

Chapter 2

Community Resources and Technologies Developed Through the NIH Roadmap Epigenomics Program

John S. Satterlee, Andrea Beckel-Mitchener, Kim McAllister, Dena C. Procaccini, Joni L. Rutter, Frederick L. Tyson, and Lisa Helbling Chadwick

Abstract

This chapter describes resources and technologies generated by the NIH Roadmap Epigenomics Program that may be useful to epigenomics researchers investigating a variety of diseases including cancer. Highlights include reference epigenome maps for a wide variety of human cells and tissues, the development of new technologies for epigenetic assays and imaging, the identification of novel epigenetic modifications, and an improved understanding of the role of epigenetic processes in a diversity of human diseases. We also discuss future needs in this area including exploration of epigenomic variation between individuals, single-cell epigenomics, environmental epigenomics, exploration of the use of surrogate tissues, and improved technologies for epigenome manipulation.

Key words Bisulfite, Cell-type, Chemo-epigenetic, ChIP-seq, Chromatin, Immunoprecipitation, CRISPR, Crotonylation, Deacetylase, Hi-C, Ectoderm, ENCODE, Endoderm, Epigenetic, Epigenome, Genome-wide, HDAC, Histone, hmC, IHEC, mC, MeDIP-Seq, Methylation, MethylC-seq, Methylome, MRE-Seq, mRNA, Nucleosomes, Nucleus, Opto-epigenetic, Pluripotent, RNA-seq, Roadmap, RRBS, Transcription

1 Introduction: Overview and Goals of Program

Epigenomics is the study of functional, and sometimes heritable, changes in the regulation of gene activity and expression that do not change the underlying DNA sequence [1–3]. Although each cell type in the human body (with some exceptions) contains the same genetic information, the interplay between transcription factor programs and epigenetic regulation enables pluripotent stem cells to give rise to the diversity of differentiated cell types (e.g., skin cells, liver cells, neurons) [4, 5]. Furthermore epigenetic processes, from gametogenesis through embryonic and neonatal stages, and continuing throughout adolescence, adulthood, and

senescence, can influence both normal development and a variety of disease processes [6–9].

Even before completion of the first human genome sequence in 2000, a grassroots scientific effort began encouraging a large-scale, organized project in Epigenomics [10]. Meetings were held by a number of NIH institutes (e.g., NCI, NIDA, NIEHS) and by other groups (e.g., American Association of Cancer Researchers) that provided recommendations on how best to help the scientific community accelerate discovery in epigenomics [11, 12].

In 2007, Epigenomics was selected as a trans-NIH Roadmap (now Common Fund) project. A 2006 portfolio analysis conducted by the NIH revealed that between 1998 and 2006, increasing numbers of funded studies in the area of epigenetics were supported by the NIH Institutes and Centers. Moreover, there was a similar escalation in the number of publications associating altered gene expression profiles and epigenetic processes with adverse disease outcomes. Additionally in 2007, a scientific meeting was hosted by NIH to garner specific recommendations for a large-scale, transformative project in this area. Recommendations from this meeting and from prior scientific meetings were used to craft the four major initiatives of the NIH Roadmap Epigenomics Program:

1. To generate reference epigenome profiles for “normal” human cells and tissues, including new ways to generate, visualize, and analyze this data
2. To discover novel epigenetic marks
3. To develop revolutionary technologies with the potential to significantly change epigenetics research (with sub-initiatives in general epigenetic technologies, epigenetic imaging, and epigenomic manipulation)
4. To transform our understanding of the epigenomic basis of disease

These goals were translated into several distinct initiatives that are illustrated in Fig. 1. These projects included the development of reference epigenomes; the identification of novel epigenetic marks; the development of new technologies for epigenomic assays, imaging, and manipulation; investigations into epigenomic processes in a variety of diseases; and the development of monoclonal antibody tools for epigenomic research. To date a total of 78 grants have been funded by this program, including 61 R01s, 7 R21s, 5 RC1 “challenge” grants, and 5 U01s (<http://commonfund.nih.gov/epigenomics/fundedresearch>). The following sections describe the technologies and community resources generated from each of these initiatives.

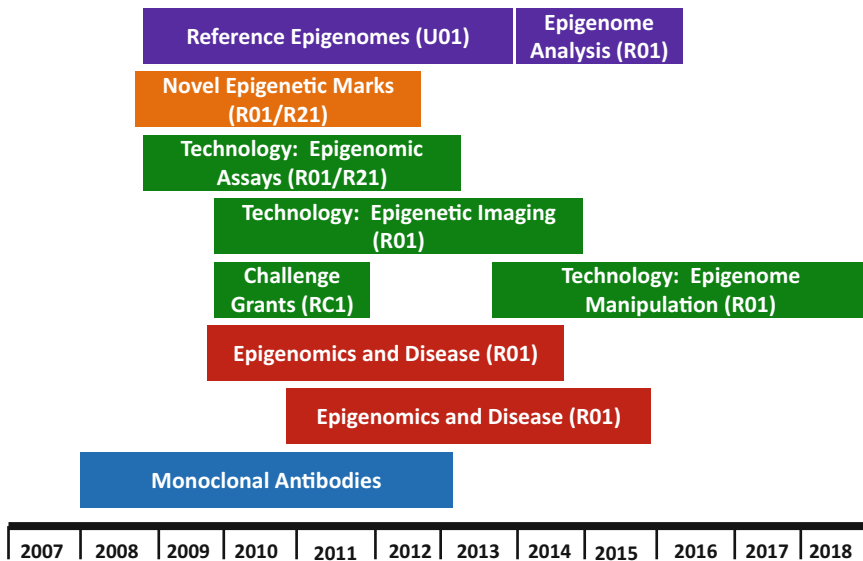


Fig. 1 Roadmap Epigenomics Program Projects and Timeline. This figure shows the major components of the program and the periods of Roadmap Epigenomics Program funding. Reference Epigenome activities are shown in *purple*, novel marks in *orange*, technology initiatives in *green*, disease studies in *maroon*, and antibodies in *blue*

2 Community Resources and Technologies Developed Through the Epigenomics Program

2.1 Comprehensive Reference Epigenome Maps

As indicated above, a key scientific recommendation was to generate comprehensive reference epigenome maps for normal human cells and tissues. These epigenomic maps would provide a public resource for the scientific community to use in investigator-initiated investigations of epigenetic regulation in biological processes and human disease. Specifically, these maps were intended to:

- Provide a normal baseline for investigators exploring environmental or disease epigenomics
- Reveal how epigenomes change during differentiation and development
- Discover how epigenomic modifications “interact” or co-occur with one another
- Reveal functional genomic elements
- Inform regenerative medicine studies by investigating the epigenomes of stem and induced pluripotent stem (iPS) cells
- Enable integration of epigenomic information with genetic information to better understand genome/epigenome interactions
- Provide a foundation for the future development of biomarkers and therapeutics.

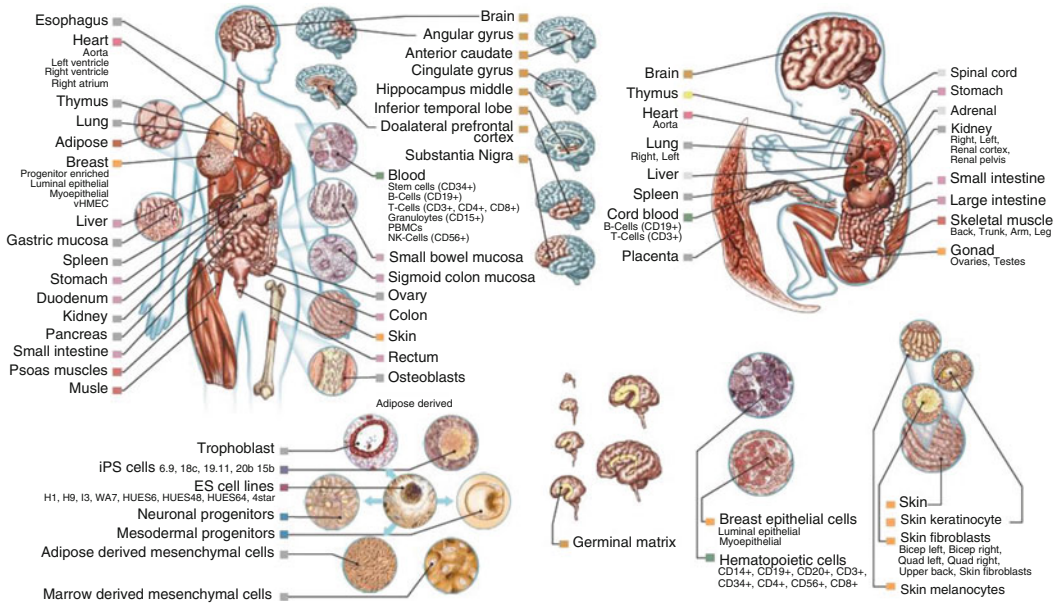


Fig. 2 Reference epigenomes from a diversity of human cells and tissues. The *top left* shows some of the characterized adult human cells and tissues including hematopoietic cells, gastrointestinal (GI) tract, adipose, muscle, heart, and several brain regions. The *top right* shows some of the mapped fetal cells and tissue types including muscle, lung, heart, GI tract, and placenta. The *bottom right* shows some of the fetal samples in more detail (brain, hematopoietic, and skin). The *bottom left* shows some of the cell lines characterized including human embryonic stem cells (hESCs), cell types differentiated into neural progenitor cells, trophoblast, etc., and induced pluripotent stem cells (iPSCs). Illustration by Rae Senarighi, University of Washington

The Roadmap Epigenome Mapping Consortium (REMC) was established to generate reference epigenome maps for normal human cells and tissues. The REMC consists of four mapping centers and one data coordination/analysis center to process the data through standardized pipelines [13]. The original goal of the REMC was to generate 25 reference epigenome maps by the end of the project. However due to rapid improvements in technology and cost reductions, the REMC has generated a total of 92 reference epigenome maps for a wide range of human cells and tissues from adults and fetuses as well as from embryonic stem (ES) cells and their derivatives (*see* Fig. 2). Normal cell or tissue types were prioritized for epigenome mapping based on their availability, the amounts of tissue available for the multiple necessary assays, to achieve a balance between purified cell types and heterogeneous tissues, and their anticipated relevance to human disease. Some of the comprehensive reference epigenomes generated were from embryonic stem (ES) cells, ES-derived cells, iPS cells, hematopoietic cell types, fetal tissues (adrenal, brain, spinal cord, heart, gastrointestinal, kidney, lung, muscle, placenta, thymus), and adult tissues (fat, breast, brain, gastrointestinal, muscle, skin, connective tissue).

The advent of next-generation sequencing coupled with development of improved or novel technologies (e.g., ChIP-seq, RNA-seq, MethylC-seq) and the requisite expansion of informatic capabilities provided an unprecedented opportunity to generate a catalog of reference epigenome maps for many human cells and tissues [8]. Of course, given the large number of known chromatin modifications and the likely existence of undiscovered chromatin modifications, it is not possible to generate “complete” epigenome maps. Thus the REMC has determined that a reference epigenome should include assays for DNA methylation, a core set of six posttranslational histone modifications (PTMs, H3K27me3, H3K36me3, H3K4me1, H3K4me3, H3K27ac, and H3K9me3), and mRNA expression analysis. Some samples were also assayed for chromatin accessibility using the DNase I hypersensitivity assay.

The reference epigenomes generated are categorized into four Classes (1–4) defined by the associated epigenomic data sets (*see* Table 1 for details). For example, Class 1 epigenomes have (1) DNA methylation maps generated by whole-genome bisulfite sequencing, (2) maps of six core histone PTMs plus up to 20 auxiliary histone PTMs, (3) chromatin accessibility data, and (4) transcript analysis by RNA-seq. As of January 1, 2014, the REMC has generated 92 Class 1–4 reference epigenomes. Additional useful data sets that do not contain all of the assays required for Class 1–4 epigenomes have also been generated by the REMC (<http://www.roadmapepigenomics.org/>). 23 of the 92 reference epigenomes are expected to be genotyped by whole-genome sequencing to help researchers probe the interactions between the genome and epigenome.

The REMC has undertaken the most comprehensive and coordinated effort to generate reference epigenome maps to date. In addition to the generation of this diverse panel of reference epigenomes, consortium investigators have delivered a number of critically important contributions to the field, a few of which are highlighted below. The consortium has developed and implemented a set of experimental and data quality standards/guidelines for these assay types (<http://www.roadmapepigenomics.org/protocols>). The Roadmap Epigenomics Program has also worked with other large projects performing epigenomic assays to develop congruent metadata standards and ontologies to maximize the utility and interoperability of data across these large projects (*see* Subheading 3.1). Early in the mapping efforts, consortium investigators developed the assay that produced the first complete mammalian methylomes [14]. Other major accomplishments include (1) identification of distinct epigenetic mechanisms regulating early and late stages of differentiation in embryonic stem cells [15]; (2) identification of epigenomic processes that transpire as ES cells give rise to the three embryonic germ layers, ectoderm, mesoderm, and endoderm [16]; (3) demonstration that chromatin

Table 1
Reference epigenome classes and progress

Assay	Class 1	Class 2	Class 3	Class 4
<i>Histone PTM ChIP-seq</i>				
H3K27me3	Required	Required	Required	Required
H3K36me3	Required	Required	Required	Required
H3K4me1	Required	Required	Required	Required
H3K4me3	Required	Required	Required	Required
H3K27ac or H3K9ac	Required	Required	Required	Required
H3K9me3	Required	Required	Required	Required
Up to 24 auxiliary histone modifications	Required	Not req.	Not req.	Not req.
<i>DNA methylation</i>				
WGBS/MethylC-seq	Required	Required	Not req.	Not req.
RRBS	Not req.	Not req.	Any assay	Any assay
MeDIP-seq/MRE-seq	Not req.	Not req.	Any assay	Any assay
<i>Chromatin accessibility</i>				
DNase I hypersensitivity	Required	If possible	Required	Not req.
<i>mRNA expression</i>				
mRNA-seq	Required	Required	Not req.	Not req.
Array based	Not req.	Not req.	Any assay	Any assay
<i>Number completed</i>	7	51	14	20

The Reference Epigenomes generated by the REMC have been classified according to the number and comprehensiveness of the assays performed. This table also indicates which assays are required or not required for a particular class of epigenome. Originally H3K9ac was a core histone PTM but it was replaced by the more informative H3K27ac. “Any assay” means that any assay of a given type (e.g., DNA methylation) was sufficient for a particular epigenome class. The most challenging assay, DNase I hypersensitivity, was encouraged if possible for Class 2 epigenomes. A total of 92 comprehensive reference epigenomes have been generated as of January 1, 2014. In addition, there are many available data sets for other tissues in which it was not possible to perform a complete set of assays (<http://www.roadmapepigenomics.org/>). Abbreviations include whole-genome bisulfite sequencing (WGBS), reduced representation bisulfite sequencing (RRBS), methylated DNA immunoprecipitation (MeDIP-seq), and methylation-sensitive restriction enzyme-sequencing (MRE-seq)

states change in response to developmental or environmental cues [17]; and (4) the discovery that, in addition to coding for amino acids, approximately 15 % of human DNA codons can specify transcription factor binding sites [18]. Consortium investigators have also developed remarkable insights into disease susceptibility by discovering that a large percentage of disease-associated variants from genome-wide association studies (GWAS) occur in DNA regulatory regions defined by DNase I hypersensitive sites [19]. Many of these regulatory DNA elements are enriched for gene variants

that have been associated with intrauterine exposure-relevant phenotypes and are frequently active during fetal development.

2.2 Reference Epigenome Data Access, Analysis, and Visualization

The data generated by the REMC constitute a rich resource that is broadly available to the scientific community. Although a brief overview is provided here, a more complete listing of different options for accessing, viewing and analyzing epigenomic data generated from these efforts is published elsewhere [20, 21]. These data can be classified into two major categories: data with associated human DNA sequence information and data without associated human DNA sequence information. In general, read density maps and data that do not include human DNA sequence information are submitted to the Gene Expression Omnibus (GEO), supported by the National Center for Biotechnology Information (NCBI), after data cleaning and can be obtained and downloaded for specific analyses (<http://www.ncbi.nlm.nih.gov/geo/roadmap/epigenomics/>). Data containing de-identified human DNA sequence information are, in many cases, also publically available through NCBI's Sequence Read Archive (SRA). In other cases, the precise language in the patient consent form and/or the Institutional Review Board of one of the epigenome mapping centers necessitated that these data be deposited into the controlled-access database of Genotypes and Phenotypes (dbGaP). Scientists may request access to use these data by completing standard dbGaP forms (<http://www.ncbi.nlm.nih.gov/gap>).

The complex nature of epigenomic data sets has stimulated the development of new integrative tools as well as multifaceted browsers to facilitate epigenome visualization. The Genboree Workbench, hosted by the Human Epigenome Atlas, allows investigators to analyze their own data with REMC-generated data using and take advantage of numerous comparative and integrative analysis tools that have been created through the REMC (<http://www.genboree.org/epigenomeatlas/index.rhtml>). Consortium data may be visualized directly on the UCSC Genome Browser using the Track Hub feature or can be browsed by cell type, mark, or a specific text query at several websites (e.g., <http://www.roadmapepigenomics.org/data>, <http://www.genboree.org/epigenomeatlas/index.rhtml>). REMC data can also be viewed in conjunction with ENCODE data (<http://encode-roadmap.org/>). Researchers interested in novel approaches for visualizing epigenomic data, including the ability to view long-range genomic interactions, should explore the Human Epigenome Browser (<http://epigenomegateway.wustl.edu/info/>).

These epigenomic data sets provide baselines for comparisons with diseased cells or tissues and have also shown great promise in elucidating the functional significance of non-coding genetic variant hits from GWAS experiments [19]. They also provide baselines

for allowing comparisons with diseased cells or tissues. Researchers may soon routinely use epigenome data sets to learn how chromatin state-defined cis-regulatory elements impact transcript distribution and abundance [8, 21]. Furthermore, with epigenome maps in hand, it may soon be possible to understand how environmental context (e.g., influences of drugs of abuse, stress, toxins, diet chemotherapy protocols, and other environmental exposures) alter the epigenome [22–24].

2.3 Discovery of Novel Epigenetic Modifications

At the time the Roadmap Epigenomics Program began, a number of histone modifications and just one DNA modification (mC) had been identified in eukaryotes. It was clear that our catalog of chromatin modifications was incomplete. Thus, an initiative was developed to discover additional epigenetic modifications in order to move toward the completion of this catalog. By 2013, researchers funded through this initiative generated numerous publications in this area and have identified and/or further characterized multiple new histone posttranslational modifications (PTMs) in yeast, *Drosophila*, and mammalian cells [25–27].

In one study alone, Dr. Yingming Zhao and his colleagues identified an additional 67 novel histone PTMs in mammalian cells including lysine crotonylation [26]. In addition, some projects have identified or implicated proteins in the specific deposition or removal of these marks on chromatin. These researchers have also discovered specific roles for the new histone PTMs, or combinations of PTMs, in the control of key biological functions such as cell cycle regulation, development, and DNA repair [28–31]. These studies move the field closer to establishing a more complete catalog of DNA and histone modifications. However, much work remains to be done to better characterize the biological functions of these DNA and histone modifications as well as the enzymes that establish, remove, or bind to these modifications.

2.4 Improved Assays for Monitoring the Epigenome

Prior to the start of the NIH Roadmap Epigenomics Program, there were several useful assays to monitor the epigenome including chromatin immunoprecipitation followed by microarray or sequencing (ChIP-chip, ChIP-seq) which detects chromatin features, reduced representation bisulfite sequencing (RRBS) which enables characterization of DNA methylation status at single base resolution at a subset of genomic regions, and the DNase I hypersensitive assay which detects the level of chromatin accessibility [8]. Although these assays were state of the art at the time, some were not sufficiently facile or comprehensive. Also there were limitations to the types of biological questions one could ask using these approaches. To address this scientific gap, an initiative on Technology Development in Epigenetics was initiated in 2008. In parallel the REMC expended great effort to improve current assay technology and develop new technologies [14, 32–34].

Four of the new or improved assays that were developed by these components of the Roadmap Epigenomics Program are described as follows. Prior to November 2009, methods for analyzing DNA methylation at single base resolution were limited and the largest existing methylome data set that existed was for the model plant *Arabidopsis thaliana* [35]. As a part of the REMC, researchers in the Ecker lab optimized their MethylC-Seq assay and applied it to two different human cell types, generating the first genome-wide single base resolution maps for any human cell type [14]. Since then, the use of MethylC-Seq and related assays for interrogating DNA methylation state has become widespread.

Most epigenomic assays measure a molecular phenotype from the chromatin of a population of cells but there is great interest in understanding how chromatin differs between individual cells. Researchers in the Soloway lab have recently published a paper that describes their bifurcated nanofluidic device for real-time detection and sorting of individual molecules of methylated DNA [36]. The DNA molecules were identified using fluorescently labeled methyl binding domain protein 1 (MDB1) which bound specifically to double-stranded, methylated DNA. They were also able to recover the sorted DNA by quantitative PCR for possible post-sorting applications. This device holds great promise for color-multiplexed epigenetic analyses and will allow recovery of genetic material for downstream studies.

Although the ability to monitor chromatin states across the genome is robust, a major obstacle has been the inability to interrogate the chromatin features that co-occur at a single genomic locus. Researchers in the Tackett and Taverna labs have developed chromatin affinity purification with mass spectrometry (ChAP-MS) that involves isolation of a single genomic locus which can then be subjected to mass spectrometry analysis to identify the proteins and associated PTMs at that locus [37]. As our ability to isolate genomic loci improves, this method may enable researchers to identify the complement of proteins that are physically associated with a given genomic locus in addition to pinpointing the precise histone modifications that are present. This technology will facilitate a deeper understanding of how these loci change during development or upon environmental exposure and also reveal in greater detail the cross-communication that occurs between epigenetic and transcriptional regulatory pathways.

Chromatin assays are frequently used to investigate a single point in time, obscuring dynamic biological processes. To address this limitation, researchers in the Henikoff lab developed a metabolic labeling strategy that enables affinity purification of tagged nucleosomes containing newly synthesized histones. This technique allows the measurement of histone turnover rates across the genome providing the ability to explore temporal chromatin changes [38].

2.5 Technologies for Improved Epigenetic Imaging

Many available epigenomic assays only probe a single time point; however it is often desirable to understand the dynamics of a biological system. Furthermore, each mammalian cell type is believed to have a distinct epigenomic profile, which may be altered in a disease state. As researchers understand more about how epigenomes are altered in specific disease states or by environmental perturbations, and develop “epigenetic therapeutics” to treat disease, our limited ability to visualize epigenetic modifying enzymes and modifications in vivo remains a significant scientific obstacle. Clinical diagnosis of diseases with a significant epigenetic component could be greatly enhanced if we improved our ability to monitor epigenetic modifying enzymes, binding proteins, or epigenetic states in specific tissues or cell types. While tissues such as blood or skin are readily available for this type of analysis, in the case of diseases impacting the brain, heart, bone, and other vital organs it may be difficult or impossible to obtain the appropriate tissue or cell type for epigenetic analysis.

Since the current technologies to determine the in vivo epigenetic state of tissues is extremely limited, a Roadmap Epigenomics Program initiative was developed to enable in vivo imaging or analysis of epigenetic changes at multiple levels ranging from a single cell to an entire organism. Several funded projects proposed a wide array of approaches including RNA aptamers, fluorescence complementation, soft X-ray tomography (SXT), magnetic resonance imaging (MRI), and positron emission tomography (PET). Two of these projects are described in more detail below.

Researchers in the Lomvardas and Larabell laboratories have been developing an in vivo fluorescence complementation approach to visualize epigenetic interactions between proteins when they physically interact. This approach will be linked to fluorescence microscopy-compatible SXT to analyze chromatin territories at high resolution in vivo. SXT analysis of olfactory nuclei indicates that downregulation of the lamin b receptor leads to the unusual nuclear architecture of the olfactory neuron. Downregulation of lamin b causes significant changes in epigenetic gene regulation in certain gene families [39]. Optimization of this technology could enable researchers to better correlate in vivo chromatin structure with chromatin features such as epigenomic marks, transcription factor binding, or chromatin-associated long noncoding RNAs.

In independent efforts, researchers in the Hooker and the Gelovani laboratories have been developing and optimizing PET radiotracers for in vivo imaging of Class I and Class III histone deacetylases (HDACs), respectively [40–42]. If robust PET ligands for these important epigenetic regulatory proteins are developed, they will be invaluable for monitoring in vivo changes in HDAC levels and/or activity over time and would be useful for monitoring difficult to access tissues such as the brain. If these ligands function

successfully in humans, they may be important for clinical diagnosis of disease as well as for monitoring the therapeutic efficacy of HDAC inhibitor treatment regimens.

2.6 Epigenomics in Health and Disease

The Reference Epigenome Maps developed by the Roadmap Epigenomics Program illustrate the critical role the epigenome plays in regulating cell type identity and function. One of the overarching goals of the Roadmap Epigenomics Program is to understand how alterations in the normal patterns of epigenetic regulation could contribute to the onset or progression of disease. The Epigenomics of Human Health and Disease initiative was developed to support epigenomic investigations of diseased, exposed, or otherwise compromised human tissues, in an effort to identify epigenetic changes that may underlie disease states.

At the time that this initiative was developed, the concept that epigenetic misregulation, rather than strictly genetic mutation, might underlie disease was not new. In 1983, Feinberg and Vogelstein published their landmark study showing that some genes became hypomethylated in human tumors relative to adjacent normal tissues, providing the first evidence that diseased cells could have different epigenetic profiles than their normal counterparts [43]. This initial finding sparked a long-lasting interest in cancer epigenetics. A 2006 portfolio analysis showed that the vast majority of NIH funding in disease epigenetics was in the area of cancer. However, it was not clear whether epigenetic misregulation is associated with other diseases and whether or not this misregulation plays a causal role. To address this knowledge gap, a Roadmap Epigenomics Program Disease initiative was implemented in 2008, which resulted in the co-funding (by the Common Fund and the relevant NIH Institutes and Centers) of 22 new grants, focused on a wide range of conditions. A second related initiative was implemented in 2010, which resulted in the funding of 11 additional grants by specific NIH Institutes and Centers. The breadth of disease research represented across these two initiatives is illustrated in Fig. 3. Below we describe the work of several investigators that used epigenomic approaches to identify key regulatory mechanisms that are altered in human disease.

The Young group used ChIP assays to define genomic regions containing large clusters of transcriptional enhancers (super-enhancers) highly enriched for transcription factors, the mediator complex, and chromatin modifications such as histone H3K27 acetylation. These super-enhancers drive transcription of key regulatory genes, such as master transcription factors, that are important in establishing cell type-specific gene expression programs during development and differentiation [44]. A subsequent study showed that a significant number of genetic variants associated with a diverse set of diseases are enriched in super-enhancers in disease-relevant cell types. This suggests that disruption of these

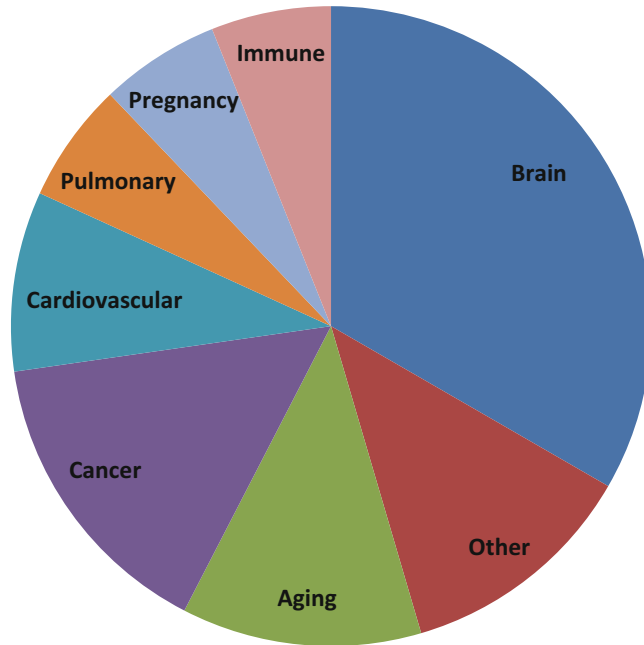


Fig. 3 The spectrum of epigenomics and disease investigations. This figure illustrates the diversity of tissues and diseases being investigated through the 33 disease projects funded as a part of the program. The major topics include brain disorders, aging, cancers, cardiovascular, pulmonary, immune research, and pregnancy

genomic elements may be a factor in the pathogenesis of a wide range of diseases [45].

The Huang group showed that in normal cells, estrogen mediates long-range epigenetic silencing through the transient formation of chromosomal loops flanked by estrogen receptor alpha (ESR1)-bound genomic regions [46]. However, in breast cancer, these loops stabilize and are associated with repressive epigenetic marks. In subsequent studies, Dr. Huang's group showed that these genomic regions, known as DEREs (distant estrogen response elements), were frequently amplified in breast carcinogenesis, and that gene expression changes associated with this phenomenon likely contribute to some breast cancers, as well as with the development of resistance to endocrine therapy.

One major issue that researchers contend with is sample origin: for example the disease affects the brain, but only peripheral blood samples are available. The Mill group took advantage of matched postmortem blood and brain samples from non-diseased individuals to compare DNA methylation patterns across several brain regions and blood [47]. As expected, they identified regions of tissue-specific methylation that differed between the brain and blood samples. These regions were mostly intragenic CpG islands

and low CpG density promoters, and were generally associated with genes that were differentially expressed across brain regions. However they also found that although the epigenomic profiles of blood and brain differed, a subset of the epigenetic variation observed between individuals within a tissue (such as the type of variation expected to contribute to phenotypic differences) was conserved across brain and blood. This suggests that although analysis of blood samples is unlikely to capture all of the disease-associated epigenetic changes found in the diseased tissue, it might have some utility as a surrogate tissue for certain human disease investigations. Furthermore, chromatin state information from specific blood cell types or other readily accessible cells or tissues may also be useful for developing biomarkers for predicting disease state or quantitating environmental exposures.

2.7 Epigenetic Antibodies

Another important resource gap at the start of the Epigenomics Program was the limited availability of high quality antibodies for probing epigenetic proteins or marks. Most of the available antibodies were polyclonal with only a finite supply of each batch. This limited the utility of these antibodies to the scientific community and made experimental validation between laboratories more challenging. Furthermore, antibodies varied dramatically in both their specificity and efficacy in ChIP experiments [48]. A minor initiative in the Roadmap Epigenomics Program aimed to develop a number of high-quality monoclonal antibodies against selected proteins and marks, and to make these antibodies available to the scientific community at low cost. As of 2013, 19 antibodies were generated and have been made available including: 5hmC, 5fC, Dicer, DNMT3L, Histone H3.3, Histone H3-pThr11, JMJD2A, NSD1, Histone H2A.Z/V, H3K56ac, Histone H2A Z-acetyl-Lys4/7/11, and Histone H3-acetyl-Lys56. These antibodies are available through NeuroMab at <http://neuromab.ucdavis.edu/catalog.cfm>.

2.8 Overall Roadmap Epigenomics Program Impact

There are many challenges in evaluating the impact of a scientific program since there is not one clear quantitative measurement that should be used. Depending on the nature of the project one could consider metrics such as web hits, data downloads, patent applications, or investigational new drug applications (INDs). Other less quantitative but salient metrics might include the number and types of spinoff projects, the extent to which the scientific community adopts protocols or reagents produced, the extent to which new protocols or assays lead to decreased costs or increased amounts of useful information, or how the program influenced the development of national and international efforts in this area. Often the effects of investment in a scientific program are not fully understood or appreciated for many years and sometimes not until well after the program ends.

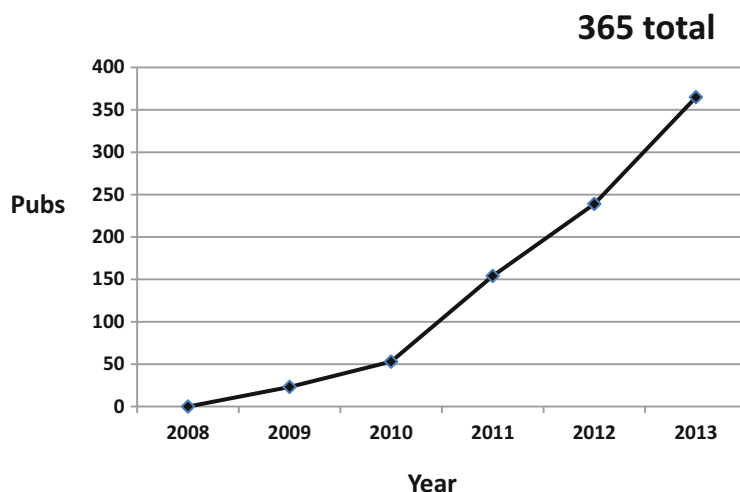


Fig. 4 Roadmap Epigenomics Program Publications Over Time. The Y-axis shows the total number of publications as of January 1, 2014 attributed to the program as a result of funding support, while the X-axis shows the year

Ultimately, scientific progress requires communication through publications. Publications can yield several useful impact metrics including (1) the absolute quantity of publications produced, (2) the quality of the publications produced, and/or (3) the number of times these publications have been cited. As shown in Fig. 4, there has been an increase in the absolute number of publications since the beginning of the Roadmap Epigenomics Program. As of January 1, 2014 there have been 365 publications attributed to the funded components of this program (<http://commonfund.nih.gov/publications?pid=5>). A significant number of Roadmap Epigenomics Program projects have ongoing funding (some began as recently as September 2013) so we expect that this number will continue to rise. Even some of the projects that have been completed 2 years ago are still publishing data obtained during Roadmap Epigenomics Program support.

With respect to publication quality, many of the results produced by projects within the program have been published in journals with high-impact factors. For example, as of January 1, 2014, there have been 23 publications in *Cell*, 12 in *Nature*, and 3 in *Science*. This is an indicator that many of the publications produced by the program are of potentially high impact.

Finally, several papers produced by the program have been highly cited suggesting they have made an important impact on the scientific community. One paper has been cited a total of 970 times as of January 1, 2014 [14]. A total of 26 papers have been cited between 100 and 401 times. Fifty-five publications have been cited between 30 and 100 times. Additionally, some of these publications have only been publically available for a short period of time and we anticipate that citations for these newer papers will increase over time.

3 Future Epigenomics Needs and Directions

3.1 *Expanding Epigenomic Data Sets*

The REMC will produce approximately 100 comprehensive epigenome maps by its conclusion. Although this is a remarkable achievement for the 5 years of this initiative, more epigenome maps are needed for uncharacterized cell types and tissues, diseased tissues, for cells and tissues with specific environmental exposures (e.g., environmental toxins, drugs of abuse), and for longitudinally-collected cells and tissues. These maps will be necessary for scientists to fully understand the role of the epigenome in human diseases. Building upon the NIH Roadmap Epigenomics Program, the International Human Epigenome Consortium (IHEC; <http://epigenomesportal.ca/ihec/>), an international consortium of funding agencies (Canada, European Union, Germany, Japan, South Korea, and the USA) aims to generate as many as 1000 epigenome maps. Hopefully, this will be a sufficient number of epigenomes to provide both a broad and comprehensive understanding of epigenomic regulation and epigenomic differences within and across disease states.

A sister project cataloguing the Encyclopedia Of DNA Elements (ENCODE; <http://www.genome.gov/encode/>) seeks to identify and catalog all of the functional elements in the human genome, along with a smaller effort to annotate the mouse genome. Although the goal of this NHGRI-funded project is not to generate epigenome maps, many of the assays used within this project are epigenome mapping assays. Roadmap Epigenomics Program and ENCODE data can be viewed together through the following web link (<http://www.encode-roadmap.org/>). The Roadmap Epigenomics Program has also worked with ENCODE and IHEC to develop congruent metadata standards and ontologies to maximize the utility and interoperability of data across these large projects. REMC, IHEC, and ENCODE data can be accessed via a common portal (<http://edcc-dev2.udes.genap.ca/edcc/ihec/index.html>). As these projects move forward, it will be important to continue to coordinate efforts such that the data from ENCODE, IHEC, the Roadmap Epigenomics Program and other entities will be as interoperable as possible, and ensure that there will be long-term support for data archiving since these rich data resources are certain to have useful synergies with one another as novel biological questions are explored.

3.2 *Epigenomic Analysis*

Although there has been a great deal of data generated by the Roadmap Epigenomics Program thus far, a deeper analysis effort is needed in order to mine this rich resource and achieve maximum utility for the scientific community. The key strength of these data is that they enable investigators to carry out analyses of already available reference epigenomic data and allow researchers to apply

this information to specific research questions. We anticipate that secondary analysis by the scientific community, especially through integration with other novel data sets, will maximize the utility of the data generated and provide models for diverse ways in which these data can be exploited to make important discoveries about biological processes and diseases. For example, analysis of epigenomic data with GWAS data was an extremely valuable use of epigenomics program data which was not foreseen when the program began [19]. To encourage further analysis projects in which REMC data is analyzed in concert with public data sets (e.g., ENCODE, TCGA, IHEC, Brainspan, dbGAP) or with investigator-generated data sets, a new Roadmap Epigenomics Program initiative was released that hopes to support around ten 2-year projects starting in 2014 that will hopefully lead to similar novel discoveries (<http://grants.nih.gov/grants/guide/rfa-files/rfa-rm-14-001.html>).

3.3 Integration with Genome Conformation Assays

Development of new genome conformation assays (e.g., Hi-C, ChIA-PET) enables researchers to probe three dimensional genome interactions in particular cell types [49, 50]. As these assays improve, it will be of great interest to correlate this data with existing epigenomic modification and transcription factor binding data to determine if/how they relate to one another. Additionally, emerging in vivo microscopy techniques may have the potential to reveal how these molecular phenotypes correspond to the three-dimensional structure of chromatin and to probe their spatial dynamics and functions within the nucleus [39].

3.4 The Utility of Surrogate Tissues for Disease Investigations

Studies investigating disease epigenomics typically focus on the diseased tissue or cell type. However, for diseases involving some organs (e.g., brain, bone), obtaining human tissue specimens can be difficult or impossible. This challenge can be overcome by obtaining post-mortem tissue or using in vivo epigenetic imaging strategies, although this imaging technology is still in its infancy. An alternative strategy could be to use readily accessible or “surrogate” tissues (e.g., blood cell types, olfactory neurons, skin) that may in part reflect epigenomic changes that have taken place in the primary diseased tissue. Although some work has been published in this area, the extent to which use of surrogate tissues is appropriate for disease investigations remains unclear [47]. To explore this issue, NIH convened a workshop in August 2012 to obtain recommendations on the use of surrogate human cells in disease epigenomics investigations. Fifteen scientists with expertise in this area converged on three major scientific recommendations. The first recommendation was to develop and explore human samples to see whether or not disease-associated epigenetic patterns from a primary tissue could also be found in readily accessible surrogate cell types and would provide information that was as useful as that provided by the primary tissue involved in the disease.

One confounding issue with human epigenetic studies is that that one cannot readily control for genotype and environmental exposures. Therefore a second recommendation was to develop and explore model organism systems in which genotype and environmental exposure can be well controlled to see how epigenetic patterns associated with a disease state compare between primary and surrogate tissues. The third recommendation was to perform a systematic study of normal human epigenetic variability in different tissues to provide information about the normal range of variability. This information will help researchers design informative epigenome-wide disease association studies. The latter recommendation may be addressed in respect to genotype part by the Genotype-Tissue Expression Program (GTEx, *see* below). However, the investigation of how the epigenome changes in different tissues over human lifespan also needs to be investigated since age may have a significant influence on the epigenome [51, 52].

3.5 Human Epigenomic Variation

A major goal of the REMC was to explore epigenomic differences between cell and tissue types. Although some pilot work was performed investigating epigenomic differences between individuals in a single cell or tissue type, this program was not designed or powered to investigate this issue rigorously. However the Common Fund-supported GTEx Program (<http://www.ncbi.nlm.nih.gov/gtex/GTEX2/gtex.cgi>) is characterizing human gene expression in multiple post-mortem and surgical tissues and correlating this information with human genotype. If this project were to expand to include additional molecular phenotypes (e.g., DNA methylation, histone modifications) this would enable for the first time a robust exploration of the relationship between genomic and epigenomic variation in specific human tissues.

3.6 Epigenomic Variation Within a Cell Type: The Promise of Single-Cell Epigenomics

Two related major challenges are to understand how the epigenome impacts gene expression in individual cells and how cell-to-cell variation influences function at the tissue, organ, and systems level. To date, the study of epigenetic modifications typically involves analyzing complex tissue samples or mixed cell populations. It is clear that distinct epigenomic signatures are associated with specific cell or tissue types and that these epigenomic signatures likely play a role in determining cellular phenotypes [16, 17]. In progenitor cell populations, the epigenome has a clear connection to fate determination and cell-type specification [53]. Even cell lines cultured in vitro are heterogeneous to varying degrees and most existing data is the result of ensemble averaging; very little is known about how the epigenome differs across individual cells within a particular cell type [54]. Technological limitations have hampered our ability to evaluate epigenetic states at the resolution of a single cell. Available evidence suggests that in certain circumstances the epigenome is modifiable. Therefore microenvironmental differences in

and around individual cells could lead to epigenomic changes in certain cells that would impact their individual cellular functions. An intriguing example of this can be found in the field of learning and memory where recent results suggest that neural activity induces learning-specific changes at the level of the epigenome [55].

The ability to perform epigenomic analysis at single-cell resolution would help us to better understand these mechanisms. Researchers in the field are improving methods that may permit the analysis of DNA methylation states in single cells [56]. Evaluating histone modifications and other chromatin marks at specific loci in single cells will likely prove to be more challenging, but the technology is evolving [57]. The Common Fund Single Cell Analysis program may help develop or improve technologies to address some of the obstacles in this area (<https://commonfund.nih.gov/Singlecell>). As new methodologies are developed and reproducibility improves, more studies will incorporate single-cell approaches and a clearer picture will emerge regarding epigenomic variability within a defined cell population and, more importantly, the functional implications of any observed differences.

3.7 Environmental Epigenomics

No Roadmap Epigenomics Program initiative explicitly addressed the effects of environmental exposures on the epigenome. Yet it has become clear that well-controlled investigations in this area are critical, since it is likely that environmental exposures (e.g., environmental toxins, substances of abuse, psychosocial stress, diet, exercise) impact our epigenomes. However, the precise mechanisms by which an environmental exposure can lead to a chromatin change are still emerging [58–60]. Furthermore, it is not well understood how long these epigenetic alterations perdure in somatic cells or how these modifications are subsequently maintained or removed [61]. Also, as is the case for many epigenetic modifications, limited work has been done to rigorously test the functional effects of modifications associated with environmental exposures. Explicit investigation into these issues for a diverse set of environmental exposures would greatly enhance our understanding of how the environment influences the epigenome and whether or not different environmental stressors impact common epigenomic processes or perhaps impact a common set of genomic loci.

3.8 Functional Epigenomics Tools and Technologies

Over the past decade there have been dramatic improvements in our ability to scrutinize chromatin. Despite these improvements in our ability to monitor epigenomes, our ability to manipulate the epigenome has remained relatively crude and largely relies on pharmacological or genetic manipulation of epigenetic regulatory proteins. Pharmacological modulators of epigenetic enzymes and processes are of great value; however these modulators typically have pleiotropic effects and rarely impact a specific tissue, cell type, or locus. Genetic manipulation of epigenetic enzymes can achieve

results similar to those obtained by pharmacological manipulation, although gene targeting approaches can sometimes limit these effects to selected tissues or cell types. In addition, whether or not tissue or cell-specific epigenomic manipulation is achieved, the technologies currently available do not have robust temporal control and typically yield changes that impact the entire genome of a cell rather than focused effects on one or a few key loci. The potential pleiotropic effects resulting from genome-wide epigenome manipulation obscure our understanding of how epigenomic changes at specific loci may influence phenotype and are an obstacle to the development of future epigenomic therapeutics.

To address this scientific barrier, a Roadmap Epigenomics Program initiative was developed to stimulate innovative research to develop novel tools and technologies to enable one or more of the following: (1) tissue- or cell-specific manipulation of epigenetic modifications or their effector molecules, (2) temporal manipulation of the epigenome, and (3) locus-specific manipulation of the epigenome. As a part of this initiative, ten projects were funded that use some of the following general strategies: (1) use of TALE, Zn finger, CRISPR, or lncRNA strategies to target epigenetic modifying enzymes to specific loci for manipulation of the epigenome in a locus-specific manner; (2) development of opto-epigenetic or chemo-epigenetic switches to enable temporal manipulation of the epigenome using light or chemical ligands; and (3) the integration of the above techniques with genetic resources to enable cell-type-specific manipulation of the epigenome. We look forward to seeing the fruits of this research in coming years.

4 Summary

We hoped to convey how some of the resources and technologies generated by the NIH Roadmap Epigenomics Program may be exploited for studying various human diseases including cancer research projects. Some of the key resources and technologies generated by the program include reference epigenome maps for a wide variety of human cells and tissues, the development of new technologies for epigenetic assays and imaging, the identification of novel epigenetic modifications, and an improved understanding of the role of epigenetic processes in a diversity of human diseases. Despite the many accomplishments of the NIH Roadmap Epigenomics Program, much future work is needed in this research area including exploration of epigenomic variation both in different individuals and within cells of the “same” type, environmental epigenomics, exploration of the use of surrogate tissues, and improved technologies for epigenome manipulation. Ultimately our collective efforts to understand the epigenome will provide the foundational knowledge necessary to elucidate the role of

epigenomic mechanisms in a wide variety of human diseases and may lead to the development of new preventative measures, diagnostics, and therapeutics for these diseases.

Acknowledgements

The views expressed in this chapter are solely those of the authors and may not necessarily reflect those of NIH. The authors would like to thank all of the researchers associated with the NIH Roadmap Epigenomics Program for their tireless efforts to help transform this scientific field. We thank NIH senior leadership for their vision in selecting Epigenomics as a Common Fund program, especially Nora Volkow (NIDA), Linda Birnbaum (NIEHS), and James Battey (NIDCD). We thank the Common Fund Office of Science Coordination for financial support and encouragement. We thank the NIH Epigenomics Work Group members for their efforts and support during the development of this program. We especially thank Olivier Blondel, Mark Caulder, Christine Colvis, Stephanie Courchesne, Elise Feingold, Michelle Freund, Astrid Haugen, Tanya Hoodbhoy, Patricia Labosky, Roger Little, Pat Mastin, Michael Pazin, Phil Smith, Randall Stewart, and Elizabeth Wilder for their efforts in helping implement this program.

References

1. Jaenisch R, Bird A (2003) Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat Genet* 33(Suppl):245–254
2. Margueron R, Reinberg D (2010) Chromatin structure and the inheritance of epigenetic information. *Nat Rev Genet* 11:285–296
3. Bernstein BE, Meissner A, Lander ES (2007) The mammalian epigenome. *Cell* 128:669–681
4. McConnell MJ, Lindberg MR, Brennand KJ, Piper JC, Voet T, Cowing-Zitron C, Shumilina S, Lasken RS, Vermeesch JR, Hall IM, Gage FH (2013) Mosaic copy number variation in human neurons. *Science* 342:632–637
5. Jung D, Giallourakis C, Mostoslavsky R, Alt FW (2006) Mechanism and control of V(D)J recombination at the immunoglobulin heavy chain locus. *Annu Rev Immunol* 24:541–570
6. Feinberg AP (2010) Epigenomics reveals a functional genome anatomy and a new approach to common disease. *Nat Biotechnol* 28:1049–1052
7. Bell CG, Beck S (2010) The epigenomic interface between genome and environment in common complex diseases. *Brief Funct Genomics* 9:477–485
8. Rivera CM, Ren B (2013) Mapping human epigenomes. *Cell* 155:39–55
9. Meissner A (2010) Epigenetic modifications in pluripotent and differentiated cells. *Nat Biotechnol* 28:1079–1088
10. Beck S, Olek A, Walter J (1999) From genomics to epigenomics: a loftier view of life. *Nat Biotechnol* 17:1144
11. The American Association for Cancer Research Human Epigenome Task Force and the European Union, Network of Excellence, Scientific Advisory Board (2008) Moving AHEAD with an international human epigenome project. *Nature* 454:711–715
12. Jones PA, Martienssen R (2005) A blueprint for a Human Epigenome Project: the AACR Human Epigenome Workshop. *Cancer Res* 65:11241–11246
13. Bernstein BE, Stamatoyannopoulos JA, Costello JF, Ren B, Milosavljevic A, Meissner A, Kellis M, Marra MA, Beaudet al, Ecker JR, Farnham PJ, Hirst M, Lander ES, Mikkelsen

- TS, Thomson JA (2010) The NIH Roadmap Epigenomics Mapping Consortium. *Nat Biotechnol* 28:1045–1048
14. Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J, Nery JR, Lee L, Ye Z, Ngo QM, Edsall L, Antosiewicz-Bourget J, Stewart R, Ruotti V, Millar AH, Thomson JA, Ren B, Ecker JR (2009) Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* 462: 315–322
 15. Xie W, Schultz MD, Lister R, Hou Z, Rajagopal N, Ray P, Whitaker JW, Tian S, Hawkins RD, Leung D, Yang H, Wang T, Lee AY, Swanson SA, Zhang J, Zhu Y, Kim A, Nery JR, Urich MA, Kuan S, Yen CA, Klugman S, Yu P, Suknuntha K, Propson NE, Chen H, Edsall LE, Wagner U, Li Y, Ye Z, Kulkarni A, Xuan Z, Chung WY, Chi NC, Antosiewicz-Bourget JE, Slukvin I, Stewart R, Zhang MQ, Wang W, Thomson JA, Ecker JR, Ren B (2013) Epigenomic analysis of multilineage differentiation of human embryonic stem cells. *Cell* 153:1134–1148
 16. Gifford CA, Ziller MJ, Gu H, Trapnell C, Donaghey J, Tsankov A, Shalek AK, Kelley DR, Shishkin AA, Issner R, Zhang X, Coyne M, Fostel JL, Holmes L, Meldrim J, Guttman M, Epstein C, Park H, Kohlbacher O, Rinn J, Gnirke A, Lander ES, Bernstein BE, Meissner A (2013) Transcriptional and epigenetic dynamics during specification of human embryonic stem cells. *Cell* 153:1149–1163
 17. Zhu J, Adli M, Zou JY, Verstappen G, Coyne M, Zhang X, Durham T, Miri M, Deshpande V, De Jager PL, Bennett DA, Houmard JA, Muoio DM, Onder TT, Camahort R, Cowan CA, Meissner A, Epstein CB, Shores N, Bernstein BE (2013) Genome-wide chromatin state transitions associated with developmental and environmental cues. *Cell* 152:642–654
 18. Stergachis AB, Haugen E, Shafer A, Fu W, Vernot B, Reynolds A, Raubitschek A, Ziegler S, LeProust EM, Akey JM, Stamatoyannopoulos JA (2013) Exonic transcription factor binding directs codon choice and affects protein evolution. *Science* 342:1367–1372
 19. Maurano MT, Humbert R, Rynes E, Thurman RE, Haugen E, Wang H, Reynolds AP, Sandstrom R, Qu H, Brody J, Shafer A, Neri F, Lee K, Kutayavin T, Stehling-Sun S, Johnson AK, Canfield TK, Giste E, Diegel M, Bates D, Hansen RS, Neph S, Sabo PJ, Heimfeld S, Raubitschek A, Ziegler S, Cotsapas C, Sotoodehnia N, Glass I, Sunyaev SR, Kaul R, Stamatoyannopoulos JA (2012) Systematic localization of common disease-associated variation in regulatory DNA. *Science* 337:1190–1195
 20. Chadwick LH (2012) The NIH Roadmap Epigenomics Program data resource. *Epigenomics* 4:317–324
 21. Karnik R, Meissner A (2013) Browsing (Epi) genomes: a guide to data resources and epigenome browsers for stem cell researchers. *Cell Stem Cell* 13:14–21
 22. Aguilar-Valles A, Vaissiere T, Griggs EM, Mikaelsson MA, Takacs IF, Young EJ, Rumbaugh G, Miller CA (2013) Methamphetamine-associated memory is regulated by a writer and an eraser of permissive histone methylation. *Biol Psychiatry* 1:1–18, <http://dx.doi.org/10.1016/j.biopsych.2013.09.014>
 23. Jirtle RL, Skinner MK (2007) Environmental epigenomics and disease susceptibility. *Nat Rev Genet* 8:253–262
 24. Ganu RS, Harris RA, Collins K, Aagaard KM (2012) Early origins of adult disease: approaches for investigating the programmable epigenome in humans, nonhuman primates, and rodents. *ILAR J* 53:306–321
 25. Zhang Z, Tan M, Xie Z, Dai L, Chen Y, Zhao Y (2011) Identification of lysine succinylation as a new post-translational modification. *Nat Chem Biol* 7:58–63
 26. Tan M, Luo H, Lee S, Jin F, Yang JS, Montellier E, Buchou T, Cheng Z, Rousseaux S, Rajagopal N, Lu Z, Ye Z, Zhu Q, Wysocka J, Ye Y, Khochbin S, Ren B, Zhao Y (2011) Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. *Cell* 146:1016–1028
 27. Yuan CC, Matthews AG, Jin Y, Chen CF, Chapman BA, Ohsumi TK, Glass KC, Kutateladze TG, Borowsky ML, Struhl K, Oettinger MA (2012) Histone H3R2 symmetric dimethylation and histone H3K4 trimethylation are tightly correlated in eukaryotic genomes. *Cell Rep* 1:83–90
 28. Dhayalan A, Tamas R, Bock I, Tattermusch A, Dimitrova E, Kudithipudi S, Ragozin S, Jeltsch A (2011) The ATRX-ADD domain binds to H3 tail peptides and reads the combined methylation state of K4 and K9. *Hum Mol Genet* 20:2195–2203
 29. Lodhi N, Tulin AV (2011) PARP1 genomics: chromatin immunoprecipitation approach using anti-PARP1 antibody (ChIP and ChIP-seq). *Methods Mol Biol* 780:191–208
 30. Peng JC, Valouev A, Swigut T, Zhang J, Zhao Y, Sidow A, Wysocka J (2009) Jarid2/Jumonji coordinates control of PRC2 enzymatic activity and target gene occupancy in pluripotent cells. *Cell* 139:1290–1302
 31. Baker SP, Phillips J, Anderson S, Qiu Q, Shabanowitz J, Smith MM, Yates JR III, Hunt

- DF, Grant PA (2010) Histone H3 Thr 45 phosphorylation is a replication-associated post-translational modification in *S. cerevisiae*. *Nat Cell Biol* 12:294–298
32. Adli M, Bernstein BE (2011) Whole-genome chromatin profiling from limited numbers of cells using nano-ChIP-seq. *Nat Protoc* 6: 1656–1668
 33. Harris RA, Wang T, Coarfa C, Nagarajan RP, Hong C, Downey SL, Johnson BE, Fouse SD, Delaney A, Zhao Y, Olshen A, Ballinger T, Zhou X, Forsberg KJ, Gu J, Echipare L, O'Geen H, Lister R, Pelizzola M, Xi Y, Epstein CB, Bernstein BE, Hawkins RD, Ren B, Chung WY, Gu H, Bock C, Gnirke A, Zhang MQ, Haussler D, Ecker JR, Li W, Farnham PJ, Waterland RA, Meissner A, Marra MA, Hirst M, Milosavljevic A, Costello JF (2010) Comparison of sequencing-based methods to profile DNA methylation and identification of monoallelic epigenetic modifications. *Nat Biotechnol* 28:1097–1105
 34. Gu H, Bock C, Mikkelsen TS, Jager N, Smith ZD, Tomazou E, Gnirke A, Lander ES, Meissner A (2010) Genome-scale DNA methylation mapping of clinical samples at single-nucleotide resolution. *Nat Methods* 7: 133–136
 35. Lister R, O'Malley RC, Tonti-Filippini J, Gregory BD, Berry CC, Millar AH, Ecker JR (2008) Highly integrated single-base resolution maps of the epigenome in *Arabidopsis*. *Cell* 133:523–536
 36. Cipriany BR, Murphy PJ, Hagarman JA, Cerf A, Latulippe D, Levy SL, Benitez JJ, Tan CP, Topolancik J, Soloway PD, Craighead HG (2012) Real-time analysis and selection of methylated DNA by fluorescence-activated single molecule sorting in a nanofluidic channel. *Proc Natl Acad Sci U S A* 109:8477–8482
 37. Byrum SD, Raman A, Taverna SD, Tackett AJ (2012) ChAP-MS: a method for identification of proteins and histone posttranslational modifications at a single genomic locus. *Cell Rep* 2:198–205
 38. Deal RB, Henikoff JG, Henikoff S (2010) Genome-wide kinetics of nucleosome turnover determined by metabolic labeling of histones. *Science* 328:1161–1164
 39. Clowney EJ, LeGros MA, Mosley CP, Clowney FG, Markenskoff-Papadimitriou EC, Myllys M, Barnea G, Larabell CA, Lomvardas S (2012) Nuclear aggregation of olfactory receptor genes governs their monogenic expression. *Cell* 151:724–737
 40. Wang Y, Zhang YL, Hennig K, Gale JP, Hong Y, Cha A, Riley M, Wagner F, Haggarty SJ, Holson E, Hooker J (2013) Class I HDAC imaging using [(3)H]CI-994 autoradiography. *Epigenetics* 8:756–764
 41. Schroeder FA, Chonde DB, Riley MM, Moseley CK, Granda ML, Wilson CM, Wagner FF, Zhang YL, Gale J, Holson EB, Haggarty SJ, Hooker JM (2013) FDG-PET imaging reveals local brain glucose utilization is altered by class I histone deacetylase inhibitors. *Neurosci Lett* 550:119–124
 42. Yeh HH, Tian M, Hinz R, Young D, Shavrin A, Mukhopadhyay U, Flores LG, Balatoni J, Soghomonyan S, Jeong HJ, Pal A, Uthamanthil R, Jackson JN, Nishii R, Mizuma H, Onoe H, Kagawa S, Higashi T, Fukumitsu N, Alauddin M, Tong W, Herholz K, Gelovani JG (2013) Imaging epigenetic regulation by histone deacetylases in the brain using PET/MRI with ¹⁸F-FAHA. *Neuroimage* 64:630–639
 43. Feinberg AP, Vogelstein B (1983) Hypomethylation distinguishes genes of some human cancers from their normal counterparts. *Nature* 301:89–92
 44. Whyte WA, Orlando DA, Hnisz D, Abraham BJ, Lin CY, Kagey MH, Rahl PB, Lee TI, Young RA (2013) Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* 153:307–319
 45. Hnisz D, Abraham BJ, Lee TI, Lau A, Saint-Andre V, Sigova AA, Hoke HA, Young RA (2013) Super-enhancers in the control of cell identity and disease. *Cell* 155:934–947
 46. Hsu PY, Hsu HK, Singer GA, Yan PS, Rodriguez BA, Liu JC, Weng YI, Deatherage DE, Chen Z, Pereira JS, Lopez R, Russo J, Wang Q, Lamartiniere CA, Nephew KP, Huang TH (2010) Estrogen-mediated epigenetic repression of large chromosomal regions through DNA looping. *Genome Res* 20:733–744
 47. Davies MN, Volta M, Pidsley R, Lunnon K, Dixit A, Lovestone S, Coarfa C, Harris RA, Milosavljevic A, Troakes C, Al-Sarraj S, Dobson R, Schalkwyk LC, Mill J (2012) Functional annotation of the human brain methylome identifies tissue-specific epigenetic variation across brain and blood. *Genome Biol* 13:R43
 48. Egelhofer TA, Minoda A, Klugman S, Lee K, Kolasinska-Zwierz P, Alekseyenko AA, Cheung MS, Day DS, Gadel S, Gorchakov AA, Gu T, Kharchenko PV, Kuan S, Latorre I, Linder-Basso D, Luu Y, Ngo Q, Perry M, Rechtsteiner A, Riddle NC, Schwartz YB, Shanower GA, Vielle A, Ahringer J, Elgin SC, Kuroda MI, Pirrotta V, Ren B, Strome S, Park PJ, Karpen GH, Hawkins RD, Lieb JD (2011) An assessment of histone-modification antibody quality. *Nat Struct Mol Biol* 18:91–93
 49. Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I,

- Lajoie BR, Sabo PJ, Dorschner MO, Sandstrom R, Bernstein B, Bender MA, Groudine M, Gnirke A, Stamatoyannopoulos J, Mirny LA, Lander ES, Dekker J (2009) Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* 326:289–293
50. Fullwood MJ, Liu MH, Pan YF, Liu J, Xu H, Mohamed YB, Orlov YL, Velkov S, Ho A, Mei PH, Chew EG, Huang PY, Welboren WJ, Han Y, Ooi HS, Ariyaratne PN, Vega VB, Luo Y, Tan PY, Choy PY, Wansa KD, Zhao B, Lim KS, Leow SC, Yow JS, Joseph R, Li H, Desai KV, Thomsen JS, Lee YK, Karuturi RK, Herve T, Bourque G, Stunnenberg HG, Ruan X, Cacheux-Rataboul V, Sung WK, Liu ET, Wei CL, Cheung E, Ruan Y (2009) An oestrogen-receptor- α -bound human chromatin interactome. *Nature* 462:58–64
 51. Zhao YQ, Jordan IK, Lunyak VV (2013) Epigenetics components of aging in the central nervous system. *Neurotherapeutics* 10:647–663
 52. Ben-Avraham D, Muzumdar RH, Atzmon G (2012) Epigenetic genome-wide association methylation in aging and longevity. *Epigenomics* 4:503–509
 53. Hong CP, Park J, Roh TY (2011) Epigenetic regulation in cell reprogramming revealed by genome-wide analysis. *Epigenomics* 3:73–81
 54. Wills QF, Livak KJ, Tipping AJ, Enver T, Goldson AJ, Sexton DW, Holmes C (2013) Single-cell gene expression analysis reveals genetic associations masked in whole-tissue experiments. *Nat Biotechnol* 31:748–752
 55. Zovkic IB, Guzman-Karlsson MC, Sweatt JD (2013) Epigenetic regulation of memory formation and maintenance. *Learn Mem* 20: 61–74
 56. Kantlehner M, Kirchner R, Hartmann P, Ellwart JW, Alunni-Fabbroni M, Schumacher A (2011) A high-throughput DNA methylation analysis of a single cell. *Nucleic Acids Res* 39:e44
 57. Hayashi-Takanaka Y, Yamagata K, Wakayama T, Stasevich TJ, Kainuma T, Tsurimoto T, Tachibana M, Shinkai Y, Kurumizaka H, Nozaki N, Kimura H (2011) Tracking epigenetic histone modifications in single cells using Fab-based live endogenous modification labeling. *Nucleic Acids Res* 39:6475–6488
 58. Robison AJ, Nestler EJ (2011) Transcriptional and epigenetic mechanisms of addiction. *Nat Rev Neurosci* 12:623–637
 59. Zhang TY, Labonte B, Wen XL, Turecki G, Meaney MJ (2013) Epigenetic mechanisms for the early environmental regulation of hippocampal glucocorticoid receptor gene expression in rodents and humans. *Neuropsychopharmacology* 38:111–123
 60. Guerrero-Bosagna C, Savenkova M, Haque MM, Nilsson E, Skinner MK (2013) Environmentally induced epigenetic transgenerational inheritance of altered Sertoli cell transcriptome and epigenome: molecular etiology of male infertility. *PLoS One* 8:e59922
 61. Berger SL, Kouzarides T, Shiekhattar R, Shilatifard A (2009) An operational definition of epigenetics. *Genes Dev* 23:781–783

Cancer Epigenetics

Risk Assessment, Diagnosis, Treatment, and Prognosis

Verma, M. (Ed.)

2015, XVI, 804 p. 73 illus., 31 illus. in color., Hardcover

ISBN: 978-1-4939-1803-4

A product of Humana Press