Chapter 4
Gene Expression, Epigenetic Regulation, and Cancer

Shuntele N. Burns

4.1 The Principles of Gene Expression

DNA (deoxyribonucleic acid) serves as a blueprint for the manufacture of different proteins that are critical in coordinating virtually every biological function performed in a cell. Genes are constructed of sequences of DNA nucleotides (each composed of a phosphate group, a five-carbon sugar, and a nitrogenous base) positioned on chromosomes that exist in the nucleus of a cell. Each gene will dictate the order of amino acids needed to make a particular protein, and differences in the sequences of amino acids contribute to the shape and function of the protein. The process by which a gene is turned on to construct a specific protein is called gene expression [1].

Proteins are important molecules that perform a variety of functions. For example, some are structural proteins, while others play enzymatic roles, serve as hormones, or protect the body against infections. Structural proteins provide the body with support; keratin, for instance, is a component of nails and hair, and collagen gives strength to tendons, ligaments, and skin. Enzymes are involved in catalyzing metabolic pathways whereby a chemical reaction can occur within a second, whereas hours may be required to complete the same reaction without enzymes. Some proteins serve as hormones, which control the cellular activities of target organs. Antibodies are plasma proteins that bind with antigens, thereby preventing foreign substances from injuring cells and causing homeostatic disruption. These and other operations make proteins essential to life [2].

Transcription and translation are the two major stages associated with gene expression. Transcription takes place in the nucleus and entails a gene operating as a template for the construction of an RNA molecule. Though most genes are

S.N. Burns (✉)
Department of Biological Sciences, Alabama State University, Montgomery, AL, USA
e-mail: sburns@alasu.edu

M.K. Mishra, K.S. Bishnupuri (eds.), *Epigenetic Advancements in Cancer*,
DOI 10.1007/978-3-319-24951-3_4

transcribed into messenger RNA (mRNA), which contains instructions that will eventually produce proteins, some genes are transcribed into two other classes of RNA: (1) ribosomal RNA (rRNA), a chief component of ribosomes—organelles on which proteins are synthesized and (2) transfer RNA (tRNA)—molecules that carry amino acids, which are subunits of proteins, to ribosomes [3].

In eukaryotes, the production of an mRNA molecule is initiated with the binding of proteins called transcription factors to the promoter, a DNA sequence situated just upstream of a gene. Subsequently, the transcription factors recruit to the promoter the enzyme RNA polymerase [1]. As RNA polymerase travels along the DNA helix associated with the gene, the enzyme unfastens the two strands of DNA and constructs together RNA nucleotides complementary to one of the DNA strands that is serving as a template; that is, when DNA bases adenine (A), thymine (T), cytosine (C), or guanine (G) are located in the DNA template, bases uracil (U), adenine (A), guanine (G), or cytosine (C) are respectively positioned into the RNA molecule. This process eventually produces a single-stranded RNA molecule that has corresponding nucleotides of the DNA sequence for the gene in question. The complementary RNA strand of the gene is referred to as messenger RNA, given that it possesses the "message" or genetic material of the gene which will be expressed [4].

The resulting strand of mRNA is then processed to remove introns, sequences of nucleotides that will not be translated into a sequence of amino acids for a protein. Once introns are deleted, the remaining nucleotide sequences—exons—are spliced together. The resulting mRNA strand is then modified with the addition of a guanine-nucleotide cap at the 5′ end of the strand and the inclusion of a tail consisting of 30–100 adenine-nucleotides (poly-A tail) at the 3′ end. The modified mRNA strand will then leave the nucleus through the nuclear envelope and then enter the cytoplasm, where the process of translation will proceed [5]. Though ordinarily the mRNA strand that is produced after transcription is composed of a sequence of all of the exons of the gene, alternative splicing may occur whereby only certain exons are used. Hence alternative splicing allows for the manufacture of different variations of a protein. For example, white blood cells are able to construct particular antibodies based on the type of antigen that enters the body [2].

The second major stage pertaining to gene expression is called translation and occurs in the cytoplasm. Translation involves the conversion of the sequence of mRNA nucleotides into a sequence of amino acids associated with a protein. Along with mRNA and amino acids, ribosomes and tRNA are some of the other components that play important roles in translation [3].

The ribosome is the organelle on which protein synthesis takes place and is constructed of two subunits, each consisting of rRNA enfolded around various proteins; the ribosomal ensemble allows for the binding of mRNA and its passage through the ribosome. A complex series of interactions ensues once mRNA is bound to the ribosome. First, small mRNA sequences called codons are exposed. A codon is three nucleotides long and typically encodes for a particular amino acid. (Sixty-one codons will encode for the 20 different amino acids; therefore, most amino acids are designated by more than one codon. The diversity of codons allows for some degree of protection against mutations that alter the sequence of bases.)

Next, a tRNA molecule physically transports an amino acid to the ribosome, which possesses three binding sites constructed for three incoming tRNA molecules. An amino acid is attached to one end of the incoming tRNA, and a sequence of nucleotides complementary to the codon, identified as the anticodon, is located at the other end of the tRNA molecule. The tRNA anticodon binds to its complementary mRNA codon that is positioned on the ribosome (e.g., codon AUG pairs with anticodon UAC). Once a second tRNA molecule docks at the adjacent codon, this newly-arriving tRNA receives the amino acid from the initial tRNA, thereby forming a peptide bond between the first amino acid and the one already coupled with the second tRNA. The initial tRNA molecule eventually leaves the ribosome once it is freed of its amino acid. Afterwards, the ribosome moves sideways across the strand of mRNA, enabling the next codon to bind to an incoming tRNA. Thus, the chain of amino acids (polypeptide) elongates one amino acid at a time. The ribosome will eventually encounter a codon (UAA, UAG, or UGA) that will not encode for an amino acid but will instead signify the termination of polypeptide synthesis. Afterwards, the polypeptide chain is released into the cytoplasm, where it will conform to the final three-dimensional shape associated with a protein, and the two ribosomal subunits will dissociate. Hence during transcription, a sequence of DNA nucleotide bases determines the sequence of mRNA nucleotide bases, and during translation, the mRNA sequence then directs the order of amino acids bonded to form the desired protein [4].

Though produced in the cytoplasm, proteins can be found in all parts of the cell. Proteins manufactured by ribosomes that float freeing in the cytoplasm remain and operate in the cytoplasm. However, proteins produced by ribosomes that are fastened to the rough endoplasmic reticulum have chemical signals that direct the proteins to their appropriate destinations within the cell, and, interestingly, these sorting signals are also encoded in the genome and are transcribed and translated [5].

Every organism has the capacity to control which genes will be expressed at any time. Organisms have to continually turn on and off genes in response to internal and external environmental signals. The regulation of gene expression is particularly important for multicellular organisms since they are composed of different types of cells, each carrying out its own unique role; some genes are turned on while others are turned off. In fact, only about 20 % of genes in a typical human cell are expressed to ultimately produce proteins required for distinctive functions, and an even smaller percentage of genes are expressed in highly specialized cells like nerve or muscle cells [6].

In eukaryotes, gene regulation can occur at different levels, ranging from pretranscriptional to posttranslational control. The regulatory mechanisms employed by cells can affect a number of operations, such as dictating whether a gene will be turned on and moderating the speed and duration at which a gene is expressed [2]. Transcription, for example, can be regulated in response to the way in which chromatin—DNA and protein—is structured in the nucleus, in that genes are typically silenced, or turned off, when chromatin is tightly packed; however, genes are more likely to be expressed when chromatin is less condensed. Histone modification and DNA methylation are two epigenetic mechanisms that have been extensively studied with regard to their

influence on chromatin structure and, subsequently, gene expression [7]. In general, epigenetics refers to any mechanism that influences gene expression without changing the nucleotide sequence of the DNA of a gene [1]. While stable epigenetic switches are critical for cell differentiation and normal development, research has shown that disruptions in normal epigenetic processes are linked to disease, including the initiation and progression of different types of cancers [8].

4.2 Histone Modification

The nucleus of a human somatic cell possesses 46 chromosomes, and when fully extended, the DNA within the nucleus measures at least 2 m. However, DNA is wrapped around proteins called histones, forming nucleosome complexes, which enable DNA to fit within the nucleus [2]. Nucleosomes can suppress transcription if they are positioned over the promoter region; the enzyme RNA polymerase and accompanying regulatory proteins will not be able to bind to the promoter. As noted earlier, RNA polymerase is essential to transcription in that it catalyzes the production of mRNA from the DNA template. However, nucleosomes can become displaced because of structural changes in the chromatin, thereby allowing room for the binding of the transcriptional proteins and hence the expression of a gene [1].

Histone modification and DNA methylation are two major types of acquired or inherited epigenetic mechanisms that can affect transcriptional activity by regulating the access of transcriptional proteins to gene promoters without altering the DNA base sequence. When observed under an electron microscope, chromatin has been described as resembling an array of beads on a string, whereby the beads represent nucleosomes, separated from one another by 10–60 base pairs of DNA. Each nucleosome contains a core of 8 histone proteins wrapped approximately 1.7 times with DNA of 147 base pairs. The histone core of a nucleosome has two copies of the following histone proteins: H2A, H2B, H3, and H4. Each of the eight histones possesses regions for histone-DNA and histone-histone interactions [9]. In addition, a histone carries an extension called a tail that is rich in positively charged amino acids, like lysine. The positive charge of the amino acids enables them to have a strong attraction to DNA, which is negatively charged because of its phosphate groups [10]. Histones H3 and H4 have been identified in particular to possess certain amino acids that are vulnerable to epigenetic modification [11]. Protein H1 serves as a linker histone that binds to DNA positioned outside the histone octamer [9]. Histone H1 is needed to create the secondary level of chromatin organization, which consists of condensed 30-nm chromatin fibers composed of several nucleosome beads; the fibers will then join to produce a tertiary structure of radial loops [12]. Histone tails are necessary for chromatin to condense into the 30-nm fibers. For instance, lysine 16 of the N-terminal tail of H4 has been shown to be very important in compacting chromatin, in that the positively charged lysine is attracted to a negatively charged area at the H2A-H2B interface of an adjacent nucleosome that is part of the condensed 30-nm fiber form [10].

Histones are subjected to a number of modifications, like acetylation, methylation, phosphorylation, and ubiquitination, and these covalent modifications can work together to alter the chromatin structure, thereby regulating transcriptional activity by controlling access to gene promoters. As noted, the positive charge of histone amino acids and the negative charge of DNA form a strong attraction. However, the addition of certain chemicals has the potential to alter this attraction. Of the various histone modifications, acetylation has been investigated most extensively. Histone acetyltransferases (HATs) are enzymes that catalyze the addition of acetyl groups to the positively charged amino acids located on the tails of the histone octamer. The addition of acetyl groups will reduce the positive charges of the histone tails, hence easing the hold of the histones to DNA and subsequently exposing the once compact nucleosome to additional proteins that serve to modify chromatin. Exposed genes can then be expressed when chromatin is loosely packed [13]. Lysine 16 in histone H4 was previously noted as a critical component in the process of condensing chromatin, in that the positively charged amino acid is attracted to a negatively charged area of a neighboring nucleosome in the 30-nm fiber form; however, the acetylation of lysine 16 tends to decrease the attraction between the amino acid and its adjacent nucleosome. Therefore, chromatin becomes less compact, allowing more space between one nucleosome bead and the next, thus enabling transcription to take place [10].

Conversely, histone deacetylases (HDACs) remove acetyl groups from histones, thereby restoring the positive charge to histone tails that contributes to the high-level compaction of chromatin and thus the repression of transcription. Furthermore, certain proteins can recruit HDACs. For example, the retinoblastoma protein (Rb), a tumor suppressor, blocks the production of molecules that are critical for numerous biological processes, including DNA replication and mitosis [14], and the influence of Rb is associated, in part, with its recruitment of HDACs to certain gene promoters [9]. HDACs have been utilized in certain treatments for cancer. In some forms of cancer, genes that normally produce proteins that inhibit cell division are not as active, and the histones close to these genes exhibit high levels of deacetylation. Therefore, drugs that function as histone deacetylase inhibitors have the potential to activate genes that typically suppress cell division, thus arresting the growth of cancer cells [13].

Along with acetylation, there are other types of histone modifications—all of which are reversible. Histone methylation, for example, may either repress or activate transcription, depending on the amino acid that is methylated and on the extent of methylation. Different classes of enzymes called histone methyltransferases (HMTs) add methyl groups to particular arginine and lysine amino acids on histone tails, while histone demethylases (HDMs) remove methyl groups. The methylation of arginine is linked with transcriptional activation, whereas the methylation of lysine is usually associated with transcriptional repression [11]. When lysine is methylated, acetylation normally does not take place, and the positive charge is maintained. In addition, lysine may be methylated more than once [10]; however, the location of the modified lysine can also determine whether transcription will occur. For instance, transcriptional inactivation is linked to the methylation of lysine 9 in histone H3, but transcriptional activation accompanies the methylation of lysine 4 in H3 [12].

Histone phosphorylation is accomplished with the assistance of kinases and involves the targeting of amino acids serine and threonine, and all four core histones can be affected. Phosphorylation takes place at the N-terminus of the histone near methylated and acetylated lysines. For example, the phosphorylation of serine 10 on histone H3 is linked with acetylation of the nearby lysine 14. The added phosphate group to serine 10 contributes a negative charge, thereby partially neutralizing the net positive charge of the modified histones and thus altering their capacity to interact with DNA, which is also negatively charged. As a result, the repellent force contributes to a less compacted chromatin state, which may allow for the transcription of genes [15]. Conversely, the phosphorylation of both serines 10 and 28 on H3 has been found to correlate with chromosome condensation during mitosis [16]. Moreover, the close association of different modifications can promote further epigenetic changes. In fact, the removal of a methyl group that is attached to lysine 9 on H3 promotes the phosphorylation of serine 10 and the acetylation of lysines 9 and 14 [15].

Another important process by which histones can be modified is ubiquitination. The polypeptide ubiquitin is composed of 76 amino acids and can be enzymatically coupled to different types of proteins. In most cells, ubiquitin is primarily used to tag damaged or old proteins for degradation by proteasomes. In fact, proteasome inhibitors have been shown to be effective in certain types of cancer therapy whereby cancer cells eventually die when proteasomal activity is suppressed. Controlled administration of proteasome inhibitors has been shown to be effective against multiple myeloma in that cancer cells are destroyed but not normal cells. Though ubiquitin may serve as a tag for the destruction of some proteins, it does not damage other proteins and can function in epigenetic genome control. For example, with regard to histone modification, ubiquitination takes place only on H2A and H2B, and only one lysine is targeted in each histone: lysine 119 in H2A and lysine 120 in H2B. Ubiquitin possesses a number of amino acids that are negatively charged, and when bound to histone tails, ubiquitin relaxes the chromatin structure by neutralizing the positively charged amino acids in the histone and thus activating transcriptional activity. Furthermore, ubiquitin can influence other modifications. For example, the ubiquitination of H2B promotes the methylation of lysines 4 and 79 on H3 [10].

The examples above illustrate some of the many complexities associated with the different types of histone modifications and their roles in altering chromatin structure, which influences the accessibility of genes for transcription.

4.3 DNA Methylation

Along with histones, the DNA itself can be tagged with external chemicals. Genes can be silenced when methyl groups are enzymatically attached to certain nucleotide bases. The methylated DNA influences the interaction of the DNA with proteins, including histones. Gene regulation involving DNA methylation is

characterized as epigenetic in that gene expression is altered without affecting the sequence of DNA. Highly methylated regions of a gene are normally associated with tightly packed DNA wrapped around histones, resulting in transcriptional inactivation [1]. DNA methylation is also heritable. Not all eukaryotes exhibit both histone modification and DNA methylation. Fruit flies and yeasts, for instance, display only histone modification [16].

During DNA methylation, the methyl group is covalently added to a cytosine situated 5′ to a guanine (CpG), forming 5-methylcytosine, and the reaction is catalyzed by DNA methyltransferase (DNMT). Methylation normally takes place in only 3–4 % of cytosines. Methylcytosines occur primarily in repetitive sequences and in the CpG islands (GpC and CpG-rich regions) of the promoters of silent genes like X-chromosome inactivated genes and imprinted genes, both of which will be discussed later in this section. During transcription and DNA synthesis, 5-methylcytosine pairs normally with guanine; however, the methylation of cytosine is reversible with the aid of the enzyme demethylase, which removes the methyl group [9].

DNA methylation may cause the cell to incur risks in some cases. Cytosines that are methylated are chemically less stable than those that are not. As a result, a methylated cytosine may undergo deamination, whereby a thymine is generated in its place, i.e., a methylated CpG is converted to TpG. Unfortunately, this error in the base sequence may not be identified by the DNA repair machinery before replication takes place [16].

Typically genes that are heavily methylated are inactive. Processes by which DNA methylation silences genes involve complex interactions with various proteins. This form of epigenetic regulation of transcription interlinks histone deacetylases and chromatin-remodeling enzymes [9]. Specialized proteins that attach to methylated cytosines are called DNA methyl-binding domains, which are usually subunits of larger complexes like histone deacetylases or methyltransferases. These enzymes then alter histone proteins located near the methylated DNA. For instance, methyl groups attached to DNA can recruit methyl-binding proteins that then recruit histone deacetylases, which change chromatin by detaching acetyl groups from histone tails. As noted in the previous section, deacetylation induces chromatin condensation by restoring the positive charge to histones. In addition, methylated DNA may be bound to proteins possessing histone methyltransferase activity; this interaction causes histone proteins in the vicinity to be methylated, resulting in chromatin condensation and, consequently, the silencing of genes. Conversely, specialized proteins that bind to histones can set in motion a chain of reactions that eventually methylates DNA with the assistance of DNA methyltransferases. For example, methylated histone tails can recruit chromodomain-containing proteins that then attract DNA methyltransferases that bind methyl groups to neighboring DNA [16].

A good illustration of how DNA methylation normally operates in compacting chromatin in a formation that represses genes is X-chromosome inactivation. A normal human somatic cell consists of two sex chromosomes, each copy inherited from a parent. The normal complement of sex chromosomes for a human female is XX, while the complement for a human male is XY. The Y chromosome is smaller than the X chromosome and does not possess all of the genes located on the X

chromosome. In theory, since the female has two X chromosomes, she should have the capacity to produce substantially more protein from her sex chromosomes than the male can produce from his. However, the female normally has only one X chromosome in the cell that contains genes that are active, and the genes of the remaining X chromosome are primarily inactive. The determination of whether an X chromosome will be inactive is random; hence within the body of a female, the maternal X chromosome may be inactive in one cell, but the paternal X chromosome may be active in a nearby cell. The inactivation of one of the two X chromosomes in a female allows the female and the male to produce relatively the same amount of X chromosome-encoded protein. And the silencing of genes on the inactive X chromosome is primarily attributed to heavy DNA methylation [17].

Genomic imprinting is a form of gene inactivation that is also associated with DNA methylation. A somatic cell possesses two copies of each gene—one copy (allele) inherited from the mother and the other allele inherited from the father. And typically a gene is expressed from both copies; however, with some genes, transcription is limited to either the maternal or the paternal copy of the gene, not both. This condition is referred to as genomic imprinting, and the repression of a parental copy of a gene can be replicated during cell division and passed on from one generation to the next [18].

Imprinting usually occurs in a limited number of genes in the mammalian body, but it is important in regulating embryonic development and growth. In some genes, the inherited maternal allele is silenced while the inherited paternal allele is expressed. In other genes, the paternal copy is silenced while the maternal copy is expressed. The silenced copy of a gene is typically highly methylated [16]. For example, *IGF2* is an imprinted gene located on chromosome 11 in humans that encodes for a protein hormone called insulin-like growth factor 2. This hormone is important for placental and fetal growth. Normally the maternal allele is repressed because of DNA methylation, but the paternal copy of the gene is actively expressed, and the resulting protein product is essential for development. Research utilizing mice as subjects has shown that when the paternal allele of *IGF2* is deleted, the placenta will be small in size, and the young will exhibit low-birth weight [19].

Prader-Willi syndrome and Angelman syndrome are examples of developmental disorders in humans related to genomic imprinting, and both diseases involve deletions in chromosome 15 that result in the silencing of one or more genes. In Prader-Willi syndrome, a small region of the paternally inherited chromosome is missing, and an equivalent region on the maternal chromosome is silent. Clinical symptoms include mental retardation, short stature, and small feet and hands. In contrast, with Angelman syndrome deletions are found in a region of the maternal chromosome, but the paternal chromosome has a silent region. Seizures and uncontrolled muscle movements are some characteristics exhibited by those inflicted with Angelman syndrome [18].

Modifications to normal patterns of DNA methylation have also been implicated in the initiation of certain cancers. Promoters of tumor suppressor genes have been found to be inactive because of the addition of methyl groups, while promoters of proto-oncogenes, which normally stimulate cells to divide, may become demethylated to the point of causing cells to multiply uncontrollably [13].

4.4 Environmental Influences on Gene Expression

Proposed in 1809 by Jean-Baptiste Lamarck, the concept of inheritance of acquired characteristics argues that environmental pressures and personal necessities can result in beneficial changes in physical appearance and body function and that eventually offspring inherit these characteristics. Lamarck employed his concept to describe why giraffes exhibit long necks. He suggested that giraffes originally possessed short necks, but in an effort to procure food located in higher areas of trees, the giraffes stretch their necks in order to reach leaves. The necks are lengthened with the aid of a nerve "fluida," and longer necks are then passed on to offspring [3]. By contrast, Charles Darwin's theory of natural selection would presume that early giraffe populations had individuals with necks of varied lengths. Natural selection due to competition caused those with longer necks to acquire more food for survival, live longer, and have more offspring. Eventually, after generations, the population would consist of giraffes with long necks because of the beneficial adaptation to the environment [2]. Darwin's approach to evolution is, of course, widely accepted today: natural selection dictates that beneficial traits are acquired by chance and are inherited through generations because of their selective advantage. However, with the discovery of epigenetic regulation, scientists are forced to take a closer look at the role of environmental influences on gene expression and how epigenetic factors are passed on to future generations [20].

Gene expression is tightly regulated in the body in order to ensure that cells manufacture the appropriate level of proteins needed for a given period and for a specific function. Cell specialization is closely tied to epigenetic regulation. Cells differentiate into various types, each performing distinctive functions, and epigenetic factors regulate how and when particular genes are turned on and off to assist the body with growth and development [21]. Liver and pancreatic cells, for instance, share the same set of genes; however, cells of the liver access the genes associated with the production of enzymes used to neutralize particular toxins but not the genes required to synthesize glucagon and insulin, glucose regulating hormones made in pancreatic alpha and beta cells, respectively [3].

In eukaryotes, gene expression can be regulated at different levels—pretranscription, transcription, posttranscription, translation, and posttranslation. And studies have shown that a variety of mechanisms used to control gene expression operate before transcription and after translation [6]. Sections above that pertain to histone modification and DNA methylation illustrate some of the epigenetic mechanisms that manipulate the structure of chromatin and thus the onset of gene expression. Furthermore, external signals, including environmental chemicals, aging, and diet, have a major impact on the operation of these epigenetic mechanisms and, ultimately, the health of an individual [1].

Both genetic mutations and epigenetic changes can be linked to the onset of disease. Some diseases are the result of genetic mutations, whereby changes exist in the sequences of nucleotides of the genome, and these mutations can develop from environmental factors and can run in families. And many diseases are influenced by

epigenetic modifications. The epigenome of an individual can continually change throughout life and can be influenced by environmental factors, including chemical cues and social interactions, which may have positive or negative effects. An accumulation of epigenetic changes is normally associated with aging; however, these changes can alter the standard operations of certain genes and lead to particular age-related diseases, like diabetes and cancer [21].

The health effects of environmental factors have been studied with monozygotic (identical) twins, which develop from one egg fertilized by a single sperm. Though identical twins share the same DNA sequences, they typically differ to some degree in appearance, behavior, and health. A thorough explanation of these differences has not been developed; however, recent studies suggest that epigenetics appears to be the cause of some of these phenotypic differences. Fraga and his associates [22], for example, found that monozygotic twins are not only identical genetically, but they are similar epigenetically with regard to DNA methylation and acetylation of histones H3 and H4. Nonetheless, the content and distribution of these epigenetic factors change considerably as twins get older, and the degree of epigenetic differences has been shown to correlate with the length of time twins live apart and variations in lifestyles, such as smoking habits, food choices, and physical activity. This strongly suggests that environmental cues play critical roles in altering one's epigenetic profile and, subsequently, the expression of genes. Epigenetic differences might explain why one identical twin is diagnosed with a disease like schizophrenia, which is genetically based, but the other twin is not [6].

Environmental exposure to certain chemicals has the capability to alter the epigenome. In fact, early exposure to certain chemicals *in utero* has been linked to epigenetic changes. Waterland and Jirtle [23] famously demonstrated in pregnant Agouti mice that dietary supplements can result in phenotypic changes in offspring. The *Agouti* gene controls the amount and distribution of pheomelanin (yellow and red) and eumelanin (brown and black) pigmentation in the coat of a mammal. A mouse carrying an *Agouti* gene with the lethal yellow mutation has a distinctive yellow coat, is obese, and is prone to develop diabetes and cancer [24]. The promoter of the mutated gene has been found to be hypomethylated, unlike the promoter of the normal *Agouti* gene [25]. In an experiment where pregnant yellow mice were fed methyl-rich supplements—folic acid, vitamin B12, betaine, and choline—most of the pups were of normal weight, displayed a brown coat, and lacked the predisposition to develop diabetes and cancer. However, pregnant yellow mice that were not given methyl supplements primarily produced pups that were also yellow and unhealthy. These findings suggest that the diet of mothers can influence the health of the young. In this example, the methyl donors supplemented in the food of pregnant mice with a mutated gene were able to make their way onto the chromosomes of the developing embryos and methylate the critical gene, hence counteracting the negative effects of the faulty gene [23]. This and many other examples show correlations between the exposure to certain chemicals and epigenetic changes. With regard to adverse health effects, prenatal introduction as well as postnatal exposure to deleterious chemicals can have a major impact on development and the susceptibility to chronic diseases [26].

If acquired environmental cues modify the chromosomes of eggs and sperm, these external influences have the potential to have long lasting effects through generations [27]. Sollars and his colleagues [28] presented a dramatic example of transgenerational inheritance involving fruit flies with a defective *Krüppel* gene, which causes abnormally small eyes. In addition, unusual outgrowths appear on the eyes under certain conditions. When the antibiotic geldanamycin was added to the food of the flies, the eye outgrowths increased substantially. It was reported that although only the first generation was directly exposed to geldanamycin, the effects of the drug were noticeable in future generations. In fact, the eye anomaly endured through 13 generations. The antibiotic was suspected of altering histone proteins since histone deacetylase inhibitors were found to suppress the development of the eye outgrowths.

In another example, Guerrero-Bosagna et al. [29] found that rats exposed *in utero* to the fungicide vinclozolin had negative reproductive consequences that lasted through future generations. Vinclozolin is an endocrine disruptor, which can lead to certain cancers and reproductive defects. When pregnant rats were injected daily with the fungicide, the resulting male offspring were found to display a decrease in sperm production. This suppression was also discovered in several generations that were not exposed to vinclozolin. Increased DNA methylation was reported in rats that were exposed to the fungicide from daily injections as well as in those of subsequent generations.

In the case of tobacco, diseases such as lung cancer, asthma, and chronic obstructive pulmonary disease have been linked to long term exposure to tobacco smoke, and substantial evidence suggests that epigenetic alterations are responsible for inducing these chronic disorders [30]. It has been proposed that epigenetic effects cause an imbalance in histone acetyltransferases and histone deacetylases of the airway immune cells. For example, studies reported by Ito and his colleagues [31] involving alveolar macrophages and bronchial biopsies from smokers and nonsmokers of similar age disclosed that in smokers, *HDAC2* gene expression is suppressed, along with general HDAC protein activity; however, the expression of inflammatory mediators *GM-CSF*, *TNF-α*, and *IL*-8 are enhanced. Furthermore, Launay et al. [32] found that tobacco smoke alters DNA methylation in a number of genes. For instance, in platelets, hypomethylation exists on the promoter of the gene that encodes for the enzyme monoamine oxidase (MAO) type B, suggesting that tobacco smoke overstimulates the expression of the gene for MAO, which has been linked to heart failure, mitochondrial dysfunction, and other abnormalities when present at elevated levels in the body [33].

The adverse effects of tobacco smoke can be reflected in the health of offspring and even subsequent generations. The respiratory function of the fetus is compromised when its mother is exposed to cigarette smoke. In fact, studies have shown that children as young as 1 year of age may exhibit asthma-like symptoms if their mothers smoked during the third trimester [34]. Furthermore, a child's increased risk of developing asthma is linked to a grandmother who smoked during pregnancy, suggesting that environmental factors can have long lasting consequences [35]. When comparing the buccal cells of children exposed *in utero* to tobacco

smoke from their mothers and those who had no such exposure, Breton et al. [36] reported that global DNA hypomethylation existed in children exposed to tobacco. It has been proposed that this occurrence may be the result of tobacco-induced oxidative stress to DNA that hinders the binding of DNA methyltransferase, thereby preventing methyl groups from binding to DNA. Another outcome of this form of early exposure to tobacco smoke was shown to be hypermethylation of certain genes. For example, the genes *AXL* and *PTPRO* were found to have an unusually high degree of methylation, though the significance of this discovery with regard to asthma pathogenesis is not yet fully understood.

These and many other studies show how the environment can have a major influence on development and on disease susceptibility. Epigenetic marks can be altered over time, and if these acquired chemical tags are passed on to sex cells, epigenetic influences can therefore be transferred from parents to offspring and even future generations. Hence chemical tags acquired from life experiences have the potential to be passed from one generation to the next.

4.5 Epigenetics and Cancer

Disruptions in normal cellular processes can be attributed to mutations, changes to the nucleotide sequences of genes, and some mutations may set the stage for the development of certain cancers. Moreover, there is increasing evidence that supports the importance of epigenetic control of gene expression with regard to carcinogenesis. DNA methylation aberrations, either as hypo- or hypermethylation, have been shown to be common in a variety of carcinogenic tumors. Chromosomal instability and oncogene activation are two processes typically associated with cancer that involve DNA hypomethylation, and tumor suppressor gene inactivation is often linked with DNA hypermethylation. Irregularities in histone modifications can also lead to notable interruptions in gene regulation. Modifications made to histone proteins can contribute to the development of disease, histone acetylation or deacetylation being the most commonly observed alterations [8].

The disruption of normal gene expression has the potential to lead to cancer. Normal cell activities are altered when genes are silenced or overly stimulated. A gene that is not typically expressed in a cell can be turned on, and abnormally high levels of a particular protein may be produced. This disruption can be the result of mutations and/or changes in gene control. Epigenetic influences on the regulation of genes have been detected in various cancer cells. Researchers are trying to understand the common changes that prompt the development of particular types of cancers and how these changes can be manipulated to ultimately destroy cancer cells [1].

There are over 100 different types of cancers, which generally comprise three main groups: carcinomas, sarcomas, and leukemias/lymphomas [37]. About 85 % of cancers are carcinomas and involve malignances of epithelial cells [9]. Sarcomas are tumors of connective tissues, like muscle and bone. Leukemia is a

malignancy of the blood cells in the bone marrow, while lymphoma is a cancer of the lymphocytes, typically in the lymph nodes. The tissue of origin (e.g., lung carcinoma) and the cell type (e.g., erythroid leukemia) can be used to further classify malignant growths [37].

Environment and heredity are the primary factors contributing to the onset of cancer, with environmental influences accounting for the vast majority of cancers [38]. Environmental factors include diet, tobacco, radiation, and infectious organisms [39].

Cancer arises from the accumulation of multiple mutations [40], and abnormalities in proto-oncogenes and tumor suppressor genes have been identified in many malignant cells. Over 100 proto-oncogenes have been identified, and at least 15 tumor suppressor genes have been discovered [9]. Proto-oncogenes normally promote cell division; however, once they mutate to become cancer-causing genes, oncogenes, excessive cell proliferation can occur. Tumor suppressor genes produce proteins that normally inhibit cell division; however, mutated tumor suppressors encode for proteins that directly or indirectly allow cells to divide uncontrollably [41].

The suppression of gene expression through epigenetic regulation has been identified in certain cancers. Both DNA methylation and histone modification have been found to silence genes. In some malignancies, cells may exhibit abnormally high levels of methyl groups on cytosines in CpG islands of promoter regions of suppressed genes. Also, histones that surround promoter regions may not be acetylated as in nonmalignant cells. Therefore, a combination of DNA methylation and histone deacetylation mechanisms may result in the suppression of gene expression [1].

Both DNA hypermethylation and hypomethylation are mechanisms that occur in healthy cells under certain conditions, but both mechanisms have also been identified in cancer. Of the two processes, hypermethylation has received more attention with regard to its role in cancer, primarily because of its occurrence on certain cancer-associated gene promoter sites. Nonetheless, global DNA hypomethylation is reported to be prominent in highly repeated DNA sequences in cells of certain cancers, like breast, colorectal, and gastric cancers. Chromosomal instability and oncogene activation have been linked with hypomethylation. The degree of demethylation varies within and between different cancer types. However, hypomethylation has the potential to be used as a biomarker for detecting the early stage and progression of certain malignancies [42].

Gene inactivation by aberrant DNA hypermethylation in promoter regions has been shown to be an important procedure in carcinogenesis, whereby key genes that normally inhibit cell division are turned off. And regions that are frequently targeted by hypermethylation are CpG islands. Normal ovarian epithelial cells, for example, possess an estrogen receptor protein, but this protein is not often present in ovarian cancer. The promoter of the *estrogen receptor-α* gene is typically hypermethylated in malignant ovarian cells, suggesting that hypermethylation is a contributing factor to the absence of the receptor protein [43]. In another example, the tumor suppressor gene *BRCA1*, which participates in DNA repair, is often mutated in inherited breast cancer. Though a mutated *BRCA1* gene is usually not identified

in non-inherited breast cancer, hypermethylation of the gene has been observed. Therefore, both epigenetic mediation and DNA mutation appear to function in the development of breast cancer [9].

Strong evidence suggests that aberrant DNA methylation patterns are responsible for adversely altering certain biological pathways. For instance, *p53*, which functions as a tumor suppressor gene and is commonly mutated in cancer cells, can be silenced indirectly through epigenetic operations. Under normal circumstances, the tumor suppressor gene *p14ARF* inhibits MDM2, an oncogenic protein that assists with the degradation of p53 protein; however, *p14ARF* is found to be inactivated through methylation. Furthermore, in leukemia, hypermethylation has been observed in gene *p73*, which is a *p53* homolog [44]. Another example of how pathways may be affected by methylation involves the inactivation of the *Rb* gene. In its active state, the Rb protein binds to the transcription factor E2F, which controls the expression of several genes necessary for the transition of the cell into the S (DNA synthesis) phase of the cell cycle. DNA synthesis cannot take place as long as Rb and E2F are bound [45]. However, hypermethylation appears to participate in the suppression of the *Rb* gene in some cancers, thereby silencing its production of protein product [46]. Also, DNA repair pathways can be thwarted when *hMLH1* is methylated. The suppression of this gene can result in microsatellite instability in gastric, endometrial, and colorectal tumors [47]. Additional examples of well-known tumor suppressor genes that have been reported to carry out methylation-mediated silencing in cancer are *APC*, *p16INK4a*, and *VHL*. These and other discoveries have encouraged the continuation of research related to DNA methylation and its influence in initiating cancer. Specialized techniques for studying methylation have also been developed, in particular, sodium bisulfate treatment and methylation-specific PCR [9].

Though not investigated as extensively as DNA methylation, histone modification has been implicated in some cancers. The amino acids of the histone tails are subject to chemical alterations—i.e., acetylation, methylation, etc. These chemical tags influence chromatin structure and function by changing the properties of the tail itself and by offering binding sites for non-histone proteins [48]. The histone is continually modified with the addition and removal of different chemical tags. Enzymes critical to these processes are HDACs, HATs, HMTs, HDMs, etc. However, irregular patterns of histone tags have been found in cancer. For example, it has been reported that genes that encode HDACs are usually over-expressed in gastric and prostate cancers [49]. In addition, aberrant deacetylation of histones has been found in certain types of leukemia, whereby gene translocations abnormally produce fusion proteins that recruit HDACs to promoters that subsequently suppress genes associated with cell differentiation [9]. The loss of both histone acetylation and histone methylation has been strongly linked with malignancy, in which the losses take place primarily at the acetylated lysine 16 and the trimethylated lysine 20 of H4. These losses were also linked to hypomethylation of repeated DNA sequences [50].

A combination of epigenetic procedures can be involved in altering gene expression. Another look at the tumor suppressor gene *Rb*, for instance, reveals that its suppression during tumorigenesis may be associated with DNA hypermethylation [46], as noted above, as well as the inactivation of HDAC1. Under normal conditions, the cell cycle is halted during the G1 (Gap 1) phase in order for the cell to make any necessary DNA repairs before the cell enters the S phase [41]. This pause in the cell cycle is accomplished, in part, when the Rb protein recruits an HDAC complex composed of HDAC1, HDAC2, or HDAC3 to the promoter-bound E2F transcription factor. HDAC activity is needed for the suppression of E2F target genes, notably the gene that encodes the cyclin E protein. This process is carried out when the enzyme deacetylates histone tails, causing chromatin to condense and, subsequently, genes to be silenced [51]. However, in cancer, a mutated Rb protein is inactivated, thus preventing its binding with HDAC so that chromatin can transform to a condensed state. As a result, the cell is no longer halted and progresses prematurely into the S phase. Hence the cell cycle is unregulated, resulting in uncontrolled proliferation of abnormal cells [52].

Research has shed light on ways in which some epigenetic pathways may be thwarted, and this knowledge has been translated into the development of novel cancer drugs that are designed to counteract negative epigenetic changes. These therapeutic strategies take into account that epigenetic modifications are potentially reversible. Scientists are examining how epigenetically inactive tumor suppressor genes and major signaling pathways can be reactivated [53].

Many genes identified as playing critical roles in carcinogenesis have been found to exhibit hypermethylation in their promoter regions; therefore, the use of DNA methylation inhibitors has the potential to become a promising treatment against tumor formation. The DNMT inhibitors 5-azacytidine and 5-aza-2′-deoxycytidine have been studied extensively. However, results have been mixed regarding their effectiveness. It has been reported that both drugs exhibit anti-leukemic activity in clinical trials, but positive results were not noted in solid tumors. A potentially adverse reaction to these drugs includes DNA instability. And, there is the possibility that abnormal methylation may return once treatment has been discontinued [54].

Inhibitors of histone deacetylases have also been examined as possible cancer drugs. HDACs typically suppress gene expression, and the abnormal recruitment of these enzymes has been found in certain cancers, like leukemia. In clinical trials, several drugs have been developed to block the binding of HDACs to their substrates. These drugs include the following: butyrates, valproic acid, depsipeptide (FR901228), and hydroxamic acid-based compounds SAHA and pyroxamide. Though these drugs have been reported to be well tolerated, alteration of gene expression has been found to be selective [9]. Therefore, further research is required to identify substrate specificities of different HDACs and drugs that are more appropriate under certain conditions [55]. Detailed understanding of the numerous and complex mechanisms involved in aberrant epigenetic regulation may lead to more effective treatments against different cancers as well as other diseases.

4.6 Conclusion

The information acquired from epigenetic studies provides a better understanding of how genes are turned on or off and the means by which epigenetic chemicals can be acquired. Though epigenetic chemicals are important for normal growth and development, they may also induce certain diseases, including cancer. Histone modification and DNA methylation are two epigenetic mechanisms that have been studied extensively, but the complexities of these processes still challenge our comprehension of their many intricacies. The study of epigenetics has not only enhanced our understanding of what makes each of us unique, but it has advanced the possibility of developing more effective ways of diagnosing and treating diseases.

References

1. Avissar Y, Choi J, DeSaix J, et al. Biology. Houston: OpenStax College. 2013. http://openstax-college.org/files/textbook_version/low_res_pdf/11/col11448-lr.pdf. Accessed 15 Oct 2013.
2. Mader S, Windelspecht M. Inquiry into life. 14th ed. New York: McGraw-Hill; 2014.
3. Starr C. Biology: concepts and applications. 5th ed. Belmont: Thomas Learning; 2003.
4. Belk C, Borden V. Biology: science for life with physiology. 2nd ed. Upper Saddle River: Prentice-Hall; 2007.
5. Cummings MR. Human heredity: principles and issues. 10th ed. Belmont: Brooks/Cole; 2014.
6. Reece JB, Meyers N, Urry LA, et al. Biology, Australian and New Zealand version. 10th ed. Melbourne: Pearson; 2015.
7. Gomez-Pinilla F, Vaynman S. Intersecting genetics with lifestyle: the role of exercise and diet. In: Dudek SM, editor. Transcriptional regulation by neuronal activity. New York: Springer; 2008. p. 337–75.
8. Tollefsbol T. Epigenetics of human disease. In: Tollefsbol T, editor. Epigenetics in human disease. New York: Academic; 2012. p. 1–6.
9. Pecorino L. Molecular biology of cancer: mechanisms, targets, and therapeutics. 2nd ed. New York: Oxford University Press; 2008.
10. Lodish H, Berk A, Kaiser CA, et al. Molecular cell biology. 6th ed. New York: W. H. Freeman and Company; 2008.
11. Strachan T, Read AP. Human molecular genetics. 3rd ed. Kentucky: Garland Publishing; 2004.
12. Popa G. DNA and protein synthesis: histones. 2014. http://www.fastbleep.com/biology-notes/40/116/1191. Accessed 7 Nov 2014.
13. Sadava D, et al. Life: the science of biology. 9th ed. Sunderland: Sinauer Associates; 2011.
14. Ishida S, Huang E, Zuzan H, et al. Role for E2F in control of both DNA replication and mitotic functions as revealed from DNA microarray analysis. J Mol Cell Biol. 2001;21:4684–99.
15. Lachman DS. Gene control. New York: Garland Science; 2010.
16. Craig NL, Cohen-Fix O, Green R, et al. Molecular biology: principles of genome function. New York: Oxford University Press; 2010.
17. Heller D, et al. Principles of life. Gordonsville: W. H. Freeman & Company; 2012.
18. Papachristodoulou D, et al. Biochemistry and molecular biology. 5th ed. New York: Oxford University Press; 2014.
19. Solomon E, et al. Biology. 10th ed. Stamford: Cengage Learning; 2015.
20. Erick T. Epigenetics: how nurture shapes our nature. Footnote. 2014. http://footnote1.com/epigenetics-reveals-how-environment-shapes-gene-expression. Accessed 1 Mar 2014.

21. National Institute of Child Health and Human Development. What role do epigenetics and developmental epigenetics play in health and disease? National Institutes of Health. 2013. http://m.nichd.nih.gov/topics/epigenetics/conditioninfo/pages/impact.aspx. Accessed 3 May 2014.
22. Fraga MF, Ballestar E, Paz MF, et al. Epigenetic differences arise during the lifetime of monozygotic twins. Proc Natl Acad Sci U S A. 2005;102:10604–9.
23. Waterland RA, Jirtle RL. Transposable elements: targets for early nutritional effects on epigenetic gene regulation. J Mol Cell Biol. 2003;23:5293–300.
24. Melmed S. The pituitary. 3rd ed. San Diego: Elsevier; 2011.
25. Piyathilake CJ, Johanning GL. Environmental influences on DNA methylation. In: Neumann HP, editor. Progress in DNA methylation research. New York: Nova Science; 2007. p. 117–32.
26. Marchlewicz EH, Anderson OS, Dolinoy DC. Early-life exposures and the epigenome: interactions between nutrients and the environment. In: Ho E, Domann F, editors. Nutrition and epigenetics. Boca Raton: Taylor & Francis; 2015. p. 3–54.
27. Rosenfeld CS. Animal models of transgenerational epigenetic effects. In: Tollefsbol T, editor. Transgenerational epigenetics: evidence and debate. San Diego: Elsevier; 2014. p. 121–37.
28. Sollars V, Lu X, Xiao L, et al. Evidence for an epigenetic mechanism by which Hsp90 acts as a capacitor for morphological evolution. Nat Genet. 2003;33:70–4.
29. Guerrero-Bosagna C, et al. Epigenetics transgeneration actions of vinclozolin on promoter regions of the sperm epigenome. 2010. http://www.ncbi.nlm.nih.gov/pubmed/20927350. Accessed 10 Nov 2014.
30. Prescott S, Saffery R. The role of epigenetic dysfunction in the epidemic of allergic disease. Clin Epigenetics. 2011;2:223–32.
31. Ito K, Lim S, Caramori G, et al. Cigarette smoking reduces histone deacetylase 2 expression, enhances cytokine expression, and inhibits glucocorticoid actions in alveolar macrophages. FASEB J. 2001;15:1110–2.
32. Launay JM, Del Pino M, Chironi G, et al. Smoking induces long-lasting effects through a monoamine-oxidase epigenetic regulation. PLOS One. 2009. http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0007959. Accessed 20 Nov 2014.
33. Kaludercic N, Carpi A, Menabo R, et al. Monoamine oxidases (MAO) in the pathogenesis of heart failure and ischemia/reperfusion injury. Mol Cell Res. 2011;7:1323–32.
34. Sears MR, Holdaway MD, Flannery EM, et al. Parental and neonatal risk factors for atopy, airway hyper-responsiveness, and asthma. Arch Dis Child. 1996;75:392–8.
35. Li YF, et al. Maternal and grandmaternal smoking patterns are associated with early childhood asthma. Chest. 2005;127:1232–41.
36. Breton CV, Byun HM, Wenten M, et al. Prenatal tobacco smoke exposure affects global and gene-specific DNA methylation. Am J Respir Crit Care Med. 2009;209:462–7.
37. Cooper GM. The cell: a molecular approach. 2nd ed. Sunderland: Sinauer Associates; 2000.
38. Weinberg RA. The biology of cancer. New York: Garland Science; 2007.
39. Anand P, Kunnumakara AB, Sundaram C, et al. Cancer is a preventable disease that requires major lifestyle changes. Pharm Res. 2008;25:2097–116.
40. Loeb KR, Loeb LA. Significance of multiple mutations in cancer. Carcinogenesis. 2000;21:379–85.
41. Sadava D, Heller C, Orlans G, et al. Life: the science of biology. 8th ed. Sunderland: Sinauer Associates; 2008.
42. Wilson AS, Power BE, Molloy PL. DNA hypomethylation and human diseases. Biochim Biophys Acta. 2007;1775:138–62.
43. O'Doherty AM, Church SW, Russell SEH, et al. Methylation status of oestrogen receptor—a gene promoter sequence in human ovarian epithelial cell lines. Br J Cancer. 2002;86:282–4.
44. Corn PG, Kuerbitz SJ, Van Noesel MM, et al. Transcriptional silencing of the p73 gene in acute lymphoblastic leukemia and Burkitt's lymphoma is associated with 5′ CpG island methylation. Cancer Res. 1999;59:3352–6.
45. Raven PH, Johnson GB. Biology. 4th ed. New York: McGraw Hill; 2001.

46. Greger V, Passarge E, Höpping W, et al. Epigenetic changes may contribute to the formation and spontaneous regression of retinoblastoma. Hum Genet. 1989;83:155–8.
47. Herranz M, Esteller M. DNA methylation and histone modifications in patients with cancer: potential prognostic and therapeutic targets. In: Sioud M, editor. Target discovery and validation reviews and protocols. Totowa: Humana Press; 2007. p. 25–62.
48. Morgan DO. The cell cycle: principles of control. Sunderland: New Science Press; 2007.
49. Song J, Noh JH, Lee JH, et al. Increased expression of histone deacetylase 2 is found in human gastric cancer. APMIS. 2005;113:264–8.
50. Fraga MF, Ballestar E, Villar-Garea A, et al. Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer. Nat Genet. 2005;37:391–400.
51. Van Quill KR, O'Brien JM. The role of the retinoblastoma protein in health, malignancy, and the pathogenesis of retinoblastoma. In: Albert DM, Polans A, editors. Ocular oncology. Boca Raton: Taylor & Francis; 2003. p. 129–88.
52. Brehm A, Miska EA, McCance DJ, et al. Retinoblastoma protein recruits histone deacetylase to repress transcription. Nature. 1998;391:597–601.
53. Nephew KP. Epigenetic gene silencing. In: Schwab M, editor. Encyclopedia of cancer. Berlin: Springer; 2012. p. 1284–6.
54. Goffin J, Eisenhauer E. DNA methyltransferase inhibitors—state of the art. Ann Oncol. 2002; 13:1699–716.
55. Saldanha SN, Tollefsbol TO. Epigenetic approaches to cancer therapy. In: Tollefsbol T, editor. Epigenetics in human disease. New York: Academic; 2012. p. 111–6.