Epigenetics

The somatic cells of the human body contain 20,000 genes. In the more than 200 cell types present in the body, different cell-specific gene sets are transcribed, while the rest of the genome is transcriptionally inactive. During development, as embryonic cells gradually become specialized cells with adult phenotypes, programs of gene expression become more and more restricted. Until recently, it was thought that most regulation of gene expression is coordinated by *cis*-regulatory elements as well as DNA-binding proteins and transcription factors, and that this regulation can occur at any of the steps in gene expression (see Chapter 17). For example, transcriptional regulation, steps in mRNA processing, and other stages of posttranscriptional regulation control the amount of gene product synthesized from a DNA template. However, as we have learned more about genome organization and the regulation of gene expression, it is clear that classical regulatory mechanisms cannot fully explain how some phenotypes arise. For example, monozygotic twins have identical genotypes but often have different phenotypes. In addition, although one allele of each gene is inherited maternally and one is inherited paternally, in some cases, only the maternal or paternal allele is expressed, while the other is transcriptionally silent.

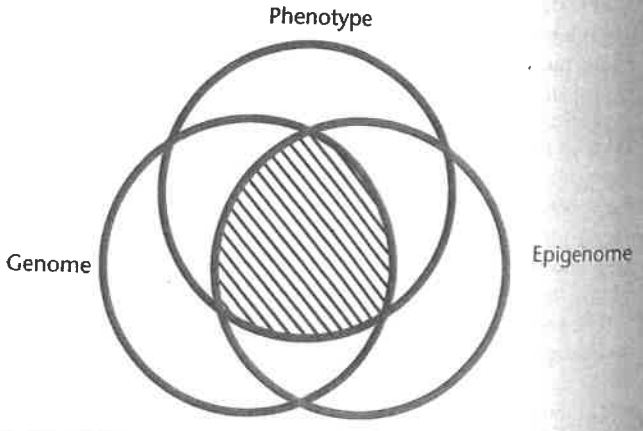
"The newly emerging field of epigenetics is providing us with a basis for understanding how heritable changes other than those in DNA sequence can influence phenotypic variation."

The newly emerging field of epigenetics is providing us with a molecular basis for understanding how heritable changes other than those in DNA sequence can influence phenotypic variation (**ST Figure 1–1**). These advances greatly extend our understanding of gene regulation and have application in wide-ranging areas including genetic disorders, cancer, and behavior.

An **epigenetic trait** is a stable, mitotically and meiotically heritable phenotype that results from changes in the pattern of gene expression without alterations of the DNA sequence. **Epigenetics** is the study of the ways in which chemical modifications to DNA and histones alter cell- and tissue-specific patterns of gene expression. These reversible modifications of DNA and chromatin structure mediate the interaction of the genome with a variety of environmental factors and generate changes in the patterns of gene expression in response to these factors. The **epigenome** refers to the pattern of epigenetic modifications present in a cell at a given time. During its life span, an organism has one genome, which can be modified in diverse cell types at different times to produce many epigenomes.

Current research efforts are focused on several aspects of epigenetics: how an epigenome arises in developing and differentiated cells and how these epigenomes are transmitted via mitosis and meiosis, making them heritable traits. In addition, because epigenetically controlled alterations to the genome are associated with common diseases such as cancer, diabetes, and asthma, efforts are also directed toward developing drugs that can modify or reverse disease-associated epigenetic changes in cells.

Here we will focus on how epigenetics is associated with some heritable genetic disorders, cancer, and environment-genome interactions. Because epigenetic changes are potentially reversible, we will also examine how knowledge of molecular mechanisms of epigenetics is being used to develop drugs and treatments for human diseases.



ST FIGURE 1–1 The phenotype of an organism is the product of interactions between the genome and the epigenome. The genome is constant from fertilization throughout life, but cells, tissues, and the organism develop different epigenomes as a result of epigenetic reprogramming of gene activity in response to environmental stimuli. These reprogramming events lead to phenotypic changes through the life cycle.

BOX 1
The Beginning of Epigenetics

C. H. Waddington coined the term *epigenetics* in the 1940s to describe how environmental influences on developmental events can affect the phenotype of the adult. He showed that by altering environmental conditions during development, organisms with identical genotypes can develop different phenotypes. Using *Drosophila melanogaster,* Waddington found that wing vein patterns could be altered by administering heat shocks during pupal development. Offspring of flies with these environmentally induced changes showed the alternative phenotype without the need for continued environmental stimulus. He called this phenomenon "genetic assimilation." In other words, interactions between the environment and the genome during certain stages of development produced heritable phenotypic changes.

In the 1970s, Holliday and Pugh extended the definition of epigenetics by proposing that changes in the program of gene expression during development depend on the methylation of specific bases in DNA, and that altering methylation patterns affects the resulting phenotype. Waddington's pioneering work, the methylation model of Holliday and Pugh, and the discovery that expression of genes from both the maternal and paternal genomes is required for normal development all helped set the stage for the birth of epigenetics and epigenomics as fields of scientific research.

Epigenetic Alterations to the Genome

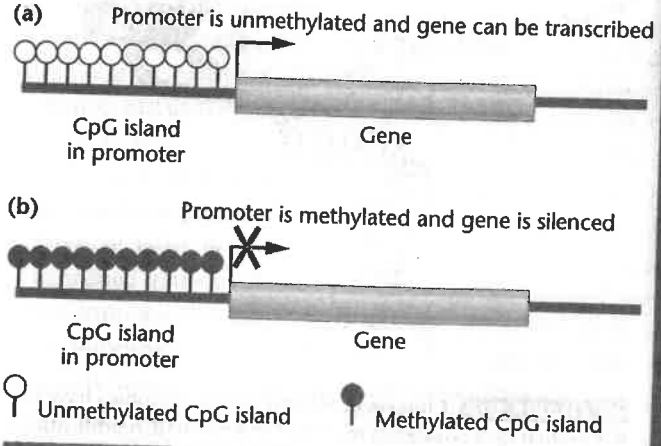
Unlike the genome, which is identical in all cell types of an organism, the epigenome is cell-type specific and changes throughout the life cycle in response to environmental cues. Like the genome, the epigenome can be transmitted to daughter cells by mitosis and to future generations by meiosis. In the following sections, we will examine mechanisms of epigenetic changes and their role in imprinting, cancer, and environment–genome interactions, providing a snapshot of the many roles played by this recently discovered mechanism of gene regulation.

There are three major epigenetic mechanisms: (1) reversible modification of DNA by the addition or removal of methyl groups; (2) chromatin remodeling by the addition or removal of chemical groups to histone proteins; and (3) regulation of gene expression by noncoding RNA molecules.

DNA Methylation

In mammals, DNA methylation takes place after replication and during cell differentiation. This process involves the addition of a methyl group ($-CH_3$) to cytosine, a reaction catalyzed by a family of enzymes called DNA methyltransferases (DNMTs). Methylation takes place almost exclusively on cytosine bases located adjacent to a guanine base, a combination called a CpG dinucleotide. Many of these dinucleotides are clustered in regions called CpG islands, which are located in and near promoter sequences adjacent to genes (**ST Figure 1–2**). CpG islands and promoters adjacent to essential genes (housekeeping genes) and cell-specific genes are unmethylated, making them available for transcription. Genes with adjacent methylated CpG islands and methylated promoters are transcriptionally silenced. The methyl groups in CpG dinucleotides occupy the major groove of DNA and silence genes by blocking the binding of transcription factors necessary to form transcription complexes.

Methylation of CpG islands is a normal process during development. For example, inactivation of one of the X chromosomes in mammalian females is associated with DNA methylation of these groups (see Chapter 7 for a detailed explanation of X-inactivation). In addition, CpG methylation during development creates a parent-specific pattern of gene transcription called genomic imprinting. Abnormal patterns of DNA methylation are associated with specific human diseases, including cancer.

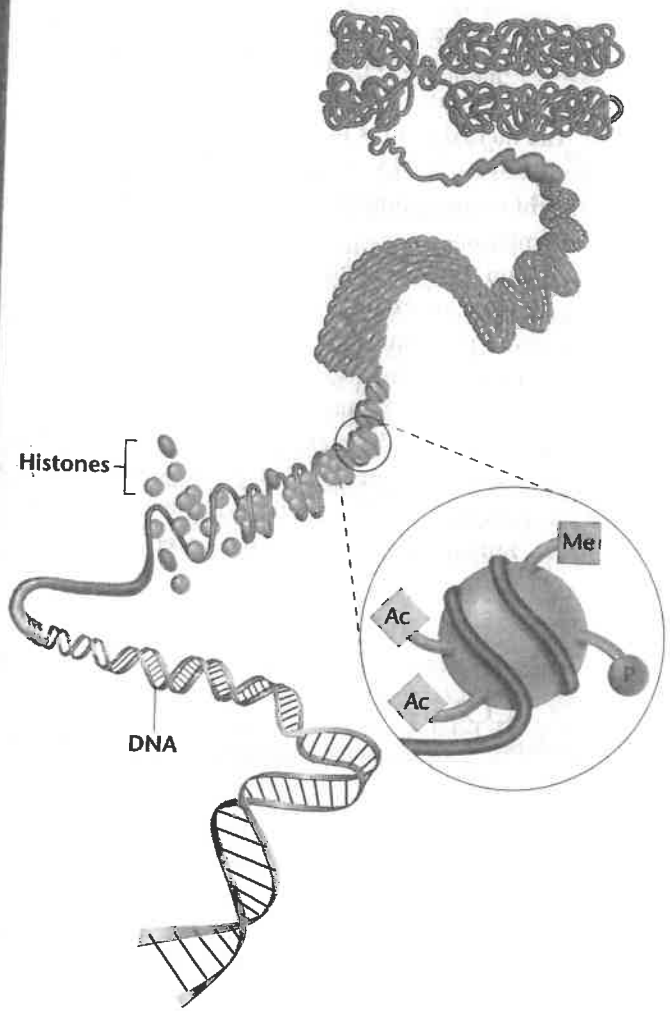


ST FIGURE 1-2 Methylation patterns of CpG dinucleotides in promoters control activity of adjacent genes.

Histone Modification and Chromatin Remodeling

In addition to DNA methylation, histone modification is an important epigenetic mechanism of gene regulation. Recall that chromatin is a dynamic structure composed of DNA wound around a core of 8 histone proteins to form nucleosomes. The N-terminal region of each histone extends beyond the nucleosome, and the amino acids in these tails can be covalently modified in several ways, including the addition of acetyl, methyl, and phosphate groups **(ST Figure 1–3)**. Several sets of proteins are involved in modifying histones. These include proteins that add chemical groups to histones ("writers"), proteins that interpret those modifications ("readers"), and proteins that remove those chemical groups ("erasers"). Some of the proteins involved in histone modification are shown in **ST Table 1.1**.

These modifications alter the structure of chromatin, making genes on nucleosomes accessible or inaccessible for transcription. Histone acetylation, for example, makes genes on these modified nucleosomes available for transcription **[ST Figure 1–4(a)]**. This modification is reversible, and removing (erasing) acetyl groups changes the chromatin configuration from an "open" to a "closed" configuration, silencing genes by making them unavailable for transcription **[ST Figure 1–4(b)]**.



ST FIGURE 1–3 Clusters of histones in nucleosomes have their N-terminal tails covalently modified in epigenetic modifications that alter patterns of gene expression. Ac = acetyl groups, Me = methyl groups, P = phosphate groups.

ST TABLE 1.1 Epigenetic Chromatin Modifier Proteins

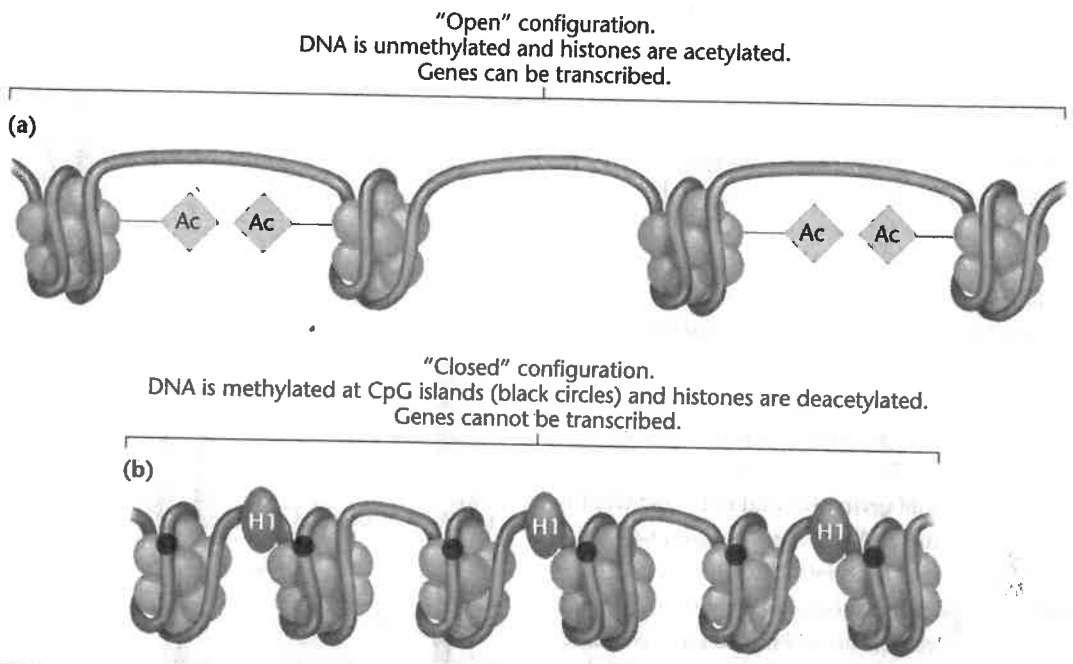
Type	Number Identified
Writers	78
Protein methyltransferases	
Histone acetyltransferases	
Readers	156
Tudor domain-containing proteins	
MBT domain-containing proteins	
Chromodomain-containing proteins	
Erasers	42
Histone deacetylases	
Lysine demethylases	

We are learning that specific combinations of histone modifications control the transcriptional status of a chromatin region. For example, whether or not lysine 9 on histone H3 will be methylated is controlled by modifications made elsewhere on this protein. On one hand, if serine 10 is phosphorylated, methylation of lysine 9 is inhibited. On the other hand, if lysine 14 is deacetylated, methylation of lysine 9 is facilitated. Many combinations of histone modifications are possible, and the sum of the complex patterns and interactions of histone modifications that alter chromatin organization and gene expression is called the **histone code**. Combinations of these changes allow differentiated cells to carry out cell-specific patterns of gene transcription and to respond to external signals that modify these patterns without any changes in DNA sequence.

MicroRNAs and Long Noncoding RNAs

In addition to messenger RNA (mRNA), genome transcription produces several classes of noncoding RNAs. Two of these, microRNAs (miRNAs) and long noncoding RNAs (lncRNAs), play important roles in epigenetic regulatory networks. miRNAs are involved in controlling pattern formation in developing embryos as well as cell signaling.

miRNAs are transcribed as precursor molecules about 70–100 nucleotides long, containing a double-stranded stem-loop and single-stranded regions. Processing removes the single-stranded regions, and the loops move to the cytoplasm where they are altered further. The resulting double-stranded RNA is incorporated into a protein complex where one RNA strand is removed and degraded, forming a mature RNA-induced silencing complex (RISC) containing the remaining



ST FIGURE 1–4 Epigenetic modifications to the genome alter the spacing of nucleosomes. In the open configuration, nucleosome positions are shifted, CpGs are unmethylated, and the genes on the DNA are available for transcription. In the closed configuration, DNA is tightly wound onto the nucleosomes, CpGs are methylated, and genes on the DNA are unavailable for transcription.

single miRNA strand. RISCs act as repressors of gene expression by binding to and destroying target mRNA molecules carrying sequences complementary to the RISC miRNA. mRNAs partially complementary to the RISC miRNA are modified, making these mRNAs less likely to be translated by ribosomes, resulting in downregulation of gene expression.

In addition to forming RISC complexes, miRNAs also associate with a different set of proteins to form RNA-induced transcriptional silencing (RITS) complexes. RITS complexes reversibly convert euchromatic chromosome regions into facultative heterochromatin, silencing the genes located within these newly created heterochromatic regions.

Long noncoding RNAs (lncRNAs) share many properties in common with mRNAs; they often have 5′ caps, 3′ poly-A tails, and are spliced. What distinguishes lncRNAs from coding (mRNA) transcripts is the lack of an extended open reading frame that codes for the insertion of amino acids into a polypeptide.

lncRNAs are found in the nucleus and the cytoplasm and through a variety of mechanisms are involved in regulating almost every stage of gene expression. As epigenetic modulators, lncRNAs bind to chromatin-modifying enzymes and direct their activity to specific regions of the genome. At these sites, the lncRNAs direct chromatin modification, altering the pattern of gene expression.

In summary, epigenetic modifications alter chromatin structure by several mechanisms including DNA methylation, histone acetylation, and action of RNAs, without changing the sequence of DNA. These epigenetic changes create an epigenome that, in turn, can regulate normal development and generate changes in gene expression as a response to environmental signals.

Epigenetics and Imprinting

Mammals inherit a maternal and a paternal copy of each autosomal gene, and usually both copies of these genes are expressed in the offspring. Imprinting is an epigenetically regulated process in which genes are expressed in a parent-of-origin pattern; that is, certain genes show expression of only the maternal allele or the paternal allele. Parent-specific patterns of allele expression are laid down in germ-line cells during gamete formation and are inherited via mitosis in somatic cells.

Differential methylation of CpG-rich regions and methylation of promoter sequences produce allele-specific imprinting and subsequent gene silencing. Once a gene has been imprinted, it remains transcriptionally silent during embryogenesis and development. Most imprinted genes direct aspects of growth during prenatal development. For example, in mice, genes on the maternal X chromosome are expressed in the placenta, while genes on the paternal X chromosome are silenced. Having only one functional allele makes imprinted genes highly susceptible to the deleterious effects of mutations. Because imprinted genes are clustered at a limited number of sites in the genome, mutation in one gene can have an impact on the function of adjacent imprinted genes, amplifying its impact on the phenotype.

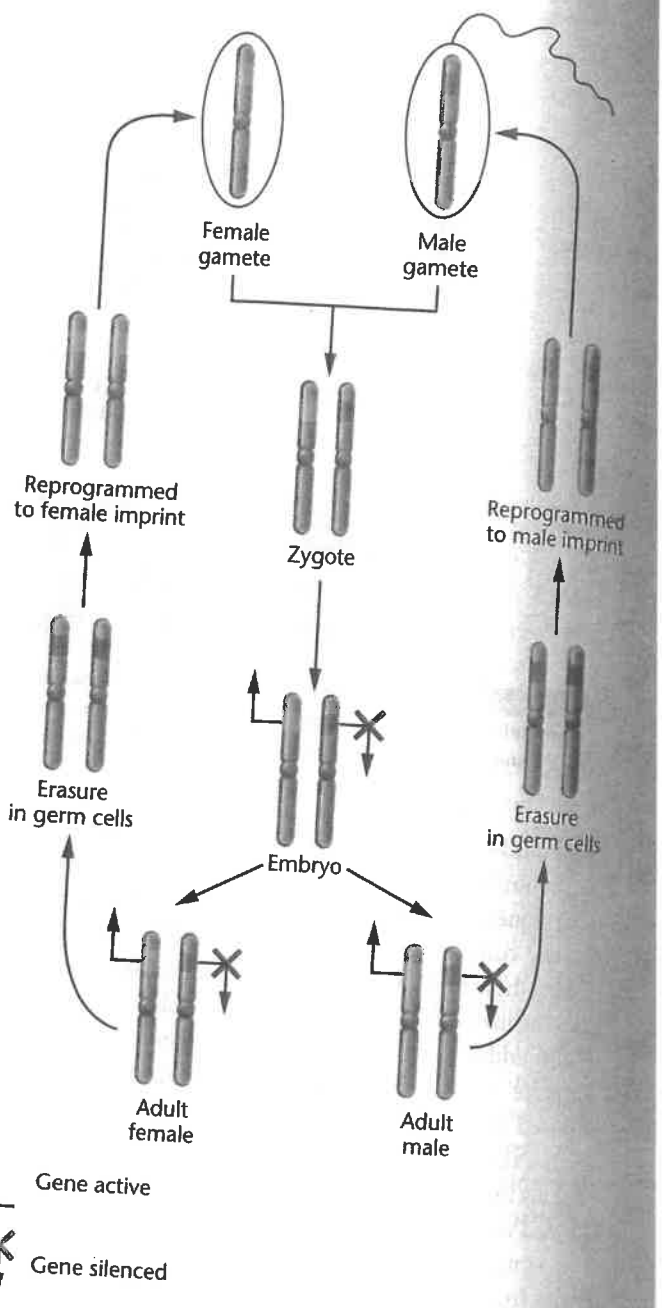
Mutations in imprinted genes can arise by changes in the DNA sequence or by dysfunctional epigenetic changes, called **epimutations**, both of which are heritable changes in the activity of a gene.

At fertilization, the mammalian embryo receives a maternal and a paternal set of chromosomes; the maternal chromosome set carries female imprints, and the paternal set contains male imprints. When gamete formation begins in female germ cells, both the maternal and paternal imprints are erased and reprogrammed to a female imprint that is transmitted to the next generation through the egg (**ST Figure 1–5**). Similarly, in male germ cells, the paternal and maternal chromosomes are reprogrammed by methylation to become a male imprinted set. Reprogramming occurs at two stages: in the parental germ cells and in the developing embryo. In stage one, erasure by demethylation and reprogramming by remethylation lay down a male- or female-specific imprinting pattern in germ cells of the parent. In stage two, large-scale demethylation occurs in the embryo sometime before the 16-cell stage of development. After implantation, differential genomic remethylation recalibrates which maternal alleles and which paternal alleles will be inactivated. It is important to remember that imprinted alleles remain transcriptionally silent in all cells, while genes silenced by epigenetic methylation can be reactivated by external signals.

Most human disorders caused by dysfunctional imprinting originate during fetal growth and development. Imprinting defects cause Prader–Willi syndrome, Angelman syndrome, Beckwith–Wiedemann syndrome, and several other diseases (**ST Table 1.2**). However, given the number of candidate genes and the possibility that additional imprinted genes remain to be discovered, the overall number of imprinting-related genetic disorders may be much higher.

In humans, most known imprinted genes encode growth factors or other growth-regulating genes. An autosomal dominant disorder of imprinting, Beckwith–Wiedemann syndrome (BWS), offers insight into how disruptions of epigenetically imprinted genes lead to an abnormal phenotype. BWS is a prenatal overgrowth disorder with abdominal wall defects, enlarged organs, large birth weight, and predisposition to cancer. BWS is not caused by mutation, nor is it associated with any chromosomal aberration. Instead it is a disorder of imprinting and is caused by abnormal methylation patterns and resulting altered patterns of gene expression.

Genes linked to BWS are located in a cluster of imprinted genes on the short arm of chromosome 11. All genes in this cluster regulate growth during prenatal development. Two important genes in this cluster are insulin-like growth factor 2 (*IGF2*), whose encoded protein plays an important role in growth and development, and *H19*. The H19 gene product is a noncoding RNA which may act as a tumor suppressor, but whose exact function remains unknown. Normally, the



ST FIGURE 1–5 Imprinting patterns are reprogrammed each generation during gamete formation. A second round of epigenetic reprogramming occurs during early embryonic development.

paternal allele of *IGF2* is expressed, and the maternal allele is imprinted and silenced. In the case of *H19*, the maternal allele is expressed, and the paternal allele is imprinted and silenced.

In BWS, the maternal *IGF2* allele is not imprinted. As a result, the maternal and paternal alleles are both transcriptionally active, resulting in the overgrowth of tissues that are characteristic of this disease.

The known number of imprinted genes represents only a small fraction (less than 1 percent) of the mammalian genome, but they play major roles in regulating growth during prenatal development. Because they act so early in

ST TABLE 1.2 **Some Imprinting Disorders in Humans**

Disorder	Locus
Albright hereditary osteodystrophy	20q13
Angelman syndrome	15q11-q15
Beckwith-Wiedemann syndrome	11p15
Prader-Willi syndrome	15q11-q15
Silver-Russell syndrome	Chromosome 7
Uniparental disomy 14	Chromosome 14

life, external or internal factors that disturb the epigenetic pattern of imprinting or the expression of imprinted genes can have serious phenotypic consequences.

Assisted Reproductive Technologies (ART) and Imprinting Defects

Several methods collectively called ART are available to help treat infertility. The use of two such methods, *in vitro* fertilization (IVF) and intracytoplasmic sperm injection (ICSI), increases the risk of two epigenetic disorders, Angelman syndrome and BWS.

Several studies have shown that children born after IVF are at risk for low or very low birth weight, a condition that may result from abnormal imprinting. The use of IVF and other ART procedures has also been associated with a six to nine-fold increased risk of BWS. In one study, more than 90 percent of children born with BWS after ART had imprinting defects. Because imprinting disorders are uncommon (BWS occurs with a frequency of about 1 in 15,000 births following normal conception), large-scale and longitudinal studies will be needed to establish a causal relationship among imprinting abnormalities, growth disorders, and ART.

Epigenetics and Cancer

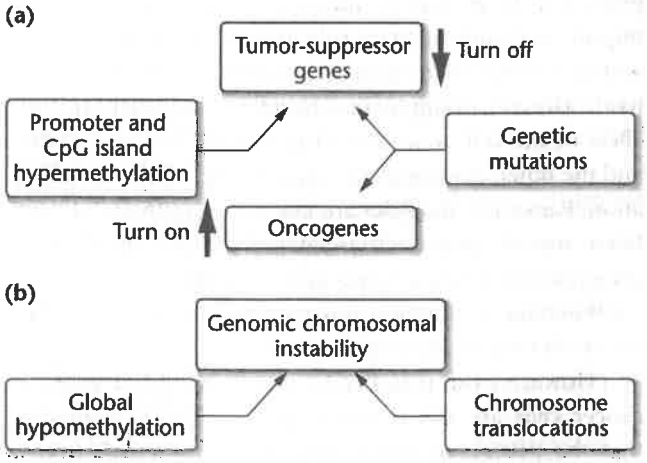
For some complex diseases, there are strong links to environmental factors such as the association between smoking and lung cancer. The discovery that epigenetics mediates changing patterns of gene expression in response to environmental signals offers a new and more direct approach to understanding the interactions between the genome and the environment in cancer and other diseases.

Following the discovery of cancer-associated genes, including those that promote (proto-oncogenes) or inhibit (tumor-suppressor genes) cell division, research into the genetics of cancer focused mainly on mutant alleles of genes involved in regulation of the cell cycle. Until recently, the conventional view has been that cancer is clonal in origin and begins in a single cell that has accumulated a suite of dominant and recessive mutations that allow it to escape control of the cell cycle. Subsequent mutations allow cells of the tumor to become metastatic, spreading the cancer to other locations in the body where new malignant tumors appear.

Converging lines of evidence are now clarifying the role of epigenetic changes in the initiation and maintenance of malignancy. These findings help researchers understand properties of cancer cells that are difficult to explain by the action of mutant alleles alone. Evidence for the role of epigenetic changes in cancer now challenges the conventional paradigm for the origin of cancer and establishes epigenomic changes as a major pathway for the formation and spread of malignant cells.

The relationship between epigenetics and cancer was first noted in the 1980s by Feinberg and Vogelstein, who observed that colon cancer cells had much lower levels of methylation than normal cells derived from the same tissue. Subsequent research by many investigators showed that during the transformation to a cancerous state, cells undergo complex changes in DNA methylation patterns. These changes fall into three categories: hypermethylation, hypomethylation, and loss of imprinting. Global genomic hypomethylation is a property of all cancers examined to date. In addition, selective hypermethylation and gene silencing are also properties of cancer cells. Cancer is now viewed as a disease that involves both epigenetic *and* genetic changes that lead to alterations in gene expression and the development of malignancy (**ST Figure 1–6**).

DNA hypomethylation reverses the inactivation of genes, leading to unrestricted transcription of many gene sets including oncogenes. It also relaxes control over imprinted genes, causing cells to acquire new growth properties. Hypomethylation of repetitive DNA sequences in heterochromatic regions increases chromosome rearrangements and changes in chromosome number, both of which are characteristic of cancer cells. In addition,



ST FIGURE 1–6 **The development and maintenance of malignant growth in cancer involves gene mutations, hypomethylation, hypermethylation, overexpression of oncogenes, noncoding RNAs, and the silencing of tumor-suppressor genes.**

ST TABLE 1.3 Some Cancer-Related Genes Inactivated by Hypermethylation in Human Cancers

Gene	Locus	Function	Related Cancers
BRCA1	17q21	DNA repair	Breast, ovarian
APC	5q21	Nucleocytoplasmic signaling	Colorectal, duodenal
MLH1	3p21	DNA repair	Colon, stomach
RB1	13q14	Cell-cycle control point	Retinoblastoma, osteosarcoma
AR	Xq11-12	Nuclear receptor for androgen; transcriptional activator	Prostate
ESR1	6q25	Nuclear receptor for estrogen; transcriptional activator	Breast, colorectal

hypomethylation of repetitive sequences leads to transcriptional activation of transposable DNA sequences such as LINEs and SINEs, further increasing genomic instability.

While widespread hypomethylation is a hallmark of cancer cells, hypermethylation at CpG islands and promoters silences tumor-suppressor genes (**ST Table 1.3**), often in a specific pattern. For example, the breast cancer gene *BRCA1* is hypermethylated and inactivated in breast and ovarian cancer, and hypermethylation of the DNA repair gene *MLH1* is a key step in the development of some forms of colon cancer. Epigenetic studies have revealed a unique pattern of abnormal DNA methylation that defines each type of cancer. Analysis of these patterns, called the CpG island methylator phenotype (CIMP), provides a way to identify tumor types and subtypes, and predict the sites to which the tumor will metastasize.

Epigenetic inactivation of tumor-suppressor genes plays a primary role in tumor formation as well as an important complementary role to mutational changes that accompany the transformation of normal cells into malignant cells. For example, in a bladder cancer cell line, one allele of the cell-cycle control gene *CDKN2A* is mutated, and the other, normal allele is inactivated by hypermethylation. Because both alleles are inactivated (although by different mechanisms), cells are able to escape control of the cell cycle and divide continuously. In many clinical cases, a combination of mutation and epigenetic hypermethylation occurs in familial forms of cancer.

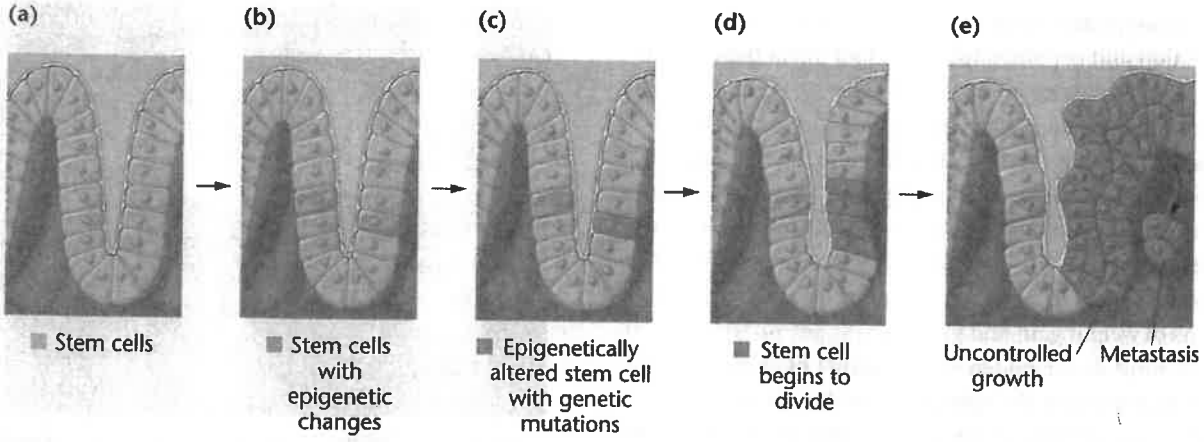
However, the majority of hypermethylated genes in cancer cells are not tumor-suppressor genes, suggesting that the pattern of hypermethylation may result from a widespread deregulation of the methylation process rather than a targeted event.

In addition to altered patterns of methylation, many cancers also have disrupted histone modification profiles. In some cases, mutations in the genes of the histone-modifying proteins such as histone acetyltransferase (HAT) and histone deacetylase (HDAC) are linked to the development of cancers. For example, individuals with Rubinstein–Taybi syndrome inherit a germ-line mutation that produces a dysfunctional HAT and have a greater than 300-fold increased risk of cancer. In other cases, HDAC complexes are selectively recruited to tumor-suppressor genes by mutated, oncogenic DNA binding proteins. Action of the HDAC complexes at these genes converts the chromatin to a closed configuration and inhibits transcription, causing the cell to lose control of the cell cycle.

Mechanisms that cause epigenetic changes in cancer cells are not well understood, partly because they take place very early in the conversion of a normal cell to a cancerous one and partly because by the time the cancer is detected, alterations in the methylation pattern have already occurred. The *MLH1* gene has an important role in genome stability, and silencing this gene by hypermethylation causes instability in repetitive microsatellite sequences, which, in turn, is an important step in the development of colon cancer and several other cancers. In some individuals with colon cancer, the *MLH1* promoter in normal cells of the colon is already silenced by hypermethylation, indicating that this epigenetic event occurs very early in tumor formation, before the development of downstream genetic mutations.

The fact that such changes occur very early in the transformation process has led to the proposal that epigenetic changes leading to cancer may occur within adult stem cells in normal tissue. Three lines of evidence support this idea: (1) epigenetic mechanisms can replace mutations as a way of silencing individual tumor-suppressor genes or activating oncogenes; (2) global hypomethylation may cause genomic instability and the large-scale chromosomal changes that are a characteristic feature of cancer; and (3) epigenetic modifications can silence multiple genes, making them more effective in transforming normal cells into malignant cells than sequential mutations of single genes. A model of cancer based on epigenetic changes in colon stem cells as initial events in carcinogenesis followed by mutational events is shown in **ST Figure 1–7**.

Analysis of data from the Cancer Genome Atlas (TCGA) shows the importance of epigenetic events in cancer. In ovarian cancer, mutations in nine specific genes are predominant, but promoter hypermethylation is observed in 168 genes. These genes are epigenetically silenced, and their reduced expression is linked to the development and maintenance of this cancer. In addition to changing ideas about the origins of cancer, the fact that epigenetic changes are potentially reversible makes it possible to develop new classes of drugs for chemotherapy. The focus of epigenetic therapy is the reactivation of genes silenced by methylation or histone modification, essentially reprogramming the pattern of gene expression in cancer cells. Several epigenetic drugs have



ST FIGURE 1-7 The epigenetic stem cell model proposes that both epigenetic changes and mutations are involved in the origins of cancer.

been approved by the U.S. Food and Drug Administration, and another 12–15 drugs are in clinical trials. One approved drug, Vidaza, is used in the treatment of myelodysplastic syndrome, a precursor to leukemia, and for treatment of acute myeloid leukemia. This drug is an analog of cytidine and is incorporated into DNA during replication during the S phase of the cell cycle. Methylation enzymes (methyltransferases) bind irreversibly to Vidaza, preventing methylation of DNA at many other sites, effectively reducing the amount of methylation in cancer cells. Other drugs that inhibit histone deacetylases (HDAC) have been approved by the FDA for use in epigenetic therapy. Experiments with cancer cell lines indicate that inhibiting HDAC activity results in the re-expression of tumor-suppressor genes. The HDAC inhibitor Zolinza is used to treat some forms of lymphoma.

The development of epigenetic drugs for cancer therapy is still in its infancy. The approved epigenetic drugs are only moderately effective on their own and are best used in combination with other anticancer drugs. To develop more effective drugs, several important questions remain to be answered: What causes cancer cells to respond to certain epigenetic drugs? What combination of epigenetic drugs and conventional anticancer drugs are most effective on specific cancers? Which epigenetic markers will be effective in predicting sensitivity or resistance to newly developed drugs? Further research into the mechanisms and locations of epigenetic genome modification in cancer cells will allow the design of more potent drugs to target epigenetic events as a form of cancer therapy.

Epigenetics and the Environment

Environmental agents including nutrition, exposure to chemicals, and physical factors, such as temperature, can alter gene expression by affecting the epigenome. In humans it is difficult to determine the relative contributions of environmental or learned behavior as factors in altering

BOX 2

What More We Need to Know about Epigenetics and Cancer

The discovery that epigenetic changes may be as important as genetic changes in the origin, maintenance, and metastasis of cancers has opened new avenues of cancer research. Key discoveries about epigenetic mechanisms include the finding of tumor-specific deregulation of genes by altered DNA methylation profiles and histone modifications, the discovery that epigenetic changes in histones or DNA methylation are interconnected, and the recognition that epigenetic changes can affect hundreds of genes in a single cancer cell. These advances were made in the span of a few years, and while it is clear that epigenetics plays a key role in cancer, many questions remain to be answered before we can draw conclusions about the relative contributions of genetics and epigenetics to the development of cancer. Some of these questions are as follows:

- Is global hypomethylation in cancer cells a cause or an effect of the malignant condition?
- Do these changes arise primarily in stem cells or in differentiated cells?
- Once methylation alterations begin, what triggers hypermethylation in cancer cells?
- Is hypermethylation a process that targets certain gene classes, or is it a random event?
- Can we develop drugs that target cancer cells and reverse tumor-specific epigenetic changes?
- Can we target specific genes for reactivation, while leaving others inactive?

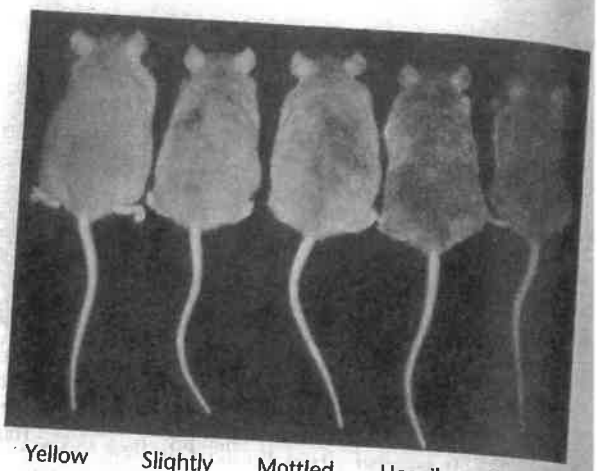
the epigenome, but there is indirect evidence that changes in nutrition and exposure to agents that affect the developing fetus can have detrimental effects during adulthood.

Women who were pregnant during the 1944–1945 famine in the Netherlands had children with increased risks for obesity, diabetes, and coronary heart disease. In addition, as adults, these individuals had significantly increased liability for schizophrenia and other neuropsychiatric disorders. Members of the F_2 generation also had abnormal patterns of weight gain and growth.

The most direct evidence for the role of environmental factors in modifying the epigenome comes from studies in experimental animals. A low-protein diet fed to pregnant rats results in permanent changes in the expression of several genes in both the F_1 and F_2 offspring. Increased expression of these genes is associated with hypomethylation of their promoters. Other evidence indicates that epigenetic changes triggered by this diet modification were gene-specific.

A dramatic example of how epigenome modifications affect the phenotype comes from the study of coat color in mice, where color is controlled by the dominant allele *Agouti* (*A*). In homozygous *AA* mice, the allele is active only during a specific time during hair development, resulting in a yellow band on an otherwise black hair shaft, producing the agouti phenotype. A nonlethal mutant allele (A^{vy}) causes yellow pigment formation along the entire hair shaft, resulting in yellow fur color. This allele is the result of the insertion of a transposable element near the transcription start site of the *Agouti* gene. A promoter element within the transposon is responsible for this change in gene expression. Researchers found that the degree of methylation in the transposon's promoter is related to the amount of yellow pigment deposited in the hair shaft and that the amount of methylation varies from individual to individual. The result is variation in coat color phenotypes even in genetically identical mice (**ST Figure 1–8**). In these mice, coat colors range from yellow (unmethylated promoter) to pseudoagouti (highly methylated promoter). In addition to a gradation in coat color, there is also a gradation in body weight. Yellow mice are more obese than the brown, pseudoagouti mice. Alleles such as A^{vy} that show variable expression from individual to individual in genetically identical strains caused by epigenetic modifications are called *metastable epialleles*. Metastable refers to the changeable nature of the epigenetic modifications, and epiallele refers to the heritability of the epigenetic status of the allele.

To evaluate the role of environmental factors in modifying the epigenome, the diet of pregnant A^{vy} mice was supplemented with methylation precursors, including folic acid, vitamin B_{12}, and choline. In the offspring, coat-color variation was reduced and shifted toward the pseudoagouti (highly methylated) phenotype. The shift in coat color was accompanied by increased methylation of the transposon's promoter. These findings have applications to epigenetic diseases in humans. For example, the risk of colorectal cancer is linked directly to dietary folate deficiency and activity differences in enzymes leading to the synthesis of methyl donors.



ST FIGURE 1–8 Variable expression of yellow phenotype in mice caused by diet-related epigenetic changes in the genome.

In addition to foods that mediate epigenetic changes in gene expression via methylation, it has recently been discovered that some foods such as rice, cabbage, wheat, and potatoes are the source of miRNAs circulating in the blood and serum of humans. Studies on one target gene showed that plant miRNAs downregulate expression of this gene and that this effect is reversible by treatment with anti-miRNAs. This preliminary but highly significant work shows that miRNAs can act across species and even across kingdoms to regulate expression of target genes, and suggests that environmental factors may play a major role in epigenetic regulation of gene expression.

Epigenome Projects

As the role of the epigenome in disease has become increasingly clear, researchers across the globe have formed multidisciplinary projects to map all the epigenetic changes that occur in the normal genome and to study the role of the epigenome in specific diseases.

These include:

- NIH Roadmap Epigenomics Project
- Human Epigenome Project (HEP)
- International Human Epigenome Consortium (IHEC)
- International Cancer Genome Consortium (ICGC)

In conclusion, we will discuss some of these projects and their goals.

The NIH Roadmap Epigenomics Project is focused on how epigenetic mechanisms controlling stem cell differentiation and organ formation generate biological responses to

external and internal stimuli that result in disease. One program of this project is the Human Epigenome Atlas, which catalogs normal human epigenomes to serve as reference standards. The Atlas provides detailed information about epigenomic modifications at specific loci, in different cell types and physiological states, as well as genotypes. These data allow researchers to perform comparative analysis of epigenomic data across genomic regions or entire genomes.

The Human Epigenome Project is a multinational, public/private consortium organized to identify, map, and establish the functional significance of all DNA methylation patterns in the human genome across all the major tissue types in the body. Analysis of these methylation patterns may show that genetic responses to environmental cues mediated by epigenetic changes are a pathway to disease.

The International Human Epigenome Consortium (IHEC) is global program established to determine how the epigenome has altered human populations in response to environmental factors. The consortium will begin by cataloging the epigenomes of 1000 individuals from different populations around the world. The catalog will also include the epigenomes of 250 different cell types.

Although these projects are in the early stages of development, the information already available strongly suggests that we are on the threshold of a new era in genetics, one in which we can study the development of disease at the genomic level, and understand impact of environmental factors on gene expression. The results of these projects may help explain how environmental settings in early life can affect predisposition to adulthood diseases.

Selected Readings and Resources

Journal Articles and Reviews

Burdge, G.C., and Lillicrop, K.A. 2010. Nutrition, epigenetics, and developmental plasticity: Implications for understanding human disease. *Annu. Rev. Nutr.* 30: 315–339.

Mercer, T.R., and Mattick, J.S. 2013. Structure and function of long noncoding RNAs in epigenetic regulation. *Nat Struct. Mol. Biol.* 20: 300–307.

Petronis, A. 2010. Epigenetics as a unifying principle in the aetiology of complex traits and diseases. *Nature* 465: 721–727.

Sandoval, J., and Esteller, M. 2012. Cancer epigenomics: Beyond genomics. *Curr. Opin. Genet. Develop.* 22: 50–55.

Soejima, H., and Higashimoto, K. 2013. Epigenetic and genetic alterations of the imprinting disorder Beckwith-Wiedemann syndrome and related disorders. *J. Hum. Genet.* 58: 402–409.

Weaver, I.C., Cervoni, N., Chanpagne, F.A., D'Allesio, A.C., Sharma, S., Seckl, J.R., Dymov, S., Szyf, M., and Meany, M.J. 2004. Epigenetic programming by maternal behavior. *Nat. Neurosci.* 7: 847–854.

Web Sites

Human Epigenome Project. www.epigenome.org

Human Epigenome Atlas. http://www.genboree.org/epigenomeatlas/index.rhtml

Computational Epigenetics Group. http://www.computational-epigenetics.de

Review Questions

1. What are the major mechanisms of epigenetic genome modification?
2. What parts of the genome are reversibly methylated? How does this affect gene expression?
3. What are the roles of proteins in histone modification?
4. Describe how reversible chemical changes to histones are linked to chromatin modification.
5. What is the histone code?
6. What is the difference between silencing genes by imprinting and silencing by epigenetic modifications?
7. Why are changes in nucleosome spacing important in changing gene expression?
8. How do microRNAs regulate epigenetic mechanisms during development?
9. What is the role of imprinting in human genetic disorders?

Discussion Questions

1. Imprinting disorders do not involve changes in DNA sequence, but only the methylated state of the DNA. Does it seem likely that imprinting disorders could be treated by controlling the maternal environment in some way, perhaps by dietary changes?
2. Should fertility clinics be required by law to disclose that some assisted reproductive technologies (ARTs) can result in epigenetic diseases? How would you and your partner balance the risks of ART with the desire to have a child?
3. How can the role of epigenetics in cancer be reconciled with the idea that cancer is caused by the accumulation of mutations in tumor-suppressor genes and proto-oncogenes?
4. Would the knowledge that plant miRNAs can affect gene expression in your body affect your food choices?