

Genes & Cancer

<http://gan.sagepub.com/>

The Decade of the Epigenomes?

Joost H.A. Martens, Hendrik G. Stunnenberg and Colin Logie
Genes & Cancer 2011 2: 680 originally published online 19 August 2011
DOI: 10.1177/1947601911417860

The online version of this article can be found at:
<http://gan.sagepub.com/content/2/6/680>

Published by:



<http://www.sagepublications.com>

Additional services and information for *Genes & Cancer* can be found at:

Email Alerts: <http://gan.sagepub.com/cgi/alerts>

Subscriptions: <http://gan.sagepub.com/subscriptions>

Reprints: <http://www.sagepub.com/journalsReprints.nav>

Permissions: <http://www.sagepub.com/journalsPermissions.nav>

The Decade of the Epigenomes?

Joost H.A. Martens, Hendrik G. Stunnenberg, and Colin Logie

Abstract

The beginning of this century was not only marked by the publication of the first draft of the human genome but also set off a decade of intense research on epigenetic phenomena. Apart from DNA methylation, it became clear that many other factors including a wide range of histone modifications, different shades of chromatin accessibility, and a vast suite of noncoding RNAs comprise the epigenome. With the recent advances in sequencing technologies, it has now become possible to analyze many of these features in depth, allowing for the first time the establishment of complete epigenomic profiles for basically every cell type of interest. Here, we will discuss the recent advances that allow comprehensive epigenetic mapping, highlight several projects that set out to better understand the epigenome, and discuss the impact that epigenomic mapping can have on our understanding of both healthy and diseased cells.

Keywords: epigenome, chromatin accessibility, DNA methylation, ChIP-seq, RNA-seq

What Is the Epigenome?

One of the enduring mysteries of biology is that we all start off as a single cell but that, instead of becoming an ever-expanding mass of identical cells, we are made up of a wide variety of specialized tissues, despite carrying basically the same DNA in each cell. Somehow, each of the 200 different kinds of cells in the human body must be reading off a different set of the hereditary instructions written into the DNA.¹ It has become clear that the availability of these hereditary instructions is mediated through a vast suite of epigenetic marks. These comprise a variety of molecular and structural modifications to DNA that do not change the underlying sequence but ensure that the right genes are expressed at the right time.² As such, each cell type, including not only healthy but also diseased cells, is thought to have a unique epigenetic state that determines its identity.³ The epigenetic state of the cells represents the sum of developmental and physiological influences. As such, the epigenome provides a crucial interface between the environment and the genome. Indeed, various factors such as nutrition, toxins, drugs, infection, disease state, and exposure to toxic agents are known to affect DNA and histone modifications.^{4,5}

Epigenetic states have been described that are meiotically heritable, whereas others can only be propagated through mitosis.⁶ Other chromatin changes are even more transient and coupled solely to active gene transcription. Although all marks are part of the cellular epigenome, the extent to which an epigenetic mark is stable determines to what extent it can embody cellular memory.^{5,7}

With the development of new techniques that allowed investigation of larger genomic regions or even an entire mammalian genome, epigenetic research has changed dramatically over the last decade, as exemplified by the vast increase in the number of publications mentioning the word “epigenome” over the last 10 years (Fig. 1). Initially, epigenetic research was focused on single loci combining, for example, chromatin immunoprecipitation with quantitative PCR analysis. This was followed by techniques based on combining chromatin immunoprecipitation (ChIP) or other chromatin enrichment methods with DNA hybridization to either custom or commercial tiled oligonucleotide microarrays (ChIP-chip). Although these techniques allowed for the first time global analysis of epigenetic marks, the sequence resolution and genomic coverage per microarray were relatively low.

Moreover, as it required multiple arrays to cover a complete genome of higher eukaryotes, the technique was very costly in order to generate comprehensive views on chromatin modifications. However, nowadays, with the widespread adoption of next-generation sequencing (NGS), genome-wide epigenetic mapping experiments can be performed with unprecedented resolution,^{5,8-10} and this has truly opened the field of epigenomics. Indeed, many projects throughout the world have been initiated to shed more light on epigenetic phenomena at the genome-wide level. Here, we will discuss the different aspects that need to be taken into account when starting large-scale epigenetic projects and will also examine the current state of epigenomic research.

What Is Worthwhile Mapping?

Several aspects need to be taken into account when expanding epigenetic research to the genome-wide level. The

Department of Molecular Biology, Faculty of Science, Nijmegen Centre for Molecular Life Sciences, Radboud University, Nijmegen, the Netherlands

Corresponding Author:

Joost H.A. Martens
 Email: j.martens@ncmls.ru.nl

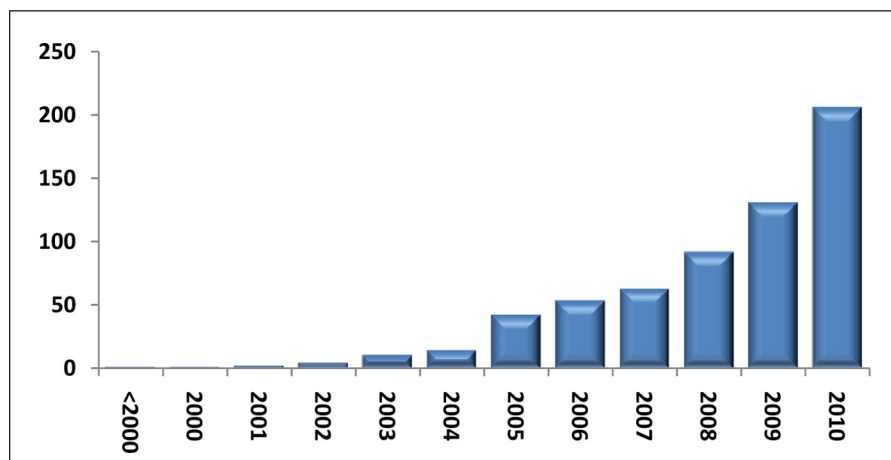


Figure 1. Scientific publications since 2000 mentioning the word “epigenome.”

first is to determine what would be most interesting to map. As every researcher probably favors his or her own model system, the answer to this question will be similar as for the DNA sequencing projects: in the end, we would like to know the full epigenome of all organisms of potential interest. Indeed, with the rapid increase in sequence technology and following the explosion of fully sequenced organisms in the last decade, it can be expected that in the coming decade, mapping of epigenomes will become routine business, ultimately resulting in establishing the epigenomic maps of most human tissues and cell types as well as those of many model organisms. However, as current resources are still limited, choices have to be made in the selection of cell types and epigenetic features that can be included into epigenomic projects.⁵ An important consideration is the amount of material that can be used. Because current technologies require substantial amounts of a rather homogenous cell population as a starting point, this is still one of the limiting steps for establishing complete epigenomic maps. Although model cell lines are ideal in terms of homogeneity, it has become clear that their epigenomic make-up is far from representative of the cells that constitute the tissue they are derived from. On the other hand, many human tissues are very heterogeneous and yield insufficient cells of a single

type to permit comprehensive epigenetic analysis. As a result, current efforts mainly focus on cell populations that can be obtained in great quantity and that are fairly homogenous such as CD4+ cells.^{11,12} To bypass this problem, pooling samples would be an option, although this again is a less optimal solution as one has to consider the variation that is being introduced within the source material. However, it has to be noted that although current techniques require substantial amounts of material, these problems are expected to be overcome at least partially by the introduction of new techniques that require less starting material.^{13,14} In summary, each epigenomic project has to carefully seek a balance between the scientific priorities, the availability of relevant (tissue) materials, and the techniques that can be utilized, in addition, of course, to the budget that can be allocated.

How to Map the Epigenome?

Current technology allows for the first time the comprehensive mapping of all different aspects of the epigenome: ChIP-seq can be used to examine histone modifications at the genome-wide level, RNA-seq for profiling the transcriptome, and a variety of techniques are available to examine global accessibility and the DNA methylome (Fig. 2). An important consideration for epigenome analysis is the establishment of

standard operating procedures (SOP) that allow data of the highest quality to be generated and compared through similar efforts among different consortia. Currently, IHEC, an international consortium that sets out to establish high-resolution reference epigenome maps for the research community, aims to define these standards, not only for data generation but also for data storage and availability. This should be the primary guideline for every effort into epigenome generation throughout the world.

ChIP-Seq

The aim of ChIP-seq experiments is to map all *in vivo* DNA sites occupied by a factor of interest. To do this, an antibody specifically recognizing a DNA-binding protein is used to immunoprecipitate the protein that has been cross-linked to its DNA-binding sites in living cells. Within epigenomic projects, the protein of interest mostly comprises a modified histone, although the technique can be extended to proteins that are part of the general transcription machinery, sequence-specific transcription factors, and other DNA-interacting proteins such as those involved in replication, recombination, or repair. Until recently, global identification of ChIPed DNA fragments was almost always done by hybridizing the ChIP DNA to microarrays that tile part or most of the genome being studied (ChIP-chip). Although valuable data have been generated this way, ChIP-chip suffered from limitations generally associated with microarrays such as the low spatial resolution and genomic coverage per microarray, its requirement for multiple arrays to cover a complete genome of higher eukaryotes, and the biases introduced in the ChIP-chip procedure.¹⁵ ChIP-seq has overcome most of the biggest problems. In ChIP-seq, massive parallel sequencing of ChIP DNA generates libraries of relatively short DNA sequences that can be mapped to an annotated genome, thereby determining the localization of proteins or chromatin modifications of interest. This approach has been shown to be of significant

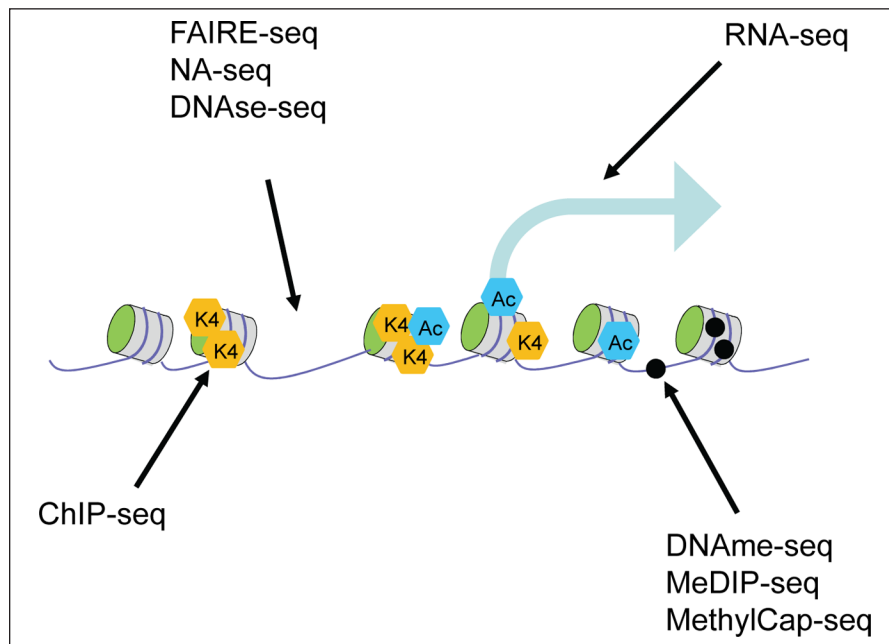


Figure 2. Current technologies used to map epigenomes.

importance as it is relatively low cost and offers the possibility to perform genome-wide analysis at single base-pair resolution.¹⁶ ChIP-seq does not suffer from the false or uncertain signals that can result from cross-hybridization, and quantification is more accurate because counting sequence reads is “digital” rather than the continuous scale of fluorescence signal obtained with microarray hybridization. Finally, ChIP-seq can focus on a binding site at higher resolution than is typical for ChIP-chip. Indeed, a size selection step on the ChIP DNA and computational features typically allow binding sites to be assigned to regions of less than 100 bp.

An important question remains, and this concerns which histone modifications should be mapped. The definition of what constitutes an epigenome of lasting value and, consequently, which marks should be mapped is affected by many parameters such as biological/clinical relevance, suitable technology/reagents, data quality, compatibility with other efforts, and again, costs. A guideline given by the IHEC consortium (<http://www.ihec-epigenomes.org/>) is to map the following 6 histone modifications: H3K4me3, H3K9me3, H3K27me3, H3K27ac, H3K4me1, and H3K36me3.

The presence of these marks is indicative of active promoters (H3K4me3, H3K27ac), active enhancers (H3K4me1, H3K27ac), actively transcribed genes (H3K36me3), or heterochromatin regions (H3K9me3, H3K27me3). This list can be extended with many marks, and indeed, another consortium operating under the umbrella of IHEC called BLUEPRINT (see below) aims at including several other informative marks such as H2A.z as a marker for open chromatin regions, H4K20me3 as an additional mark for heterochromatin, p300 and/or CBP as additional enhancer markers, H3K79me3 as a transcription initiation marker, H3 and H4 acetylation as markers of promoters and transcription, and RNA polymerase II as a marker of transcription activity. Undoubtedly, this list can and will be extended with the development of new reagents as well as with discovery of other relevant molecular chromatin features.

RNA-Seq

For profiling mRNA populations, microarrays have dominated for more than a decade, bringing us most of what we know about entire transcriptomes from yeast and bacteria to mouse and man.¹⁵

However, throughout the research community, this method has recently been replaced with RNA-seq. Unlike the array-based methods, RNA-seq is quantitative and will detect exon-specific expression, hence providing important information on mRNA isoforms. In addition, RNA-seq can provide a wealth of information about gene and transcript structure as sequence reads from within exons as well as reads spanning exons will be mapped (“junction reads”).¹⁷ Also, the seq-based methods can be allele specific through detection of sequence polymorphisms and, by high sampling, detect very low abundance RNAs unambiguously, for example, from a rare subpopulation of cells present in the sample. Finally, extending RNA-seq to paired-end reads permits systematic discovery of alternative splicing events.¹⁸

Genome-Wide DNA Methylation Mapping

Since the advent of ultra high-throughput sequencing, several groups have begun to use this technique to help measure the methylation status of DNA at CpG clusters in the human genome. These methods use methylated DNA immunoprecipitation (MeDIP) assays coupled to sequencing (MeDIP-seq)¹⁹ or capturing of methylated DNA followed by sequencing (MethylCap-seq).²⁰ These assays are aimed at producing a reduced genome sample of CpG island fragments, which in some methods are treated with bisulfite and then sequenced. Although these techniques do not allow a complete coverage of all CpGs in the human genome, they all provide a massive increase in the number of regions that can be assayed in a single, relatively cheap experiment compared to even the most efficient available assays in the pre-NGS era. The increases in sequencing capacity resulted in 2009 in the description of the first human DNA methylome at base-pair resolution,²¹ marking a major step forward in uncovering the DNA methylome and setting a new standard for DNA methylome analysis. Apart from conventional DNA methylation, recently, another DNA modification was identified, DNA hydroxymethylation

(5hmC),^{22,23} adding a new layer of epigenetic information but also adding an additional challenge for developing genome-wide technologies that can detect this mark and also discriminate between the 2 modifications.

Chromatin Accessibility Mapping

Chromosomal DNA is occluded by virtue of being wrapped into nucleosomes. This can be detected by endonuclease accessibility, and high accessibility is a distinct hallmark of the epigenome that is strongly associated with regulatory regions. Chromatin accessibility is a strong indicator of transcription factor occupancy and nucleosome remodeling that occurs at genomic regions participating in gene regulation or other nuclear processes. Therefore, over the past years, several techniques have emerged to assess the accessibility of the epigenome, including DNase I hypersensitivity,¹¹ FAIRE,²⁴ and nuclease accessibility²⁵ (NA) mapping. Mapping DNase I-hypersensitive (HS) sites has throughout many years been a valuable tool for identifying a wide variety of different types of regulatory elements, such as promoters, enhancers, silencers, insulators, and locus control regions.²⁶ The method utilizes the DNase I enzyme to selectively digest DNA depleted (presumably through the actions of transcription factors) of nucleosomes, whereas DNA regions tightly wrapped in nucleosome and higher order structures are more resistant. DNase-seq is a high-throughput method that identifies DNase I HS sites across the entire genome by capturing and sequencing DNase-digested fragments.

An alternative to the enzyme-based DNase-seq and NA-seq detection of functional *in vivo* regulatory elements in mammalian cells is FAIRE (formaldehyde-assisted isolation of regulatory elements). Apart from the lack of an enzymatic component, advantages of FAIRE include that it does not require additional treatments before the open chromatin state is captured and that it

positively selects chromatin regions that are accessible.

Data Analysis

Apart from data generation, an important aspect includes the data storage and the primary data analysis, such as alignment of sequencing tags. Large-scale epigenomic projects generate such massive amounts of data that only a processed output can be stored. Furthermore, the data volume is only expected to increase over the coming years. Therefore, data need to be efficiently stored, and sufficient capacity needs to be in place to allow for this. Moreover, it is still common in large-scale projects for data to undergo primary analysis in disparate ways, whereby individual pipelines are utilized to generate data sets.⁵ Such varieties in data handling impede integrated comparative analyses by introducing additional variability. Therefore, standard pipelines should be in place for primary analysis of the data sets that are generated wherever possible so as to allow optimal comparability of data sets and reduce the extra efforts that need to be made to align data sets from different projects. It can be anticipated that the IHEC consortium will have a leading role in coordinating these data-structuring efforts. Ultimately, alignment of NGS data sets will allow comparison of the different epigenomes, resulting in a better comprehension of the role of genetic and epigenetic variation in cell physiology. Although the field of epigenomic comparisons is not yet fully matured, conceptual and computational approaches are already being developed that address the heterogeneity and context dependence of these genome-wide epigenetic data sets.^{8,27}

Recent, Ongoing, and Future Epigenomic Projects

The field of epigenetics has expanded over the last decade, and the initial small-scale epigenetic efforts at the beginning of the century were soon followed by the initiation of large

epigenetic consortia that set out to understand the wealth of information stored in the epigenome. Most of these consortia had a strong epigenetic/epigenomic component and focused on subsets of epigenetic modifications (such as DNA methylation) as well as on other biological aspects. With the arrival of new sequencing technologies, the aim of more recent projects has shifted toward mapping full epigenomes as the major objective. Clearly, these efforts require substantial resources, and as a result, the need for collaboration has grown. Indeed, initially, most projects were initiated at a national or continental level, while more recently, efforts have extended toward intercontinental collaborations (Fig. 3). As such, it is clear that mapping the epigenome is really a global effort.

One of the first projects with a major epigenetic component was the **ENCODE** (<http://www.genome.gov/10005107>) project.²⁸ The Encyclopedia of DNA Elements (ENCODE) project launched in 2003 by the United States National Human Genome Research Institute aimed to identify all functional elements in the human genome sequence, whereas the modENCODE²⁹ project has the same goal with respect to model organisms. Both projects use a wide array of different assays to identify functional elements, and as such, an important aspect of the program includes the analysis of large-scale epigenetic changes. In the same period, the Epigenome Network of Excellence (**NOE**; www.epigenome-noe.net/), one of the first European projects fully focused on epigenetics, was launched. Its major aims are to provide clear and visible benefits for the entire field of epigenetic research and to support the epigenetic community with conferences, workshops, training visits, and shared resources. Throughout the years, it has been developed into a powerful framework that assimilates existing synergies within Europe and has attracted widespread attention not only from the European epigenetic community but also from outside. A similar effort was initiated in 2008 in Australia where

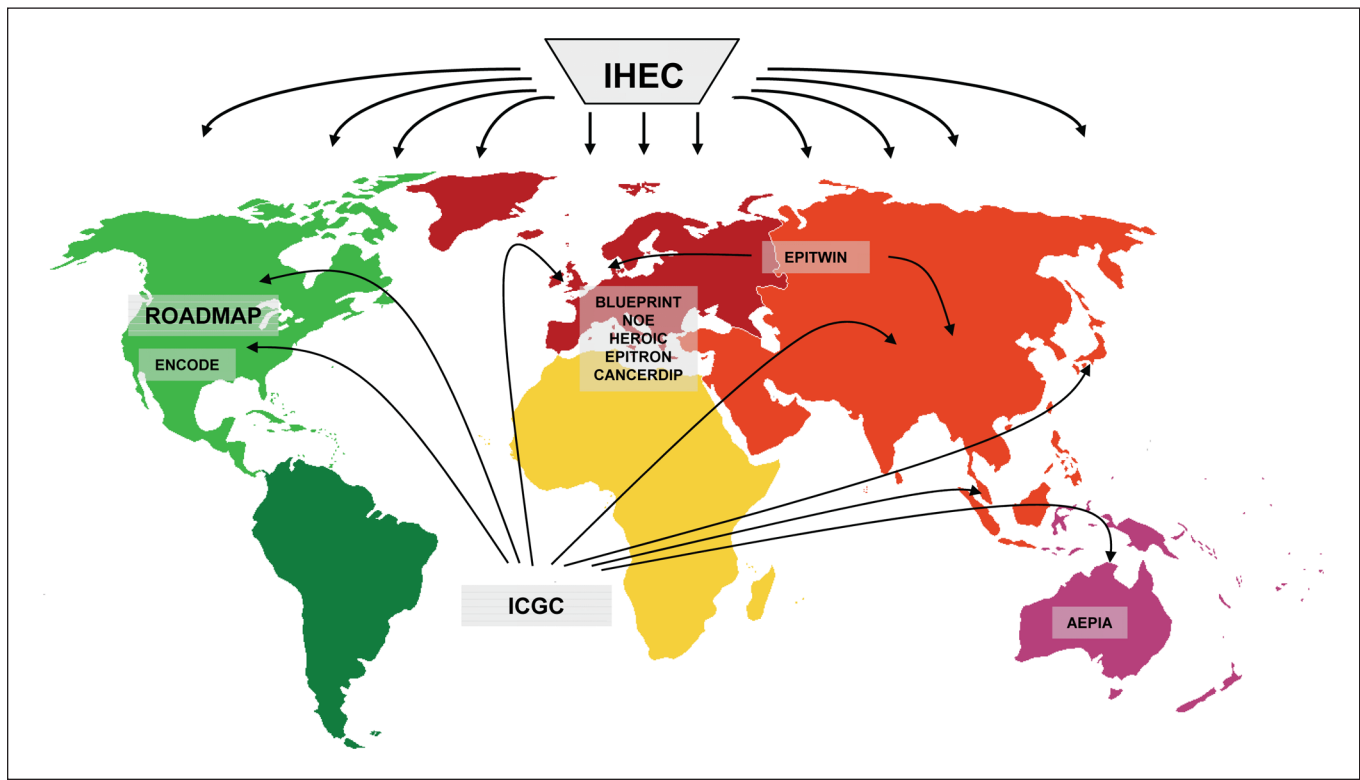


Figure 3. Examples of consortia that set out to map epigenomes.

researchers formed the Australian Alliance for Epigenetics (**AEPIA**; <http://www.epialliance.org.au/>).

Disease Epigenomic Projects

Many diseases are associated with alterations of the epigenome³⁰ and can include abnormalities caused by direct genetic changes in epigenetic enzymes as well as indirect changes through aberrant regulation of proteins regulating the epigenetic enzymes. For example, in leukemia, genetic changes in epigenetic enzymes such as the TET2³¹ and EZH2³² in myelodysplastic syndromes (MDS), MLL in acute lymphocytic leukemia (ALL) and ALL, PHF23, DNMT3A, and JARID1A in acute myeloid leukemia^{33,34} have been reported. In addition, in a subtype of leukemia called acute promyelocytic leukemia (APL), it was reported that a large set of epigenetic genes was targeted by PML-RAR α in addition to aberrant recruitment of

HDACs by this oncofusion protein,³⁵ showing that genetic alterations that do not directly involve an epigenetic enzyme can have a profound effect on the epigenome also. In addition, as epigenetic changes such as aberrant DNA methylation and histone modifications may be mediated by exposure to environmental or lifestyle factors, comprehensive epigenetic analysis of disease genomes is suggested to be one of the most effective ways to identify causative changes involved in tumorigenesis, irrespective of whether they are inherited or acquired. Projects such as Epitron, CancerDIP, and ICGC are examples of consortia that set out to gain more insight into the connections between aberrant regulation of epigenetic modifications and disease at a large-scale level. The **EPITRON** (<http://www.epitron.eu/>) consortium's primary goal is to define the epigenetic alterations by which cancer cells deviate from their normal counterparts and to identify novel targets

that can be subjected to treatment with "epigenetic" drugs. The **CANCERDIP** (<http://www.cancerdip.eu/>) consortium's primary focus is the identification of DNA methylation patterns in different tumor types as well as the identification of the epigenetic machinery involved in the establishment of these aberrant patterns, the major aim being to understand the pathway that leads to cancer but also to generate markers with clinical value. Similarly, the International Cancer Genome Consortium (**ICGC**; <http://www.icgc.org/>),³⁶ consisting of groups from the US, Canada, Europe, Australia, and Asia, was launched to coordinate large-scale cancer genome studies in tumors from 50 different cancer types and/or subtypes that are of clinical and societal importance across the globe. In terms of the epigenome, the main goal is to examine global DNA methylation, and within this project, the integration of DNA methylome data with genomic data and expression is expected to provide

novel insights into the interplay between the genome and the epigenome.

The Epigenomes of Healthy Cells

Epigenetic modifications constitute a complex layer of information on top of the genome sequence and are thought to be cell type specific. As such, every cell type and tissue that arise during development are expected to be uniquely identifiable through its epigenome. Comparisons that include varieties of pluripotent and differentiated cells or both provide powerful systems for investigating how the epigenetic code influences cellular fate. The **NIH Roadmap Epigenomics Mapping Consortium** (<http://www.roadmapepigenomics.org/>),³⁷ which was launched in 2008, was one of the first projects focusing primarily on establishing epigenomes in normal cells and tissues. The consortium aims to utilize next-generation sequencing technologies to map DNA methylation, histone modifications, chromatin accessibility, and small RNA transcripts in stem cells and primary *ex vivo* tissues that are selected to represent the normal counterparts of tissues and organ systems frequently involved in human disease. It is expected to deliver a collection of normal epigenomes that will provide a reference framework for comparison and integration of a broad array of future studies.

The **Heroic** project was one of the first European projects that adapted next-generation sequencing to examine epigenetic phenomena in well-characterized progenitor differentiation systems such as pluripotent mouse ES cells and paradigm-silencing models from genomic imprinting and X inactivation. This project set the stage for a follow-up project called **BLUEPRINT**, which represents the European branch of the **IHEC** consortium. The **BLUEPRINT** consortium is the first European consortium solely devoted to epigenomic mapping. It aims to generate at least 100 reference epigenomes and focuses on distinct types of hematopoietic cells from healthy

individuals as well as on their malignant leukemic counterparts.

Epitwin, a joint project of research groups in the United Kingdom and China, will capture the subtle epigenetic signatures that mark the differences between 5,000 twins. The consortium plans to look at the methylation patterns of 20 million CpGs in the DNA of each twin and to compare them. Rather than looking at similarities, the consortium will search for differences that explain why many identical twins do not develop the same disease, and as such, this will provide important insights into the environmentally induced aspects of the epigenome. Initially, the consortium will focus on obesity, diabetes, allergies, heart disease, osteoporosis, and longevity, but the methods that will be developed are in principle applicable to every common trait or disease.

Toward a Map of All Human Healthy and Diseased Epigenomes

To coordinate all the above epigenomic efforts, a major role is set aside for the International Human Epigenome Consortium (IHEC; <http://www.ihecepigenomes.org/>). Its primary aim is to provide high-resolution reference epigenome maps to the research community. For this, it will coordinate epigenomic mapping and characterization worldwide to avoid redundant research efforts; to implement high data quality standards; to coordinate data storage, management, and analysis; and to provide free access to the epigenomes produced. The maps that are generated under the umbrella of IHEC should contain detailed information on DNA methylation, histone modification, nucleosome occupancy, and corresponding coding and noncoding RNA expression in different normal and disease cell types. As such, it will allow integration of different layers of epigenetic information for a wide variety of distinct cell types and thus provide a powerful resource for both basic and applied research.

What Will Be the Impact of Epigenomes?

Last February, in a series of articles published in *Nature and Science* to celebrate the 10-year anniversary of the human genome project (HGP), its legacy was discussed. At the time of the HGP publication, it was regarded to be the first step toward understanding and treating many human diseases. Although the HGP and follow-up sequencing efforts have yielded significant advances in many areas of basic research such as on genome anatomy, genomic variation, and human history, it is still unclear how many additional steps are needed before we actually do understand most diseases. Mapping the full epigenome of a human will be a similar accomplishment, but again, one should be careful to overestimate the short-term impact of epigenomic knowledge on understanding and curing human disease. So, on that note, what can be expected from the various epigenomic profiling projects that are currently ongoing? First and foremost, the current epigenome-wide maps will provide comprehensive lists of chromatin features that can serve as a launch point for “upstream” investigations to identify the transcription factors, regulatory molecules, and pathways that initiate, modulate, or maintain epigenomic features. Similarly, the epigenomic maps will allow pursuit of “downstream” investigations to identify genes with similar patterns of epigenetic features that suggest coordinated regulation of gene expression in particular cell types.^{5,8,38} Second, current epigenomic projects will establish new standards for future initiatives to produce, analyze, and integrate large data sets related to health and diseases in human and in model organisms. These include methods to optimize comparability between different data sets, avoiding duplication of effort, exploiting the economic benefits of increased mapping efforts, and developing standard operating procedures for epigenomic studies.⁵ Third, because the epigenomic efforts are spread over different continents, it will be possible to

compare different human populations that are subjected to different environmental settings and different nutritional behaviors. As environment and nutrition are believed to be key players in establishing the epigenome,³⁹ these studies will allow for the first time a comprehensive examination of the actual impact of these factors on the epigenome. Also, epigenomic mapping will have an important impact on our current understanding of many biological processes such as cellular differentiation, reprogramming, and development, as well as basic gene regulatory mechanisms. For example, mapping of DNA methylation, histone modifications, chromatin accessibility, and various species of RNA in the same cell type will provide insight into the cross-talk between different epigenetic phenomena,⁴⁰ not only at single loci, but also across larger chromosomal territories. Hence, it can be expected that these projects will reveal new principles in the regulation of genome structure and function. In terms of understanding disease and therapy design, the epigenomic maps are expected to enhance the identification of upstream factors and pathways that might contribute to a given disease state as well as downstream genes affected by the disease state.^{3,30} Moreover, epigenomes have great promise as discovery engines for potential biomarkers for disease states or environmental factors (e.g., toxin exposure, viral or bacterial infections, drug usage, and psychosocial stress) and thus may be useful for diagnosis or to monitor disease progression and remission. Finally, knowing what is epigenetically altered in a disease cell opens the possibility to revert the aberrant epigenetic patterns using specific drugs that would erase or override the aberrant epigenetic modification.

Still, it should be noted that, most likely, these epigenomic maps will not directly disclose all the biological functions of the genomic elements and that, in most cases, functional assays will be needed to determine the precise function and activity associated with each epigenome. Similar to the HGP project,

mapping epigenomes will be a step forward toward understanding human diseases but will definitely not comprise the final step.

Outlook

It is clear that with the initiation of worldwide efforts to tackle epigenomes and the ongoing developments in sequencing, miniaturizing, and other technologies, the coming 10 years will yield information on numerous epigenomes, initially mostly focusing on human healthy and diseased tissues but gradually growing toward cells of various model organisms. These efforts will provide important new insights into the question as to the impact of cell differentiation and environmentally conditioned physiology on the epigenome. Furthermore, the availability of these data sets will provide loci where to test the causal relationship between defined epigenome features and physiology. Does the epigenome merely reflect a cell lineage's developmental and environmental history, or does it also condition its capacity to adapt to different environments? And if so, how? Hence, the epigenomic maps will certainly advance our current understanding of normal and aberrant cellular states and will comprise another important hurdle that needed to be surmounted in the quest to acquire a full understanding of human diseases.

Acknowledgments

The authors thank members of the BLUEPRINT consortium for valuable input.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

- Bell CG, Beck S. The epigenomic interface between genome and environment in common complex diseases. *Brief Funct Genomics*. 2010;9:477-85.
- Bernstein BE, Meissner A, Lander ES. The mammalian epigenome. *Cell*. 2007;128:669-81.
- Maunakea AK, Chepelev I, Zhao K. Epigenome mapping in normal and disease states. *Circ Res*. 2010;107:327-39.
- Zhang TY, Meaney MJ. Epigenetics and the environmental regulation of the genome and its function. *Annu Rev Psychol*. 2010;61:439-66, C1-3.
- Satterlee JS, Schubeler D, Ng HH. Tackling the epigenome: challenges and opportunities for collaboration. *Nat Biotechnol*. 2010;28:1039-44.
- Youngson NA, Whitelaw E. Transgenerational epigenetic effects. *Annu Rev Genomics Hum Genet*. 2008;9:233-57.
- Ng RK, Gurdon JB. Epigenetic inheritance of cell differentiation status. *Cell Cycle*. 2008;7:1173-7.
- Hawkins RD, Hon GC, Lee LK, *et al*. Distinct epigenomic landscapes of pluripotent and lineage-committed human cells. *Cell Stem Cell*. 2010;6:479-91.
- Bentley DR, Balasubramanian S, Swerdlow HP, *et al*. Accurate whole human genome sequencing using reversible terminator chemistry. *Nature*. 2008;456:53-9.
- Hirst M, Marra MA. Next generation sequencing based approaches to epigenomics. *Brief Funct Genomics*. 2010;9:455-65.
- Boyle AP, Davis S, Shulha HP, *et al*. High-resolution mapping and characterization of open chromatin across the genome. *Cell*. 2008;132:311-22.
- Barski A, Cuddapah S, Cui K, *et al*. High-resolution profiling of histone methylations in the human genome. *Cell*. 2007;129:823-37.
- Adli M, Zhu J, Bernstein BE. Genome-wide chromatin maps derived from limited numbers of hematopoietic progenitors. *Nat Methods*. 2010;7:615-8.
- Goren A, Ozsolak F, Shores N, *et al*. Chromatin profiling by directly sequencing small quantities of immunoprecipitated DNA. *Nat Methods*. 2010;7:47-9.
- Wold B, Myers RM. Sequence census methods for functional genomics. *Nat Methods*. 2008;5:19-21.
- Robertson G, Hirst M, Bainbridge M, *et al*. Genome-wide profiles of STAT1 DNA association using chromatin immunoprecipitation and massively parallel sequencing. *Nat Methods*. 2007;4:651-7.
- Maher CA, Kumar-Sinha C, Cao X, *et al*. Transcriptome sequencing to detect gene fusions in cancer. *Nature*. 2009;458:97-101.
- Richard H, Schulz MH, Sultan M, *et al*. Prediction of alternative isoforms from exon expression levels in RNA-Seq experiments. *Nucleic Acids Res*. 2010;38:e112.
- Down TA, Rakyan VK, Turner DJ, *et al*. A Bayesian deconvolution strategy for immunoprecipitation-based DNA methylome analysis. *Nat Biotechnol*. 2008;26:779-85.
- Bock C, Tomazou EM, Brinkman AB, *et al*. Quantitative comparison of genome-wide DNA methylation mapping technologies. *Nat Biotechnol*. 2010;28:1106-14.
- Lister R, Pelizzola M, Dowen RH, *et al*. Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature*. 2009;462:315-22.
- Kriaucionis S, Heintz N. The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. *Science*. 2009;324:929-30.
- Tahiliani M, Koh KP, Shen Y, *et al*. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine

- in mammalian DNA by MLL partner TET1. *Science*. 2009;324:930-5.
24. Giresi PG, Kim J, McDaniel RM, Iyer VR, Lieb JD. FAIRE (formaldehyde-assisted isolation of regulatory elements) isolates active regulatory elements from human chromatin. *Genome Res*. 2007;17:877-85.
 25. Gargiulo G, Levy S, Bucci G, *et al.* NA-Seq: a discovery tool for the analysis of chromatin structure and dynamics during differentiation. *Dev Cell*. 2009;16:466-81.
 26. Song L, Crawford GE. DNase-seq: a high-resolution technique for mapping active gene regulatory elements across the genome from mammalian cells. *Cold Spring Harb Protoc*. 2010;2010:prot5384.
 27. Milosavljevic A. Putting epigenome comparison into practice. *Nat Biotechnol*. 2010;28:1053-6.
 28. The ENCODE (ENCyclopedia Of DNA Elements) Project. *Science*. 2004;306:636-40.
 29. Muers M. Functional genomics: the modENCODE guide to the genome. *Nat Rev Genet*. 2010;12:80.
 30. Ting AH, McGarvey KM, Baylin SB. The cancer epigenome: components and functional correlates. *Genes Dev*. 2006;20:3215-31.
 31. Langemeijer SM, Kuiper RP, Berends M, *et al.* Acquired mutations in TET2 are common in myelodysplastic syndromes. *Nat Genet*. 2009;41:838-42.
 32. Nikoloski G, Langemeijer SM, Kuiper RP, *et al.* Somatic mutations of the histone methyltransferase gene EZH2 in myelodysplastic syndromes. *Nat Genet*. 2010;42:665-7.
 33. Chi P, Allis CD, Wang GG. Covalent histone modifications: miswritten, misinterpreted and mis-erased in human cancers. *Nat Rev Cancer*. 2010;10:457-69.
 34. Ley TJ, Ding L, Walter MJ, *et al.* DNMT3A mutations in acute myeloid leukemia. *N Engl J Med*. 2010;363:2424-33.
 35. Martens JH, Brinkman AB, Simmer F, *et al.* PML-RARalpha/RXR alters the epigenetic landscape in acute promyelocytic leukemia. *Cancer Cell*. 2010;17:173-85.
 36. Hudson TJ, Anderson W, Artez A, *et al.* International network of cancer genome projects. *Nature*. 2010;464:993-8.
 37. Bernstein BE, Stamatoyannopoulos JA, Costello JF, *et al.* The NIH Roadmap Epigenomics Mapping Consortium. *Nat Biotechnol*. 2010;28:1045-8.
 38. Ernst J, Kellis M. Discovery and characterization of chromatin states for systematic annotation of the human genome. *Nat Biotechnol*. 2010;28:817-25.
 39. Suter MA, Aagaard-Tillery KM. Environmental influences on epigenetic profiles. *Semin Reprod Med*. 2009;27:380-90.
 40. Lee JS, Smith E, Shilatifard A. The language of histone crosstalk. *Cell*. 2010;142:682-5.