippocampus. *Sci Rep.* 2017;7(1):8696.
55. Marabita F, Almgren M, Sjöholm LK, et al. Smoking induces DNA methylation changes in multiple sclerosis patients with exposure-response relationship. *Sci Rep.* 2017;7(1):14589.
56. GBD 2015 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet (London, England).* 2016;388(10053):1545–1602.
57. Dayeh T, Volkov P, Salö S, et al. Genome-wide DNA methylation analysis of human pancreatic islets from type 2 diabetic and non-diabetic donors identifies candidate genes that influence insulin secretion. *PLoS Genet.* 2014;10(3):e1004160.
58. Wu Y, Cui W, Zhang D, Wu W, Yang Z. The shortening of leukocyte telomere length relates to DNA hypermethylation of LINE-1 in type 2 diabetes mellitus. *Oncotarget.* 2017;8(43):73964–73973.
59. Soriano-Tárraga C, Jiménez-Conde J, Giralte-Steinhauer E, et al. Epigenome-wide association study identifies TXNIP gene associated with type 2 diabetes mellitus and sustained hyperglycemia. *Hum Mol Genet.* 2016;25(3):609–619.

60. Dick KJ, Nelson CP, Tsaprouni L, et al. DNA methylation and body-mass index: a genome-wide analysis. *Lancet (London, England)*. 2014;383(9933):1990–1998.
61. Burris HH, Baccarelli AA, Byun H-M, et al. Offspring DNA methylation of the aryl-hydrocarbon receptor repressor gene is associated with maternal BMI, gestational age, and birth weight. *Epigenetics*. 2015;10(10):913–921.
62. Houde A-A, Légaré C, Biron S, et al. Leptin and adiponectin DNA methylation levels in adipose tissues and blood cells are associated with BMI, waist girth and LDL-cholesterol levels in severely obese men and women. *BMC Med Genet*. 2015;16:29.
63. de Toro-Martín J, Guénard F, Tchernof A, et al. A CpG-SNP located within the ARPC3 gene promoter is associated with hypertriglyceridemia in severely obese patients. *Ann Nutr Metab*. 2016;68(3):203–212.
64. Guay S-P, Voisin G, Brisson D, et al. Epigenome-wide analysis in familial hypercholesterolemia identified new loci associated with high-density lipoprotein cholesterol concentration. *Epigenomics*. 2012;4(6):623–639.
65. Guay S-P, Légaré C, Brisson D, et al. Epigenetic and genetic variations at the TNNT1 gene locus are associated with HDL-C levels and coronary artery disease. *Epigenomics*. 2016;8(3):359–371.
66. Braun KVE, Dhana K, de Vries PS, et al. Epigenome-wide association study (EWAS) on lipids: the Rotterdam Study. *Clin Epigenet*. 2017;9:15.
67. Verloigne M, Loyen A, Van Hecke L, et al. Variation in population levels of sedentary time in European children and adolescents according to cross-European studies: a systematic literature review within DEDIPAC. *Int J Behav Nutr Phys Act*. 2016;13:69.
68. Iossifov I, O’Roak BJ, Sanders SJ, et al. The contribution of de novo coding mutations to autism spectrum disorder. *Nature*. 2014;515(7526):216–221.
69. Amir RE, Van den Veyver IB, Wan M, Tran CQ, Francke U, Zoghbi HY. Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nat Genet*. 1999;23(2):185–188.
70. Cuddapah VA, Pillai RB, Shekar KV, et al. Methyl-CpG-binding protein 2 (MECP2) mutation type is associated with disease severity in Rett syndrome. *J Med Genet*. 2014;51(3):152–158.
71. Chen L, Chen K, Lavery LA, et al. MeCP2 binds to non-CG methylated DNA as neurons mature, influencing transcription and the timing of onset for Rett syndrome. *Proc Natl Acad Sci U S A*. 2015;112(17):5509–5514.
72. Correction for Chen et al. MeCP2 binds to non-CG methylated DNA as neurons mature, influencing transcription and the timing of onset for Rett syndrome. *Proc Natl Acad Sci U S A*. 2015;112(22):E2982.
73. Fasolino M, Zhou Z. The Crucial Role of DNA Methylation and MeCP2 in Neuronal Function. *Genes (Basel)*. 2017;(5):8.
74. Liu Z, Li X, Zhang J-T, et al. Autism-like behaviours and germline transmission in transgenic monkeys overexpressing MeCP2. *Nature*. 2016;530(7588):98–102.
75. Jin X-R, Chen X-S, Xiao L. MeCP2 deficiency in neuroglia: new progress in the pathogenesis of Rett syndrome. *Front Mol Neurosci*. 2017;10:316.
76. Schmidt RJ, Schroeder DI, Crary-Dooley FK, et al. Self-reported pregnancy exposures and placental DNA methylation in the MARBLES prospective autism sibling study. *Environ Epigenetics*. 2016;4(2).
77. Nardone S, Sams DS, Zito A, Reuveni E, Elliott E. Dysregulation of cortical neuron DNA methylation profile in autism spectrum disorder. *Cereb Cortex*. 2017;27(12):5739–5754.
78. Andrews SV, Ellis SE, Bakulski KM, et al. Cross-tissue integration of genetic and epigenetic data offers insight into autism spectrum disorder. *Nat Commun*. 2017;8(1):1011.
79. Nour El Huda AR, Norsidah KZ, Mohd Nabil Fikri R, Hanisah MN, Kartini A, Norlelawati AT. DNA methylation of MB-COMT in Malaysian schizophrenia patients. *Psychiatry Clin Neurosci*. November 2017. <https://doi.org/10.1111/pcn.12622>.
80. Shorter KR, Miller BH. Epigenetic mechanisms in schizophrenia. *Prog Biophys Mol Biol*. 2015;118(1-2):1–7.
81. Ovenden ES, McGregor NW, Emsley RA, Warnich L. DNA methylation and antipsychotic treatment mechanisms in schizophrenia: progress and future directions. *Prog Neuro-Psychopharmacol Biol Psychiatry*. 2018;81:38–49.
82. Dashtipour K, Tafreshi A, Adler C, et al. Hypermethylation of Synphilin-1, alpha-synuclein-interacting protein (SNCAIP) gene in the cerebral cortex of patients with sporadic Parkinson’s disease. *Brain Sci*. 2017;(7):7.
83. Chuang Y-H, Paul KC, Bronstein JM, Bordelon Y, Horvath S, Ritz B. Parkinson’s disease is associated with DNA methylation levels in human blood and saliva. *Genome Med*. 2017;9(1):76.
84. Eryilmaz IE, Cecener G, Erer S, et al. Epigenetic approach to early-onset Parkinson’s disease: low methylation status of SNCA and PARK2 promoter regions. *Neurol Res*. 2017;39(11):965–972.
85. Bulterijis S, Hull RS, Björk VCE, Roy AG. It is time to classify biological aging as a disease. *Front Genet*. 2015;6:205.
86. Berdyshev GD, Korotaev GK, Boiarskikh GV, Vaniushin BF. Nucleotide composition of DNA and RNA from somatic tissues of humpback and its changes during spawning. *Biokhimiia*. 1967;32(5):988–993.
87. Vanyushin BF, Nemirovsky LE, Klimenko VV, Vasiliev VK, Belozersky AN. The 5-methylcytosine in DNA of rats. Tissue and age specificity and the changes induced by hydrocortisone and other agents. *Gerontologia*. 1973;19(3):138–152.
88. Wilson VL, Smith RA, Ma S, Cutler RG. Genomic 5-methyldeoxycytidine decreases with age. *J Biol Chem*. 1987;262(21):9948–9951.
89. Horvath S. DNA methylation age of human tissues and cell types. *Genome Biol*. 2013;14(10):R115.
90. Horvath S. Erratum to: DNA methylation age of human tissues and cell types. *Genome Biol*. 2015;16:96.
91. Horvath S, Ritz BR. Increased epigenetic age and granulocyte counts in the blood of Parkinson’s disease patients. *Aging (Albany NY)*. 2015;7(12):1130–1142.
92. Horvath S, Langfelder P, Kwak S, et al. Huntington’s disease accelerates epigenetic aging of human brain and disrupts DNA methylation levels. *Aging (Albany NY)*. 2016;8(7):1485–1512.
93. Levine ME, Lu AT, Bennett DA, Horvath S. Epigenetic age of the pre-frontal cortex is associated with neuritic plaques, amyloid load, and Alzheimer’s disease related cognitive functioning. *Aging (Albany NY)*. 2015;7(12):1198–1211.
94. Horvath S, Levine AJ. HIV-1 infection accelerates age according to the epigenetic clock. *J Infect Dis*. 2015;212(10):1563–1573.
95. Bakulski KM, Fallin MD. Epigenetic epidemiology: promises for public health research. *Environ Mol Mutagen*. 2014;55(3):171–183.
96. Hannon E, Lunnon K, Schalkwyk L, Mill J. Interindividual methylomic variation across blood, cortex, and cerebellum: implications for epigenetic studies of neurological and neuropsychiatric phenotypes. *Epigenetics*. 2015;10(11):1024–1032.
97. Farré P, Jones MJ, Meaney MJ, Emberly E, Turecki G, Kobor MS. Concordant and discordant DNA methylation signatures of aging in human blood and brain. *Epigenet Chromatin*. 2015;8:19.
98. Hong X, Hao K, Ladd-Acosta C, et al. Genome-wide association study identifies peanut allergy-specific loci and evidence of epigenetic mediation in US children. *Nat Commun*. 2015;6:6304.

99. Joubert BR, Felix JF, Yousefi P, et al. DNA methylation in newborns and maternal smoking in pregnancy: genome-wide consortium meta-analysis. *Am J Hum Genet.* 2016;98(4):680–696.
100. Cardenas A, Allard C, Doyon M, et al. Validation of a DNA methylation reference panel for the estimation of nucleated cells types in cord blood. *Epigenetics.* 2016;11(11):773–779.
101. Bakulski KM, Feinberg JI, Andrews SV, et al. DNA methylation of cord blood cell types: applications for mixed cell birth studies. *Epigenetics.* 2016;11(5):354–362.
102. Guintivano J, Aryee MJ, Kaminsky ZA. A cell epigenotype specific model for the correction of brain cellular heterogeneity bias and its application to age, brain region and major depression. *Epigenetics.* 2013;8(3):290–302.
103. Zou J, Lippert C, Heckerman D, Aryee M, Listgarten J. Epigenome-wide association studies without the need for cell-type composition. *Nat Methods.* 2014;11(3):309–311.
104. Rahmani E, Zaitlen N, Baran Y, et al. Sparse PCA corrects for cell type heterogeneity in epigenome-wide association studies. *Nat Methods.* 2016;13(5):443–445.
105. Houseman EA, Molitor J, Marsit CJ. Reference-free cell mixture adjustments in analysis of DNA methylation data. *Bioinformatics.* 2014;30(10):1431–1439.
106. Maksimovic J, Gagnon-Bartsch JA, Speed TP, Oshlack A. Removing unwanted variation in a differential methylation analysis of Illumina HumanMethylation450 array data. *Nucleic Acids Res.* 2015;43(16):e106.
107. Leek JT, Storey JD. Capturing heterogeneity in gene expression studies by surrogate variable analysis. *PLoS Genet.* 2007;3(9):1724–1735.
108. Teschendorff AE, Zhuang J, Widschwendter M. Independent surrogate variable analysis to deconvolve confounding factors in large-scale microarray profiling studies. *Bioinformatics.* 2011;27(11):1496–1505.
109. Teschendorff AE, Zheng SC. Cell-type deconvolution in epigenome-wide association studies: a review and recommendations. *Epigenomics.* 2017;9(5):757–768.
110. Pogribny IP, Rusyn I. Environmental toxicants, epigenetics, and cancer. *Adv Exp Med Biol.* 2013;754:215–232.
111. Barrès R, Yan J, Egan B, et al. Acute exercise remodels promoter methylation in human skeletal muscle. *Cell Metab.* 2012;15(3):405–411.
112. Pulecio J, Verma N, Mejía-Ramírez E, Huangfu D, Raya A. CRISPR/Cas9-Based engineering of the epigenome. *Cell Stem Cell.* 2017;21(4):431–447.
113. Hemani G, Tilling K, Davey Smith G. Orienting the causal relationship between imprecisely measured traits using GWAS summary data. *PLoS Genet.* 2017;13(11):e1007081.
114. Dekkers KF, van Iterson M, Slieker RC, et al. Blood lipids influence DNA methylation in circulating cells. *Genome Biol.* 2016;17(1):138.
115. Morales E, Vilahur N, Salas LA, et al. Genome-wide DNA methylation study in human placenta identifies novel loci associated with maternal smoking during pregnancy. *Int J Epidemiol.* 2016;45(5):1644–1655.
116. Teschendorff AE, Relton CL. Statistical and integrative system-level analysis of DNA methylation data. *Nat Rev Genet.* 2018;19(3):129–147.