Cell

Epigenetics: A Landscape Takes Shape

Aaron D. Goldberg,[1] C. David Allis,[1,*] and Emily Bernstein[1,*]
[1]Laboratory of Chromatin Biology, The Rockefeller University, New York, NY 10021, USA
*Correspondence: alliscd@rockefeller.edu (C.D.A.), bernste@rockefeller.edu (E.B.)
DOI 10.1016/j.cell.2007.02.006

Epigenetics has recently evolved from a collection of diverse phenomena to a defined and far-reaching field of study. In this Essay, we examine the epistemology of epigenetics, provide a brief overview of underlying molecular mechanisms, and suggest future challenges for the field.

Historically, the word "epigenetics" was used to describe events that could not be explained by genetic principles. Conrad Waddington (1905–1975), who is given credit for coining the term, defined epigenetics as "the branch of biology which studies the causal interactions between genes and their products, which bring the phenotype into being" (Waddington, 1942). Over the years, numerous biological phenomena, some considered bizarre and inexplicable, have been lumped into the category of epigenetics. These include seemingly unrelated processes, such as paramutation in maize (an interaction between two alleles in which one allele causes heritable changes in the other allele); position effect variegation in the fruit fly *Drosophila* (in which the local chromatin environment of a gene determines its expression); and imprinting of specific paternal or maternal loci in mammals. Although mysteries abound, the field is now beginning to uncover common molecular mechanisms underlying epigenetic phenomena. We have recently witnessed an explosion of research efforts, meetings and symposia, international initiatives, internet resources, commercial enterprises, and even a recent textbook dedicated to epigenetics, all of which lead us to this year's special review issue in *Cell*. What underlies this swell of interest in epigenetics? Whether it is the enigma of epigenetic processes or their fundamental importance in myriad biological contexts, one thing is clear—the field of epigenetics is gaining respect.

Epigenetics, in a broad sense, is a bridge between genotype and phenotype—a phenomenon that changes the final outcome of a locus or chromosome without changing the underlying DNA sequence. For example, even though the vast majority of cells in a multicellular organism share an identical genotype, organismal development generates a diversity of cell types with disparate, yet stable, profiles of gene expression and distinct cellular functions. Thus, cellular differentiation may be considered an epigenetic phenomenon, largely governed by changes in what Waddington described as the "epigenetic landscape" rather than alterations in genetic inheritance (Waddington, 1957; Figure 1). More specifically, epigenetics may be defined as the study of any potentially stable and, ideally, heritable change in gene expression or cellular phenotype that occurs without changes in Watson-Crick base-pairing of DNA.

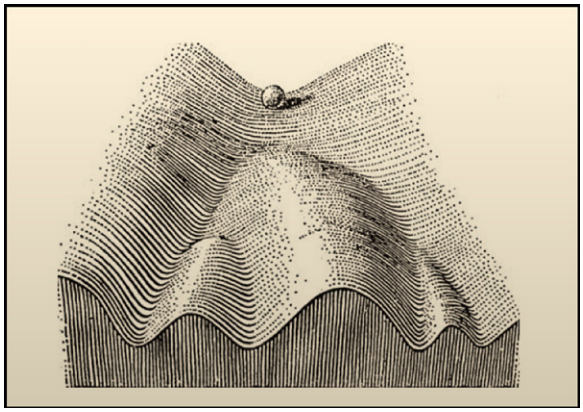

Figure 1. Waddington's Classical Epigenetic Landscape
In 1957, Conrad Waddington proposed the concept of an epigenetic landscape to represent the process of cellular decision-making during development. At various points in this dynamic visual metaphor, the cell (represented by a ball) can take specific permitted trajectories, leading to different outcomes or cell fates. Figure reprinted from Waddington, 1957.

Here, we aim to briefly introduce some of the core molecular actors that play upon the epigenetic stage and touch upon concepts of epigenetic heritability and stability. Despite the field's recent progress, significant and fundamental questions remain to be answered, many of which center on the propagation of epigenetic information through cellular division and differentiation. We highlight some of these questions as challenges to the emerging field. We also refer readers to the review articles appearing in this special issue, as well as a new textbook entitled *Epigenetics* (Allis et al., 2007; see Book Review by Y. Shi, page 639 of this issue).

Epigenetic Mechanisms at Work

Much of today's epigenetic research is converging on the study of covalent and noncovalent modifications of DNA and histone proteins and the mechanisms by which such modifications influence overall chromatin structure. Chromatin, the complex of DNA and its intimately associated proteins, provides an attractive candidate for shaping the features of a cell's epigenetic landscape (see Review by B.E. Bernstein et al., page 669 of this issue). Diverse

epigenetic phenomena have recently led researchers to conserved molecular mechanisms involving chromatin modification, a theme reinforced throughout this special issue. We favor the view that the macromolecular entities described below all significantly contribute to the physiologically relevant organization of most eukaryotic genomes. These entities, and possibly others yet unknown, should be considered collectively when exploring epigenetic mechanisms.

DNA Methylation

DNA methylation is perhaps the best characterized chemical modification of chromatin. In mammals, nearly all DNA methylation occurs on cytosine residues of CpG dinucleotides. Regions of the genome that have a high density of CpGs are referred to as CpG islands, and DNA methylation of these islands correlates with transcriptional repression (Goll and Bestor, 2005). Genomic patterns of cytosine methylation in mammals, whether donated by de novo or maintenance DNA methyltransferases (DNMTs), play a critical role in gene regulation and chromatin organization during embryogenesis and gametogenesis (Goll and Bestor, 2005; see Review by M.A. Surani et al., page 747 of this issue). Maintenance methyltransferases add methyl groups to hemi-methylated DNA during DNA replication, whereas de novo DNMTs act after DNA replication (see Review by A. Groth et al., page 721 of this issue). Moreover, the formation of heterochromatin in many organisms is mediated in part by DNA methylation and its binding proteins in combination with RNA and histone modifications characteristic of silent chromatin (see Review by M. Zaratiegui et al., page 763 of this issue). DNA methylation plays a role in many cellular processes including silencing of repetitive and centromeric sequences from fungi to mammals; X chromosome inactivation in female mammals; and mammalian imprinting, all of which can be stably maintained (see Review by P.K. Yang and K.I. Kuroda, page 777 of this issue). Taken together, DNA methylation provides a stable, heritable, and critical component of epigenetic regulation.

Chromatin Variation: Covalent and Noncovalent Mechanisms

The list of covalent modifications to histone proteins continues to grow (see Review by T. Kouzarides, page 693 of this issue). However, only recently has genetic and biochemical evidence converged to clearly connect covalent histone modifications with longstanding epigenetic phenomena. Genetic screens for *su*ppressors of position effect *var*iegation [*Su(var)*] in *Drosophila*, for example, revealed over 100 genes that encode vital constituents of heterochromatin. Many of these genes are conserved from flies to humans, including heterochromatin protein 1 (HP1) and the histone H3K9 methyltransferase *Su(var)3-9* (Schotta et al., 2003). *Drosophila* genetics also provides another link between epigenetics and histone modifications, in the form of two evolutionarily conserved families of proteins that regulate homeotic genes antagonistically during development: the Polycomb Group (PcG) and the Trithorax Group (TrxG). Further molecular characterization of PcG proteins reveals that they contribute to two distinct protein complexes that are responsible for "writing" (PRC2, *P*olycomb *R*epressive *C*omplex 2) and "reading" (PRC1) methylation of H3K27 and facilitating chromatin condensation. Meanwhile, TrxG proteins mediate methylation of H3K4 and promote transcriptionally active chromatin (Ringrose and Paro, 2004; Ruthenburg et al., 2007; Review by B. Schuettengruber et al., page 735 of this issue)

How might covalent histone modifications modify chromatin structure to influence patterns of gene expression (Jenuwein and Allis, 2001)? A longstanding literature suggests that charge-altering modifications such as acetylation and phosphorylation can directly alter the physical properties of the chromatin fiber, leading to changes in higher-order structures. More recently, effector-mediated functions have been well documented, whereby histone modifications recruit or stabilize the localization of specific binding partners to chromatin. In the first example of an emerging epigenetic paradigm, the catalytic enzyme *Su(var)3-9* serves as a "writer" of histone H3K9 methylation, whereas HP1 serves as a "reader" or "effector" recognizing H3K9 methylation and directing biological processes (such as heterochromatin stabilization) to particular areas of the chromatin fiber (Lachner et al., 2001). In addition, effector-repulsive functions have been demonstrated in which histone modifications abrogate the binding of particular effectors (Fischle et al., 2005; T.H. Bestor, personal communication).

Noncovalent mechanisms such as chromatin remodeling and the incorporation of specialized histone variants provide the cell with additional tools for introducing variation into the chromatin template. ATP-dependent chromatin remodeling complexes are thought to modify chromatin accessibility by altering histone-DNA interactions, perhaps by sliding or ejecting nucleosomes (Smith and Peterson, 2005). In addition, histone variants such as H3.3 and H2A.Z, often carrying their own modification patterns, are exchanged within chromosomal domains by dedicated chaperone and exchange machinery (Polo and Almouzni, 2006). Mechanistic links between covalent and noncovalent mechanisms have also been uncovered, as effectors can include subunits of nucleosome remodeling complexes (Wysocka et al., 2006). Covalent modification, nucleosome remodeling, and histone variants work together to introduce meaningful variation into the chromatin fiber, and their collective contribution to epigenetics is only now being rigorously explored.

Noncoding RNA

Recently, it has become evident that RNA, particularly noncoding RNAs, have a hand in controlling multiple epigenetic phenomena (Bernstein and Allis, 2005). Clear examples of RNA involvement range from dosage compensation mechanisms in *Drosophila* and mammals mediated by the *rox* and *XIST* RNAs, respectively, to the silencing of both genes and repetitive DNA sequences by posttranscriptional

(PTGS) and transcriptional (TGS) RNA interference (RNAi)-related pathways, respectively, in almost all eukaryotes. These RNAs often act in concert with various components of the cell's chromatin and DNA methylation machinery to achieve stable silencing. Although we might not consider PTGS-inducing RNAs (e.g., microRNAs, siRNAs, etc.) to be epigenetic in nature, TGS-evoking RNAs (e.g., repeat-associated siRNAs, *Xist* RNA, and small RNAs in *S. pombe*) are more clearly epigenetic, as they can induce long-term silencing effects that can be inherited through cell division (Bernstein and Allis, 2005).

Epigenetic Pathways Intersect

It is becoming clear that significant crosstalk exists between different epigenetic pathways. In *S. pombe*, for example, biochemical purification of the RNAi machinery revealed an interaction between small RNAs and chromatin regulators such as the H3K9me-reading chromodomain-containing protein Chp1 (Grewal and Jia, 2007). Silencing of the inactive X chromosome is another prime example, where *Xist* RNA, DNA methylation, histone modifications, and their writers and readers all play a role (Heard, 2005). Moreover, DNMTs act in part by interacting directly or indirectly with chromatin-modifying enzymes such as histone deacetylases (Dobosy and Selker, 2001). New studies reveal that DNMT3L, an essential accessory protein to the de novo DNMTs in the germline, interprets patterns of histone modifications, adding weight to an emerging theme that DNA methylation and histone modifications functionally interact (Freitag and Selker, 2005; T.H. Bestor, personal communication).

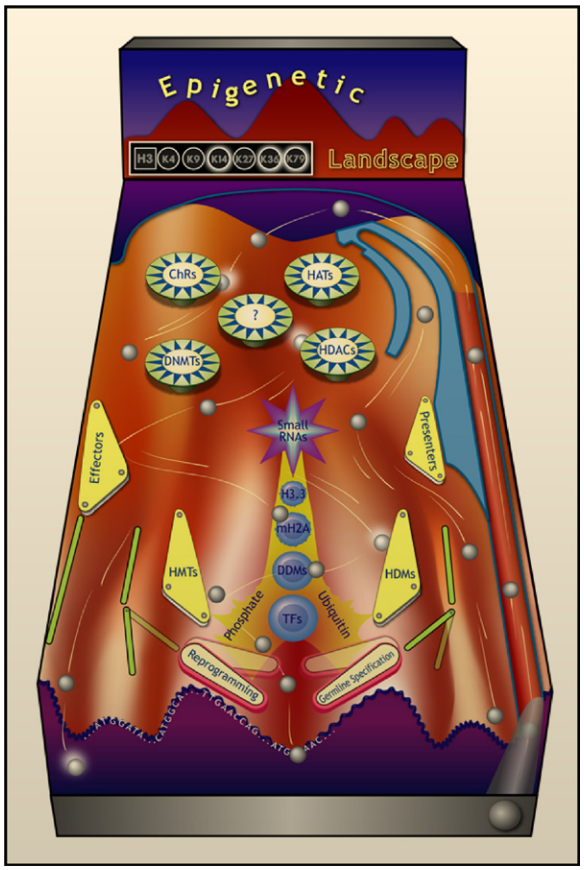

Figure 2. A Current View of the Epigenetic Molecular Machinery

Known factors that regulate epigenetic phenomena are shown directing the complex movement of pinballs (cells) across the elegant landscape proposed by Waddington (see Figure 1). No specific order of molecular events is implied; as such a sequence remains unknown. Effector proteins recognize specific histone modifications, while presenters are proposed to impart substrate specificity for histone-modifying enzymes (Ruthenburg et al., 2007). H3.3 and macroH2A are shown only as representative histone variants involved in transcriptional activation or repression, respectively. For simplicity, other histone (and nonhistone) proteins are not shown. (ChR, chromatin remodelers; DNMTs, DNA methyltransferases; HATs, histone acetyltransferases; HDACs, histone deacetylases; HMTs, histone methyltransferases; HDMs, histone demethylases; DDMs, DNA demethylases [unidentified in mammals to date]; and TFs, transcription factors [reflecting the genetic component of the epigenetic process].) Illustration by Sue Ann Fung-Ho.

Questions and Challenges

If modifications to DNA and chromatin along with the expression of specific noncoding RNAs help to define epigenetic identity, how are these factors inherited through cell division? What mechanisms enable epigenetic stability in a defined cellular lineage while allowing epigenetic flexibility during cellular differentiation and development? Mechanisms exist for the faithful propagation of DNA sequences and even DNA methylation patterns during replication, yet we understand comparatively little about non-DNA-based epigenetic inheritance.

Key questions regarding chromatin-based inheritance surprisingly remain unanswered. For example, are histone modifications inherited through replication, and if so, how does this occur (see Review by A. Groth et al., page 721 of this issue)? How do newly incorporated histones "learn" from parental chromatin? Is there an active mechanism for templating histone modifications during DNA replication? Is templating of new histone modifications simply a consequence of spreading to adjacent nucleosomes, as in position effect variegation? Or is templating of parental modifications onto new histones even necessary? Are half of the modifications enough? If histone modifications help to define cellular identity, mechanisms should exist for templating of new histone modifications, or such marks would be quickly diluted out in rapidly dividing cells.

Numerous related questions remain. Throughout development, a multipotent stem cell can become many different cell types, each genetically identical but unique in cellular phenotype (Figure 1). After a lineage transition, how is the epigenetic identity of specific differentiated cell types stably inherited? How does RNA fit into the issue of chromatin inheritance, if at all? With the recent discoveries of histone lysine demethylases, how do these "erasers" of covalent modifications function in epigenetic phenomena? Are all modifications actively erased? What factors are responsible for the reversal of DNA methylation patterns during gametogenesis and early development?

How do we reconcile proposed DNA demethylases and known histone demethylases with the notion of methylation as a stable epigenetic indexing system? How are demethylases regulated? The emerging dialectic of epigenetics, including the marks, writers, presenters, readers, and erasers, promises to be a rich conversation (see Review by M.A. Surani et al., page 747 of this issue).

Finally, we need a better molecular understanding of epigenetic phenomena in human development and disease. Genomic reprogramming during germ cell specification and early embryonic development, as well as during somatic cell nuclear transfer and therapeutic cloning, is by definition an epigenetic process. Yet, despite its importance and impact in basic biology and regenerative medicine, the molecular mechanisms of reprogramming remain poorly understood (see review by M.A. Surani et al., page 747 of this issue). Moreover, DNA methylation systems and the chromatin machinery that writes and interprets histone modifications are clearly involved in disease states, particularly cancer (Sparmann and van Lohuizen, 2006; Yoo and Jones, 2006; see Review by P.A. Jones and S.B. Baylin, page 683 of this issue). Are alterations in epigenetic pathways the cause or effect of the diseases in question? Despite these uncertainties, therapeutic inhibition of histone deacetylation and DNA methylation is already proving useful in clinical trials of cancer therapies (Yoo and Jones, 2006). As we venture into human biology, how do we move the field of epigenetics out of the cell and into the organism? While the field of chromatin biology has been revolutionized by biochemical approaches, we should welcome studies that apply the powerful tools of genetics to validate newly discovered molecular pathways as genuine players in epigenetics and developmental biology.

Epigenetics Today and Tomorrow

We are excited by the emergence of epigenetics as an accepted field of inquiry, by the recent convergence of epigenetic molecular mechanisms on chromatin modification, and by the promise of continued epigenetic research in human development and disease. As with any emerging field, many challenges remain, particularly with regard to the propagation of epigenetic information through cellular division and differentiation. Indeed, questions regarding the stability and inheritance of chromatin modifications, during and outside of DNA replication, must be resolved before considering histones and associated nonhistone proteins as true carriers of epigenetic information.

Modifications to chromatin have clearly emerged as a core mechanism for regulating the transcriptional status of a genetic locus, whether a small element within an individual gene, a chromosomal domain, or even an entire chromosome. However, while transcriptional control and chromatin-related mechanisms are intense areas of research (see review by B. Li et al., page 707 of this issue), we should keep an open mind about non-chromatin and even non-nuclear epigenetic mechanisms, as new phenomena continue to be explored. A good example is the regulation of the *HOTHEAD* gene in *Arabidopsis*, where an inherited RNA cache has been proposed as a possible explanation for non-Mendelian inheritance (Lolle et al., 2005).

Epigenetics as a research area is steeped in a rich literature, but its acceptance as a field has been hindered by the view that it is a mere collection of curious but scattered phenomena documented in a wide range of organisms (Allis et al., 2007). Now the molecular mechanisms underlying diverse epigenetic processes are being defined at a remarkable pace; key players and pathways are being uncovered with new tools and approaches (Figure 2). Many challenges remain, but we agree with James Watson that "the major problem, I think, is chromatin… you can inherit something beyond the DNA sequence. That's where the real excitement of genetics is now" (Watson, 2003).

ACKNOWLEDGMENTS

We thank M. Lachner, A. Ruthenburg, and D. Shechter for helpful comments and Sue Ann Fung-Ho for Figure 2. A.D.G. is supported by NIH MSTP grant GM07739 and C.D.A. by an NIH Merit Award.

REFERENCES

Allis, C.D., Jenuwein, T., Reinberg, D., and Caparros, M.L., eds. (2007). Epigenetics (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press).

Bernstein, E., and Allis, C.D. (2005). Genes Dev. *19*, 1635–1655.

Dobosy, J.R., and Selker, E.U. (2001). Cell. Mol. Life Sci. *58*, 721–727.

Fischle, W., Tseng, B.S., Dormann, H.L., Ueberheide, B.M., Garcia, B.A., Shabanowitz, J., Hunt, D.F., Funabiki, H., and Allis, C.D. (2005). Nature *438*, 1116–1122.

Freitag, M., and Selker, E.U. (2005). Curr. Opin. Genet. Dev. *15*, 191–199.

Goll, M.G., and Bestor, T.H. (2005). Annu. Rev. Biochem. *74*, 481–514.

Grewal, S.I., and Jia, S. (2007). Nat. Rev. Genet. *8*, 35–46.

Heard, E. (2005). Curr. Opin. Genet. Dev. *15*, 482–489.

Jenuwein, T., and Allis, C.D. (2001). Science *293*, 1074–1080.

Lachner, M., O'Carroll, D., Rea, S., Mechtler, K., and Jenuwein, T. (2001). Nature *410*, 116–120.

Lolle, S.J., Victor, J.L., Young, J.M., and Pruitt, R.E. (2005). Nature *434*, 505–509.

Polo, S.E., and Almouzni, G. (2006). Curr. Opin. Genet. Dev. *16*, 104–111.

Ringrose, L., and Paro, R. (2004). Annu. Rev. Genet. *38*, 413–443.

Ruthenburg, A.J., Allis, C.D., and Wysocka, J. (2007). Mol. Cell *25*, 15–30.

Schotta, G., Ebert, A., Dorn, R., and Reuter, G. (2003). Semin. Cell Dev. Biol. *14*, 67–75.

Smith, C.L., and Peterson, C.L. (2005). Curr. Top. Dev. Biol. *65*, 115–148.

Sparmann, A., and van Lohuizen, M. (2006). Nat. Rev. Cancer *6*, 846–856.

Waddington, C.H. (1942). Endeavour *1*, 18–20.

Waddington, C.H. (1957). The Strategy of the Genes; a Discussion of Some Aspects of Theoretical Biology (London: Allen & Unwin).

Watson, J.D. (2003). Sci. Am. *288*, 66–69.

Wysocka, J., Swigut, T., Xiao, H., Milne, T.A., Kwon, S.Y., Landry, J., Kauer, M., Tackett, A.J., Chait, B.T., Badenhorst, P., et al. (2006). Nature *442*, 86–90.

Yoo, C.B., and Jones, P.A. (2006). Nat. Rev. Drug Discov. *5*, 37–50.