现代生物技术前沿

Epigenetics

〔美〕C．D．艾利斯
〔奥〕T．詹内怀恩
〔美〕D．赖因伯格
〔英〕M.-L．卡珀若斯
编著

表观遗传学

（影印版）

科学出版社
www.sciencep.com

现代生物技术前沿

Epigenetics

〔美〕C．D．艾利斯
〔奥〕T．詹内怀恩
〔美〕D．赖因伯格
〔英〕M.-L．卡珀若斯
编著

表观遗传学

（影印版）

科学出版社
北京

图字:01-2007-5841

内 容 简 介

表观遗传学是研究非DNA序列变化的、可遗传的表达改变的科学,表观遗传学机制参与个体发育、成体的许多生物学进程中基因表达的调节,这一过程的异常可导致人类癌症和其他疾病的发生。本书是基础研究专著,24章均由知名专家撰写,分别从不同角度探讨表观遗传学问题。首先介绍表观遗传学简史、基本概念;然后详细阐释表观遗传调节的分子机制及与之相关的细胞生物学过程;介绍了组氨酸和DNA甲基化,siRNA和基因沉默,X染色体失活,剂量补偿和印记,以及微生物、植物、昆虫和哺乳动物中的表观遗传机制;还讨论了细胞分裂和分化中的表观遗传作用,以及这些途径中发生的错误对癌症和人类肿瘤的影响;此外,也阐述了表观遗传在动物克隆中可能发生的影响。

本书可供发育生物学、细胞生物学、遗传学、基因组学、神经生物学和分子生物学等生命科学相关领域的教学研究人员,以及本科生、研究生参考。

书名原文:Epigenetics
原著正文内彩图集中列于书中所附光盘。

图书在版编目(CIP)数据

表观遗传学=Epigenetics:英文/(美)艾利斯(Allis, C. D.)等编著.—影印本.—北京:科学出版社,2008

(现代生物技术前沿)

ISBN 978-7-03-020774-6

Ⅰ.表… Ⅱ.艾… Ⅲ.发育遗传学-英文 Ⅳ.Q344

中国版本图书馆CIP数据核字(2007)第200842号

责任编辑:莫结胜/责任印制:钱玉芬/封面设计:王 浩

科学出版社出版
北京东黄城根北街16号
邮政编码:100717
http://www.sciencep.com

铭浩彩色印装有限公司印刷

科学出版社发行 各地新华书店经销

*

2008年1月第 一 版 开本:889×1194 1/16
2008年1月第一次印刷 印张:32
印数:1—2 000 字数:991 000

定价:96.00元(含光盘)

(如有印装质量问题,我社负责调换〈环伟〉)

Long before epigenetics changed from little more than a diverse collection of bizarre phenomena to a well-respected field covered by its own textbook, a talented group of foresighted molecular biologists laid a rich foundation upon which the modern era of chromatin biology and epigenetics is based. This group includes Vince Allfrey, Wolfram Hörz, Hal Weintraub, Alan Wolffe, and Abe Worcel. This book is dedicated to their collective memory. Their passion and commitment to the study of chromatin biology inspired all of us who followed their work, and we now profit from their many insights.

Preface

This advanced textbook on "Epigenetics" is truly a reflection of many talented colleagues and individuals, all of whom made this book possible and a rewarding experience. However, without hesitation, the editors want to thank Marie-Laure Caparros (London), without whom this project would have never materialized. Early in the process, it became evident that the editorial team needed help in coordinating such a large project, particularly for keeping the dialogue and editorial feedback with the >40 colleagues who agreed to provide outstanding chapter contributions, only to realize that we wanted more than their expert reviews and attention to detail. Marie-Laure has been instrumental in keeping the momentum moving forward, has bravely exchanged critical comments when needed, has informed all of us on the many deadlines, and has provided necessary coherence to make embryonic chapters come to life. Without her, this book would not have been possible. We are also grateful to our individual assistants, who forever kept us on our toes: Elizabeth Conley (David Allis), Christopher Robinson (Thomas Jenuwein), and Shelli Altman (Danny Reinberg). All of them are the unsung heroes of this book. We thank all of them for their innumerable contributions, large and small, and their unending patience with each of us and our quirky styles and shortcomings as editors.

Discussions for such a book took initial form on the coattails of the outstanding 69th Cold Spring Harbor Symposium on *Epigenetics* in the summer of 2004, but were seeded in early 2003 and formally commissioned by CSHL Press through Alex Gann and other colleagues. This was followed by formulating an editorial team between David Allis, Thomas Jenuwein, and Danny Reinberg. The first concrete outline for this project, including the brainstorming of various chapters and potential contributing authors, was done on the picnic bench at the FASEB meeting on *Chromatin and Transcription* in Snowmass, Colorado, July 2003. We were then very fortunate to confirm the lineup of contributing colleagues who are the leaders in their field.

In the early planning stages, a vision crystallized for a different concept. Ideally, we sought to ask not for a compilation of expert reviews which might soon be outdated. Rather, we wanted to compile a set of conceptual chapters, from pairs of experts, that highlight important discoveries for students in chromatin biology and for colleagues outside the epigenetics field. In keeping to a conceptual outline, we hoped to have a more long-lasting impact. Also, by including many diagrams and illuminating figures, and appendices, we hoped to list most of the systems and epigenetic marks currently known. The General Summaries were aimed as a stand-alone précis of the topics covered in each chapter, preceded by "teaser" images to entice the reader to investigate.

The figures have been another important hallmark for this book; particularly, the examples provided in the Overview and Concepts chapter. Here, Stefan Kubicek, a Ph.D. student from the Jenuwein lab at the IMP (Vienna), and Marie-Laure Caparros have been the masters of the diagrams. They honed draft upon draft of figures (sometimes only from sketches) for the chapters, such that we could gain a more coherent presentation. Several postdocs and Ph.D. students (Gabriella Farkas, Fatima Santos, Heike Wöhrmann, and others) in the labs of several authors also kindly contributed to the excellent illustrations in this book. However, we were unable to convert all of the contributions, and some figures have remained as submitted. We are also particularly grateful to Monika Lachner, Mario Richter, Roopsha Sengupta, Patrick Trojer, and other Ph.D. students and Postdocs in the Allis, Jenuwein, and Reinberg laboratories for amending, proofreading, and finalizing the tables and summaries that are displayed in the appendices. Here, Dr. Steven Gray (St. James Hospital, Dublin) has been

particularly instrumental in validating and providing additional information for the table that lists all the currently known histone modifications.

Where appropriate, submitted chapters were sent out for comments from other colleagues who provided important input for streamlining and clarifying some of the complex concepts. Not all of this input could be converted into the revised and final versions, but the comments and suggestions helped to shape many of the chapters and the overall framework of the book. Here, we are indebted to G. Almouzni, P. Becker, H. Cedar, V. Chandler, W. Dean, R. Feil, A. Ferguson-Smith, M. Gartenberg, S. Grewal, M. Hampsey, E. Heard, R. Metzenberg, V. Pirrotta, F. Santos, T. Schedl, D. Solter, R. Sternglanz, S. Tilghman, and others.

Finally, we acknowledge the intellectual and, in some cases, emotional contributions made by all of our colleagues in the field who provided the chapters to make this book what it is. Their contributions, by way of written chapters and drawings, stand by themselves. But what may not be obvious is the feedback and cross-fertilization that all of them had with the editorial team to help shape and guide the book as it took form. The Overview and Concepts chapter itself reflects their feedback, as in early drafts, we put too much of our own colors and bias into the sentences. For their wisdom and for bringing us a deeper perspective and balance, we thank them, and we admit that any deficiencies and mistakes there are ours.

Financial support for this book has come from CSHL Press (New York), the *Epigenome* FP6 NoE (European Union), IMP (Vienna), the Rockefeller University (New York), and the Howard Hughes Medical School-Robert Wood Johnson Medical School (Piscataway, New Jersey). Critical contributions were also made by Upstate Serologicals (Lake Placid, New York) and AbCam (Cambridge, UK), leading suppliers of epigenetic-based reagents and tools.

CDA, TJ, DR

Epigenetic Mechanisms That Operate in Different Model Organisms

		S. cerevisiae	*S. pombe*	*N. crassa*	*C. elegans*	*Drosophila*	Mammals	*A. thaliana*
	GENOMIC FEATURES							
	Genome size	12 Mb	14 Mb	40 Mb	100 Mb	180 Mb	3,400 Mb	150 Mb
	Number of genes	6,000	5,000	10,000	20,000	14,000	~25,000	25,000
	Average size of genes	1.45 kb	1.45 kb	1.7 kb	2 kb	5 kb	35–46 kb	2 kb
	Average number of introns/gene	≤1	2	2	5	3	6–8	4–5
	% Genome as protein coding	70	60	44	25	13	1–1.5 (Hs)	26
	EPIGENETIC FEATURES							
ON	Histone acetylation	+	+	+	+	+	+	+
ON	H3K4 methylation	+	+	+	+	+	+	+
ON	H3K36 methylation	+	+	+	+	+	+	+
ON	H3K79 methylation	+	+	+	+	+	+	+
ON	H3.3 histone variant	+	+	+	+	+	+	+
ON/OFF	SWI/SNF ATPase complexes	+	+	+	+	+	+	+
ON	CHD1 ATPase family	+	(+)[a]	(+)[a]	(+)[a]	+	+	+
ON	SWR1 ATPase	+	(+)[a]	(+)[a]	(+)[a]	+	+	(+)[a]
ON/OFF	ISWI ATPase	+	+	+	+	+	+	+
ON/OFF	INO80 ATPase	+	+	+	+	+	+	+
OFF	MI-2 ATPase	–	(+)[a]	(+)[a]	+	+	+	+
OFF	CENP-A centromeric histone variant	+	+	+	+	+	+	+
OFF	H3K9 methylation[b]	–	+	+	+	+	+	+
OFF	HP1-like proteins	–	+	+	+	+	+	+
OFF	RNA interference	–	+	+	+	+	+	+
OFF	H4K20 methylation[c]	–	+	+	+	+	+	+
OFF	H3K27 methylation	–	–	+	+	+	+	+
OFF	Polycomb repressive complexes	–	–	–	+	+	+	+
OFF	DNA methylation	–	–	+	–	(+)[d]	+	+
OFF	DNA methylation binding proteins	–	+[e]	+	+[f]	+[g]	+	+
OFF	Imprinting	–	–	–	–	+[h]	+	+

Abbreviation: (Hs) *Homo sapiens*.

[a] Epigenetic feature considered to be present based on sequence homology but no functional data.

[b] There is evidence that H3K9 methylation is found at active chromatin regions; however, the functional significance of this is unknown.

[c] H4K20 tri-methylation is not present in *S. cerevisiae*, whereas all three H4K20 methylation states are present in multicellular organisms.

[d] *Drosophila* possess very low levels of DNA methylation.

[e] Mutated Dnmt2.

[f] Dnmt2 (Pp) and MBD-domain proteins (Ce, Cb, Pp).

[g] Dnmt2 and MBD-domain proteins (Dm).

[h] Chromosome- or genome-wide rather than gene-specific.

The *Brno Nomenclature* for Histone Modifications

Modifying group	Amino acid(s) modified	Level of modification	Abbreviation for modification[a]	Examples of modified residues[b]
Acetyl-	lysine	mono-	ac	H3K9ac
ADP ribosyl-	glutamate	mono-	ar1	H2BE2ar1
	glutamate	poly-	arn	H2BE2arn[d]
Biotin-	lysine	mono-	bio	H4K12bio
Methyl-	arginine	mono-	me1	H3R17me1
	arginine	di-, symmetrical	me2s	H3R2me2s
	arginine	di-, asymmetrical	me2a	H3R17me2a
Methyl-	lysine	mono-	me1	H3K4me1
	lysine	di-	me2	H3K4me2
	lysine	tri-	me3	H3K4me3
Phosphoryl-	serine or threonine	mono-	ph	H3S10ph
Ubiquityl-	lysine	mono-[c]	ub1	H2BK120ub1
Sumoyl-	lysine	mono-	su	H4K5su[d]

[a] The use of lowercase letters for the modifications helps distinguish them from either amino acids (identified by their single-letter codes) or histones (such as H2A), for which letters are always uppercase.

[b] The nomenclature starts from the left with the histone, then the residue, then the modification. In cases where the modified residue is not known, or not relevant, the modification should follow the histone, for example H4ac and H2Bar1. Multiple modifications can be accommodated by simply extending the listing (for example, H3K4me3K9acS10ph...) for as long as necessary. Because each individual modified residue begins with the uppercase letter specifying the amino acid, and because the modifications themselves are all designated by lowercase letters, the use of commas or dots to separate the individual modified residues in a 'word' specifying multiple modifications is not necessary. On occasion, the presence of an unmodified residue may be an essential component of an information-bearing combination of residues; in this case, the residue should be inserted without additions (for example, H3K9S10ph) to indicate H3 unmodified at Lys-9 and phosphorylated at Ser-10.

[c] Polyubiquitylated histones are designated ubn.

[d] Hypothetical at present.

Adapted, with permission, from Turner B.M., *Nat. Struct. Mol. Biol.* **12**: 110–112. (2005).

Contents

CHAPTER 1

Epigenetics: From Phenomenon to Field

Daniel E. Gottschling

Fred Hutchinson Cancer Research Center, Seattle, Washington 98109

CONTENTS

1 Introduction

In the summer of 2004, the 69th Cold Spring Harbor Symposium on Quantitative Biology covered the topic of "Epigenetics," and many of the authors of this book were in attendance. As an observer at this Symposium, I knew this was going to be an interesting meeting. It started simply enough by trying to define epigenetics. After a week of querying participants about this, it became clear that such a request was akin to asking someone to define "family values"—everyone knew what it meant, but it had a different meaning for each person. Part of the reason for the range of opinions may be understood from the etymology of "epigenetics" as explained by David Haig: The word had two distinct origins in the biological literature in the past century, and the meaning has continued to evolve. Waddington first coined the term for the study of "causal mechanisms" by which "the genes of the genotype bring about phenotypic effects" (see Haig 2004). Later, Nanney used it to explain his realization that cells with the same genotype could have different phenotypes that persisted for many generations.

I define an epigenetic phenomenon as a change in phenotype that is heritable but does not involve DNA mutation. Furthermore, the change in phenotype must be switch-like, "ON" or "OFF," rather than a graded response, and it must be heritable even if the initial conditions that caused the switch disappear. Thus, I consider epigenetic phenomena to include the lambda bacteriophage switch between lysis and lysogeny (Ptashne 2004), pili switching in uropathogenic *Escherichia coli* (Hernday et al. 2003), position-effect variegation in *Drosophila* (Henikoff 1990), heritable changes in cortical patterning of *Tetrahymena* (Frankel 1990), prion diseases (Wickner et al. 2004a), and X-chromosome inactivation (Lyon 1993).

The 69th Symposium came on the 100th anniversary of genetics as a field of study at Cold Spring Harbor Laboratory, making it very timely to consider epigenetics. Given this historical context, I thought it appropriate to provide an examination of epigenetics through the portal of previous Cold Spring Harbor Symposia. Although the 69th Symposium was the first dedicated to the topic, epigenetic phenomena and their study have been presented throughout the history of this distinguished series. The history I present is narrowed further by my limitations and likings. For a more complete and scholarly portrayal, I can recommend the more than 1000 reviews on epigenetics that have been written in the past five years.

In presenting this chronological account, I hope to convey a sense of how a collection of apparently disparate phenomena coalesced into a field of study that affects all areas of biology, and that the study of epigenetics is founded upon trying to explain the unexpected—perhaps more than any other field of biological research.

2 A History of Epigenetics at Cold Spring Harbor Symposia

In 1941 during the 9th Symposium, the great *Drosophila* geneticist H.J. Muller described developments on his original "eversporting displacement," in which gross chromosomal rearrangements resulted in the mutant mosaic expression of genes near the breakpoint (Muller 1941). By the time of this meeting, he referred to it as "position effect variegation." It was well established that the affected genes had been transferred "into the neighborhood of a heterochromatic region," that the transferred euchromatic regions had been "partly, but variably, transformed into a heterochromatic condition—'heterochromatized'," and that *addition* of extra copies of heterochromatic chromosomes "allowed the affected gene to become more normal in its functioning." This latter observation was an unexpected quandary at the time, which we now know to be the result of a titration of limiting heterochromatin components.

At the 16th Symposium (1951), a detailed understanding of the gene was of high priority. This may explain why little progress had been made on understanding position-effect variegation (PEV), although more examples were being discovered. However, the opening speaker noted that PEV would be an exciting area for future research (Goldschmidt 1951). Barbara McClintock noted that chromosomal position effects were the basis of differences in "mutable loci" of maize, and she speculated that the variation of mutability she observed likely had its roots in the same mechanisms underlying PEV in *Drosophila* (McClintock 1951).

By the time of the 21st Symposium, McClintock's ideas about "controlling elements" had developed (McClintock 1956). Two were particularly poignant with regard to epigenetics. In the *Spm* controlling element system, she had uncovered variants that allowed her to distinguish between *trans*-acting factors that could "suppress" a gene (reduce or eliminate its phenotypic expression) rather than mutate it. She also noted that some controlling elements could suppress gene action not only at the locus where it had inserted, but also at loci that were located some distance on either side of it. Others were discovering this "spreading effect" as well. J. Schultz presented a biochemical and physical characterization of whole *Drosophila* that contained

different amounts of heterochromatin (Schultz 1956). Although the work was quite primitive and the conclusions drawn were limited, the work represented early attempts to dissect the structure of heterochromatin and demonstrated just how difficult the problem would be.

Two talks at the 23rd Symposium were landmarks with respect to our present-day Symposium. First, R.A. Brink described his stunning observations of "paramutation" at the *R* locus in maize. If two alleles (R^{st} and R^{r}) with distinct phenotypes as homozygotes are combined to form a heterozygote, and this R^{st}/R^{r} plant is in turn crossed again, the resulting progeny that contain the Rr allele will *always* have an R^{st} phenotype, even though the R^{st} is no longer present (Brink 1958). However, this phenotype is metastable—in subsequent crosses the phenotype reverts to the normal R^{r} phenotype. He meant for the word paramutation "to be applied in this context in its literal sense, as referring to a phenomenon distinct from, but not wholly unlike, mutation." Second, D.L. Nanney went to great lengths to articulate "conceptual and operational distinctions between genetic and epigenetic systems" (Nanney 1958). In essence, he defined epigenetics differently from how it had been originally intended by Waddington (for details, see Haig 2004). He found it necessary to do so in order to describe phenomena he observed in *Tetrahymena*. He found evidence that the cytoplasmic history of conjugating parental cells influenced the mating-type determination of resulting progeny. His definition encompassed observations made by others as well, including Brink's work on the *R* locus and McClintock's work noted in the 21st Symposium.

Mary Lyon's recently proposed hypothesis of X-chromosome inactivation in female mammals (Lyon 1961) was of considerable interest at the 29th Symposium. S. Gartler, E. Beutler, and W.E. Nance presented further experimental evidence in support of it (Beutler 1964; Gartler and Linder 1964; Nance 1964). Beutler reviewed multiple examples of mosaic expression of X-linked genes in women, supporting the random nature of X inactivation. From careful quantitative analysis of an X-linked gene product, Nance deduced that X inactivation occurred before the 32-cell stage of the embryo.

The 38th Symposium on "Chromosome Structure and Function" represented a return to examining eukaryotic chromosomes—significant progress had been made studying prokaryotic and phage systems, and consequently, bacterial gene expression had dominated much of the thinking in the burgeoning field of molecular biology. However, an appreciation for chromatin (DNA with histones and nonhistone proteins) in eukaryotes was building, but it was unclear whether it played a role in chromosome structure or function, or both (Swift 1974). Nevertheless, several groups began to speculate that posttranslational modification of chromatin proteins, including histones, was associated with gene transcription or overall chromosome structure (Allfrey et al. 1974; Louie et al. 1974; Weintraub 1974). There was only a hint of epigenetic phenomena in the air. It had been hypothesized that repetitive DNA regulated most genes in eukaryotes, partly based on the fact that McClintock's controlling elements were repeated in the genome. However, it was reported that most repeated DNA sequences were unlinked to genes (Peacock et al. 1974; Rudkin and Tartof 1974). From these observations, the idea that repeated elements regulated gene expression lost significant support from those in attendance. More importantly, however, these same studies discovered that most of the repetitive DNA was located in heterochromatin.

The 42nd Symposium demonstrated that in four years, an amazing number of technical and intellectual advances had transformed the study of eukaryotic chromosomes (Chambon 1978). This included the use of DNA restriction enzymes, development of recombinant DNA technology, routine separation of proteins and nucleic acids, the ability to perform Southern and northern analysis, rapid DNA and RNA sequencing, and immunofluorescence on chromosomes. The nucleosome hypothesis had been introduced, and mRNA splicing had been discovered. Biochemical and cytological differences in chromatin structure, especially between actively transcribed and inactive genes, comprised the primary interest at this meeting. However, most relevant to epigenetics, Hal Weintraub and colleagues presented ideas about how chromatin could impart variegated gene expression in an organism (Weintraub et al. 1978).

The 45th Symposium was a celebration of Barbara McClintock's discoveries—movable genetic elements (Yarmolinsky 1981). Mechanistic studies of bacterial transposition had made enormous progress and justifiably represented about half the presentations, whereas others presented evidence that transposition and regulated genomic reorganization occurred not only in maize, but also in other eukaryotes—including flies, snapdragons, *Trypanosomes*, *Ascobolus*, and budding yeast. In the context of this meeting, all observed variegated expression events were ascribed to transposition. Moreover, there was a reticence to seriously consider that controlling elements were responsible for most gene regulation (Campbell 1981), which led some to suggest that "the sole function of these elements is to promote genetic variability." In essence, the idea that heterochromatin was responsible for the regulated expression in position-effect

variegation was called into question. With respect to future epigenetic studies, perhaps the most noteworthy discussion was the firm establishment of "silent mating cassettes" in *Saccharomyces cerevisiae* (Haber et al. 1981; Klar et al. 1981; Nasmyth et al. 1981; Rine et al. 1981).

Leading up to the 47th Symposium, a general correlation had been established in vertebrate systems that the overall level of cytosine methylation in CpG DNA sequences was lower for genes that were transcribed than for those that were not. However, there were exceptions to this generalization, and more detailed analysis was presented that methylation of a specific area of a gene's promoter was most important (Cedar et al. 1983; Doerfler et al. 1983; La Volpe et al. 1983). On the basis of restriction/modification systems of bacteria, it was thought that DNA methylation prevented binding of key regulatory proteins. Furthermore, it had been shown that DNA methylation patterns could be mitotically inherited in vertebrates, which led to the hypothesis that DNA methylation could serve as a means of transcriptional "memory" as cells divided through development (Shapiro and Mohandas 1983). Another major epigenetic-related finding was the identification of DNA sequences on either side of the "silent mating cassettes" in budding yeast that were responsible for transcriptional repression of genes within the cassettes—these defined the first DNA sequences required for chromosomal position effects (Abraham et al. 1983).

"The Molecular Biology of Development" was the topic for the 50th Symposium, and it too encompassed a number of important advances. Perhaps one of the most exciting developments was the overall awareness that fundamental molecular properties were conserved throughout evolution—e.g., human RAS functioned in budding yeast, homeo box proteins were conserved between flies and humans (Rubin 1985). New efforts to understand chromosome imprinting began with the development of nuclear transfer in mice (Solter et al. 1985). These studies revealed that parent-of-origin information was stored within the paternal and maternal genomes of a new zygote; it was not just the DNA that was important, but the chromosomes contained additional information about which parent they had passed through, and the information was required for successful development of an embryo. Part of the answer was thought to lie in the fact that differential gene expression was dependent on the parental origin of a chromosome (Cattanach and Kirk 1985).

There were a number of studies aimed at understanding the complex regulation of the bithorax complex, but notably, E.B. Lewis made special mention of the curious nature of known *trans* regulators of the locus; nearly all were repressors of the locus (Lewis 1985). Thus, maintaining a gene in a silenced state for many cell doublings was imperative for normal development. This contrasted with much of the thinking at the time—that gene activation/induction was where the critical regulatory decisions of development would be.

DNA transformation and insertional mutagenesis techniques had recently been achieved for a number of organisms. One particularly creative and epigenetic-related use of this technology came in *Drosophila*. A P-element transposon with the *white* eye-color gene on it was created and "hopped" throughout the genome (Rubin et al. 1985). This provided a means to map sites throughout the *Drosophila* genome where PEV could occur.

This meeting also highlighted the first genetic approaches to dissecting sex determination and sex chromosome dosage compensation in *Drosophila* (Belote et al. 1985; Maine et al. 1985) and *Caenorhabditis elegans* (Hodgkin et al. 1985; Wood et al. 1985).

The 58th Symposium highlighted the celebration of the 40th anniversary of Watson and Crick's discovery. Part of the celebration was a coming-out party for epigenetic phenomena: There was identification of new phenomena, beginnings of molecular analysis of other phenomena, and sufficient progress had been made in a number of systems to propose hypotheses and to test them.

In trypanosomes, the family of *Variable Surface antigen Genes (VSG)* located near telomeres are largely silenced, with only one *VSG* expressed at a time. Although this organism does not appear to contain methylated DNA, it was reported that the silenced *VSG* genes contained a novel minor base: β-D-glucosylhydroxymethyluracil (Borst et al. 1993). This base appeared to be in place of thymidine in the DNA. Parallels between this base and cytosine methylation in other organisms were easy to draw—the modifications were important for maintaining a silenced gene. But how the base was introduced into the DNA, or how it imparted such a function, was unclear.

Progress had also been made in vertebrate epigenetic phenomena, including chromosomal imprinting and X inactivation (Ariel et al. 1993; Li et al. 1993; Tilghman et al. 1993; Willard et al. 1993). It had become clear by this time that numerous loci were subject to imprinting in mammals; only one allele was expressed in diploid cells, and expression was dependent on parental origin. The *Igf2-H19* locus was of particular interest, primarily because it contained two nearby genes that were regulated in opposing fashion. *Igf2* is expressed from the paternal

chromosome while the maternal copy is repressed, whereas the paternal allele of *H19* is repressed and its maternal allele is expressed. Interestingly, methylated CpG was observed just upstream of both genes on the paternal chromosome. It was proposed that the differential methylation regulated access of the two genes to a nearby enhancer element—the enhancer was closer to, and just downstream of, *H19* (Tilghman et al. 1993). A mutually exclusive competition between the two genes for the enhancer was envisioned; when the *H19* gene was methylated, the enhancer was free to activate the more distant *Igf2* gene. Support for the idea that DNA methylation played a regulatory role in this process came from mouse studies. Mutation of the first vertebrate gene encoding a 5-methyl-cytosine DNA methyltransferase in ES cells showed that as embryos developed, the paternal copy of *H19* became hypomethylated and the gene became transcriptionally active (Li et al. 1993).

An important step in the way in which ^{5Me}CpG mediated its effects came from the purification of the first ^{5Me}CpG DNA-binding complex (MeCP1) (Bird 1993). Not only did it bind DNA, but when tethered upstream of a reporter gene, MeCP1 caused the gene to be repressed. Although this did not explain regulation at the *Igf2-H19* locus, it did provide a potential mechanism to explain the general correlation between DNA methylation and gene repression.

Genetic mapping over a number of years had identified a portion of the human X chromosome as being critical for imparting X inactivation. Molecular cloning studies of this X-inactivation center led to the discovery of the *Xist* gene (Willard et al. 1993), an ~17-kb noncoding RNA that was expressed only on the inactive X chromosome. The mouse version of *Xist* was surprisingly homologous in structure and sequence and held the promise of being an excellent model system to dissect the way in which this RNA functioned to repress most of the X chromosome.

Two notable findings were described in *Neurospora* (Selker et al. 1993). First, it was shown that cytosine DNA methylation was not limited to CpG dinucleotides but could occur in seemingly any DNA context. Second was the amazing description of the phenomenon of repeat-induced point mutation (RIP). Sequences become "RIP'd" when there is a sequence duplication (linked or unlinked) in a haploid genome and the genome is put through the sexual cycle via conjugation. Two events occur: Both copies of the duplicated DNA pick up G:C → A:T mutations, and DNA within a few hundred base pairs of the RIP'd sequences becomes methylated. This double attack on the genome is quite efficient—50% of unlinked loci succumb to RIP, whereas tightly linked loci approach 100%—and readily abolishes gene function.

The *brown* gene in *Drosophila*, when translocated near heterochromatin, displays dominant PEV; the translocated copy can cause repression of the wild-type copy. In searching for enhancers and suppressors of this *trans*-inactivation phenomenon, Henikoff discovered that duplication of the gene located near heterochromatin *increased* the level of repression on the normal copy (Martin-Morris et al. 1993). Although the mechanism underlying this event remained mysterious, it was postulated that the phenomenon might be similar to RIP in *Neurospora*, although it had to occur in the absence of DNA methylation, which does not occur in *Drosophila*.

Paul Schedl elucidated the concept of chromosomal "boundary elements" (Vazquez et al. 1993). The first were located on either side of the "puff" region at a heat shock locus in *Drosophila* and were defined by their unusual chromatin structure—an ~300-bp nuclease-resistant core bordered by nuclease hypersensitive sites. It was postulated that such elements separated chromatin domains along the chromosome. Two in vivo assays supported this hypothesis: (1) When bordering either side of a reported gene, boundary elements effectively eliminated chromosomal position effects when the construct was inserted randomly throughout the genome. (2) The boundary element was also defined by its ability to block enhancer function. When inserted between a gene promoter and its enhancer, the boundary element blocked the gene's expression. Although not as well defined, the concept of boundary elements was also developing in other organisms, especially at the globin locus in mammals (Clark et al. 1993).

Budding yeast shone the light on a mechanistic inroad to chromatin-related epigenetic phenomena. It had already been established that the silencers at the silent mating-type loci were sites for several DNA-binding proteins. Their binding appeared to be context-dependent, as exemplified by the Rap1 protein, which not only was important in silencing, but also bound upstream of a number of genes to activate transcription (for review, see Laurenson and Rine 1992).

Over the years, numerous links had been made between DNA replication and transcriptionally quiescent regions of the genome. The inactive X chromosome, heterochromatin, and silenced imprinted loci had all been reported to replicate late in S phase relative to transcriptionally active regions of the genome. In addition, it had been shown that the establishment of silencing at the silent mating-type loci required passage through S phase, suggesting that silent chromatin had to be built on newly repli-

cated DNA. Thus, there was great interest when one of the silencers was found to be an origin of DNA replication, and its origin activity could not be separated from silencing function (Fox et al. 1993). Furthermore, mutants in the recently identified origin recognition complex (ORC) were found to cripple silencing (Bell et al. 1993; Fox et al. 1993).

The discovery that telomeres in *Saccharomyces cerevisiae* exerted PEV, just like that seen in *Drosophila*, brought another entrée into dissecting heterochromatic structure and its influence on gene expression. Reporter genes inserted near telomeres give variegated expression in a colony. The repressed state is dependent on many of the same genes (*SIR2*, *SIR3*, *SIR4*) as those required for silencing at the silent mating-type loci. Several key aspects about the silent chromatin structure and the regulation of the variegated expression were described. It is worth noting that heterochromatin is defined cytologically as condensed chromatin, but silent chromatin in *S. cerevisiae* has never been visualized in this way. Nevertheless, because of similarities to PEV in *Drosophila*, there was enthusiasm to consider silent chromatin in yeast to be a functional equivalent of heterochromatin (described in Weintraub 1993).

From the yeast studies, a number of fundamental concepts began to come to light. First, the importance of histone H3 and H4 became evident. In particular, the amino-terminal tail of histones H3 and H4 appeared to be directly involved in the formation of silent heterochromatin (Thompson et al. 1993). Specific mutants in the tails of these histones alleviated or crippled silencing and led to the notion that both the net charge of the residues on the tails and specific residues within the tails contributed to silencing. In addition, these early days of chromatin immunoprecipitation (ChIP) demonstrated that the lysines in the amino-terminal tail of histone H4 were hypoacetylated in regions of silent chromatin relative to the rest of the genome. Moreover, one of the histone mutants identified histone H4 K16, which could be acetylated, as critical for forming silent chromatin.

Telomeres appeared to provide the simplest system in which to develop an understanding of how Sir proteins mediated silencing. The concept of recruiting silencing proteins was being developed. Briefly, the telomeric DNA-binding protein, Rap1, was found to interact with Sir3 and Sir4 by two-hybrid methods (described in Palladino et al. 1993). Thus, Rap1 could "recruit" these Sir proteins to the telomeric region of the genome. There was evidence that Sir3 and Sir4 could bind to one another, and most importantly, Sir3 and perhaps Sir4 interacted with the tails of histones H3 and H4 (Thompson et al. 1993). Furthermore, overexpression for Sir3 caused it to "spread" inward along the chromatin fiber from the telomere, suggesting that it was a limiting component of silent chromatin and could "polymerize" along the chromatin (Renauld et al. 1993). Taken together, there appeared to be a large interaction network important for silencing—the Sir proteins initiated assembly at telomeric DNA, due to their interaction with Rap1, and then polymerized from the telomere along the chromatin fiber, presumably by binding to the tails of histones H3 and H4.

Switching between transcriptional states in variegated telomeric expression appeared to be the result of a competition between silent and active gene expression (Aparicio and Gottschling 1994; described in Weintraub 1993). If the transcriptional activator for a telomeric gene was deleted, the gene's basal transcriptional machinery was insufficient for expression and the gene was constitutively silenced. Conversely, overexpression of the activator caused the telomeric gene to be expressed continuously—the gene was never silenced. In the absence of *SIR3* (or *SIR2* or *SIR4*), basal gene expression was sufficient, whereas increased dosage of *SIR3* increased the fraction of cells that were silenced. Although a transcriptional activator could overcome silencing throughout the cell cycle, it was most effective when cells were arrested in S phase, presumably when chromatin was being replicated and, hence, most susceptible to competition. Somewhat surprisingly, cells arrested in G_2/M also could be easily switched, suggesting that silent chromatin had not yet been fully assembled by this time.

Silent chromatin in yeast was shown to be recalcitrant to nucleases and DNA modification enzymes, suggesting that the underlying DNA was much less accessible relative to most of the genome (described in Thompson et al. 1993).

It also appeared that there was a hierarchy of silencing within the yeast genome: The telomeres were the most sensitive to perturbation, *HML* was next, and *HMR* was the least sensitive. In fact, when the *SIR1* gene was mutated, the normally completely silenced *HM* loci displayed variegated expression (Pillus and Rine 1989).

Finally, Sir3 and Sir4 were localized to the nuclear periphery, as were the telomeres. It was proposed that the nucleus was organized such that the nuclear envelope provided a special environment for silencing (Palladino et al. 1993).

Schizosaccharomyces pombe also has silent mating cassettes that were suspected to behave similarly to those in *S. cerevisiae*. However, in *S. pombe*, there was an added twist to the story of mating-type switching. In an elegant set of experiments, Amar Klar proposed how a "mark" is

imprinted on one strand of DNA in a cell (Klar and Bonaduce 1993). The mark is manifested, after two cell divisions in one of the four granddaughter cells, as a double-stranded break that facilitates mating-type switching. This yeast does not have any known DNA modifications (methylation, etc.), hence, a different type of mark was postulated to be left on the DNA strand.

The topic of the 59th Symposium was "The Molecular Genetics of Cancer." The concept of epigenetic regulation in oncogenesis had begun to develop after the idea of tumor suppressor genes became established. There had been a couple of studies supporting such a notion, but an interesting twist to the story came in studies of Beckwith-Wiedemann syndrome and Wilms' tumor patients. Mutations in both types of patients had been mapped to a locus that included the imprinted *H19-IGF2* genes. Feinberg et al. (1994) discovered "loss of imprinting" (LOI) for these genes in affected patients—the maternal locus lost its imprint, *H19* was repressed, and *IGF2* was expressed. Thus LOI, which in principle could occur elsewhere in the genome, could cause either biallelic expression and/or extinction of genes critical in oncogenesis.

In the couple of years leading up to the 63rd Symposium on "Mechanisms of Transcription," several important developments occurred that would affect the molecular understanding of several epigenetic phenomena. Histone-modifying enzymes were identified—specifically, histone acetylases and deacetylases. Some of these enzymes played critical roles in regulating gene expression and provided an entry into gene products that directly affected PEV and silencing. The tip of this iceberg was presented at the Symposium (see Losick 1998). Molecular dissection of the Sir3 and Sir4 silencing proteins in yeast revealed the polyvalent nature of their interactions and revealed how the network of interactions between all the Sir proteins, the histones, and various DNA-binding factors set up silent chromatin. In addition, the molecular details of how various loci (telomeres, the rDNA, *HM* loci, and double-stranded breaks) could compete for the limited supply of Sir proteins were shown. By crippling the ability of a specific locus to recruit silencing factors, Sir protein levels were increased at the other loci (Cockell et al. 1998). This provided direct evidence that principles of mass action were at work and that silencing at one locus could affect the epigenetic silencing at other loci—an idea originally put forth in studies on PEV in *Drosophila*, but not yet tested (Locke et al. 1988).

Another finding explained how DNA methylation could regulate gene expression through chromatin. This came with the identification of protein complexes composed of MeCP2, which bind both methylated DNA and histone deacetylases (Wade et al. 1998). Methylated DNA could serve as a point of recruiting deacetylases to a locus and thus facilitate silencing of nearby genes.

The concept of boundary elements was extended from *Drosophila* to mammals, with clear evidence provided at the β-globin locus, thus indicating that chromatin boundaries were indeed likely conserved in metazoans and perhaps all eukaryotes (Bell et al. 1998).

The 64th Symposium on "Signaling and Gene Expression in the Immune System" provided evidence about how monoallelic expression arose, and that it might be more widespread than previously thought. Monoallelic expression at the immunoglobulin loci had been obvious in lymphocytes for some time—it guaranteed the production of a single receptor type per lymphoid cell (Mostoslavsky et al. 1999). The allele to be expressed was chosen early in development, apparently at random: Both alleles began in a repressed state, but over time one became demethylated. It was unclear how a single allele was chosen, but the phenomenon appeared at other loci, too, where the necessity of monoallelism was not obvious. For instance, only one allele of genes encoding the cytokines IL-2 and IL-4 was expressed (Pannetier et al. 1999).

The most significant epigenetics-related talk at the 65th Symposium concerned the discovery that the Sir2 protein was a histone deacetylase (Imai et al. 2000). This was the only Sir protein that had clear homologs in all other eukaryotes and that regulated PEV. It seemed to be the enzyme primarily responsible for removing acetyl moieties from histones in silent chromatin. Furthermore, because it was an NAD-dependent enzyme, it linked the regulation of silencing (heterochromatin) to cellular physiology.

The 68th Symposium on "The Genome of *Homo sapiens*" was an important landmark in genetics, and although there is still much genetic work to be done, the complete sequencing of this and other genomes signified that it was time to move "above genetics"—a literal meaning of epigenetics.

This historical account highlights several themes shared with many other areas of research. First, it demonstrates the episodic nature of advances in epigenetics. Second, as molecular mechanisms underlying epigenetic phenomena began to be understood, it made it easier to connect epigenetics to biological regulation in general. Third, it showed that people whom we now consider to be scientific luminaries had made these connections early on—it just took a while for most others to "see" the obvious.

3 The 69th Symposium

A few general principles have been identified over the years that are common to all epigenetic phenomena, and they serve to guide experimental approaches in the search for a detailed understanding. First, the differences between the two phenotypic states ("OFF" and "ON") always have a corresponding difference in structure at a key regulatory point—form translates into function. Hence, identifying the two distinct structures, the components that compose them, and the compositional differences between them have been the primary tasks. Second, the distinct structures must have the ability to be maintained and perpetuated in a milieu of competing factors and entropic forces. Thus, each structure requires self-reinforcement or positive feedback loops which ensure that it is maintained and propagated over many cellular divisions; in some cases, such as X-chromosome inactivation, this appears to be on the order of a lifetime.

Many of the mechanistic principles defined in the earlier symposia continued to be refined in the 69th Symposium, but there were also new developments. To put these new developments in context, it is important to note that two other discoveries had a major impact on epigenetics. One was the discovery of RNA interference and related RNA-based mechanisms of regulation. The other was the discovery of mechanisms underlying the prion hypothesis. Both of these fields have advanced rapidly in the past decade, with some of the studies contributing to knowledge about chromatin-based epigenetics and others providing new perspectives about heritable transmission of phenotypes.

Many of the accomplishments reported at the Symposium are detailed in the chapters of this book, so I eschew discussing these topics here. However, I will touch upon a few advances that caught my fancy and are not covered within these pages. At the end, I will try to distill the most important concepts I took away from the meeting.

3.1 The Histone Code Hypothesis

In considering histone modifications and their potential information content, there were many discussions about the "histone code hypothesis" (Jenuwein and Allis 2001). Most of those I participated in, or overheard, were informal and rather lively. The proponents of the "code" cite examples such as tri-methylation of histone H3 at K9 and its greater affinity for the HP1 class of heterochromatin proteins (Jenuwein and Allis 2001). Those on the other side cite biochemical and genetic evidence that the net charge on the amino-terminal tail of histone H4, irrespective of which position the charge is at, has dramatic effects on DNA binding or phenotype (Megee et al. 1995; Zheng and Hayes 2003).

Grunstein presented data that included genome-wide analysis of histone acetylation modifications and chromatin-associated proteins using specific antibodies and ChIP-Chip in *S. cerevisiae* (Millar et al. 2004). His focus was on the epigenetic switch associated with H4K16 acetylation for binding, or not binding, particular chromatin proteins—thus supporting the histone code hypothesis. Although not discussed, some of his data appeared to support reports from others that for much of the genome, there is no correlation between specific histone modifications and gene expression (i.e., all active genes have the same marks, and these marks are not present on inactive genes) (Schubeler et al. 2004; Dion et al. 2005). Taking all the results together, I suspect that both specific modifications *and* general net charge effects will be used as mechanisms for regulating chromatin structure and gene expression.

3.2 Dynamic Silent Chromatin

I must confess that, on the basis of static images of heterochromatin and the refractory nature of silent chromatin, I was convinced that once established, a heterochromatic state was as solid as granite. Only when it was time for DNA replication would the impervious structure become relaxed. In thinking this way, I foolishly ignored principles of equilibrium dynamics I had learned in undergraduate chemistry. However, these lessons were brought home again by studies of silent chromatin and heterochromatin, where it was shown that silencing proteins of yeast (Sir3), and heterochromatin proteins in mammalian cells (HP1), were in a dynamic equilibrium—proteins were rapidly exchanged between heterochromatin and the soluble compartment—even when the chromatin was in its most impervious state (Cheng and Gartenberg 2000; Cheutin et al. 2003). The realization of its dynamic qualities forced a different view of how an epigenetic chromatin state is maintained and propagated. It suggests that in some systems the epigenetic state can be reversed at any time, not just during DNA replication. Hence, we can infer that mechanisms of reinforcement and propagation for silenced chromatin must function constantly.

Methylation of histones was widely held to be the modification that would indeed impart a "permanent" mark on the chromatin (for review, see Kubicek and Jenuwein 2004). In contrast to all other histone modification (e.g.,