

Preface

SINCE PUBLISHING THE FIRST Cold Spring Harbor Laboratory Press edition of *Epigenetics* in 2007, significant advances have been made by researchers worldwide working in multiple fields that touch on epigenetics. For those new to the field, we would point them first to the overview chapter (Chapter 3) for an introduction to the basic concepts and a synthesis of what the field encompasses. The editorial team acknowledges that, although numerous exciting findings could be cited here, several are particularly noteworthy and worth some expanded comments as we chart our course into this edition in 2014.

First, owing to dramatic innovation in sequencing technologies, often referred to as massively parallel or deep sequencing (e.g., genome-wide RNA-Seq or ChIP-Seq approaches), the textbook notion that the flow of genetic information is from DNA to protein via messenger RNA has undergone a remarkable paradigm shift. It is now widely accepted that RNA alone can perform many diverse roles and that a remarkably large fraction of the genome is transcribed, with some estimates as high as >90%. Interestingly only ~2% of these transcripts fall into the messenger RNA category, with a high percentage (~70%) accounting for divergently transcribed noncoding RNAs, be they long or short (see the essay by Rinn [Ch. 2]; also reviewed in Darnell 2011; Guttman and Rinn 2012). The function of these noncoding RNAs remains one of the intense areas of investigation, with emerging models suggesting that these RNAs may work to integrate or provide a scaffold for the chromatin-remodeling and -modifying enzyme complexes or to bring about critical changes in the nuclear architecture through *cis* or *trans* mechanisms, as well as to allow recruitment of factors that silence a chromatin domain (e.g., Polycomb) or facilitate transcription (e.g., eRNA recruiting mediator, elaborated in the Kim et al. essay). We also point out the intriguing links between defined histone modifications and the splicing of pre-messenger RNA (i.e., intronic and exonic definition) to underscore the concept that the RNA world is expanding and intimately linked to chromatin states (Huff et al. 2010).

Second, remarkable progress has been made in documenting the discovery and structures of chromatin-binding modules that “read” one or more histone modifications (see the new chapters in this book by Cheng, Patel, Marmorstein and Zhou, and Seto and Yoshida). How can one make sense

of all of this staggering posttranslational modification complexity? Carrying epigenomics to a genome-wide scale, Zhong and co-workers (Xiao et al. 2012) have introduced what they refer to as “comparative epigenomics,” wherein an impressive collection of epigenetic marks (histone modifications, genomic distributions of cytosine methylation, histone variants, transcription factors, etc.) have been mapped in human, mouse, and porcine cells, drawing upon evolution as a useful guide for highlighting the functional importance of various marks. Importantly, comparative epigenomics has revealed regulatory features of the genome that cannot be ascertained by sequence comparisons alone. Outside of better-known co-associating marks, such as those associated with bivalent domains (i.e., H3K4me3 and H3K27me3) at promoters of developmentally regulated genes, other highly conserved co-marks have been identified. For example, H3K27ac+H3K4me1/2 and H3K27ac+H3K4me2/3 mark active enhancer and promoter elements, respectively. The authors of these findings conclude that the general problem of “having too many epigenetic mark combinations and not knowing how to distinguish random versus functional co-localizations can be overcome by using evolutionary conservation.” We applaud this study as it provides a fresh approach to the complexities of epigenomes, and we look forward to other studies that draw on the insights gleaned from using evolution as a guide.

Third, the fundamental question remains as to how any epigenetic marks are inherited, with our understanding being much more complete for how cytosine methylation marks in DNA are templated during replication. With respect to histone marks, an emerging literature suggests novel mechanisms that include allosteric regulation of the key histone-modifying enzyme complexes, wherein modifications on one histone tail, such as histone H2B ubiquitination (McGinty et al. 2008) or histone H3K27me3 (Margueron et al. 2009), can stimulate downstream activating (e.g., DOT1L [KMT4]) or inactivating (e.g., PRC2) histone methyltransferases (KMTs), respectively. Taken together, these groundbreaking studies suggest that new covalent modifications can be introduced to naïve chromatin templates, providing a potential mechanism of inheritance from unmodified (in some cases, newly synthesized histones) to newly modified states during replication and

chromatin assembly, that can be passed on to future generations. We look forward to future studies along this line, especially when addressed by *in vivo* (i.e., mutants of histones and chromatin machinery; see Rando 2012a) and *in vitro* (i.e., the use of “designer chromatin” templates; see Fierz and Muir 2012) systems. Complexities of this “language” include elaborating the cross-talk relationships between histone marks, with the added complication of number and type of covalent modifications in both histone proteins (e.g., mono- vs. di- vs. trilyserine methylation, lysine acetylation vs. crotonylation, arginine symmetric vs. asymmetric dimethylation) and DNA (e.g., methylation vs. hydroxymethylation on cytosine residues). There is no doubt that deciphering the links between histone modifications, DNA methylation, and noncoding RNAs promises to stimulate and challenge the next generation of scientists entering the general field of epigenetics.

Fourth, histone variants provide cells with the means to tailor chromatin assembly pathways to create distinct chromatin states at distinct genomic locations. We envision that the evolution of histone variants has given the cell a regulatory option to remodel the chromatin template, even outside of the classical notion of coupling histone synthesis to DNA replication during S phase (i.e., replication-independent histone deposition; see the chapter by Henikoff and Smith). It is not surprising then that histone variants, especially the replication-independent types, would require a dedicated machinery and energy to accomplish their task of “courting” and “escorting” their histones into place in the genome. Quite recently, a remarkable series of papers, spearheaded largely by physician scientists using exome sequencing, have identified mutations in “epigenetic regulators” in a remarkably wide range of human cancers. For example, DAXX, ATRX, and the H3.3 variant have been linked to tumorigenesis (pancreatic neuroendocrine tumors or panNETs for short; Jiao et al. 2011), strongly suggesting that DAXX-mediated, H3.3-specific chromatin assembly constitutes the tumor-suppressor function of the ATRX-DAXX complex, likely leading to chromosomal abnormalities that include dysfunctional telomeres. Perhaps the biggest surprise came with the finding that cancer-causing mutations exist in histone-encoding genes themselves (reviewed in Dawson and Kouzarides 2012; You and Jones 2012; Shen and Laird 2013). One of us (C.D.A.) has been known to say, “Every amino acid in histones matters,” but this contention is difficult to test in organisms where histone genetics is not readily feasible. Given that oncogenic mutations have now been mapped to H3 amino termini at two “hot spots”—K27 and G34—in distinct groups of pediatric glioblastoma patients (interestingly those with stem vs. cortex tumors, respectively) (described in the Liu et al. essay [Ch. 2]), we look forward to

insights here that will help to diagnose a devastating and lethal set of childhood cancers (see Rheinbay et al. 2012 for review and references). A high-frequency H3 mutation at K36 has also been linked to other pediatric cancers (e.g., chondroblastoma; Behjati et al. 2013), underscoring the functional importance of lysine-based covalent modifications in histone proteins. Examples are being uncovered in other components of the epigenetic machinery that have disease links that lie outside of cancer (e.g., pathways linked to neurological functions and mental retardation; see Schaefer et al. 2011; Lotsch et al. 2013). Cancer and other disease links (covered in the Baylin and Jones, Audia and Campbell, and Zoghbi and Beaudet chapters; see also the Qi, Schaefer, and Liu et al. chapters) promise to fuel the continued interest in epigenetics well into the future editions beyond this one.

Fifth, the idea that chromatin-remodeling pathways might provide therapeutically useful targets, which may permit mis-silenced or mis-activated genes to be reversed as the genes themselves are not altered by mutations, has led to the general acceptance that developing drugs against chromatin-based targets is a viable new route for treatment in clinical oncology. Specifically, the identification of altered DNA methylation and histone acetylase (HAT) activity in a range of human cancers, coupled with the use of histone deacetylase (HDAC) and DNA methylation inhibitors in the treatment of human cancer, make this a compelling argument, as do the well-documented genetic lesions in histone lysine methyltransferases such as EZH2 (KMT6A), MMSET, etc. Given the genetic links to these key epigenetic-based enzymes, small molecule inhibitors have been designed and tested with positive therapeutic outcomes. Some of these inhibitors are FDA-approved and in widespread use in clinical trials. It is clear that the regulatory signals provided by chromatin modifications will revolutionize our view of cancer as new models of “epigenetic carcinogenesis” are advanced (see also the chapter by Audia and Campbell).

Catalytic enzymes are not the only class of epigenetic regulators that have proven to be worth drugging. In late 2010, a pair of back-to-back papers (Filippakopoulos et al. 2010; Nicodeme et al. 2010) revealed that histone acetyl-lysine binding pockets, or bromodomains, are druggable by small molecules, with useful clinical outcomes (see the essays by Schaefer and by Qi [Ch. 2] and the chapters by Busslinger and Tarakhovskiy and by Marmorstein and Zhou). Moreover, this work laid the foundation for an equally remarkable study wherein large-scale structural analysis of the human bromodomain family was performed, providing remarkable insights into the molecular discrimination by which the different histone acetyllysine reading modules discriminate different chromatin contexts

(Filippakopoulos et al. 2012). We look forward to extending these types of studies to other chromatin-reading “pockets,” with added specificity holding promise for a new frontier for drug discovery (Arrowsmith et al. 2012). Last, with regard to potential therapeutic targets, we stress that *histones* are not the only physiologically relevant recipients of this covalent “language.” Large cohorts of *nonhistone* proteins are now well known to be modified by what were originally described as histone-modifying enzymes (e.g., the acetylation and methylation of p53 by p300 [KAT3B] and Set7/9 [KMT7], respectively, originally reported by Gu and Roeder 1997; Chuikov et al. 2004). Histone “mimicry” has been well documented by Tarakhovsky and others (Sampath et al. 2007; Marazzi et al. 2012), suggesting that these mechanisms extend well beyond histone proteins (Sims and Reinberg 2008).

Finally, the very roots of epigenetics are grounded in problems of developmental biology, as articulated by Waddington and others (see the Felsenfeld chapter). The chromatin packaging system has evolved to make certain genes less or more accessible to transcription factors and other machinery that must engage the true genetic template (see the closing Pirrotta chapter). Although there can be little doubt that we are entering a “postgenomic” or “epigenomic” era, we acknowledge that transcriptional networks likely lie at the heart of reprogramming differentiated cell types from more pluripotent embryonic cell types. Nowhere is that better illustrated than with the generation of induced pluripotent stem cells (iPS cells or iPSCs) by Yamanaka and colleagues in 2006, wherein a set of master gene transcription factors encoded by key pluripotency genes (e.g., Oct-3/4, Sox2, c-Myc, and Klf4) were introduced into nonpluripotent cells, such as mouse adult fibroblasts, and shown to reprogram (or dedifferentiate) them backward to more pluripotent or totipotent states (Takahashi and Yamanaka 2006). These groundbreaking studies build nicely upon pioneering studies by Gurdon and others, who demonstrated early on that somatic, adult nuclei could be reprogrammed, provided that they were transplanted into an egg (oocyte) environment (Gurdon et al. 1958). Although the importance of transcription “master regulators” cannot be questioned, the low efficiency of reprogramming, the stability of the induced states, and the tendency for reprogrammed cells to take a turn toward a more neoplastic state suggest chromatin underpinnings or “barriers” to the reprogramming process have yet to be fully understood (Soufi et al. 2012; Chen et al. 2013). The editors are pleased that the discovery of induced pluripotency is described in the Takahashi essay (Ch. 2), and the general topic of “reprogramming” is covered in this textbook in the Hochedlinger and Jaenisch chapter. As well, topics that closely align epigenetics with developmental

biology issues include the chapters by Grossniklaus and Paro, Kingston and Tamkun, and Reik and Surani.

In closing, this preface highlights only a few of the exciting areas that have come to light since the first edition was published. Our overview and the chapters that follow will not only develop these areas further, but also touch on many more. As well, a novel collection of short essays by junior scientists, who made important discoveries that have already set the field of epigenetics on new and exciting courses, appears in this edition. These essays touch on the history of how these discoveries were made. A quick comparison between the first and second edition of this book underscores the remarkable progress made by the field between these editions. Twelve new chapters have been added with a significant updating of all of the earlier chapters. For example, Figure 3 of Ch. 3 (also in the first edition) suggested that epigenetic alterations, as compared to true genetics, might not be stable or part of true germline inheritance. However, the long-standing debate contrasting the distinction between innate and acquired characteristics (Lamarckian theory) is being revisited in light of new research indicating that environmental factors can provide adaptive responses, via noncoding RNAs in somatic and germline lineages (Ashe et al. 2012; Lee et al. 2012; Rando 2012b). Clearly, new discoveries, likely fueled by readers of this edition, will form the foundation of other editions that take us beyond our current understanding. The developmental biologists of the past must be looking down with great pleasure.

Our goal here, as with the first edition, is to educate newcomers and seasoned veterans alike as to the key concepts that shape and guide the broad field of epigenetics. Words of others underscore the most general problem that our textbook hopes to address: “*We are more than the sum of our genes*” (Klar 1998); “*You can inherit something beyond the DNA sequence. That’s where the real excitement is now*” (Watson 2003); or *Time* magazine’s 2010 cover story headline “*Why your DNA isn’t your destiny*” (Cloud 2010). The field of epigenetics does not appear to be slowing down; remarkably the slope of its citation index in the literature continues to climb. We hope that the readers of this textbook will share in our excitement and yet be inspired to tackle the many problems that remain unsolved or poorly understood. We remain grateful to all of those who have turned this edition into a long-awaited reality.

REFERENCES

Arrowsmith CH, Bountra C, Fish PV, Lee K, Schapira M. 2012. Epigenetic protein families: A new frontier for drug discovery. *Nat Rev Drug Discov* 11: 384–400.

Preface

- Ashe A, Sapetschnig A, Weick EM, Mitchell J, Bagijn MP, Cording AC, Doebley AL, Goldstein LD, Lehrbach NJ, Le Pen J, et al. 2012. piRNAs can trigger a multigenerational epigenetic memory in the germline of *C. elegans*. *Cell* **150**: 88–99.
- Behjati S, Tarpey PS, Presneau N, Scheipl S, Pillay N, Van Loo P, Wedge DC, Cooke SL, Gundem G, Davies H, et al. 2013. Distinct H3F3A and H3F3B driver mutations define chondroblastoma and giant cell tumor of bone. *Nat Genet* **45**: 1479–1482.
- Chen J, Liu H, Liu J, Qi J, Wei B, Yang J, Liang H, Chen Y, Chen J, Wu Y, et al. 2013. H3K9 methylation is a barrier during somatic cell reprogramming into iPSCs. *Nat Genet* **45**: 34–42.
- Chuikov S, Kurash JK, Wilson JR, Xiao B, Justin N, Ivanov GS, McKinney K, Tempst P, Prives C, Gambelin SJ, et al. 2004. Regulation of p53 activity through lysine methylation. *Nature* **432**: 353–360.
- Cloud J. 2010. Why genes aren't destiny. *Time* **175**: 48–53.
- Darnell JE. 2011. *RNA: Life's indispensable molecule*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Dawson MA, Kouzarides T. 2012. Cancer epigenetics: From mechanism to therapy. *Cell* **150**: 12–27.
- Fierz B, Muir TW. 2012. Chromatin as an expansive canvas for chemical biology. *Nat Chem Biol* **8**: 417–427.
- Filippakopoulos P, Qi J, Picaud S, Shen Y, Smith WB, Fedorov O, Morse EM, Keates T, Hickman TT, Felletar I, et al. 2010. Selective inhibition of BET bromodomains. *Nature* **468**: 1067–1073.
- Filippakopoulos P, Picaud S, Mangos M, Keates T, Lambert JB, Barsyte-Lovejoy D, Felletar I, Volkmer R, Muller S, Pawson T, et al. 2012. Histone recognition and large-scale structural analysis of the human bromodomain family. *Cell* **149**: 214–231.
- Gu W, Roeder RG. 1997. Activation of p53 sequence-specific DNA binding by acetylation of the p53 C-terminal domain. *Cell* **90**: 595–606.
- Gurdon JB, Elsdale TR, Fischberg M. 1958. Sexually mature individuals of *Xenopus laevis* from the transplantation of single somatic nuclei. *Nature* **182**: 64–65.
- Guttman M, Rinn JL. 2012. Modular regulatory principles of large non-coding RNAs. *Nature* **482**: 339–346.
- Huff JT, Plocik AM, Guthrie C, Yamamoto KR. 2010. Reciprocal intronic and exonic histone modification regions in humans. *Nat Struct Mol Biol* **17**: 1495–1499.
- Jiao Y, Shi C, Edil BH, de Wilde RF, Klimstra DS, Maitra A, Schulick RD, Tang LH, Wolfgang CL, Choti MA, et al. 2011. DAXX/ATRX, MEN1, and mTOR pathway genes are frequently altered in pancreatic neuroendocrine tumors. *Science* **331**: 1199–1203.
- Klar AJ. 1998. Propagating epigenetic states through meiosis: Where Mendel's gene is more than a DNA moiety. *Trends Genet* **14**: 299–301.
- Lee HC, Gu W, Shirayama M, Youngman E, Conte D Jr, Mello CC. 2012. *C. elegans* piRNAs mediate the genome-wide surveillance of germline transcripts. *Cell* **150**: 78–87.
- Lotsch J, Schneider G, Reker D, Parnham MJ, Schneider P, Geisslinger G, Doehring A. 2013. Common non-epigenetic drugs as epigenetic modulators. *Trends Mol Med* **19**: 742–753.
- Marazzi I, Ho JS, Kim J, Manicassamy B, Dewell S, Albrecht RA, Seibert CW, Schaefer U, Jeffrey KL, Prinjha RK, et al. 2012. Suppression of the antiviral response by an influenza histone mimic. *Nature* **483**: 428–433.
- Margueron R, Justin N, Ohno K, Sharpe ML, Son J, Drury WJ 3rd, Voigt P, Martin SR, Taylor WR, De Marco V, et al. 2009. Role of the Polycomb protein EED in the propagation of repressive histone marks. *Nature* **461**: 762–767.
- McGinty RK, Kim J, Chatterjee C, Roeder RG, Muir TW. 2008. Chemically ubiquitylated histone H2B stimulates hDot1 L-mediated intranucleosomal methylation. *Nature* **453**: 812–816.
- Nicodeme E, Jeffrey KL, Schaefer U, Beinke S, Dewell S, Chung CW, Chandwani R, Marazzi I, Wilson P, Coste H, et al. 2010. Suppression of inflammation by a synthetic histone mimic. *Nature* **468**: 1119–1123.
- Rando OJ. 2012a. Combinatorial complexity in chromatin structure and function: Revisiting the histone code. *Curr Opin Genet Dev* **22**: 148–155.
- Rando OJ. 2012b. Daddy issues: Paternal effects on phenotype. *Cell* **151**: 702–708.
- Rheinbay E, Louis DN, Bernstein BE, Suva ML. 2012. A tell-tail sign of chromatin: Histone mutations drive pediatric glioblastoma. *Cancer Cell* **21**: 329–331.
- Sampath SC, Marazzi I, Yap KL, Sampath SC, Krutchinsky AN, Mecklenbrauker I, Viale A, Rudensky E, Zhou MM, Chait BT, et al. 2007. Methylation of a histone mimic within the histone methyltransferase G9a regulates protein complex assembly. *Mol Cell* **27**: 596–608.
- Schaefer A, Tarakhovskiy A, Greengard P. 2011. Epigenetic mechanisms of mental retardation. *Prog Drug Res* **67**: 125–146.
- Shen H, Laird PW. 2013. Interplay between the cancer genome and epigenome. *Cell* **153**: 38–55.
- Sims RJ III, Reinberg D. 2008. Is there a code embedded in proteins that is based on post-translational modifications? *Nat Rev Mol Cell Biol* **9**: 815–820.
- Soufi A, Donahue G, Zaret KS. 2012. Facilitators and impediments of the pluripotency reprogramming factors' initial engagement with the genome. *Cell* **151**: 994–1004.
- Takahashi K, Yamanaka S. 2006. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**: 663–676.
- Watson JD. 2003. Celebrating the genetic jubilee: A conversation with James D. Watson. Interviewed by John Rennie. *Sci Am* **288**: 66–69.
- Xiao S, Xie D, Cao X, Yu P, Xing X, Chen CC, Musselman M, Xie M, West FD, Lewin HA, et al. 2012. Comparative epigenomic annotation of regulatory DNA. *Cell* **149**: 1381–1392.
- You JS, Jones PA. 2012. Cancer genetics and epigenetics: Two sides of the same coin? *Cancer Cell* **22**: 9–20.

Acknowledgments

AS IS LIKELY WITH EVERY MAJOR textbook undertaking, the project seems to grow beyond boundaries with many reasons why the book ever sees the light of day. Nowhere is this truer than with this second edition of the textbook *Epigenetics*. Here the boundaries grew: The number of chapters increased, as did the size of our overview and concepts chapter. Why is this? Here we can only suggest that part of it lies with all of the exciting science that is the collective field of epigenetics.

We are hugely grateful to all of the authors contributing to this second edition—some from the first edition of this book, but also many newcomers. Special thanks need to go to the junior scientists who stopped their benchwork to share with us, in highlighted essays contained in the second chapter, some of what was behind all of the exciting discoveries made firsthand by them and their colleagues, helping to make the field what it is today. It is the care and attention of all these authors that form the heart of this book; their knowledge and expertise makes this textbook what it we intended it to be—the latest detailed textbook on the exciting topic of epigenetics. With each chapter and essay, we have consulted with outside experts who weighed in with their constructive comments, making the written pieces as accurate and timely as possible. We thank them all.

As with the first edition, we are grateful for John Inglis's instigation and support for the second edition and to all of the staff at CSHL Press (Inez Sialiano, Kathy Bubbeo, Richard Sever, Jan Argentine, and Denise Weiss), who have been key players in this effort. With this second edition, the editors owe them special thanks as we pushed them to the breaking point with endless delays and behind-schedule demands. We also appreciate that CSHL Press will make each of the chapters of the textbook available online at *CSH Perspectives in Biology*, so that each contribution can be cited as a full publication. All of the editors' assistants (Marisa Cerio [C.D.A.], Marcela Mare [T.J.], and Michele Giunta [D.R.]) have also shown remarkable patience

trying their best to arrange the countless calls, meetings, and calendars to make it all happen.

Special gratitude needs to be directed to two individuals, who, as much as anyone involved with this second edition, know every page, sentence, and word included in this textbook, inside and out. Put simply, Marie-Laure Caparros and Monika Lachner have made this textbook a reality. From text to tension, references to rants, figures to frustration, appendices to anxiety, they have done it all. Where does their patience come from? No one knows, but the three of us guess that some form of very special genetics and epigenetics are in both of them. Thank you beyond words for your amazing work, care, and attention to the fine details that make a good textbook, hopefully, a great one.

Finally, the three of us acknowledge that any mistakes or omissions in this textbook lie with us. Many folks have wondered what took this second edition so long to come out? What dark hole did those deadlines fall into? Slowness on our part has likely caused these delays. Please accept our apologies. Even so, we have enjoyed the challenge of trying to bring the many advances made in the field of epigenetics into a form where all of the readers of this textbook will join us in the genuine excitement that promises to come.

Financial support for this book was provided by CSHL Press (New York), Rockefeller University (New York), the NYU School of Medicine (New York), and the Max Planck Institute of Immunobiology and Epigenetics (Freiburg, Germany) provided further funds for the book's developmental editing.

C. David Allis, *The Rockefeller University, New York*

Thomas Jenuwein, *Max Planck Institute of Immunobiology and Epigenetics, Freiburg*

Danny Reinberg, *NYU School of Medicine Smilow Research Center*

September 15, 2014

Acknowledgments



(Left to right) Monika Lachner, Thomas Jenuwein, Danny Reinberg, Marie-Laure Caparros, and David Allis at an editorial meeting in New York.