New insights and updated guidelines for epigenome-wide association studies ☆

Lisa H. Chadwick [a], Akira Sawa [b], Ivana V. Yang [c], Andrea Baccarelli [d], Xandra O. Breakefield [e], Hong-Wen Deng [g], Dana C. Dolinoy [f], M. Daniele Fallin [h], Nina T. Holland [i], E. Andres Houseman [j], Stavros Lomvardas [k], Mahendra Rao [l], John S. Satterlee [m], Frederick L. Tyson [n], Pandurangan Vijayanand [o], John M. Greally [p,*]

[a] Division of Extramural Research and Training, National Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, NC, USA
[b] Department of Psychiatry, Johns Hopkins University, Baltimore, MD, USA
[c] Department of Medicine, University of Colorado Denver, Aurora, CO, USA
[d] Department of Environmental Health, Harvard School of Public Health, Boston, MA, USA
[e] Department of Neurology, Massachusetts General Hospital, Boston, MA, USA
[f] Department of Environmental Health Sciences, University of Michigan, Ann Arbor, MI, USA
[g] Department of Biostatistics and Bioinformatics, Tulane University School of Public Health and Tropical Medicine, New Orleans, LA, USA
[h] Department of Mental Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA
[i] Division of Environmental Health Sciences, School of Public Health, University of California Berkeley, Berkeley, CA, USA
[j] Department of Biostatistics, College of Public Health and Human Sciences, Oregon State University, Corvallis, OR, USA
[k] Department of Anatomy, University of California San Francisco, San Francisco, CA, USA
[l] Laboratory of Stem Cell Biology, National Institute of Arthritis and Musculoskeletal and Skin Diseases, National Institutes of Health, Bethesda, MD, USA
[m] Division of Basic Neuroscience and Behavioral Research, National Institute on Drug Abuse, National Institutes of Health, Bethesda, MD, USA
[n] Division of Intramural Research, National Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, NC, USA
[o] La Jolla Institute for Allergy and Immunology, La Jolla, CA, USA
[p] Center for Epigenomics, Albert Einstein College of Medicine, Bronx, NY, USA

ARTICLE INFO

ABSTRACT

Epigenetic dysregulation in disease is increasingly studied as a potential mediator of pathophysiology. The epigenetic events are believed to occur in somatic cells, but the limited changes of DNA methylation in studies to date indicate that only subsets of the cells tested undergo epigenetic dysregulation. The recognition of this subpopulation effect indicates the need for care in design and execution of epigenome-wide association studies (EWASs), paying particular attention to confounding sources of variability. To maximize the sensitivity of the EWASs, ideally, the cell type mediating the disease should be tested, which is not always practical or ethical in human subjects. The value of using accessible cells as surrogates for the target, disease-mediating cell type has not been rigorously tested to date. In this review, participants in a workshop convened by the National Institutes of Health update EWAS design and execution guidelines to reflect new insights in the field.

Epigenetic regulatory mechanisms play a crucial role in normal human development, in part by establishing and maintaining the gene expression programs necessary for cells to perform their unique functional roles. In recent years, there has been growing interest in investigating whether changes in these epigenetic programs contribute to the development of a variety of complex human diseases. Several lines of evidence suggest that this might be the case. These include the potential for the epigenome to mediate environmental influences (reviewed in Cortessis et al., 2012; Hou et al., 2012; Perera & Herbstman, 2011; Reamon-Buettner et al., 2008) or the maintenance of memory of events that occurred in the past, including prenatal exposures influencing adult disease susceptibility (reviewed in Barouki et al., 2012; Gluckman et al., 2011; Warner & Ozanne, 2010; Waterland & Michels, 2007). The field of cancer epigenomics has established a precedent for the silencing or activation of genes being causally involved in neoplasia (Dawson & Kouzarides, 2012; Esteller, 2007), somatic events that are usually limited to the cells or tissue in which the cancer arose, with a few notable exceptions (Cui et al., 2003; Gaston et al., 2001). Unlike the genome, the

epigenome is inherently malleable from a biochemical perspective, and the potential to reverse deleterious epigenetic events has been seen as a major opportunity, especially in the treatment of cancer (Griffiths & Gore, 2013; Popovic & Licht, 2012; Rius & Lyko, 2012).

The field of epigenome-wide association studies (EWASs) is now very active, testing a wide variety of human diseases and other phenotypes. These EWASs have almost all been based on the study of DNA methylation, an epigenetic regulator that is less demanding in terms of clinical sample acquisition than alternatives such as sequencing of DNA from chromatin immunoprecipitation, which maps histone modifications or other chromatin constituents. A consistent outcome of the EWAS studies to date has been the observation of moderate changes in DNA methylation between disease and normal groups, and not a switch between fully unmethylated and methylated states at a given genomic locus. As DNA methylation exists in a binary state at a specific location on an individual allele, moderate changes have to reflect allelic and cell subpopulation changes between the tested groups, an epigenetic mosaicism that may be of pathophysiological significance if subsets of cells can reasonably be proposed to mediate the organ's disease state.

This emerging observation of cell subpopulation effects has forced a re-evaluation of how we should approach EWAS design and execution. With limited effect sizes associated with the disease or phenotype, it is necessary to pay increased attention to other sources of variability potentially affecting the study. It is also essential that we invest our efforts in a cell type that is likely to manifest the differences sought. Epigenomic dysregulation associated with human disease is generally thought of in terms of somatic rather than germline events, raising the issue whether the cell type(s) mediating the disease (target cell type) needs to be sampled, or whether a more accessible surrogate cell type can yield sufficiently useful information.

The influence of cell type on epigenetic variability is highlighted by large-scale epigenomic mapping efforts such as that led by the National Institutes of Health (NIH) Roadmap Epigenomics Program (Bernstein et al., 2010). There are other influences on epigenome-wide studies to consider, such as technical influences, DNA sequence polymorphism, and human subject characteristics (age, sex, exposure history), all potentially confounding the ability to recognize genuine effects associated with a disease or phenotype and leading, in some cases, to misinterpretation of the results obtained.

In this review, we describe the broader issue of rigorous EWAS design and execution, updating prior excellent EWAS overviews (Bell & Spector, 2011; Mill & Heijmans, 2013; Rakyan et al., 2011; Satterlee et al., 2010; Verma, 2012), with an emphasis on the specific issue of the value of studies based on surrogate cell use, and define 3 areas of research priority that could further improve our design and interpretation of EWASs.

Updated guidelines for rigorous EWASs

Taking a cue from the history of genome-wide association studies (GWASs), it is essential that rigorous standards are developed for EWASs, which are likely to be more complex than GWASs, involving many different types of epigenetic regulatory mechanisms, cell types, and likely confounding influences. As a starting point, we provide an overview of updated suggested best practices for EWASs, described below and in Fig. 1, building upon prior recommendations for how best to design and perform such studies (Bell & Spector, 2011; Rakyan et al., 2011; Satterlee et al., 2010; Verma, 2012).

Begin with an explicit biological hypothesis: Linking epigenetic changes to disease or phenotypic causation, or as a marker or mediator of environmental exposure, should be stated in terms of an explicit hypothesis. Although there is clearly value to exploratory or pilot studies to test whether a certain phenotype or condition has any evidence for nonrandomly associated epigenetic changes, when performing a definitive study, it is essential to define the underlying hypothetical mechanism involving epigenetic perturbation. Having such a clear hypothesis allows the experimental design and, in particular, the analysis and interpretation to be focused productively. For example, the timing of sample collection or the choice of cross-sectional versus longitudinal studies and the rationale for choosing a specific surrogate cell type for study will be dictated to a major extent by the underlying hypothesis.

Purified cell types are preferable for epigenetic studies: It is highly desirable that single cell types are used where possible. Although it has been appreciated for some time that different cell types have distinctive epigenetic profiles (Shen et al., 2012; Varley et al., 2013), it has recently been confirmed that the presence of different proportions of cell subtypes in mixed populations of cells can generate distinctive DNA methylation profiles (Houseman et al., 2012), which has been found for some loci to explain up to 40% of DNA methylation differences between individuals (Adalsteinsson et al., 2012). If there is a systematic bias in cell subpopulation composition between the groups of individuals being tested (e.g., between disease states, exposures, or phenotypes), DNA methylation assays performed will identify differences between groups, but these will not necessarily represent altered epigenetic patterns within each cell type associated with the disease. It is therefore preferable that pure cell samples be tested when feasible, often a difficult issue when cell numbers from purified cell samples yield less material than can usually be tested in genome-wide assays. It should be recognized that "pure" does not mean homogeneous in terms of function, so purification does not eliminate the possibility of cell subtype composition exerting an influence. However, the selection of purified cells ensures that as similar cell types as possible are compared between groups and makes it easier for other groups to reproduce the experiment, both valuable considerations.

If it is not possible to purify cells, it should be attempted to account for subpopulation effects in statistical models, either through analysis of the samples collected (for example, histological studies to quantify cell proportions or measurement of differential white blood cell count in leukocyte-based studies) or through the development of new analytical techniques that use DNA methylation signatures to measure cell subpopulation proportions (Houseman et al., 2012). Examples of this kind of approach were published recently (Guintivano et al., 2013; Liu et al., 2013), demonstrating a major reduction in association signals after adjustment for the estimated cell proportions in the blood samples tested, emphasizing both the potential for variability in cell subpopulations to exert strong effects on the DNA methylation signals as well as the potential for sophisticated analytical approaches to account for these effects.

Target versus surrogate cell types: An underlying assumption in many EWASs is that epigenetic changes associated with a particular disease are likely acquired in somatic cells during development or during aging (as opposed to through the germline or extremely early in development). It follows that these epigenetic changes may not be observed in all cell types in the body. The choice of cell type is thus of great importance in human disease studies. The ideal situation would be to acquire those cells directly affected by or mediating the disease, cells we refer to as the *target* cell type, purified to the greatest extent possible. For many diseases, obtaining such samples can be very challenging. As an example, target cell types in disorders of the central nervous system generally can only be studied in postmortem specimens.
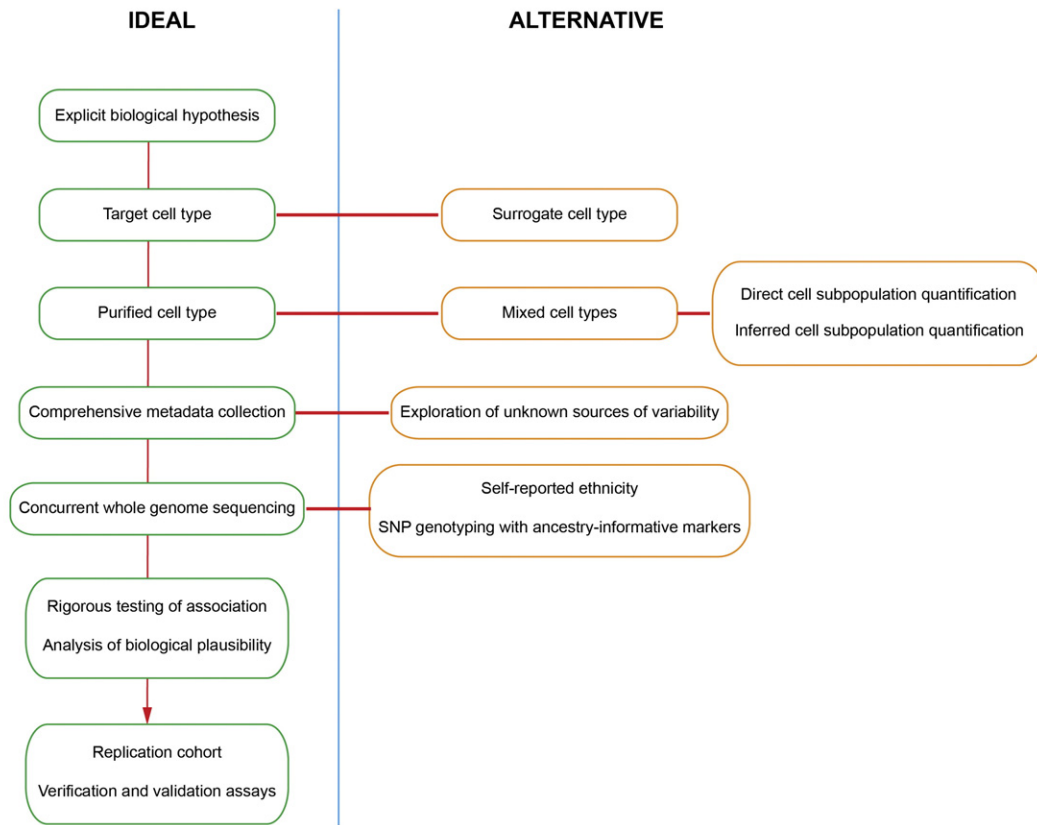
Epigenome-wide asociation study design

Fig. 1. On the left are shown components of an ideal EWAS design, with alternative approaches on the right that represent some of the compromises that often have to be made when dealing with studies involving human subjects.

Although there are some preliminary indications that chromatin-modifying enzymes maintain their activities in postmortem brain (Monoranu et al., 2011), the experience with expression profiling indicates that we need, at a minimum, to be alert to issues like sample acquisition and storage, and for information about tissue pH, agonal factors (Atz et al., 2007), and pharmacological treatments. It should also be noted that even with more accessible organs, procedures that are performed on adults to collect target cells are often less frequently performed on children, a good example being bronchoscopy of the lung and sampling of airway epithelium.

Prompted by these difficulties and by the availability of existing samples such as blood from cohorts of subjects with diseases of interest and/or well-characterized environmental exposures, the question arises whether surrogate cells, defined as "nontarget" readily accessible cell types, share the epigenetic dysregulation of the target cell type in a disease (or other) state of interest and are thus potentially informative for disease studies. There is no clear answer to this question at present. What is apparent is that there is a need for studies to test whether such an approach is feasible, potentially informative, and capable of generating enough of a "signal" of disease association that can be discriminated from the "noise" of the multiple variables that can confound these studies.

Sources of variability need to be understood: There are many potential sources of variability that can affect EWAS data interpretation, including patient or disease issues (age (Heyn et al., 2012), sex (Sarter et al., 2005), and medications (Gonzalez-Fierro et al., 2011; Junien, 2006) or exposure histories), sample collection

issues, nucleic acid purification protocols (Soriano-Tarraga et al., 2013), influences inherent to the experimental assays performed, and even the version and type of software analytical tools used to process and interpret the data generated. To account for these influences, it is recommended that metadata (data describing the information collected and generated) also be collected systematically and comprehensively. At a minimum, variables known to affect epigenetic marks such as age and sex must be collected as a part of the metadata. In addition, technical variables associated with experimental protocols also must be collected to account for technical variability or "batch effects" that are present in both array- and sequence-based data (Leek et al., 2010). If the timing of an exposure is potentially critical in terms of a susceptibility window, this information needs also to be collected. Finally, extensive disease phenotype data collected at this stage can provide valuable insight into the epigenetic regulation of that disease state. In the data analysis stage, these potential sources of variability can be explored and their influences can be accounted for in statistical modeling (Leek et al., 2010) to avoid misleading, artifactual results.

The role of underlying DNA sequence variability has to be determined: It is now appreciated that DNA sequence is not completely inert in terms of the capability to establish epigenetic patterns locally. Some of these influences are reasonably intuitive, such as the polymorphism of the target sequence for a DNA-binding protein being associated with failure of its binding to the site (Spivakov et al., 2012), whereas others are mechanistically less clear, such as single nucleotide polymorphisms (SNPs) being associated with generally increased or decreased DNA

methylation in the flanking hundreds of base pairs (Gertz et al., 2011) or over even greater genomic distances (Gibbs et al., 2010; Liu et al., 2013). Unrecognized polymorphisms of the CG dinucleotides tested in a DNA methylation study can also lead to misleading results, prompting some analytical approaches to filter out overlapping common SNPs (Wang et al., 2012). At a minimum, self-reported ethnicity should be part of the metadata captured and balanced between groups as part of the study design, or handled through analytic adjustment, recognizing that this information is frequently misleading (Choudhry et al., 2007). Some ancestral information can be gleaned from the SNPs included explicitly, as well as SNPs in probes, of microarray data such as those from the Illumina 450 K microarray. Direct genotyping approaches using SNP microarrays (including ancestry-informative markers) or even whole genome sequencing are preferable, as this will allow more robust testing of the effect of sequence variability.

Attention needs to be paid to the analytical resources and approaches: Given the sizes of the data sets generated, the complexity of the information when combined with a full range of metadata, and the broad range of alternatives in exploring these data, it is essential that a comprehensive data analysis plan is assembled and put in place with adequate hardware and expertise resources before initiating these studies.

Define in advance how results will be interpreted: As mentioned earlier, it also needs to be defined in advance of a project's initiation what will constitute a meaningful outcome to the study. Although it is essential that the associations between epigenetic variability and phenotype (disease or environmental factor) survive rigorous statistical testing, it is also recommended that the results be evaluated in terms of biological plausibility. Statistically robust results may reasonably constitute a new insight but lack biological plausibility in terms of the starting hypothesis. Such cases should be treated with a degree of skepticism and should be the focus of especially careful testing of the results generated, as described next.

Test the conclusions rigorously: All studies, whether positive or negative in terms of the criteria above, need to undergo validation following completion and analysis of the epigenome-wide assays. It should be emphasized that epigenome-wide assays are inherently less quantitative and accurate than single-locus studies, so until such single-locus studies are performed, the results of the epigenome-wide assays can only be regarded as preliminary and suggestive. The values generated by the epigenome-wide study should be technically analyzed to verify that the assays were consistently generating values reflecting the underlying biological reality. Such analysis should involve the use of a gold-standard approach for a range of values, such as a sodium bisulphite sequencing approach for DNA methylation, exploring a range of values generated by the epigenome-wide assay. The loci found to be distinctive between experimental groups by the epigenome-wide assay should also be tested individually using the gold-standard technique to confirm that these candidates are indeed distinctive epigenetically between groups.

Finally, it is desirable that the results be replicated in a separate cohort of individuals, whether as part of the original study (Joubert et al., 2012) or as a separate project. These measures will lend the study the necessary confidence that the results will continue to be reproducible and are reflective of genuine underlying biological changes. Although independent replication is generally favored over a split sample approach, such independent replication cohorts will be inherently more challenging to assemble in EWASs than for GWASs, given the potential for the confounding variables of environment, genotype, and others mentioned earlier to influence the results. This being the case, a second, validation cohort will have to be studied as comprehensively as the initial test cohort, blurring the lines between the split and independent cohort paradigms.

Experiences need to be shared: Finally, in this young field, it is essential that all insights and experience be shared fully so that the best practices, biological insights, and analytical approaches are disseminated effectively and contribute to the continuous improvement of research design, analysis, and interpretation in the field. Even negative results or the effects of strong confounding variables need to be documented to help guide subsequent studies and to ensure that subsequent epigenomic projects are designed to be as rigorous as possible. Beginning at the stage of obtaining consent for study participation, attention should be paid to ensuring that DNA sequence data can be deposited into appropriate NIH or other public databases (e.g., database of Genotypes and Phenotypes, Gene Expression Omnibus) to make them accessible to other investigators.
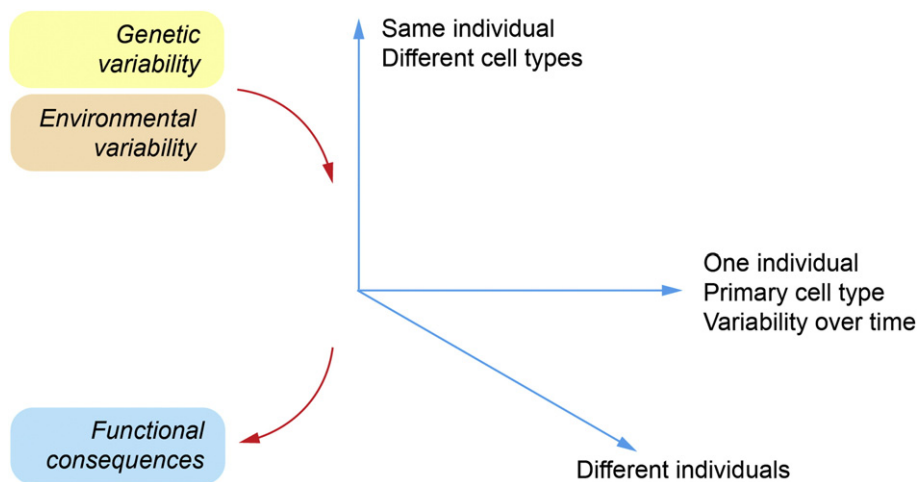

Fig. 2. It needs to be determined whether epigenetic variability occurs over time in the same individual in the same cell type, whether different cells from that same individual will have comparable patterns of variability, and whether different individuals will have comparable or distinctive patterns in terms of degree and the loci involved. This all has to be considered in parallel with the genetic variability between individuals that may be influential, as well as environmental influences that can potentially modify the epigenome. Finally, the functional consequences of this epigenetic variability need also to be measured to understand their significance.

The need for pilot studies

The discussion above highlights deficiencies in our understanding of influences that could affect EWASs and indicates that some foundational studies may be needed to allow rational EWAS design. There is a clear need to test whether surrogate cells can be used in EWASs. A different cell type to the target cell type affected by the disease may have distinctive patterns of epigenetic regulatory marks such as DNA methylation (Varley et al., 2013) and chromatin constituents (Shen et al., 2012), probably related causally to differences in transcriptional patterns between the cell types. The issue to be tested is whether, even in this different epigenetic background, the surrogate cell type has distinctive epigenetic patterns in subjects with a disease of interest compared with subjects who do not have the disease and whether these epigenetic changes are reflective of those occurring in the target cell type. Even in a situation when the disease-associated epigenetic differences in the surrogate cells are not identical to those found in the target tissues, the finding of differences may still be of significant value as a clinical application, such as a diagnostic test or a biomarker of outcome or exposure. This is easier to test in animal models, in which longitudinal studies can be performed; different cell types can be readily sampled at necropsy; and control can be exerted over environment, age, genotype, and the timing of exposures. There is, however, no guarantee that the physiological response of the animal predicts that of humans, so whether the findings have human disease significance remains uncertain even after successful execution of the animal-based project (Seok et al., 2013). These caveats aside, animal models have the potential to provide us with valuable insight into how specific, controlled exposures affect epigenetic marks in target and surrogate tissues.

The third area of emphasis is the need to understand epigenetic variability, an emerging issue that needs to be taken into account in EWASs, including those testing cell surrogacy. There are several types of epigenetic variability to consider (Fig. 2), of which interindividual variability for epigenomic patterns for the same cell type is only recently being recognized and studied (Gemma et al., 2013; Yuen et al., 2009). Variability may occur over time with repeated sampling of a specific cell type from an individual using a longitudinal approach (intraindividual variability or "epigenetic drift" across time). Temporal variability is a complicating factor in EWASs that is not encountered in GWASs and because of the typical case-control design represents a cross-sectional approach that, in individuals already manifesting the disease of interest, could be complicated by issues of reverse causality.

At present, we do not know the extent of epigenomic variability, the proportion of loci affected, where in the genome variability tends to occur, whether early environmental exposures affect variability, whether this variability changes during the life of the subject, or whether different cell types have differing patterns of variability. It is clear that these issues need to be addressed so that researchers can design EWASs appropriately, and to interpret reference data accurately. For example, the goal of the NIH Roadmap Epigenomics Program was to develop a resource of reference maps to characterize cell type–defined variability but not population-level variability or individual variability across time. Consequently, loci with low variability cannot be distinguished from those with high variability, with resulting problems interpreting these reference maps in terms of confidence for each locus.

Conclusions

The field of EWASs is growing, with a broad range of diseases and other phenotypes being studied in human subjects. Just as GWASs evolved by consensus to adopt best practices, a similar process needs to occur for epigenomic studies. In many respects, EWASs are more complicated than GWASs in terms of the types and magnitude of data generated, the somatic as opposed to typically germline nature of the events, and the broader range of potential confounding variables. The challenge for the field of epigenomics and human disease is to define the best practices that will allow studies to be performed that will prove robust to replication and reveal underlying biological events involved in diseases, whether in directly affected or surrogate cells.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgments

References

Adalsteinsson, B.T., Gudnason, H., Aspelund, T., Harris, T.B., Launer, L.J., Eiriksdottir, G., Smith, A.V., Gudnason, V., 2012. Heterogeneity in white blood cells has potential to confound DNA methylation measurements. PLoS One 7 (10), e46705.

Atz, M., Walsh, D., Cartagena, P., Li, J., Evans, S., Choudary, P., Overman, K., Stein, R., Tomita, H., Potkin, S., Myers, R., Watson, S.J., Jones, E.G., Akil, H., Bunney Jr., W.E., Vawter, M.P., Members of National Institute of Mental Health Conte, C., Pritzker Neuropsychiatric Disorders Research, C., 2007. Methodological considerations for gene expression profiling of human brain. J. Neurosci. Methods 163 (2), 295–309.

Barouki, R., Gluckman, P.D., Grandjean, P., Hanson, M., Heindel, J.J., 2012. Developmental origins of non-communicable disease: implications for research and public health. Environ. Health 11, 42.

Bell, J.T., Spector, T.D., 2011. A twin approach to unraveling epigenetics. Trends Genet. 27, 116–125.

Bernstein, B.E., Stamatoyannopoulos, J.A., Costello, J.F., Ren, B., Milosavljevic, A., Meissner, A., Kellis, M., Marra, M.A., Beaudet, A.L., Ecker, J.R., Farnham, P.J., Hirst, M., Lander, E.S., Mikkelsen, T.S., Thomson, J.A., 2010. The NIH Roadmap Epigenomics Mapping Consortium. Nat. Biotechnol. 28 (10), 1045–1048.

Choudhry, S., Seibold, M.A., Borrell, L.N., Tang, H., Serebrisky, D., Chapela, R., Rodriguez-Santana, J.R., Avila, P.C., Ziv, E., Rodriguez-Cintron, W., Risch, N.J., Burchard, E.G., 2007. Dissecting complex diseases in complex populations: asthma in Latino Americans. Proc. Am. Thorac. Soc. 4 (3), 226–233.

Cortessis, V.K., Thomas, D.C., Levine, A.J., Breton, C.V., Mack, T.M., Siegmund, K.D., Siegmund, K.D., Haile, R.W., Laird, P.W., 2012. Environmental epigenetics: prospects for studying epigenetic mediation of exposure-response relationships. Hum. Genet. 131 (10), 1565–1589.

Cui, H., Cruz-Correa, M., Giardiello, F.M., Hutcheon, D.F., Kafonek, D.R., Brandenburg, S., Wu, Y., He, X., Powe, N.R., Feinberg, A.P., 2003. Loss of IGF2 imprinting: a potential marker of colorectal cancer risk. Science 299 (5613), 1753–1755.

Dawson, M.A., Kouzarides, T., 2012. Cancer epigenetics: from mechanism to therapy. Cell 150, 12–27.

Esteller, M., 2007. Cancer epigenomics: DNA methylomes and histone-modification maps. Nat. Rev. Genet. 8, 286–298.

Gaston, V., Le Bouc, Y., Soupre, V., Burglen, L., Donadieu, J., Oro, H., Audry, G., Vazquez, M.P., Gicquel, C., 2001. Analysis of the methylation status of the KCNQ1OT and H19 genes in leukocyte DNA for the diagnosis and prognosis of Beckwith-Wiedemann syndrome. Eur. J. Hum. Genet. 9 (6), 409–418.

Gemma, C., Ramagopalan, S.V., Down, T.A., Beyan, H., Hawa, M.I., Holland, M.L., Hurd, P.J., Giovannoni, G., David Leslie, R., Ebers, G.C., Rakyan, V.K., 2013. Inactive or moderately active human promoters are enriched for inter-individual epialleles. Genome Biol. 14 (5), R43.

Gertz, J., Varley, K.E., Reddy, T.E., Bowling, K.M., Pauli, F., Parker, S.L., Kucera, K.S., Willard, H.F., Myers, R.M., 2011. Analysis of DNA methylation in a three-generation family reveals widespread genetic influence on epigenetic regulation. PLoS Genet. 7 (8), e1002228.

Gibbs, J.R., van der Brug, M.P., Hernandez, D.G., Traynor, B.J., Nalls, M.A., Lai, S.L., Arepalli, S., Dillman, A., Rafferty, I.P., Troncoso, J., Johnson, R., Zielke, H.R., Ferrucci, L., Longo, D.L., Cookson, M.R., Singleton, A.B., 2010. Abundant quantitative trait loci exist for DNA methylation and gene expression in human brain. PLoS Genet. 6 (5), e1000952.

Gluckman, P.D., Hanson, M.A., Low, F.M., 2011. The role of developmental plasticity and epigenetics in human health. Birth Defects Res. C Embryo Today 93, 12–18.

Gonzalez-Fierro, A., Vasquez-Bahena, D., Taja-Chayeb, L., Vidal, S., Trejo-Becerril, C., Perez-Cardenas, E., de la Cruz-Hernandez, E., Chavez-Blanco, A., Gutierrez, O., Rodriguez, D., Fernandez, Z., Duenas-Gonzalez, A., 2011. Pharmacokinetics of hydralazine, an antihypertensive and DNA-demethylating agent, using controlled-release formulations designed for use in dosing schedules based on the acetylator phenotype. Int. J. Clin. Pharmacol. Ther. 49 (8), 519–524.

Griffiths, E.A., Gore, S.D., 2013. Epigenetic therapies in MDS and AML. Adv. Exp. Med. Biol. 754, 253–283.

Guintivano, J., Aryee, M., Kaminsky, Z., 2013. A cell epigenotype specific model for the correction of brain cellular heterogeneity bias and its application to age, brain region and major depression. Epigenetics 8.

Heyn, H., Li, N., Ferreira, H.J., Moran, S., Pisano, D.G., Gomez, A., Diez, J., Sanchez-Mut, J.V., Setien, F., Carmona, F.J., Puca, A.A., Sayols, S., Pujana, M.A., Serra-Musach, J., Iglesias-Platas, I., Formiga, F., Fernandez, A.F., Fraga, M.F., Heath, S.C., Valencia, A., Gut, I.G., Wang, J., Esteller, M., 2012. Distinct DNA methylomes of newborns and centenarians. Proc. Natl. Acad. Sci. U. S. A. 109 (26), 10522–10527.

Hou, L., Zhang, X., Wang, D., Baccarelli, A., 2012. Environmental chemical exposures and human epigenetics. Int. J. Epidemiol. 41, 79–105.

Houseman, E.A., Accomando, W.P., Koestler, D.C., Christensen, B.C., Marsit, C.J., Nelson, H.H., Wiencke, J.K., Kelsey, K.T., 2012. DNA methylation arrays as surrogate measures of cell mixture distribution. BMC Bioinform. 13, 86.

Joubert, B.R., Haberg, S.E., Nilsen, R.M., Wang, X., Vollset, S.E., Murphy, S.K., Huang, Z., Hoyo, C., Midttun, O., Cupul-Uicab, L.A., Ueland, P.M., Wu, M.C., Nystad, W., Bell, D.A., Peddada, S.D., London, S.J., 2012. 450K epigenome-wide scan identifies differential DNA methylation in newborns related to maternal smoking during pregnancy. Environ. Health Perspect. 120 (10), 1425–1431.

Junien, C., 2006. Impact of diets and nutrients/drugs on early epigenetic programming. J. Inherit. Metab. Dis. 29, 359–365.

Leek, J.T., Scharpf, R.B., Bravo, H.C., Simcha, D., Langmead, B., Johnson, W.E., Geman, D., Baggerly, K., Irizarry, R.A., 2010. Tackling the widespread and critical impact of batch effects in high-throughput data. Nat. Rev. Genet. 11 (10), 733–739.

Liu, Y., Aryee, M.J., Padyukov, L., Fallin, M.D., Hesselberg, E., Runarsson, A., Reinius, L., Acevedo, N., Taub, M., Ronninger, M., Shchetynsky, K., Scheynius, A., Kere, J., Alfredsson, L., Klareskog, L., Ekstrom, T.J., Feinberg, A.P., 2013. Epigenome-wide association data implicate DNA methylation as an intermediary of genetic risk in rheumatoid arthritis. Nat. Biotechnol. 31 (2), 142–147.

Mill, J., Heijmans, B.T., 2013. From promises to practical strategies in epigenetic epidemiology. Nat. Rev. Genet. 14, 585–594.

Monoranu, C.M., Grunblatt, E., Bartl, J., Meyer, A., Apfelbacher, M., Keller, D., Michel, T. M., Al-Saraj, S., Schmitt, A., Falkai, P., Roggendorf, W., Deckert, J., Ferrer, I., Riederer, P., 2011. Methyl- and acetyltransferases are stable epigenetic markers postmortem. Cell Tissue Bank. 12 (4), 289–297.

Perera, F., Herbstman, J., 2011. Prenatal environmental exposures, epigenetics, and disease. Reprod. Toxicol. 31, 363–373.

Popovic, R., Licht, J.D., 2012. Emerging epigenetic targets and therapies in cancer medicine. Cancer Discov. 2, 405–413.

Rakyan, V.K., Down, T.A., Balding, D.J., Beck, S., 2011. Epigenome-wide association studies for common human diseases. Nat. Rev. Genet. 12, 529–541.

Reamon-Buettner, S.M., Mutschler, V., Borlak, J., 2008. The next innovation cycle in toxicogenomics: environmental Epigenetics. Mutat. Res. 659, 158–165.

Rius, M., Lyko, F., 2012. Epigenetic cancer therapy: rationales, targets and drugs. Oncogene 31, 4257–4265.

Sarter, B., Long, T.I., Tsong, W.H., Koh, W.P., Yu, M.C., Laird, P.W., 2005. Sex differential in methylation patterns of selected genes in Singapore Chinese. Hum. Genet. 117, 402–403.

Satterlee, J.S., Schubeler, D., Ng, H.H., 2010. Tackling the epigenome: challenges and opportunities for collaboration. Nat. Biotechnol. 28, 1039–1044.

Seok, J., Warren, H.S., Cuenca, A.G., Mindrinos, M.N., Baker, H.V., Xu, W., Richards, D.R., McDonald-Smith, G.P., Gao, H., Hennessy, L., Finnerty, C.C., Lopez, C.M., Honari, S., Moore, E.E., Minei, J.P., Cuschieri, J., Bankey, P.E., Johnson, J.L., Sperry, J., Nathens, A.B., Billiar, T.R., West, M.A., Jeschke, M.G., Klein, M.B., Gamelli, R.L., Gibran, N.S., Brownstein, B.H., Miller-Graziano, C., Calvano, S.E., Mason, P.H., Cobb, J.P., Rahme, L.G., Lowry, S.F., Maier, R.V., Moldawer, L.L., Herndon, D.N., Davis, R.W., Xiao, W., Tompkins, R.G., Inflammation and L. S. C. R. P. Host Response to Injury, 2013. Genomic responses in mouse models poorly mimic human inflammatory diseases. Proc. Natl. Acad. Sci. U. S. A. 110 (9), 3507–3512.

Shen, Y., Yue, F., McCleary, D.F., Ye, Z., Edsall, L., Kuan, S., Wagner, U., Dixon, J., Lee, L., Lobanenkov, V.V., Ren, B., 2012. A map of the cis-regulatory sequences in the mouse genome. Nature 488 (7409), 116–120.

Soriano-Tarraga, C., Jimenez-Conde, J., Giralt-Steinhauer, E., Ois, A., Rodriguez-Campello, A., Cuadrado-Godia, E., Fernandez-Cadenas, I, Montaner, J., Lucas, G., Elosua, R., Roquer, J., GeneStroke, C., The Spanish Stroke Genetics, 2013. DNA isolation method is a source of global DNA methylation variability measured with LUMA. Experimental analysis and a systematic review. PLoS One 8 (4), e60750.

Spivakov, M., Akhtar, J., Kheradpour, P., Beal, K., Girardot, C., Koscielny, G., Herrero, J., Kellis, M., Furlong, E.E., Birney, E., 2012. Analysis of variation at transcription factor binding sites in Drosophila and humans. Genome Biol. 13 (9), R49.

Varley, K.E., Gertz, J., Bowling, K.M., Parker, S.L., Reddy, T.E., Pauli-Behn, F., Cross, M.K., Williams, B.A., Stamatoyannopoulos, J.A., Crawford, G.E., Absher, D.M., Wold, B.J., Myers, R.M., 2013. Dynamic DNA methylation across diverse human cell lines and tissues. Genome Res. 23 (3), 555–567.

Verma, M., 2012. Epigenome-Wide Association Studies (EWAS) in Cancer. Curr. Genomics 13 (4), 308–313.

Wang, D., Yan, L., Hu, Q., Sucheston, L.E., Higgins, M.J., Ambrosone, C.B., Johnson, C.S., Smiraglia, D.J., Liu, S., 2012. IMA: an R package for high-throughput analysis of Illumina's 450K Infinium methylation data. Bioinformatics 28 (5), 729–730.

Warner, M.J., Ozanne, S.E., 2010. Mechanisms involved in the developmental programming of adulthood disease. Biochem. J. 427, 333–347.

Waterland, R.A., Michels, K.B., 2007. Epigenetic epidemiology of the developmental origins hypothesis. Annu. Rev. Nutr. 27, 363–388.

Yuen, R.K., Avila, L., Penaherrera, M.S., von Dadelszen, P., Lefebvre, L., Kobor, M.S., Robinson, W.P., 2009. Human placental-specific epipolymorphism and its association with adverse pregnancy outcomes. PLoS One 4 (10), e7389.