

Strategies for dissecting epigenetic mechanisms in the mouse

Jesse Mager & Marisa S Bartolomei

Epigenetics generally refers to heritable changes in gene expression that are independent of nucleotide sequence. With complete genome sequences in hand, understanding the epigenetic control of genomes is the next step towards comprehending how the same DNA sequence gives rise to different cells, lineages and organs. Epigenetics also contributes to individual variation in normal biology and in disease states. The mouse provides a unique opportunity to understand how epigenetic differences contribute to both development and disease in a tractable mammalian system. Here we discuss current approaches and protocols used to study epigenetics in the mouse, including loss-of-function studies, mutagenesis screens, somatic cell nuclear transfer, genomics and proteomics.

With the completion of genome sequences, understanding the epigenetic control of genomes is the next step toward basic biological knowledge of both normal and disease states. Originally used by Conrad Waddington to describe interactions of genes and their environment¹, the term epigenetics is now more commonly used to refer to heritable changes in gene expression that are independent of nucleotide sequence. Epigenetics can be loosely defined as a situation in which the same DNA sequence is treated differently, and this variable treatment can occur at the level of the individual, organ, cell lineage or differentiation state. Even the two parental alleles of a gene can behave differently in the same cell nucleus, as happens with imprinted genes. Epigenetic regulation is paramount to achieve the biological complexity that occurs in multicellular organisms; it is also becoming apparent that the complexity of epigenetic regulation increases with genome size.

Here we describe current approaches to studying epigenetic regulation using the mouse as a model system. We limit our focus to what is possible *in vivo* (whole mouse), and we address the benefits and caveats of various techniques. Because our laboratory studies the mechanisms that regulate genome imprinting and X chromosome inactivation (XCI), we use these epigenetic phenomena as examples throughout. Before discussing the methods used, it is worthwhile summarizing

briefly the known epigenetic states and the genes and proteins that confer them. *In vivo*, DNA is wrapped around histone octamers, and these nucleosomes can be packaged and ordered to differing degrees of condensation, resulting in various local conformations of chromatin. Therefore, changes in nucleosome composition, as well as biochemical modifications to core histones and DNA itself, are implicated in epigenetic differences.

Epigenetic modifications

Epigenetic modifications can be subdivided into five categories, most of which are interdependent.

(i) Biochemical modification of DNA. In higher eukaryotes, only the methylation of cytosine residues in CpG dinucleotides has been identified as an important modification.

(ii) Biochemical modifications of core histones. With the development of chromatin immunoprecipitation (ChIP), modifications of specific residues in histone tails have come to the forefront of epigenetic regulation. Furthermore, biochemical studies are beginning to define both a hierarchy of these modifications and mutually exclusive modifications on particular histone residues^{2,3}.

(iii) Variant histones. Although few histone variants (such as protamines or macro-H2A) have been shown to correlate with epigenetic states in the mouse, recent findings from other systems suggest that dynamic exchange of histone or histone variants in chromatin correlates with the transcriptional status of a given locus^{4,5}.

(iv) Looping and local conformation. In DNase hypersensitivity assays, local chromatin structure has been defined by its accessibility to transcriptional machinery. DNA looping is thought to be important for the interaction of regulatory elements in a given locus.

(v) Higher-order chromatin structure. Long-held concepts of the existence of large regions of euchromatic and heterochromatic chromatin along chromosomes are proving to be true, and technical advances are facilitating the epigenetic characterization of each state^{6,7}. There is also much interest in the sequences and mechanisms that define the boundaries between these regions and in how these boundaries are established and maintained^{8,9}.

Mediators of epigenetic modifications

Given the breadth of epigenetic alterations that have been identified, it is not surprising that numerous proteins and RNAs have been implicated in epigenetic control. Six functional classes, based largely on the five types of modification listed above, can be used to categorize most of the genes encoding these epigenetic regulators.

Howard Hughes Medical Institute and Department of Cell and Developmental Biology, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania 19104, USA. Correspondence should be addressed to M.S.B. (bartolom@mail.med.upenn.edu).

Published online 27 October 2005; doi:10.1038/ng1664

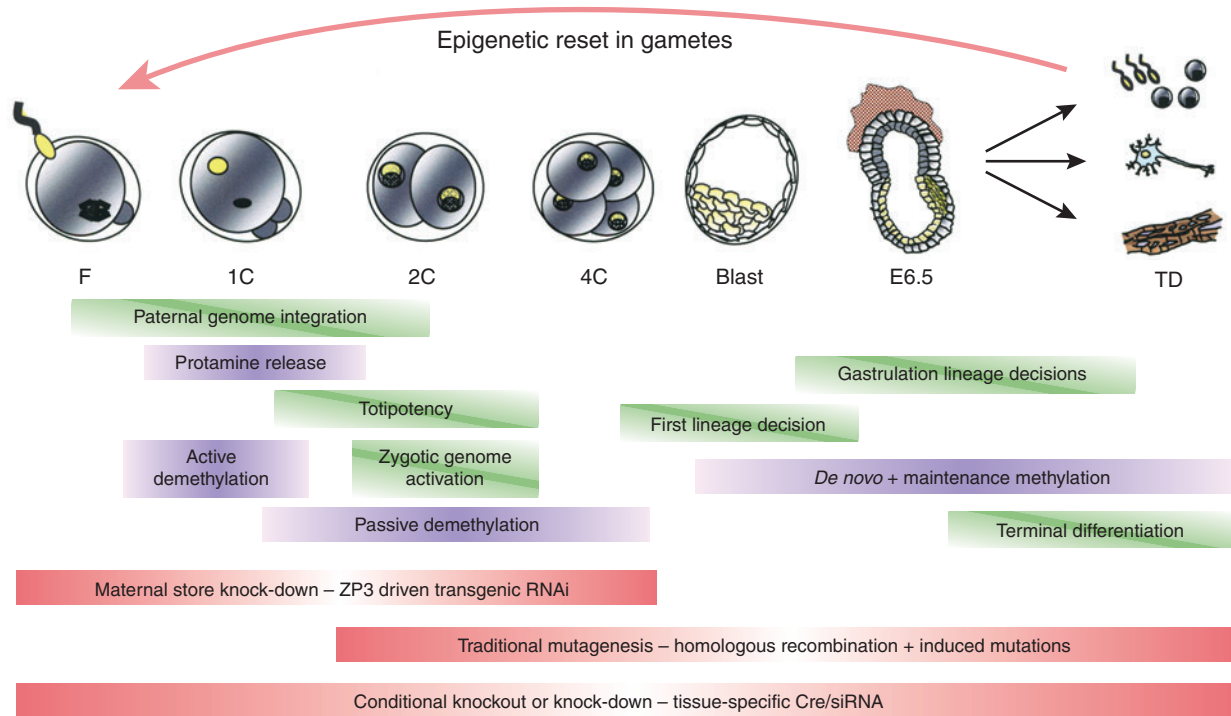


Figure 1 Developmentally defined epigenetic events. Shown is mouse development from fertilization (F) to terminal differentiation (TD). Purple bars indicate epigenetic events and modifications; green bars indicate important developmentally regulated epigenetic landmarks; and red bars indicate experimental approaches to examine epigenetic events *in vivo* during different developmental time windows. We expect that advances in the coming years will enable specific histone modification events to be added to this scheme. The timing of some of these events, such as protamine release and histone replacement, has not been defined precisely. C, cell; E, embryonic day; siRNA, short interfering RNA.

(i) DNA methyltransferases (DNMTs) and methyl CpG-binding domain (MBD) proteins. There are at least four DNMTs and six MBD proteins in the mouse. Some (or specific isoforms) of these proteins have dynamic temporal and spatial expression patterns early in development, findings that have helped to elucidate their roles in establishing or maintaining patterns of DNA methylation^{10–15}.

(ii) Histone-modifying proteins. Too numerous to list here, these proteins include histone methyltransferases, acetyltransferases and deacetylases. Some of the proteins that confer these histone modifications remain unknown in the mouse.

(iii) Histone exchange chaperones. Also not well characterized in the mouse, these proteins facilitate the exchange of core and variant histones. In addition, there are indications that some of these proteins interact with the transcriptional machinery itself, adding to the complexity of how specific loci are recognized for histone exchange¹⁶.

(iv) Insulator proteins. Enhancer-blocking proteins, such as the CCCTC-binding protein (CTCF), are known to bind to specific DNA elements and to prevent promoter and enhancer interactions *in cis*¹⁷. Other insulator proteins, such as USF, function as barriers that block the spread of silencing and activating modifications at the sites where they are bound¹⁸. Proteins in this class are also thought to facilitate (or to block) looping of local chromatin and interactions of regulatory elements.

(v) Chromatin-modifying complexes. ATP-dependent chromatin-remodeling complexes have been studied in several organisms including the mouse¹⁹. These large complexes (such as NuRD, Brahma, RSC and HP1 complexes) generally contain proteins from several of the functional categories listed here, and their composition is regulated

both temporally and tissue-specifically. In addition, the classically defined Polycomb group (PcG) protein complexes have recently received much attention owing to their involvement in the epigenetic regulation of some loci and the identification of histone-modifying activity in these complexes^{20–23}.

(vi) Genes with unknown activities. There are a growing number of genes that have a role in epigenetic regulation, but whose precise activity remains unclear. Genes in this class are largely noncoding RNAs, such as *Xist* or *Kcnq1ot1*, and recently identified microRNAs²⁴. Although much is known about the behavior of some of these RNA molecules, it has proved difficult to define their molecular functions *in vivo*.

Finally, it is important to note that various genes, such as those encoding tissue-specific transcription factors, contribute to epigenetic status but have functions that are not necessarily required to establish or to maintain this status *per se*. These genes are part of our evolving knowledge, but their discussion is beyond the scope of this perspective.

Endogenous retrotransposons

Some of the first mammalian epigenetic phenomena to be studied involve loci where integration of the intracisternal A-particle (IAP) retrotransposon alters the expression of nearby genes. For example, epigenetically inherited kink-tail phenotypes have been studied since the 1930s (ref. 25), even though the Axin-fused allele (*Axin1^{Fu}*) responsible has been only recently characterized²⁶. The DNA methylation status of IAPs at the *Axin1* and *agouti* (*A*) loci correlate with expression patterns and phenotypes (kink-tails and coat color, respectively), and these methylation patterns are heritable through the male and female germ

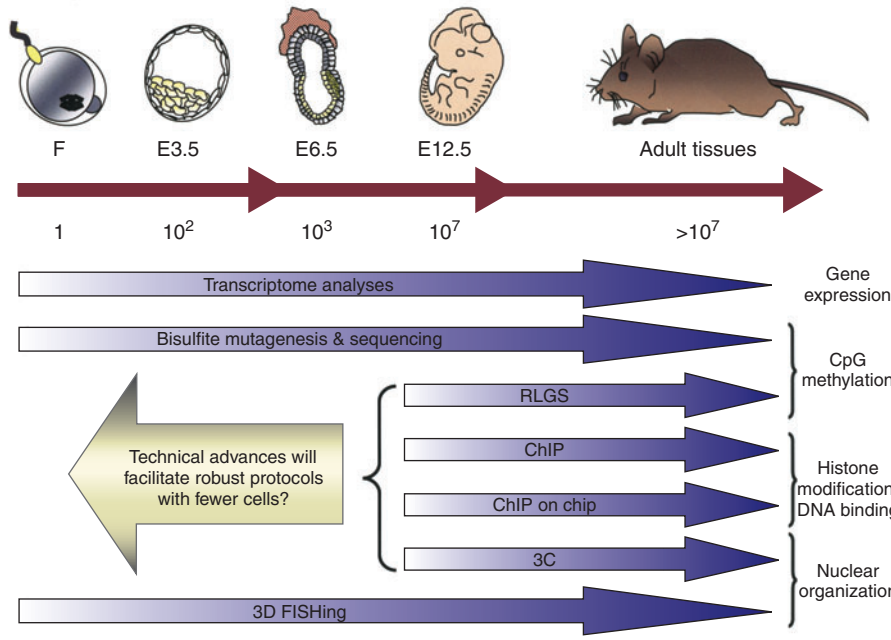


Figure 2 Current assays for epigenetic analysis. Approximate cell numbers are shown below the red arrow and scheme of mouse development. Blue arrows represent different types of analysis and their relative positions indicate the approximate cell number input that current protocols require. 3D FISHing, three-dimensional nuclear imaging combined with fluorescence *in situ* hybridization; E, embryonic day.

line^{27,28}. Notably, in addition to imprinted genes, these loci provide endogenous elements in the mouse with which to study the molecular mechanisms responsible for heritable epigenetic states that escape reprogramming in the early embryo.

Loss-of-function studies

Loss-of-function alleles generated through gene targeting, RNA-mediated interference (RNAi) and classical mutagenesis have supplied the most information about *trans*-acting factors that confer epigenetic states. It is only through such mouse studies that the roles of DNMTs, MBDs, CTCF and the PcG proteins EED and EZH2 have been elucidated. Prime examples of the importance of loss-of-function experiments lie in the many *Dnmt1* deletion alleles that have helped to establish their roles in epigenetic regulation. *Dnmt1*^{-/-} mice die at embryonic day 9.5 and show abnormal expression of imprinted genes, expression of normally silent IAP elements and defects in XCI²⁹⁻³¹. Methylation is properly established in mutant embryos but is gradually lost, indicating that DNMT1 has a role in maintenance—an observation that is supported by *in vitro* findings. In contrast to *Dnmt1*, inactivation of *Dnmt3a* and *Dnmt3b* blocks *de novo* methylation in embryonic stem cells and early embryos³². It has been more recently shown that DNMT3A and DNMT3L methylate transposable elements and imprinted genes in germ cells^{33,34}.

Whereas loss-of-function analysis has helped to define distinct roles for DNMTs, genetic analysis of the MBD proteins has proved more challenging¹³. Mice that have null mutations in *Mecp2* or *Mbd2* are viable and fertile and show surprisingly mild phenotypes and no observable defects in epigenetic phenomena such as XCI and imprinting^{35,36}. More detailed studies and the use of evolving technology have facilitated the functional analysis of these proteins. For example, MBD2 has a role in regulating T-cell lineage determination³⁷ and MECP2 is implicated in chromatin looping³⁸.

Traditional gene targeting is the standard for generating null alleles, but other technology is proving effective in isolating hypomorphic or neomorphic alleles that are useful for assigning gene function. For example, an *N*-ethyl-*N*-nitrosourea (ENU)-based mutagenesis screen for embryonic lethal phenotypes produced loss-of-function alleles that showed that EED has a role in maintaining genomic imprinting and XCI^{21,22}. More recently, transgenic RNAi has facilitated the production of knock-down mice in which the targeted RNA is depleted to various levels, depending on the transgenic line, in an ‘epiallelic series’^{39,40}. Transgenic RNAi in which a double-stranded RNA targeting CTCF mRNA has been used specifically during oocyte maturation to show that CTCF has a role in protecting the imprinted gene *H19* from acquiring methylation⁴¹. In addition, targeted mutations resulting in subtle base substitutions have been informative at the widely studied *H19-Igf2* region and other loci^{42,43}. Although modern advances are leading to high-throughput epigenome analyses, *in vivo* genetic engineering will remain the standard for assessing the function of a specific gene (Fig. 1).

Mutagenesis screens

For several decades, mutagenesis screens have been successfully implemented in many model systems. Both large-scale X ray-induced deletions and chemically induced point mutations (particularly those induced by ENU) have yielded much information in the mouse^{44,45}. Given the power of genetic screens to identify genes not previously implicated in a particular pathway or developmental process, why has this approach not been used to identify factors that function in epigenetic regulation? In part, the answer may lie in the lack of a robust, easily assayed epigenetic reporter or phenotype. Whole-mouse screens are expensive, laborious and lengthy, and the lack of a robust phenotype to screen makes the process risky. Nevertheless, investigators have invested in screens designed to identify epigenetic regulators, and the effort seems to be working.

A recently reported ENU-based mutagenesis screen for dominant phenotypes showing variegated expression of a green fluorescent protein (GFP) transgene resulted in the isolation of six lines with heritable alterations in GFP expression⁴⁶. Other epigenetic abnormalities were observed in these mutant lines, including defects in XCI, indicating that this screen for epigenetic regulatory proteins was successful. An ENU-based mutagenesis screen for new alleles of the ATP-dependent chromatin remodeler, BRG1, is also proving fruitful (S. Bultman and T. Magnuson, unpublished data).

On the basis of genome-wide RNAi screens in worms and flies⁴⁷, we and others have undertaken RNAi screens to identify genes involved in imprinting and other developmentally regulated epigenetic processes. The generation of an imprinted reporter (*e.g.*, a paternally silenced GFP transgene) would facilitate a relatively easy, whole-mouse mutagenesis screen for loss of imprinting mutations. So far, however, the derivation of faithfully imprinted transgenes that incorporate a sequence that is not normally imprinted, such as GFP, has been problematic. As large consortiums are formed to undertake genome-wide projects such as the generation of a loss-of-function mutation for every gene⁴⁸, it will be desirable to establish relatively high-throughput assays to assess

epigenetic regulation. This task, however, will be difficult: given the wide array of proteins and modifications contributing to heritable gene expression, it is unlikely that any one reporter will assay more than a few epigenetic regulators.

Biochemical and chromatin studies

In recent years, ChIP has become the accepted approach to study the epigenetic state of histones in a given locus. The technique requires one to guess which histone modification to examine. Under the right circumstances, ChIP facilitates a precise and complex analysis of particular loci, or even of whole genomes when combined with microarray technologies. The production of specific antibodies directed against particular histone modifications (e.g., di- versus trimethylation of Lys9 in histone H3) has enabled the histone code to be elaborated and correlated with states of gene expression⁴⁹. For example, Umlauf *et al.*⁵⁰ have recently used ChIP to examine parental allele-specific histone modifications at the imprinted *Kcnq1* domain on chromosome 7 and have determined that the placenta, but not the embryo, maintains imprinted gene expression exclusively through repressive histone methylation marks.

ChIP is most accurate when there is an abundance of material and only one cell type in the sample. For these reasons, extensive genome-wide histone profiles can be reliably established for single-celled organisms such as yeast⁵¹. Applying this approach to the mouse, however, has numerous potential difficulties. Although current protocols enable ChIP to be done on modest numbers of cells ($>1 \times 10^6$), many epigenetic processes occur during early development when only a few cells comprise a single embryo (Fig. 1), currently ruling out ChIP as a reasonable approach (Fig. 2). Even when using cells from an adult mouse, one needs to consider that different types of cell in a given tissue or organ may or may not share the same epigenetic profile. Another problem is the quantitative comparison of various histone modifications, because subtle differences could be explained by differing antibody affinities. Nevertheless, ChIP is a powerful way to examine whether specific histone modifications are present at a particular locus. Given that there may be as many as 150 different histone modifications, ChIP will remain an invaluable approach for epigenetic studies.

Another technique that is proving informative is chromatin conformation capture (3C; ref. 52 and Figs. 2 and 3). This approach tests whether two DNA sequences of interest are located near each other *in vivo*. Similar to ChIP, 3C requires a hypothesis of interacting regions. In the mouse, the β -globin locus was the first to be examined extensively with 3C. Using this approach, Tolhuis *et al.*⁵³ have shown that chromatin conformation differs greatly between nonexpressing brain tissue and erythroid lineages in which β -globin is expressed: chromatin loops that bring together the β -globin locus control region and active genes form exclusively in tissues where the genes are expressed. Recently, 3C has been used successfully to show that allele-specific looping interactions occur at the imprinted *H19* locus⁵⁴. 3C is a useful assay with which to test

long-range chromatin interactions. In addition, technological advances such as high-resolution three-dimensional nuclear imaging combined with fluorescence *in situ* hybridization are beginning to facilitate analysis of the spatial organization of chromosomes and chromatin domains in the nucleus⁵⁵ (Fig. 2).

Somatic cell nuclear transfer

Somatic cell nuclear transfer (SCNT) has undoubtedly attracted the most headlines in the popular press and continues to hold promise for therapeutic cloning. The successful production of cloned mice (and other animals) resolves some fundamental issues, the most important being that it is possible for the nucleus of a differentiated cell to give rise to a whole organism. That is, the differentiation process and all accompanying epigenetic alterations can be reversed, at least for some types of cell and at a low frequency (~5% of transferred SCNT embryos give rise to live-born mice). SCNT has also been used to show that odorant receptor choice in olfactory sensory neurons is epigenetically determined^{56,57}. In these studies, the use of marked olfactory sensory neurons as nuclear donors in SCNT resulted in cloned mice that showed the wild-type range of odorant receptor gene expression, indicating that receptor gene choice is conferred by reversible changes.

Although a somatic cell nucleus can be epigenetically reprogrammed to produce a viable mouse, the analysis of SCNT embryos has created a murky picture. Cloned mice can appear normal, but they may have

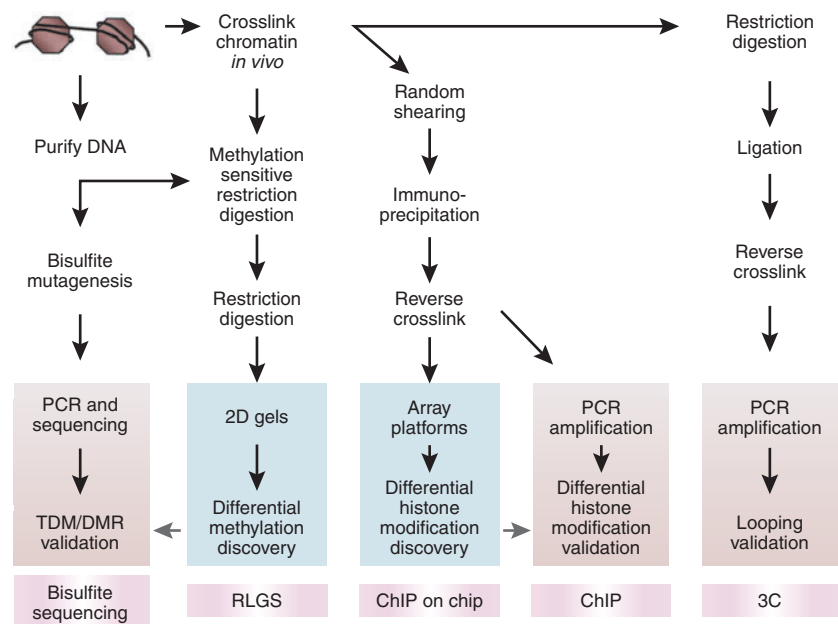


Figure 3 Summary of protocols used in epigenome analysis. Outlined are the protocols for five common approaches (specified in the purple boxes at the bottom) used to examine the epigenome. Red boxes indicate that a hypothesis is required before experimentation (e.g., a suspected looping interaction or differentially methylated locus). Blue boxes indicate the discovery of previously unsuspected epigenetic differences (biased only by microarray platform). In bisulfite mutagenesis, genomic DNA is treated with sodium bisulfite, which converts unmethylated cytosines residues to uracil residues but leaves methylated cytosines unaltered. Subsequent PCR amplification and sequencing provide the methylation status of every CpG in the amplified region. In 3C analysis, crosslinked chromatin is digested with a restriction enzyme, religated, reverse crosslinked and then amplified by PCR using one primer in each suspected interacting region to assay interaction between the two segments of DNA. In RLGS, precise two-dimensional (2D) agarose gels of radioactively labeled methyl-sensitive restriction fragments facilitate comparison of cleavage fragments among different tissues (TDMs). *In silico* prediction of restriction fragment sizes identifies the TDMs, which are confirmed by DNA sequencing. DMR, differentially methylated region.

both genetic and epigenetic abnormalities including inappropriate DNA methylation and gene expression profiles^{58,59}. It is important to keep in mind that these studies can assay only epigenetic abnormalities at known loci. Given the inconsistency in SCNT mice that show an adverse phenotype^{60,61}, it is impossible to know whether those mice that appear normal have epigenetic defects at other sites. SCNT is the ultimate test of nuclear reprogramming, and efforts to reverse epigenetic marks that occur through development and differentiation will continue to rely on this approach.

Genome-wide approaches

The complete sequencing of many organisms is facilitating the development of genome-wide epigenetic approaches, and large consortiums are being established to characterize the human epigenome⁶². Currently, there are no consortiums examining the mouse epigenome that are comparable to those formed for the human epigenome. There are, however, a few European-based initiatives that integrate mouse epigenomics, including the Austrian GEN-AU initiative and the recently formed Epigenome Network of Excellence (<http://www.epigenome-noe.net/>), which is designed to build more coherent European research areas in epigenetics. Assuming that it is possible to define an epigenome for a particular tissue or differentiation state, it will be invaluable to compare the epigenetic signatures of mouse and human (and other organisms) to look for conservation, as is now routinely done with genome sequences.

Current approaches to examine the epigenome fall into two categories: those that analyze DNA methylation and those that examine histone composition (Fig. 3). Although several human genome-wide methylation studies have been published^{63–65}, few have been reported in the mouse. Of note is a recent study that has identified tissue-specific differentially methylated regions (TDMs) using restriction landmark genome scanning (RLGS)⁶⁶. This approach relies on a restriction enzyme that cleaves only unmethylated recognition sites (Fig. 3). Using this approach, Song *et al.*⁶⁶ have identified 150 TDMs among liver, kidney, brain, colon, muscle and testis. In addition, despite the widely accepted view that CpG islands are unmethylated, these experiments suggest that 5% of CpG islands will be TDMs. The main drawback of RLGS is that it identifies differential methylation on the basis of the methylation status of only one or two CpGs in a restriction fragment; thus, many TDMs will be missed or misassigned. RLGS is, however, a reasonably high-throughput method for identifying differences in CpG methylation between two tissues or developmental time points and has potential for further optimization. Song *et al.*⁶⁶ also correlated gene expression with the TDMs identified. Such combined analyses are necessary to enhance our understanding of epigenetic regulation.

The most common approach to analyze CpG methylation is bisulfite sequencing (Figs. 2 and 3). Currently, bisulfite sequencing is the only routinely used assay that offers resolution of each CpG dinucleotide in a given region. In a pilot study for the Human Epigenome Project (<http://www.epigenome.org/>), Rakan *et al.*⁶⁵ reported successful optimization for high-throughput bisulfite sequencing analysis. They examined DNA methylation levels in regulatory, exonic and intronic regions associated with 90 genes in the human major histocompatibility complex in seven tissues. Although the amplicons were designed to assay only defined regions (as opposed to the complete locus), this large-scale effort sets the stage for more thorough analyses. Owing to the comprehensive nature of the data resulting from bisulfite sequencing, this approach will undoubtedly be used to examine the mouse epigenome.

To assay the genome for various histone modifications, DNA isolated from ChIPs can be hybridized to a microarray chip (Fig. 3). This 'ChIP on chip' protocol has been widely used for various experiments

and facilitates the discovery of histone modifications or transcription factor-binding sites throughout the genome. Many studies have used ChIP on chip to identify downstream targets of DNA-binding proteins of interest, and the same approach is now being used to examine enrichment of histone modifications^{67,68}. For example, using ChIP on chip, Bernstein *et al.*⁶⁹ generated histone modification maps of di- and trimethylation of Lys4 and acetylation of Lys9 and Lys14 in histone H3 across the nonrepetitive portions of human chromosomes 21 and 22, and compared methylation of Lys4 in several mouse and human orthologous regions. A notable observation is the conservation of histone methylation sites even when the underlying nucleotide sequence is not highly conserved between the two species. Another study using ChIP to define maps of repressive histone lysine methylation for repetitive elements across four mouse epigenomes has suggested that tandem repeats and double-stranded RNA function as triggers for more stable chromatin imprints⁷⁰. These types of analysis will increase our understanding of epigenome organization and function.

In addition to the technical aspects of ChIP, another consideration of ChIP on chip is the need to amplify the immunopurified DNA before hybridization, an experimental detail that has many different protocols, each with its own benefits and caveats. Until recently, ChIP on chip studies have been biased as they are limited by the microarray design (generally limited to cDNAs, promoters and 5' sequences of known genes only). Advances in microarray technology are now allowing much higher density oligonucleotide arrays to be made that contain perhaps as many as 6,000,000 features per chip⁷¹. These DNA 'tiling' arrays will facilitate an unbiased assessment of the whole genome.

In addition to ensuring complete coverage, the short overlapping oligonucleotides (25–60 bases) used in tiling arrays permit relatively fine mapping of observed hybridization differences, although, as in any ChIP on chip experiment, resolution is limited largely by the size of crosslinked fragments (generally 500–1,000 bp). Not only do tiling arrays hold promise for advancing ChIP on chip techniques, they are already being applied to other genome-wide epigenetic analyses⁷¹. For example, they have been used to examine the transcriptional profile of all nonrepetitive DNA in the human genome⁷². Bertone *et al.*⁷² report that many intergenic sequences are transcribed. These findings are similar to other recent results in human, rice and *Arabidopsis thaliana*^{73–75}, indicating that there is as much as ten times more transcribed sequence than was previously identified by expressed sequence tag sequencing and gene prediction algorithms, suggesting that our current understanding of the transcriptome is far from complete.

Defining normal: the epigeneticist's toolbox

Regardless of the approach used, there is an inherent problem in any attempt to characterize the epigenome. What is the ground state? Is there a tissue type or developmental time point that should be used to determine what is different about the epigenome in the tissue or individual being studied? In theory, such a tissue or time point would represent the 'totipotent' state. The appropriate stage corresponding to this state may be the two-cell embryo immediately after the early wave of active demethylation (Fig. 1). By all measures, two-cell blastomeres are still totipotent and theoretically have not yet acquired epigenetic differences. Thus, these cells could act as the epigenetic template on which to overlay findings from tissue- or lineage-specific studies.

Comparison of a differentiated state to a totipotent state would facilitate an understanding of what epigenetic changes have occurred along a given developmental pathway. In addition, comparing two differentiated states to each other, as well as to the totipotent two-cell epigenome, would identify differences between the two that have been acquired specifically in one lineage but not the other. To understand

what epigenetic changes contribute to disease states, we must first identify the battery of modifications that normally defines each tissue type. A chief obstacle, however, is that most of the approaches discussed above require large amounts of starting material. Although it is not currently possible, applying these genome-wide techniques to early embryos will be facilitated by technological advances that require smaller amounts of input material (Fig. 2).

Assuming that epigenetic technologies advance at a rate similar to that of DNA sequencing in the past 10 years, it is easy to imagine the next era of epigenomic analysis. With refined protocols and more efficient techniques, we may be able to take one sample and obtain a transcription profile, a genomic methylation profile and a genome-wide map of histone modifications. When this is possible, the combined information from these various approaches will provide a comprehensive view of the epigenome from any sample. As increasing amounts of epigenetic data accrue, there will be a need to overlay these different types of data. Ideally, it will be possible to view and to compare easily the various epigenetic modifications that occur at a given locus in different tissues and at distinct developmental time points. Here again, it is likely that tools for the mouse epigenome will follow models that have been initiated by the Encyclopedia of DNA Elements (<http://www.genome.gov/encode/>), Human Epigenome Project and Epigenome Network of Excellence and current genome browsers (e.g., Ensembl and National Center for Biotechnology Information).

What will be the next wave of functional epigenetic analysis? Because many of the genes known to be involved in these processes produce embryonic lethal phenotypes when deleted, a more subtle approach may be required. The development of conditional transgenic RNAi has much to offer in this regard. Work from our laboratory has shown that RNAi transgenic lines have target RNA levels that vary widely from 1% to 100% of wild-type levels, probably due to transgene copy number and position effects at the integration sites (ref. 41 and K.J. Reese & M.S.B., unpublished data). It would therefore be possible to generate 'epigenetically sensitive' lines of mice that are knocked-down for several epigenetic regulatory proteins (e.g., 25% DNMT, 25% histone methyltransferase and 25% MBD activity). Various combinations of knock-down alleles may uncover phenotypes that are due to interactions among these proteins or the modifications that they confer. In addition, comparing the epigenomic profiles from various knock-down combinations may elucidate hierarchies and redundancies among epigenetic modifications. Such sensitized lines may also provide a suitable background on which to examine other candidate epigenetic regulatory genes by circumventing functional redundancies.

Much of the current excitement surrounding epigenomics is due to the strong epigenetic phenotypes of many human disease states. If our goal is to model human conditions in the mouse, then we need a two-pronged comprehensive approach: a thorough knowledge of normally occurring epigenetic modifications (genome-wide observation), coupled with functional epigenetic annotation of the genome (whole-animal loss-of-function or RNAi studies), which will enable us to identify and to understand epigenetic abnormalities that contribute to disease states.

ACKNOWLEDGMENTS

We thank S. Bultman and T. Magnuson for communication of unpublished results; D. Barlow and T. Jenuwein for discussion of European-based initiatives in epigenetics; and S. Reiner and members of the laboratory for critical reading of the manuscript. We apologize to those authors whose primary work could not be referenced owing to space limitations. This work was supported by the US National Institutes of Health and the Howard Hughes Medical Institute. J.M. is supported by the Damon Runyon Cancer Research Foundation.

COMPETING INTERESTS STATEMENT

The authors declare that they have no competing financial interests.

Published online at <http://www.nature.com/naturegenetics/>
Reprints and permissions information is available online at <http://npg.nature.com/reprintsandpermissions/>

- Waddington, C. The epigenotype. *Endeavor* **1**, 18–20 (1942).
- Cowell, I.G. *et al.* Heterochromatin, HP1 and methylation at lysine 9 of histone H3 in animals. *Chromosoma* **111**, 22–36 (2002).
- Sun, Z.W. & Allis, C.D. Ubiquitination of histone H2B regulates H3 methylation and gene silencing in yeast. *Nature* **418**, 104–108 (2002).
- Costanzi, C. & Pehrson, J.R. Histone macroH2A1 is concentrated in the inactive X chromosome of female mammals. *Nature* **393**, 599–601 (1998).
- Sarma, K. & Reinberg, D. Histone variants meet their match. *Nat. Rev. Mol. Cell Biol.* **6**, 139–149 (2005).
- Luger, K. Structure and dynamic behavior of nucleosomes. *Curr. Opin. Genet. Dev.* **13**, 127–135 (2003).
- Zheng, C. & Hayes, J.J. Structures and interactions of the core histone tail domains. *Biopolymers* **68**, 539–546 (2003).
- Labrador, M. & Corces, V.G. Setting the boundaries of chromatin domains and nuclear organization. *Cell* **111**, 151–154 (2002).
- Goetze, S. *et al.* Performance of genomic bordering elements at predefined genomic loci. *Mol. Cell. Biol.* **25**, 2260–2272 (2005).
- Rousseaux, S. *et al.* Establishment of male-specific epigenetic information. *Gene* **345**, 139–153 (2005).
- Rangwala, S.H. & Richards, E.J. The value-added genome: building and maintaining genomic cytosine methylation landscapes. *Curr. Opin. Genet. Dev.* **14**, 686–691 (2004).
- Chen, T. & Li, E. Structure and function of eukaryotic DNA methyltransferases. *Curr. Top. Dev. Biol.* **60**, 55–89 (2004).
- Hendrich, B. & Tweedie, S. The methyl-CpG binding domain and the evolving role of DNA methylation in animals. *Trends Genet.* **19**, 269–277 (2003).
- Jorgensen, H.F. & Bird, A. MeCP2 and other methyl-CpG binding proteins. *Ment. Retard. Dev. Disabil. Res. Rev.* **8**, 87–93 (2002).
- Nakao, M. *et al.* Regulation of transcription and chromatin by methyl-CpG binding protein MBD1. *Brain Dev.* **23** Suppl 1, S174–S176 (2001).
- Henikoff, S., Furuyama, T. & Ahmad, K. Histone variants, nucleosome assembly and epigenetic inheritance. *Trends Genet.* **20**, 320–326 (2004).
- Bell, A.C., West, A.G. & Felsenfeld, G. The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. *Cell* **98**, 387–396 (1999).
- West, A.G., Huang, S., Gaszner, M., Litt, M.D. & Felsenfeld, G. Recruitment of histone modifications by USF proteins at a vertebrate barrier element. *Mol. Cell* **16**, 453–463 (2004).
- Gebuhr, T.C., Bultman, S.J. & Magnuson, T. Pc-G/trx-G and the SWI/SNF connection: developmental gene regulation through chromatin remodeling. *Genesis* **26**, 189–197 (2000).
- Plath, K. *et al.* Role of histone H3 lysine 27 methylation in X inactivation. *Science* **300**, 131–135 (2003).
- Mager, J., Montgomery, N.D., de Villena, F.P. & Magnuson, T. Genome imprinting regulated by the mouse Polycomb group protein Eed. *Nat. Genet.* **33**, 502–507 (2003).
- Wang, J. *et al.* Imprinted X inactivation maintained by a mouse *Polycomb* group gene. *Nat. Genet.* **28**, 371–375 (2001).
- Hernandez-Munoz, I. *et al.* Stable X chromosome inactivation involves the PRC1 Polycomb complex and requires histone MACROH2A1 and the CULLIN3/SPOP ubiquitin E3 ligase. *Proc. Natl. Acad. Sci. USA* **102**, 7635–7640 (2005).
- Huttenhofer, A., Schattner, P. & Polacek, N. Non-coding RNAs: hope or hype? *Trends Genet.* **21**, 289–297 (2005).
- Roberts, E. & Quisenberry, J.H. Linkage of the genes for non-yellow (y) and pink-eye (p2) in the house mouse (*Mus musculus*). *Am. Nat.* **69**, 181–183 (1935).
- Vasicek, T.J. *et al.* Two dominant mutations in the mouse fused gene are the result of transposon insertions. *Genetics* **147**, 777–786 (1997).
- Michaud, E.J. *et al.* Differential expression of a new dominant agouti allele (Aiapy) is correlated with methylation state and is influenced by parental lineage. *Genes Dev.* **8**, 1463–1472 (1994).
- Rakyan, V. & Whitelaw, E. Transgenerational epigenetic inheritance. *Curr. Biol.* **13**, R6 (2003).
- Li, E., Bestor, T.H. & Jaenisch, R. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* **69**, 915–926 (1992).
- Li, E., Beard, C. & Jaenisch, R. Role for DNA methylation in genomic imprinting. *Nature* **366**, 362–365 (1993).
- Sado, T. *et al.* X inactivation in the mouse embryo deficient for Dnmt1: distinct effect of hypomethylation on imprinted and random X inactivation. *Dev. Biol.* **225**, 294–303 (2000).
- Okano, M., Bell, D.W., Haber, D.A. & Li, E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for *de novo* methylation and mammalian development. *Cell* **99**, 247–257 (1999).
- Bourc'his, D., Xu, G.L., Lin, C.S., Bollman, B. & Bestor, T.H. Dnmt3L and the establishment of maternal genomic imprints. *Science* **294**, 2536–2539 (2001).
- Kaneda, M. *et al.* Essential role for *de novo* DNA methyltransferase Dnmt3a in paternal and maternal imprinting. *Nature* **429**, 900–903 (2004).
- Guy, J., Hendrich, B., Holmes, M., Martin, J.E. & Bird, A. A mouse Mecp2-null mutation causes neurological symptoms that mimic Rett syndrome. *Nat. Genet.* **27**, 322–326 (2001).
- Hendrich, B., Guy, J., Ramsahoye, B., Wilson, V.A. & Bird, A. Closely related proteins

- MBD2 and MBD3 play distinctive but interacting roles in mouse development. *Genes Dev.* **15**, 710–723 (2001).
37. Hutchins, A.S. *et al.* Gene silencing quantitatively controls the function of a developmental trans-activator. *Mol. Cell* **10**, 81–91 (2002).
 38. Horike, S., Cai, S., Miyano, M., Cheng, J.F. & Kohwi-Shigematsu, T. Loss of silent-chromatin looping and impaired imprinting of DLX5 in Rett syndrome. *Nat. Genet.* **37**, 31–40 (2005).
 39. Hemann, M.T. *et al.* An epi-allelic series of p53 hypomorphs created by stable RNAi produces distinct tumor phenotypes *in vivo*. *Nat. Genet.* **33**, 396–400 (2003).
 40. Kunath, T. *et al.* Transgenic RNA interference in ES cell-derived embryos recapitulates a genetic null phenotype. *Nat. Biotechnol.* **21**, 559–561 (2003).
 41. Fedoriw, A.M., Stein, P., Svoboda, P., Schultz, R.M. & Bartolomei, M.S. Transgenic RNAi reveals essential function for CTCF in H19 gene imprinting. *Science* **303**, 238–240 (2004).
 42. Engel, N., West, A.G., Felsenfeld, G. & Bartolomei, M.S. Antagonism between DNA hypermethylation and enhancer-blocking activity at the H19 DMD is uncovered by CpG mutations. *Nat. Genet.* **36**, 883–888 (2004).
 43. Schoenherr, C.J., Levorse, J.M. & Tilghman, S.M. CTCF maintains differential methylation at the Igf2/H19 locus. *Nat. Genet.* **33**, 66–69 (2003).
 44. Russell, L.B. & Bangham, J.W. The paternal genome in mouse zygotes is less sensitive to ENU mutagenesis than the maternal genome. *Mutat. Res.* **248**, 203–209 (1991).
 45. Rinchik, E.M., Tonjes, R.R., Paul, D. & Potter, M.D. Molecular analysis of radiation-induced albino (c)-locus mutations that cause death at preimplantation stages of development. *Genetics* **135**, 1107–1116 (1993).
 46. Blewitt, M.E. *et al.* An N-ethyl-N-nitrosourea screen for genes involved in variegation in the mouse. *Proc. Natl. Acad. Sci. USA* **102**, 7629–7634 (2005).
 47. Carpenter, A.E. & Sabatini, D.M. Systematic genome-wide screens of gene function. *Nat. Rev. Genet.* **5**, 11–22 (2004).
 48. Austin, C.P. *et al.* The knockout mouse project. *Nat. Genet.* **36**, 921–924 (2004).
 49. Wang, Y. *et al.* Beyond the double helix: writing and reading the histone code. *Novartis Found. Symp.* **259**, 3–17; discussion 17–21, 163–169 (2004).
 50. Umlauf, D. *et al.* Imprinting along the *Kcnq1* domain on mouse chromosome 7 involves repressive histone methylation and recruitment of Polycomb group complexes. *Nat. Genet.* **36**, 1296–1300 (2004).
 51. Roh, T.Y., Ngau, W.C., Cui, K., Landsman, D. & Zhao, K. High-resolution genome-wide mapping of histone modifications. *Nat. Biotechnol.* **22**, 1013–1016 (2004).
 52. Dekker, J., Rippe, K., Dekker, M. & Kleckner, N. Capturing chromosome conformation. *Science* **295**, 1306–1311 (2002).
 53. Tolhuis, B., Palstra, R.J., Splinter, E., Grosveld, F. & de Laat, W. Looping and interaction between hypersensitive sites in the active β -globin locus. *Mol. Cell* **10**, 1453–1465 (2002).
 54. Murrell, A., Heeson, S. & Reik, W. Interaction between differentially methylated regions partitions the imprinted genes *Igf2* and *H19* into parent-specific chromatin loops. *Nat. Genet.* **36**, 889–893 (2004).
 55. Bolzer, A. *et al.* Three-dimensional maps of all chromosomes in human male fibroblast nuclei and prometaphase rosettes. *PLoS Biol.* **3**, e157 (2005).
 56. Eggan, K. *et al.* Mice cloned from olfactory sensory neurons. *Nature* **428**, 44–49 (2004).
 57. Li, J., Ishii, T., Feinstein, P. & Mombaerts, P. Odorant receptor gene choice is reset by nuclear transfer from mouse olfactory sensory neurons. *Nature* **428**, 393–399 (2004).
 58. Mann, M.R. *et al.* Disruption of imprinted gene methylation and expression in cloned preimplantation stage mouse embryos. *Biol. Reprod.* **69**, 902–914 (2003).
 59. Boiani, M., Eckardt, S., Scholer, H.R. & McLaughlin, K.J. Oct4 distribution and level in mouse clones: consequences for pluripotency. *Genes Dev.* **16**, 1209–1219 (2002).
 60. Ogonuki, N. *et al.* Early death of mice cloned from somatic cells. *Nat. Genet.* **30**, 253–254 (2002).
 61. Eggan, K. *et al.* Hybrid vigor, fetal overgrowth, and viability of mice derived by nuclear cloning and tetraploid embryo complementation. *Proc. Natl. Acad. Sci. USA* **98**, 6209–6214 (2001).
 62. Murrell, A., Rakan, V.K. & Beck, S. From genome to epigenome. *Hum. Mol. Genet.* **14** Suppl 1, R3–R10 (2005).
 63. Costello, J.F. *et al.* Aberrant CpG-island methylation has non-random and tumor-type-specific patterns. *Nat. Genet.* **24**, 132–138 (2000).
 64. Grunau, C., Hindermann, W. & Rosenthal, A. Large-scale methylation analysis of human genomic DNA reveals tissue-specific differences between the methylation profiles of genes and pseudogenes. *Hum. Mol. Genet.* **9**, 2651–2663 (2000).
 65. Rakan, V.K. *et al.* DNA methylation profiling of the human major histocompatibility complex: a pilot study for the human epigenome project. *PLoS Biol.* **2**, e405 (2004).
 66. Song, F. *et al.* Association of tissue-specific differentially methylated regions (TDMs) with differential gene expression. *Proc. Natl. Acad. Sci. USA* **102**, 3336–3341 (2005).
 67. Buck, M.J. & Lieb, J.D. ChIP-chip: considerations for the design, analysis, and application of genome-wide chromatin immunoprecipitation experiments. *Genomics* **83**, 349–360 (2004).
 68. Taverner, N.V., Smith, J.C. & Wardle, F.C. Identifying transcriptional targets. *Genome Biol.* **5**, 210 (2004).
 69. Bernstein, B.E. *et al.* Genomic maps and comparative analysis of histone modifications in human and mouse. *Cell* **120**, 169–181 (2005).
 70. Martens, J.H. *et al.* The profile of repeat-associated histone lysine methylation states in the mouse epigenome. *EMBO J.* **24**, 800–812 (2005).
 71. Mockler, T.C. & Ecker, J.R. Applications of DNA tiling arrays for whole-genome analysis. *Genomics* **85**, 1–15 (2005).
 72. Bertone, P. *et al.* Global identification of human transcribed sequences with genome tiling arrays. *Science* **306**, 2242–2246 (2004).
 73. Cawley, S. *et al.* Unbiased mapping of transcription factor binding sites along human chromosomes 21 and 22 points to widespread regulation of noncoding RNAs. *Cell* **116**, 499–509 (2004).
 74. Yamada, K. *et al.* Empirical analysis of transcriptional activity in the *Arabidopsis* genome. *Science* **302**, 842–846 (2003).
 75. Jiao, Y. *et al.* A tiling microarray expression analysis of rice chromosome 4 suggests a chromosome-level regulation of transcription. *Plant Cell* **17**, 1641–1657 (2005).