

CHAPTER 19

THE ORGANIZATION AND CONTROL OF EUKARYOTIC GENOMES

OUTLINE

- I. The Structure of Chromatin
 - A. Chromatin structure is based on successive levels of DNA packing
- II. Genome Organization at the DNA Level
 - A. Repetitive DNA and other noncoding sequences account for much of a eukaryotic genome
 - B. Gene families have evolved by duplication of ancestral genes
 - C. Gene amplification, loss, or rearrangement can alter a cell's genome
- III. The Control of Gene Expression
 - A. Each cell of a multicellular eukaryote expresses only a small fraction of its genome
 - B. The control of gene expression can occur at any step in the pathway from gene to functional protein
 - C. Chromatin modifications affect the availability of genes for transcription
 - D. Transcription initiation is controlled by proteins that interact with DNA and with each other
 - E. Posttranscriptional mechanisms play supporting roles in the control of gene expression
- IV. The Molecular Biology of Cancer
 - A. Cancer results from genetic changes that affect the cell cycle
 - B. Oncogene proteins and faulty tumor-suppressor proteins
 - C. Multiple mutations underlie the development of cancer

OBJECTIVES

After reading this chapter and attending lecture, the student should be able to:

1. Compare the organization of prokaryotic and eukaryotic genomes.
2. Describe the current model for progressive levels of DNA packing.
3. Explain how histones influence folding in eukaryotic DNA.
4. Distinguish between heterochromatin and euchromatin.
5. Using the Barr body as an example, describe the function of heterochromatin in interphase cells.
6. Describe where satellite DNA is found and what role it may play in the cell.
7. Describe the role of telomeres in solving the end-replication problem with the lagging DNA strand.
8. Using the genes for rRNA as an example, explain how multigene families of identical genes can be advantageous for a cell.

9. Using α -globin and β -globin genes as examples, describe how multigene families of nonidentical genes probably evolve, including the role of transposition.
10. Explain the potential role that promoters and enhancers play in transcriptional control.
11. Explain why the nuclear envelope in eukaryotes offers a level of post-transcriptional control beyond that found in prokaryotes.
12. Explain why the ability to rapidly degrade mRNA can be an adaptive advantage for prokaryotes.
13. Describe the importance of mRNA degradation in eukaryotes, describe how it can be prevented.
14. Explain how gene expression may be controlled at the translational and post-translational level.
15. Compare the arrangement of coordinately controlled genes in prokaryotes and eukaryotes.
16. Explain how eukaryotic genes can be coordinately expressed and give some examples of coordinate gene expression in eukaryotes.
17. Provide evidence from studies of polygene chromosomes, that eukaryotic gene expression is controlled at transcription and that gene regulation responds to chemical signals such as steroid hormones.
18. Describe the key steps of steroid hormone action on gene expression in vertebrates.
19. In general terms, explain how genome plasticity can influence gene expression.
20. Describe the effects of gene amplification, selective gene loss and DNA methylation.
21. Explain how rearrangements in the genome can activate or inactivate genes.
22. Explain the genetic basis for antibody diversity.
23. Explain how DNA methylation may be a cellular mechanism for long-term control of gene expression and how it can influence early development.
24. Describe the normal control mechanisms that limit cell growth and division.
25. Briefly describe the four mechanisms that can convert proto-oncogenes to oncogenes.
26. Explain how changes in tumor-suppressor genes can be involved in transforming normal cells into cancerous cells.
27. Explain how oncogenes are involved in virus-induced cancers.

KEY TERMS

histones	multigene family	genomic imprinting	proteasomes
nucleosome	pseudogene	histone acetylation	oncogenes
heterochromatin	gene amplification	control elements	proto-oncogenes
euchromatin	retrotransposons	enhancers	tumor-suppressor genes
repetitive DNA	immunoglobulins	activator	<i>ras</i> gene
satellite DNA	differentiation	DNA-binding domain	p53 gene
<i>Alu</i> elements	DNA methylation	alternative splicing	

LECTURE NOTES

Eukaryotic gene regulation is more complex than in prokaryotes, because eukaryotes:

- Have larger, more complex genomes. This requires that eukaryotic DNA be more complexly organized than prokaryotic DNA.
- Require cell specialization or *differentiation*.

I. The Structure of Chromatin

A. Chromatin Structure is Based on Successive Levels of DNA Packing

Prokaryotic and eukaryotic cells both contain double-stranded DNA, but their genomes are organized differently.

Prokaryotic DNA is:

- Usually circular
- Much smaller than eukaryotic DNA; it makes up a small nucleoid region only visible with an electron microscope
- Associated with only a few protein molecules
- Less elaborately structured and folded than eukaryotic DNA; bacterial chromosomes have some additional structure as the DNA-protein fiber forms loops that are anchored to the plasma membrane

Eukaryotic DNA is:

- Complexed with a large amount of protein to form *chromatin*
- Highly extended and tangled during interphase
- Condensed into short, thick, discrete *chromosomes* during mitosis; when stained, chromosomes are clearly visible with a light microscope

Eukaryotic chromosomes contain an enormous amount of DNA, which requires an elaborate system of DNA packing to fit all of the cell's DNA into the nucleus.

B. Nucleosomes, or "beads on a string"

Histone proteins associated with DNA are responsible for the first level of DNA packing in eukaryotes.

Histones = Small proteins that are rich in basic amino acids and that bind to DNA, forming chromatin.

- Contain a high proportion of positively charged amino acids (arginine and lysine), which bind tightly to the negatively charged DNA
- Are present in approximately equal amounts to DNA in eukaryotic cells
- Are similar from one eukaryote to another, suggesting that histone genes have been highly conserved during evolution. There are five types of histones in eukaryotes.

Partially unfolded *chromatin* (DNA and its associated proteins) resembles beads spaced along the DNA string. Each beadlike structure is a nucleosome (see Campbell, Figure 19.1a).

Nucleosome = The basic unit of DNA packing; it is formed from DNA wound around a protein core that consists of two copies each of four types of histone (H2A, H2B, H3, H4). A fifth histone (H1) attaches near the bead when the chromatin undergoes the next level of packing.

- Nucleosomes may control gene expression by controlling access of transcription proteins to DNA.
- Nucleosome heterogeneity may also help control gene expression; nucleosomes may differ in the extent of amino acid modification and in the type of nonhistone proteins present.

C. Higher levels of DNA packing

The 30-nm *chromatin fiber* is the next level of DNA packing (see Campbell, Figure 19.1b).

- This structure consists of a tightly wound coil with six nucleosomes per turn.
- Molecules of histone H1 pull the nucleosomes into a cylinder 30nm in diameter.

In the next level of higher-order packing, the 30-nm chromatin fiber forms *looped domains*, which:

- Are attached to a nonhistone protein scaffold
- Contain 20,000 to 100,000 base pairs
- Coil and fold, further compacting the chromatin into a mitotic chromosome characteristic of metaphase

Interphase chromatin is much less condensed than mitotic chromatin, but it still exhibits higher-order packing.

- Its nucleosome string is usually coiled into a 30-nm fiber, which is folded into looped domains.
- Interphase looped domains attach to a scaffolding inside the nuclear envelope (nuclear lamina); this helps organize areas of active transcription.
- Chromatin fibers of different chromosomes do not become entangled as they occupy restricted areas within the nucleus.

Portions of some chromosomes remain highly condensed throughout the cell cycle, even during interphase. Such heterochromatin is not transcribed.

Heterochromatin = Chromatin that remains highly condensed during interphase and that is not actively transcribed

Euchromatin = Chromatin that is less condensed during interphase and is actively transcribed; euchromatin becomes highly condensed during mitosis

What is the function of heterochromatin in interphase cells?

- Since most heterochromatin is not transcribed, it may be a coarse control of gene expression.
- For example, Barr bodies in mammalian cells are *X* chromosomes that are mostly condensed into heterochromatin. In female somatic cells, one *X* chromosome is a Barr body, so the other *X* chromosome is the only one transcribed.

II. Genome Organization at the DNA Level

An organism's genome is plastic, or changeable, in ways that affect the availability of specific genes for expression.

- Genes may be available for expression in some cells and not others, or at some time in the organism's development and not others.
- Genes may, under some conditions, be amplified or made more available than usual.
- Changes in the physical arrangement of DNA, such as levels of DNA packing, affect gene expression. For example, genes in heterochromatin and mitotic chromosomes are not expressed.

The structural organization of an organism's genome is also somewhat plastic; movement of DNA within the genome and chemical modification of DNA influence gene expression.

A. Repetitive DNA and noncoding sequences account for much of a eukaryotic genome

DNA in eukaryotic genomes is organized differently from that in prokaryotes.

- In prokaryotes, most DNA codes for protein (mRNA), tRNA or rRNA, and coding sequences are uninterrupted. Small amounts of noncoding DNA consist mainly of control sequences, such as promoters.
- In eukaryotes, most DNA does *not* encode protein or RNA, and coding sequences may be interrupted by long stretches of noncoding DNA (introns). Certain DNA sequences may be present in multiple copies.

1. Tandemly repetitive DNA

About 10–25% of total DNA in higher eukaryotes is *satellite DNA* that consists of short (five to 10 nucleotides) sequences that are tandemly repeated thousands of times.

Satellite DNA = In eukaryotic chromosomes, highly repetitive DNA consisting of short unusual nucleotide sequences that are tandemly repeated thousands of times.

- Called satellite DNA because its unusual nucleotide ratio gives it a density different from the rest of the cell's DNA. Thus, during ultracentrifugation, satellite DNA separates out in a cesium chloride gradient as a “satellite” band separate from the rest of the DNA.
- Is not transcribed and its function is not known. Since most satellite DNA in chromosomes is located at the tips and the centromere, scientists speculate that it plays a structural role during chromosome replication and chromatid separation in mitosis and meiosis.

It is known that short tandem repeats called telomeres—at the ends of eukaryotic chromosomes—are important in maintaining the integrity of the lagging DNA strand during replication.

Telomere = Series of short tandem repeats at the ends of eukaryotic chromosomes; prevents chromosomes from shortening with each replication cycle

- Before an Okazaki fragment of the lagging DNA strand can be synthesized, RNA primers must be produced on a DNA template ahead of the sequence to be replicated.
- Since such a template is not possible for the end of a linear DNA molecule, there must be a mechanism to prevent DNA strands from becoming shorter with each replication cycle.
- This end-replication problem is solved by the presence of special repeating telomeric sequences on the ends of linear chromosomes.
- To compensate for the loss of telomeric nucleotides that occurs each replication cycle, the enzyme *telomerase* periodically restores this repetitive sequence to the ends of DNA molecules.
- Telomeric sequences are similar among many organisms and contain a block of G nucleotides. For example, human chromosomes have 250–1500 repetitions of the base sequence TTAGGG (AATCCC on the complementary strand).

There are other highly repetitive sequences in eukaryotic genomes. For example,

- Some are transposons; generally regarded as nonfunctional, they are associated with some diseases (e.g., neurofibromatosis-1 or elephant man's disease and some cancers).
- Mutations can extend the repetitive sequences normally found within the boundary of genes and cause them to malfunction. (e.g., fragile X syndrome and Huntington's disease.)

2. Interspersed repetitive DNA

Eukaryotes also possess large amounts (25–40% in mammals) of repeated units, hundreds or thousands of base pairs long, dispersed at random intervals throughout the genome.

B. Gene families have evolved by duplication of ancestral genes

Most eukaryotic genes are *unique sequences* present as single copies in the genome, but some genes are part of a multigene family.

Multigene family = A collection of genes that are similar or identical in sequence and presumably of common ancestral origin; such genes may be clustered or dispersed in the genome.

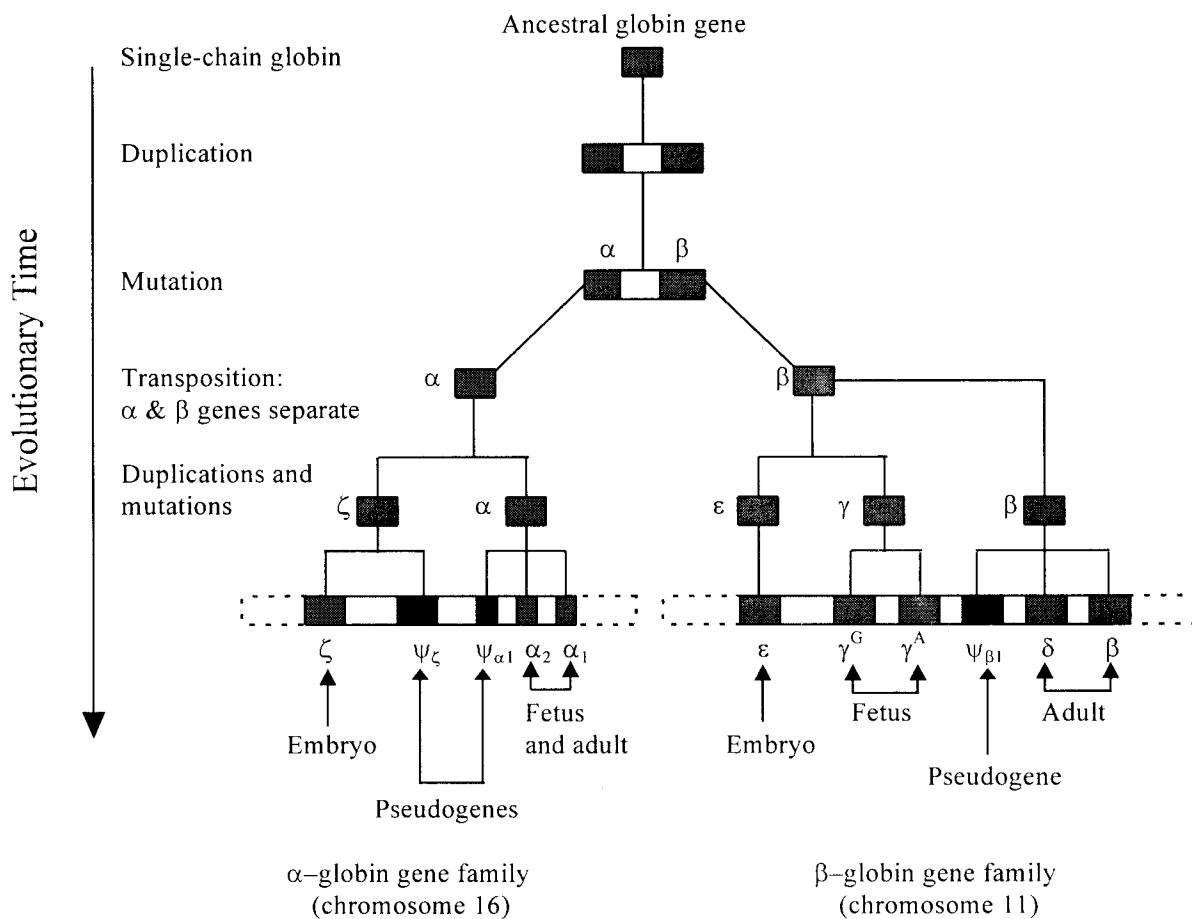
Families of *identical* genes:

- Probably arise from a single ancestral gene that has undergone repeated duplication. Such *tandem gene duplication* results from mistakes made during DNA replication and recombination.
- Are usually clustered and almost exclusively genes for RNA products. (One exception is the gene family coding for histone proteins.)
- Include genes for the major rRNA molecules; huge tandem repeats of these genes enable cells to make millions of ribosomes during active protein synthesis (see Campbell, Figure 19.2).

Families of *nonidentical* genes:

- Arise over time from mutations that accumulate in duplicated genes.
- Can be clustered on the same chromosome or scattered throughout the genome. (Note that for various reasons, gene sequences in tandem arrays on the same chromosome tend to stay very similar to one another. Transposition events that translocate variants of duplicated genes to different chromosomes, help stabilize their differences and thus promote diversity.)
- May include pseudogenes or nonfunctional versions of the duplicated gene.

Pseudogene = Nonfunctional gene that has a DNA sequence similar to a functional gene; but as a consequence of mutation, lacks sites necessary for expression.



A good example of how multigene families can evolve from a single ancestral gene is the globin gene family—actually two related families of genes that encode globins, the α and β polypeptide subunits of hemoglobin.

Based on amino acid homologies, the evolutionary history has been reconstructed as follows:

- The original α and β genes evolved from duplication of a common ancestral globin gene. Gene duplication was followed by mutation.
- Transposition separated the α globin and β globin families, so they exist on different chromosomes.
- Subsequent episodes of gene duplication and mutation resulted in new genes and pseudogenes in each family.

The consequence is that each globin gene family consists of a group of similar, but not identical genes clustered on a chromosome.

- In each gene family, the genes are arranged in order of their expression. During development, genes are turned on or off in response to the organism's changing environment as it develops from an embryo into a fetus, and then into an adult.
- At all times during development, functional hemoglobin consists of two α -like and two β -like polypeptides.
 - During embryogenesis, the ζ and ϵ forms predominate.
 - About 10 weeks after fertilization, the products of the α -globin genes replace that of the ζ gene, and the γ -chain gene products— γ^G and the γ^A globins—become more prevalent.
 - Prior to birth, the β -gene product begins to replace the γ^G and γ^A globins, so by six months of age, the adult β -like globins, δ and β , are present.
- In humans, embryonic and fetal hemoglobins have a higher affinity for oxygen than the adult forms, allowing efficient oxygen exchange between mother and developing fetus.

C. Gene amplification, loss, or rearrangement can alter a cell's genome

1. Gene amplification and selective gene loss

Gene amplification may temporarily increase the number of gene copies at certain times in development.

Gene amplification = Selective synthesis of DNA, which results in multiple copies of a single gene.

- For example, amphibian rRNA genes are selectively amplified in the oocyte, which:
 - Results in a million or more additional copies of the rRNA genes that exist as extrachromosomal circles of DNA in the nucleoli.
 - Permits the oocyte to make huge numbers of ribosomes that will produce the vast amounts of proteins needed when the egg is fertilized.
- Gene amplification occurs in cancer cells exposed to high concentrations of chemotherapeutic drugs.
 - Some cancer cells survive chemotherapy, because they contain amplified genes conferring drug resistance.
 - Increased drug resistance can be created experimentally by exposing a cell population to increasing drug doses and artificially selecting for surviving cells that have amplified drug-resistance genes.

Genes may also be selectively lost in certain tissues by elimination of chromosomes.

Chromosome diminution = Elimination of whole chromosomes or parts of chromosomes from certain cells early in embryonic development.

- For example, chromosome diminution occurs in gall midges during early development; all but two cells lose 32 of their 40 chromosomes during the first mitotic division after the 16-cell stage.
- The two cells that retain the complete genome are germ cells that will produce gametes in the adult. The other 14 cells become somatic cells with only eight chromosomes.

2. Rearrangements in the genome

Substantial stretches of DNA can be re-shuffled within the genome; these rearrangements are more common than gene amplification or gene loss.

a. Transposons

All organisms probably have transposons that move DNA from one location to another within the genome (see Campbell, Chapter 17). Transposons can rearrange the genome by:

- Inserting into the middle of a coding sequence of another gene; it can prevent the interrupted gene from functioning normally (see Campbell, Figure 19.4).
- Inserting within a sequence that regulates transcription; the transposition may increase or decrease a protein's production.
- Inserting its own gene just downstream from an active promoter that activates its transcription.

Retrotransposons = Transposable elements that move within a genome by means of an RNA intermediate (see Campbell, Figure 19.5).

Retrotransposons insert at another site by utilizing reverse transcriptase to convert back to DNA.

b. Immunoglobulin genes

During cellular differentiation in mammals, permanent rearrangements of DNA segments occur in those genes that encode antibodies, or *immunoglobulins*.

Immunoglobulins = A class of proteins (antibodies) produced by B lymphocytes that specifically recognize and help combat viruses, bacteria, and other invaders of the body. Immunoglobulin molecules consist of:

- Four polypeptide chains held together by disulfide bridges
- Each chain has two major parts:
 - A *constant region*, which is the same for all antibodies of a particular class
 - A *variable region*, which gives an antibody the ability to recognize and bind to a specific foreign molecule

B lymphocytes, which produce immunoglobulins, are a type of white blood cell found in the mammalian immune system.

- The human immune system contains millions of subpopulations of B lymphocytes that produce different antibodies.
- B lymphocytes are very specialized; each differentiated cell and its descendants produce only one specific antibody.

Antibody specificity and diversity are properties that emerge from the unique organization of the antibody gene, which is formed by a rearrangement of the genome during B cell development (see Campbell, Figure 19.6).

- As an unspecialized cell differentiates into a B lymphocyte, its antibody gene is pieced together randomly from several DNA segments that are physically separated in the genome.

- In the genome of an embryonic cell, there is an intervening DNA sequence between the sequence coding for an antibody's constant region and the site containing hundreds of coding sequences for the variable regions.
- As a B cell differentiates, the intervening DNA is deleted, and the DNA sequence for a variable region connects with the DNA sequence for a constant region, forming a continuous nucleotide sequence that will be transcribed.
- The primary RNA transcript is processed to form mRNA that is translated into one of the polypeptides of an antibody molecule.
- Antibody variation results from:
 - Different combinations of variable and constant regions in the polypeptides
 - Different combinations of polypeptides

III. The Control of Gene Expression

A. Each cell of a multicellular eukaryote expresses only a small fraction of its genes

Cellular differentiation = Divergence in structure and function of different cell types, as they become specialized during an organism's development

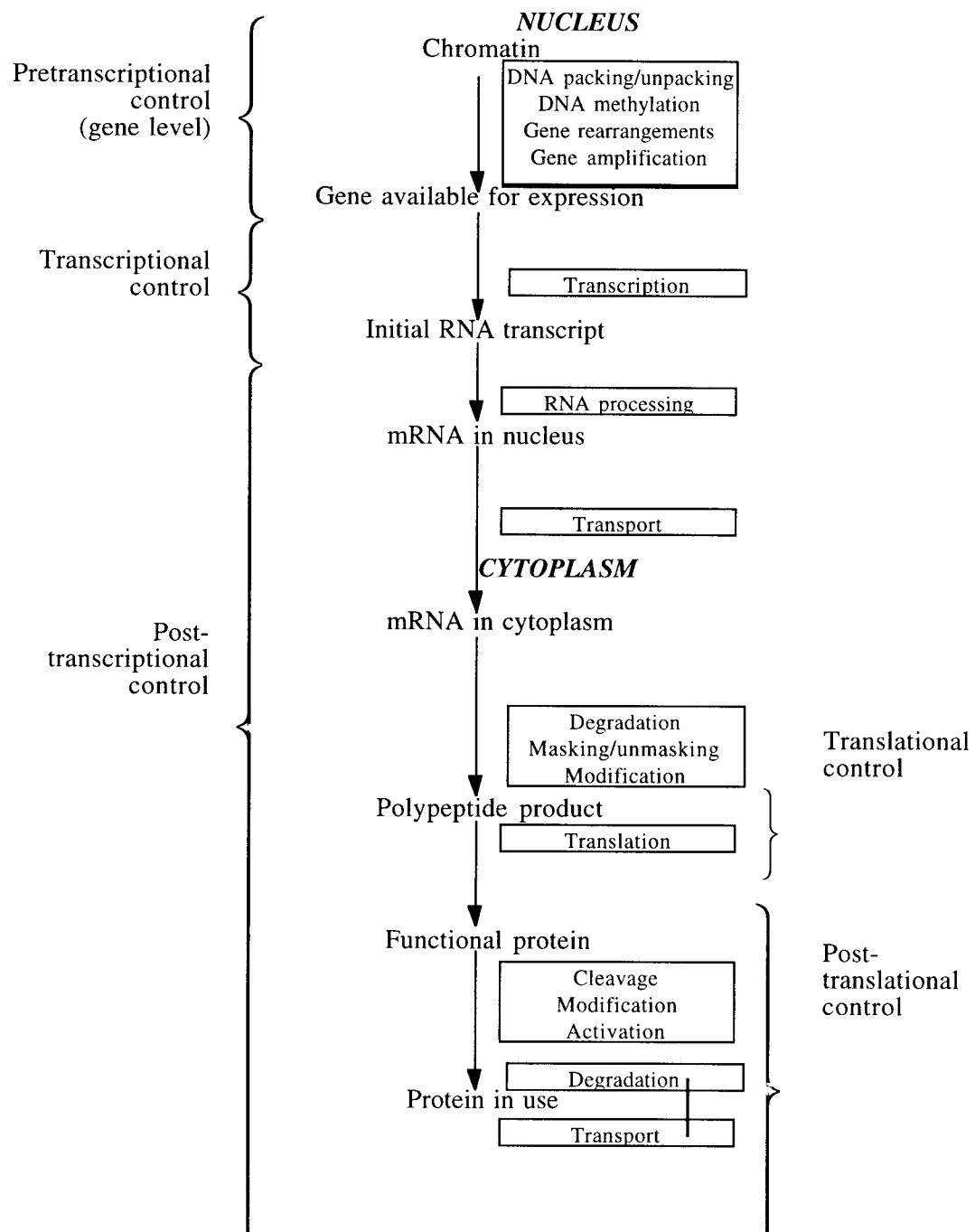
- Cell differentiation requires that gene expression must be regulated on a long-term basis.
- Highly specialized cells, such as muscle or nerve, express only a small percentage of their genes, so transcription enzymes must locate the right genes at the right time.
- Uncontrolled or incorrect gene action can cause serious imbalances and disease, including cancer. Thus, eukaryotic gene regulation is of interest in medical as well as basic research.

DNA-binding proteins regulate gene activity in all organisms—prokaryotes as well as eukaryotes.

- Usually, it is DNA transcription that is controlled.
- Eukaryotes have more complex chromosomal structure, gene organization and cell structure than prokaryotes, which offer added opportunities for controlling gene expression.

B. The control of gene expression can occur at any step in the pathway from gene to functional protein: *an overview*

Complexities in chromosome structure, gene organization and cell structure provide opportunities for the control of gene expression in eukaryotic cells. The steps of gene expression where gene regulation can occur are outlined below (see also Campbell, Figure 19.7).



C. Chromatin modifications affect the availability of genes for transcription

Chromatin organization:

- Packages DNA into a compact form that can be contained by the cell's nucleus.
- Controls which DNA regions are available for transcription.
 - Condensed heterochromatin is not expressed.
 - A gene's location relative to nucleosomes and to scaffold attachment sites influences its expression.

Chemical modifications of chromatin play key roles in both chromatin structure and the regulation of transcription.

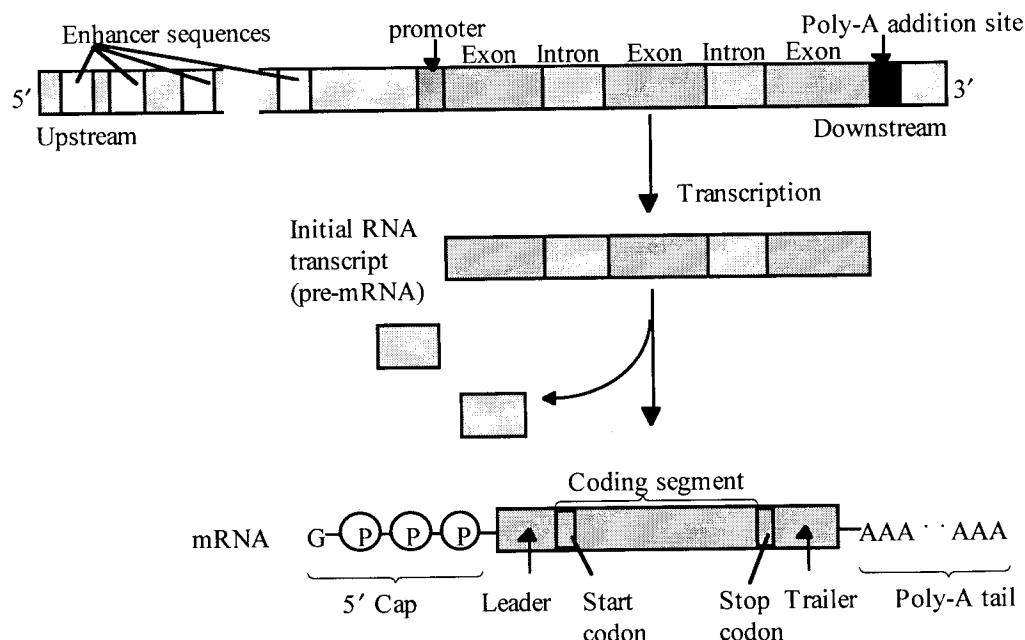
1. DNA methylation

DNA methylation = The addition of methyl groups ($-\text{CH}_3$) to bases of DNA, after DNA synthesis

- Most plant and animal DNA contains methylated bases (usually cytosine); about 5% of the cytosine residues are methylated.
- May be a cellular mechanism for long-term control of gene expression. When researchers examine the same genes from different types of cells, they find:
 - Genes that are not expressed (e.g., Barr bodies) are more heavily methylated than those that are expressed.
 - Drugs that inhibit methylation can induce gene reactivation, even in Barr bodies.
- In vertebrates, DNA methylation reinforces earlier developmental decisions made by other mechanisms.
 - For example, genes must be selectively turned on or off for normal cell differentiation to occur. DNA methylation ensures that once a gene is turned off, it stays off.
 - DNA methylation patterns are inherited and thus perpetuated as cells divide; clones of a cell lineage forming specialized tissues have a chemical record of regulatory events that occurred during early development.

2. Histone acetylation

- Acetylation enzymes attach $-\text{COCH}_3$ groups to certain amino acids of histone proteins



- Acetylated histone proteins have altered conformation and bind to DNA less tightly; as a result, transcription proteins have easier access to genes in the acetylated region.

D. Transcription initiation is controlled by proteins that interact with DNA and with each other

1. Organization of a typical eukaryotic gene

The following is a brief review of a eukaryotic gene and its transcript (see also Campbell, Figure 19.8).

Eukaryotic genes:

- Contain introns, noncoding sequences that intervene within the coding sequence
- Contain a promoter sequence at the 5' upstream end; a transcription initiation complex, including RNA polymerase, attaches to a promoter sequence and transcribes introns along with the coding sequences, or exons
- May be regulated by control elements, other noncoding control sequences that can be located thousands of nucleotides away from the promoter

Control element = Segments of noncoding DNA that help regulate the transcription of a gene by binding specific proteins (transcription factors).

The primary RNA transcript (pre-mRNA) is processed into mature mRNA by:

- Removal of introns
- Addition of a modified guanosine triphosphate cap at the 5' end
- Addition of a poly-A tail at the 3' end

2. The roles of transcription factors

In both prokaryotes and eukaryotes, transcription requires that RNA polymerase recognize and bind to DNA at the promoter. However, transcription in eukaryotes requires the presence of proteins known as *transcription factors*; transcription factors augment transcription by binding:

- Directly to DNA (protein-DNA interactions)
- To each other and/or to RNA polymerase (protein-protein interactions)

Eukaryotic RNA polymerase cannot recognize the promoter without the help of a specific transcription factor that binds to the TATA box of the promoter.

Transcription factors and the TATA box were discussed with the material in Chapter 17. This is another point in the course where students, if they compartmentalize their learning, may not relate your discussion of transcriptional control to what they have already learned about transcription. Experiment with organizing the presentation of this material. It may be more prudent to cover transcriptional control of gene expression (both prokaryotic and eukaryotic) with the lectures on protein synthesis. If not, you should probably begin your lecture with a review of transcription, especially RNA polymerase binding and initiation of transcription.

Associations between transcription factors and control elements (specific segments of DNA) are important transcriptional controls in eukaryotes.

- Proximal control elements are close to or within the promoter; distal control elements may be thousands of nucleotides away from the promoter or even downstream from the gene.
- Transcription factors known as activators bind to enhancer control elements to stimulate transcription.
- Transcription factors known as repressors bind to silencer control elements to inhibit transcription

How do activators stimulate transcription?

- One hypothesis is that a hairpin loop forms in DNA, bringing the activator bound to an enhancer into contact with other transcription factors and polymerase at the promoter (see Campbell, Figure 19.9).
- Diverse activators may selectively stimulate gene expression at appropriate stages in cell development.

The involvement of transcription factors in eukaryotes offers additional opportunities for transcriptional control. This control depends on selective binding of specific transcription factors to specific DNA sequences and/or other proteins; the highly selective binding depends on molecular structure.

- There must be a complementary fit between the surfaces of a transcription factor and its specific DNA-binding site.
- Hundreds of transcription factors have been discovered; and though each of these proteins is unique, many recognize their DNA-binding sites with only one of a few possible structural motifs or *domains* containing α helices or β sheets (see Campbell, Figure 19.10).

3. Coordinately controlled genes

Coordinately controlled genes are arranged differently in a eukaryotic chromosome than in prokaryotic genomes.

- Prokaryotic genes that are turned on and off together are often clustered into operons; these adjacent genes share regulatory sites located at one end of the cluster. All genes of the operon are transcribed into one mRNA molecule and are translated together.
- Eukaryotic genes coding for enzymes of a metabolic pathway are often scattered over different chromosomes. Even functionally related genes on the same chromosome have their own promoters and are individually transcribed.

Eukaryotic genes can be coordinately expressed, even though they may be scattered throughout the genome.

- Coordinately controlled genes are each associated with specific regulatory DNA sequences or enhancers. These sequences are recognized by a single type of transcription factor that activates or represses a group of genes in synchrony.
- Examples of coordinate gene expression in eukaryotes include:
 - *Heat shock response.* Exposure to high temperature activates genes coding for heat shock proteins, which help stabilize and repair heat-denatured proteins in the cell.
 - *Steroid hormone action.* Steroids activate protein receptors, and the protein-receptor complex, in turn, activates genes. In a secondary response, proteins produced this way can activate another group of genes (see Campbell Chapter 45).
 - *Cellular differentiation.* During cellular differentiation, coordinately controlled genes producing particular sets of proteins are switched on and off.

E. Posttranscriptional mechanisms play supporting roles in the control of gene expression

Transcription produces a primary transcript, but gene expression—the production of protein, tRNA, or rRNA—may be stopped or enhanced at any posttranscriptional step. Because eukaryotic cells have a nuclear envelope, translation is segregated from transcription. This offers additional opportunities for controlling gene expression.

1. Regulation of mRNA degradation

Protein synthesis is also controlled by mRNA's lifespan in the cytoplasm.

- Prokaryotic mRNA molecules are degraded by enzymes after only a few minutes. Thus, bacteria can quickly alter patterns of protein synthesis in response to environmental change.
- Eukaryotic mRNA molecules can exist for several hours or even weeks.
- The longevity of a mRNA affects how much protein synthesis it directs. Those that are viable longer can produce more of their protein.
- For example, long-lived mRNAs for hemoglobin are repeatedly translated in developing vertebrate red blood cells.

2. Control of translation

Gene expression can also be regulated by mechanisms that control translation of mRNA into protein. Most of these translational controls repress initiation of protein synthesis; for example.

- Binding of translation repressor protein to the 5'-end of a particular mRNA can prevent ribosome attachment.
- Translation of all mRNAs can be blocked by the inactivation of certain initiation factors. Such global translational control occurs during early embryonic development of many animals.
- Prior to fertilization, the ovum produces and stores inactive mRNA to be used later during the first embryonic cleavage.
- The inactive mRNA is stored in the ovum's cytosol until fertilization, when the sudden activation of an initiation factor triggers translation.
- Delayed translation of stockpiled mRNA allows cells to respond quickly with a burst of protein synthesis when it is needed.

3. Protein processing and degradation

Posttranslational control is the last level of control for regulating gene expression.

- Many eukaryotic polypeptides must be modified or transported before becoming biologically active. Such modifications include:
 - Adding phosphate groups
 - Adding chemical groups, such as sugars
 - Dispatching proteins targeted by signal sequences for specific sites
- Selective degradation of particular proteins and regulation of enzyme activity are also control mechanisms of gene expression.
 - Cells attach ubiquitin to proteins to mark them for destruction
 - Proteasomes recognize the ubiquitin and degrade the tagged protein (see Campbell, Figure 19.11)
 - Mutated cell-cycle proteins that are impervious to proteasome degradation can lead to cancer

IV. The Molecular Biology of Cancer

A. Cancer results from genetic changes that affect the cell cycle

Cancer is a variety of diseases in which cells escape from the normal controls on growth and division—the cell cycle—and it can result from mutations that alter normal gene expression in somatic cells. These mutations:

- Can be random and spontaneous
- Most likely occur as a result of environmental influences, such as:
 - Infection by certain viruses
 - Exposure to carcinogens

Carcinogens = Physical agents such as X-rays and chemical agents that cause cancer by mutating DNA

Whether cancer is caused by physical agents, chemicals or viruses, the mechanism is the same—the activation of *oncogenes* that are either native to the cell or introduced in viral genomes.

Oncogene = Cancer-causing gene

- Discovered during the study of tumors induced by specific viruses
- Harold Varmus and Michael Bishop won a Nobel Prize for their discovery of oncogenes in RNA viruses (retroviruses) that cause uncontrolled growth of infected cells in culture.

Researchers later discovered that some animal genomes, including human, contain genes that closely resemble viral oncogenes. These proto-oncogenes normally regulate growth, division and adhesion in cells.

Proto-oncogenes = Gene that normally codes for regulatory proteins controlling cell growth, division and adhesion, and that can be transformed by mutation into an oncogene.

Three types of mutations can convert proto-oncogenes to oncogenes:

1. *Movement of DNA within the genome.* Malignant cells frequently contain chromosomes that have broken and rejoined, placing pieces of different chromosomes side-by-side and possibly separating the oncogene from its normal control regions. In its new position, an oncogene may be next to active promoters or other control sequences that enhance transcription. Abnormal expression of an oncogene may occur if the oncogene is transposed to a new locus that has a highly active promoter.
2. *Gene amplification.* Sometimes more copies of oncogenes are present in a cell than is normal.
3. *Point mutation.* A slight change in the nucleotide sequence might produce a growth-stimulating protein that is more active or more resistant to degradation than the normal protein.

In addition to mutations affecting growth-stimulating proteins, changes in *tumor-suppressor genes* coding for proteins that normally *inhibit* growth can also promote cancer.

- The protein products of tumor-suppressor genes have several functions:
 - Cooperate in DNA repair (helping obviate cancer-causing mutations)
 - Control cell anchorage (cell-cell adhesion; cell interaction with extracellular matrix)
 - Components of cell-signaling pathways that inhibit the cell-cycle

B. Oncogene proteins and faulty tumor-suppressor proteins interfere with normal signaling pathways

Mutation in the *ras* proto-oncogene and the *p53* tumor suppressor gene are very common in human cancers; the frequency of mutation for *ras* is about 30 % and close to 50% for *p53*.

The *ras* protein is a G protein that relays a growth signal from a growth factor receptor to a cascade of protein kinases. The cellular response is the synthesis of a protein that stimulates the cell cycle (see Campbell Fig 19.13a).

- Under normal conditions, the pathway will not operate unless triggered by the appropriate growth factor.
- A mutated *ras* gene can produce a hyperactive version of the *ras* protein that stimulates the signal transduction cascade on its own, leading to excessive cell division.

The tumor-suppressor protein encoded by the wild-type *p53* gene is a transcription factor of several genes that promotes the synthesis of growth-inhibiting proteins. Expression of this protective “guardian angel” protein prevents a cell from passing on mutations due to DNA damage in three ways:

- Activates the *p21* gene, whose product allosterically binds to cyclin-dependent kinases, halting the cell cycle.
- Activates genes directly involved in DNA repair
- When DNA damage is irreparable, activates “suicide” genes, whose products cause cell death (apoptosis)

Mutations in the *p53* gene can lead to excessive cell growth and cancer.

- For example, mutant tumor-suppressor genes are associated with inherited forms of colorectal cancer, Wilm’s tumor, and breast cancer.

C. Multiple mutations underlie the development of cancer

More than one somatic mutation is probably needed to transform normal cells into cancerous cells. One of the best understood examples is colorectal cancer (see Campbell, Figure 19.14).

- Development of metastasizing colorectal cancer is gradual, and the first sign is unusually rapid cell division of apparently normal cells in the colon lining; a benign tumor (polyp) appears, and eventually a malignant tumor may develop.
- During this process, mutations in oncogenes and tumor-suppressor genes gradually accumulate. After a number of genes have changed, the tumor becomes malignant.
- About half a dozen changes must occur at the DNA level for the cell to become fully cancerous: usually, the appearance of at least one active oncogene and the mutation or loss of several tumor-suppressor genes.

Viruses play a role in about 15% of human cancer cases worldwide, e.g., some types of leukemia, liver cancer, cervical cancer.

- Viruses might add oncogenes to cells, disrupt tumor-suppressor genes DNA, or convert proto-oncogenes to oncogenes.

Breast cancer, the second most common type of cancer in women, is associated with somatic mutations of tumor-suppressor genes.

- Inherited breast cancer accounts for 5–10% of all breast cancer cases.
 - Mutations in either the BRCA1 or BRCA2 (stands for BReast CAncer) gene increase the risk of developing breast cancer. BRCA1 mutations also increase the risk of ovarian cancer.
 - Another locus, accounts for most of the remaining breast cancer cases linked to family history.

The study of genes associated with inherited cancer may lead to early diagnosis and treatment.

REFERENCES

Alberts, B., D. Bray, J. Lewis, M. Raff, K. Roberts and J.D. Watson. *Molecular Biology of the Cell*. 3rd ed. New York: Garland, 1994.

Campbell, N., et al. *Biology*. 5th ed. Menlo Park, California: Benjamin/Cummings, 1998.

Cavenee, W.K., and R.L. White. “The Genetic Basis of Cancer.” *Scientific American*, March, 1995.