Welcome to Part 3 of Bio 219

Lecturer – David Ray

Contact info:
Office hours – 1:00-2:00 pm MTW
Office location – LSB 5102
Office phone – 293-5102 ext 31454
E-mail – david.ray@mail.wvu.edu

Lectures and other resources are available online at
http://www.as.wvu.edu/~dray.
Go to ‘Courses’ link
Chapter 12:
The Cell Nucleus and Control of Gene Expression
The Cell Nucleus...

- Every somatic cell contains the same genetic information regardless of whether it is expressed the same way.
- Just as in a building under construction, all workers have access to the complete set of blueprints.
- However, the workers constructing the floors only use the information related to their word and the workers installing the wiring only use the information related to their project.
- Brain cells only express genes related to brain activity and liver cells only express the genes related to their activity.
The nucleus of most cells is typically amorphous and consists of:

- The chromosomes (in the form of chromatin)
- Nucleoli
- Nucleoplasm
- The nuclear matrix

All surrounded by the nuclear envelope
The Cell Nucleus...

• The nuclear envelope is complex; consists of several distinct components: 2 cellular membranes arranged parallel to one another & separated by 10 - 50 nm
  – Separates genetic material in nucleus from cytoplasm - important distinction between prokaryotes & eukaryotes; its appearance is an evolutionary landmark
• 1. Serves as barrier; keeps ions, solutes, macromolecules from passing between nucleus & cytoplasm
The Cell Nucleus...

- The nuclear envelope...
  - B. Fused at sites forming circular pores that contain complex assemblies of proteins
    - 1. Average mammalian cell: contains ~3000 nuclear pores
    - 2. Pore density: ~2-4/μm² (metabolically inactive bird erythrocyte) to >60/μm² (active oocyte)
  - C. Outer membrane is generally studded with ribosomes & is often seen to be continuous with RER membrane; the space between the membranes is continuous with the ER lumen
The Cell Nucleus...

• The nuclear envelope...
  - D. Inner surface of the nuclear envelope of animal cells is bound by integral membrane proteins to a thin filamentous meshwork (nuclear lamina)
    • 1. Provides mechanical support to nuclear envelope & serves as a site of attachment for chromatin fibers
    • 2. Different mutations in 1 lamin gene (LMNA) are responsible for several distinct human diseases, e.g., a rare form of muscular dystrophy and progeria
The Cell Nucleus…

• The nuclear pore complexes…
  – the gateways across the barrier of the nuclear envelope
    • Many molecules, including RNAs & proteins, are transported in both directions across the nuclear envelope
      – 1. Proteins involved in replication & transcription
      – 2. Messenger RNAs, tRNAs, & ribosomal subunits are made in nucleus & moved through nuclear pores to cytoplasm
      – 3. snRNAs move in both directions: made in nucleus, assembled into RNP particles in cytoplasm & shipped back to nucleus where they function in mRNA processing
      – 4. One HeLa cell nucleus must import ~560,000 ribosomal proteins & export ~14,000 ribosomal subunits per minute
The Cell Nucleus...

- The nuclear pore complexes...
  - Particles of varying sizes injected into a cell can be observed passing into the nucleus through nuclear envelope with EM
    - 1. They pass through center of nuclear pores in single-file
    - 2. Even large ribosomal subunits may pass
The Cell Nucleus…

• The nuclear pore complexes…
  - a complex, basketlike apparatus, the NPC; projecting into both the cytoplasm & nucleoplasm
    • 1. NPC is a huge, supramolecular complex with octagonal symmetry
    • 2. Depending on species, NPCs contain ~30 different proteins
    • 3. Most nuclear pore proteins (nucleoporins) are symmetrically positioned & present on both cytoplasmic & nuclear side of structure
The Cell Nucleus…

• The nuclear pore complexes…
  – Proteins to be imported to the nucleus have a stretch of amino acids that acts as a nuclear localization signal (NLS)
    • The sequence enables a protein to pass through nuclear pores & enter nucleus
    • Best studied or classical NLS have 1 or 2 short stretches of positively charged amino
    • Modifying the signal can prevent entry of polypeptides or allow entry of proteins not typically located in the nucleus
The Cell Nucleus...

• The nuclear pore complexes...
  – Transport receptors are proteins that move molecules in and out of the nucleus
    • Importins move macromolecules into the nucleus
    • Exportins move them out
The Cell Nucleus...

- **Nuclear import**
  1. NLS protein binds to *importin* dimer
  2. Trimer binds to cytoplasmic filament
  3. Whole complex is imported through the pore
  4. Ran-GTP causes dissociation of NLS from *importin* complex
  5. Ran-GDP and the *importin* proteins are exported to the cytoplasm
  6. Ran-GDP is reimported to the nucleus and converted back to Ran-GTP
The Cell Nucleus…

• Chromosomes and chromatin…
  – The average human cell has ~6 billion bp divided among 46 chromosomes (0.34 nm/bp x 6 billion bp; 2 meters long) for a diploid, unreplicated number of chromosomes
  • A. An unreplicated chromosome is a single, continuous DNA strand; the larger the chromosome, the longer is the DNA it contains
    – 1. How does it fit into a cell nucleus (10 μm in diameter) & still perform its functions by remaining accessible to enzymes and regulatory proteins?
    – 2. How is the single DNA molecule of each chromosome organized so that it does not become hopelessly tangled with the molecules of the other chromosomes?
    – 3. The answer to both of the above questions is packaging
The Cell Nucleus…

• Chromosomes and chromatin…
  – Chromosomes are made of DNA & associated proteins, which together are called **chromatin** - highly extended nucleoprotein fibers
  – Two major groups of proteins - histones & nonhistone
    • 1. Histones - small, well-defined basic protein group; very high lysine and/or arginine content
    • 2. Nonhistone proteins - many widely diverse structural, enzymatic & regulatory proteins
  – Understanding how histones and DNA interact is the first step in understanding how DNA is packaged.
The Cell Nucleus...

• Chromosomes and chromatin...
  – The orderly packaging of eukaryotic DNA depends on histones,
  – **Nucleosomes**: repeating subunits of DNA and histones
  – Histones are divided into 5 distinct classes distinguished by arginine/lysine ratio & posttranslational (after the proteins are synthesized) modifications (phosphorylation & acetylation)

<table>
<thead>
<tr>
<th>Table 12.1</th>
<th>Calf Thymus Histones</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Histone</strong></td>
<td><strong>Number of Residues</strong></td>
</tr>
<tr>
<td>H1</td>
<td>215</td>
</tr>
<tr>
<td>H2A</td>
<td>129</td>
</tr>
<tr>
<td>H2B</td>
<td>125</td>
</tr>
<tr>
<td>H3</td>
<td>135</td>
</tr>
<tr>
<td>H4</td>
<td>102</td>
</tr>
</tbody>
</table>

*Unit evolutionary period: the time for a protein's amino acid sequence to change by 1 percent after two species have diverged.
The Cell Nucleus…

• Chromosomes and chromatin…
  – Histone amino acid sequences particularly H3 & H4 are very conserved & have changed very little over long periods of evolutionary time - H4 of peas & cows vary in only 2 amino acids out of 102
  – Why?
The Cell Nucleus...

• Chromosomes and chromatin...
  – 1. Histones interact with the DNA backbone, which is identical in all organisms
  – 2. Nearly all amino acids in a histone molecule interact with either DNA or another histone; thus, very few amino acids in a histone can be replaced with another without severely affecting its function
The Cell Nucleus...

• Chromosomes and chromatin...
  – Nucleosomes
    • 1. Each nucleosome contains a nucleosome core particle - 146 bp of supercoiled DNA
    • 2. DNA wraps almost 2X around disc-shaped, 8 histone complex, the histone core
    • 3. The histone core of each nucleosome consists of 2 molecules each of H2A, H2B, H3 & H4; assembled into an octamer; human cell contains ~300 million histones
    • 4. The remaining histone, type H1, normally resides outside the nucleosome core particle; it binds to part of the linker DNA that connects adjacent nucleosome core particles; thus called linker histone
The Cell Nucleus…

• Chromosomes and chromatin…
  – **Nucleosomes**
    • 5. Together H1 protein & the histone octamer interact with ~168 bp of DNA
    • 6. When H1-depleted chromatin is observed under EM, nucleosome core particles & naked linker DNA can be seen as separate elements, which together appear like beads on a string.
  – This first level of packaging condenses DNA by a 7:1 ratio
The Cell Nucleus…

• Chromosomes and chromatin…
  – Higher level structure
    • Nucleosomes are further organized into 30 nm fibers
    • The assembly of the 30 nm fiber increases packing ratio another 6-fold (~40-fold all together)
The Cell Nucleus…

• Chromosomes and chromatin…
  – Maintenance of 30-nm fiber depends on the interaction between histone molecules of neighboring nucleosomes; linker histones & core histones
    • 1. Extract H1 linker histones from compacted chromatin —> 30-nm fibers uncoil to form thinner, more extended beaded filament
    • 2. Core histones of adjacent nucleosomes may interact with one another via their long, flexible tails
The Cell Nucleus…

• Chromosomes and chromatin…
  • 3. H4 histone N-terminal tail from 1 nucleosome core particle can reach out & make extensive contact with both linker DNA between nucleosome particles & H2A/H2B dimer of adjacent particles
  • 4. Such interactions may mediate folding of nucleosomal filament into a thicker fiber
  • 5. Chromatin fibers prepared with histones lacking their tails cannot fold into higher-order fibers
The Cell Nucleus...

• Chromosomes and chromatin...
  – Next level of packaging – 30-nm chromatin fiber is gathered into series of large, supercoiled loops or domains that may be compacted into even thicker (80 – 100 nm) fibers
  • DNA loops begin & end with AT-rich sequences tethered to proteins of an organized nuclear scaffold or matrix
The Cell Nucleus…

- Chromosomes and chromatin…
  - Chromatin loops are normally spread out within nucleus & cannot be visualized, their presence can be revealed under certain circumstances
  - Treat mitotic chromosomes with solutions that extract histones —> histone-free DNA can be seen to extend outward as loops from a protein scaffold
The Cell Nucleus…

• Chromosomes and chromatin…
  – The **mitotic chromosome** is the organization level during cellular reproduction
  – 1 μm of chromosome length has ~1 cm of DNA; the ultimate in chromosome compactness with a packing ratio of 10,000:1; the compaction occurs by poorly understood processes
  – *movie*

Figure 12-14 Cell and Molecular Biology, 4/e (© 2005 John Wiley & Sons)
The Cell Nucleus…

• Heterochromatin and Euchromatin…
  – During normal cell function chromatin is less organized; ~10% of chromatin, however, remains condensed
    • A. Heterochromatin – chromatin that remains compacted during interphase; this compacted chromatin is found at nuclear periphery
    • B. Euchromatin – chromatin that returns to dispersed state after each mitosis
The Cell Nucleus…

• Heterochromatin and Euchromatin
  – Heterochromatin- divided into 2 classes depending on whether it's permanently or transiently compacted
    • A. Constitutive heterochromatin - stays condensed in all cells at all times; permanently silenced DNA
      – 1. In mammalian cells - most is found in regions flanking centromeres in a few other sites (Y chromosome distal arm in males)
      – 2. In many plants, telomeres (chromosome ends) also consist of blocks of constitutive heterochromatin
      – 3. Consists primarily of repeated DNA sequences & contains relatively few genes
The Cell Nucleus…

- Heterochromatin and Euchromatin
  - Heterochromatin…
    - A. Constitutive heterochromatin …
      - 4. If normally active genes move into site adjacent to constitutive heterochromatin (change position via transposition or translocation) —> tend to become transcriptionally silenced (position effect)
      - 5. May have components whose influence spreads out a certain distance, affecting nearby genes
      - 6. The spread of heterochromatin along a chromosome is normally blocked by specialized barrier sequences in genome
The Cell Nucleus…

• Heterochromatin and Euchromatin
  – Heterochromatin…
    • B. Facultative heterochromatin - specifically inactivated during certain phases of organism's life or in certain types of differentiated cells;
    • one example: X chromosome in female mammals
      – 1. Male cells have a tiny Y chromosome & much larger X chromosome; since X & Y chromosomes have only a few genes in common, males have a single copy of most genes carried on sex chromosomes
      – 2. In female mammals, only one X chromosome is transcriptionally active; why?
The Cell Nucleus…

• Heterochromatin and Euchromatin
  – Heterochromatin…
    • Facultative heterochromatin -
      – 3. The other X chromosome is condensed as heterochromatic clump (Barr body) early in embryonic development
      – 4. Barr body ensures that cells of both males & females have the same number of active X chromosomes & thus synthesize equivalent amounts of products encoded by X-linked genes
The Cell Nucleus…

• Heterochromatin and Euchromatin
  – X chromosome inactivation – The Lyon hypothesis:
    • A. Heterochromatization of X chromosome in female mammals occurs during early embryonic development & leads to inactivation of genes on that chromosome
The Cell Nucleus…

• Heterochromatin and Euchromatin
  – X chromosome inactivation – The Lyon hypothesis:
    • B. Heterochromatization in embryo is random process in the sense that the paternally-derived & maternally-derived X chromosomes have equal chances of being inactivated *in any given cell*
      – Once X chromosome is inactivated, same X chromosome is inactivated in all of the cell's descendants
The Cell Nucleus…

• Heterochromatin and Euchromatin
  – X chromosome inactivation – The Lyon hypothesis:
    • C. Heterochromatized X chromosome is reactivated in germ cells before meiosis (the creation of gametes), so both X chromosomes are active during oogenesis; all gametes get a euchromatic X chromosome
The Cell Nucleus…

• Heterochromatin and Euchromatin
  – The consequences of X-inactivation
    • Adult mammalian females are **genetic mosaics** (with different alleles functioning in different cells)
      – A. This is true since paternal & maternal X chromosomes may have different alleles for same trait
      – B. X-linked pigment genes in cats – calico
      – C. Pigmentation genes in humans are not found on X chromosome so there are no calico women, **but**….
The Cell Nucleus…

• The Histone Code
  – Cells contain a wide array of enzymes that can add or remove chemical groups to or from amino acid residues in the histone tails
    • A. The amino acid residues are most susceptible to acetylation, methylation or phosphorylation
    • B. A hypothesis called the histone code has emerged in the last few years
The Cell Nucleus...

• The Histone Code
  – B. The histone code hypothesis
    – 1. The state & activity of a particular region of chromatin depends upon the specific modifications, or combination of modifications, to the histone tails in that region
    – 2. The pattern of modifications on the tails of the core histones contains encoded information governing the properties of the nucleosomes containing them
The Cell Nucleus...

• The Histone Code
  – C. Two interrelated chromatin properties were shown to depend upon histone modification patterns
    – 1. The degree of compaction – most importantly, whether a region of chromatin is heterochromatic or euchromatic
    – 2. The likelihood that a gene or cluster of genes will be transcribed

Acetylation decreases the positive charge in the N-terminal tails of histones H3 and H4.

A positive charge is replaced by a neutral acetyl group. This reduces the interaction strength of the nucleosome with the negatively charged phosphate backbone of DNA.

A more open chromatin that can be more easily unwound and is more accessible to the transcriptional machinery.
The Cell Nucleus…

• Mitotic Chromosomes…
  – Interphase chromosome DNA is very dispersed so it can be accessed for replication and transcription
    • A. Mitotic chromosome DNA is in its most highly condensed state & favors delivery of an intact DNA package to each daughter cell
    • B. Mitotic chromosomes are also useful for research because they contain a complete set of a cell's genetic material & are made readily visible by simple techniques
The Cell Nucleus...

• Mitotic Chromosomes...
  – Have characteristic shapes determined by DNA length & centromere position; can be visualized with staining procedures; use karyotypes to screen for chromosomal abnormalities
  – **karyotype** – matched homologous pairs placed in order of decreasing size
The Cell Nucleus…

• Chromosomal aberrations…
  – Inversions – portions of chromosomes that have been reversed in orientation
    • Most of the pericentric (around the centromere) inversions observed in humans do not in themselves give rise to any specific phenotypic abnormalities.
    • However pericentric inversion has been found to be associated with infertility (problems in meiosis)
The Cell Nucleus…

• Chromosomal aberrations…
  – Translocations – portions of chromosomes that are transferred to different chromosomes

  ![Balanced Translocation](image1)
  ![Unbalanced Translocation](image2)

  – Emanuel syndrome – an unbalanced 11;22 translocation
    • cleft palate, heart defects, ear anomalies, genital anomalies in males, muscular hypotonia (low muscle tone) and moderate to severe mental deficiency.
The Cell Nucleus…

- Chromosomal aberrations…
  - Deletions – just what it suggests
    - ~ 50 to 60 children are born with 5p- Syndrome (five p minus) in the U.S. each year
    - aka *Cri du Chat* Syndrome.
    - 5p- Syndrome is characterized at birth by a high pitched cry, low birth weight, poor muscle tone, microcephaly, and potential medical complications.
The Cell Nucleus…

• Chromosomal aberrations…
  – Duplications – repeated portions of chromosomes
    • Individuals may be affected differently depending on the duplicated region and the extent of the duplication
  – Trisomy, monosomy
The Cell Nucleus...

- **Telomeres** - unusual stretches of repeated sequences at DNA molecule tips forming a cap at each end of the chromosome's single, continuous, double-stranded DNA molecule; specific proteins bind there
The Cell Nucleus…

• Telomeres …
  – The sequence below repeated ~500 - 5,000 times in humans; same one seen in all vertebrates & it is similar in most other organisms
    TTAGGG
    AATCCC
  • 1. The high degree of similarity suggests that telomeres have highly conserved function in diverse organisms
  • 2. A number of DNA-binding proteins bind specifically to telomere sequence & are essential for telomere function
The Cell Nucleus…

- Telomeres and the *end-replication problem*…
  - DNA polymerases do not initiate DNA synthesis; they can only add DNA to the 3' end of an existing strand (a primer); they need a primer to start replication at the 5' end of new DNA strand
The Cell Nucleus...

- Telomeres and the **end-replication problem**...
  - 1. Replication is started by synthesis of an RNA primer at 5' end of the new strand;

![Diagram of DNA replication with RNA primers](image)
The Cell Nucleus...

- Telomeres and the end-replication problem...
  - 2. The primer is then removed so that the new strand's 5' end is missing a short piece (~8-12 nucleotides). As a result, the template strand with the unreplicated 3' end overhangs the newly made strand's 5' end.
The Cell Nucleus...

- Telomeres and the **end-replication problem**...
  
  - 3. Rather than existing as an unprotected, single-stranded terminus, the overhanging strand is tucked back into the double-stranded portion of the telomere to form a loop
  
  - 4. This conformation may protect the telomeric end of DNA from proteins that normally recognize single-stranded DNA & trigger a DNA repair response

- If cells were not able to replicate the ends of their DNA, the chromosomes would be expected to become shorter & shorter with each round of cell division
The Cell Nucleus...

• Telomeres and telomerase...
  – **Telomerase** is an enzyme that can add new repeat units to chromosome ends
  – A ribonucleoprotein (made up of RNA and protein)
  – Telomerase is a *reverse transcriptase* that synthesizes DNA using an RNA template, but unlike most reverse transcriptases, the RNA serving as the template is an integral part of enzyme (unusual)
  – Animation
    http://faculty.plattsburgh.edu/donald.slish/Telomerase.html
The Cell Nucleus...

• Telomeres and telomerase...
  – Telomere functions
    • A. Required for complete replication of the chromosome
    • B. Form caps that protect chromosomes from nucleases & other destabilizing influences
    • C. Also prevents chromosome ends from fusing with one another
      – 1. In mice genetically engineered to lack telomerase, many chromosomes undergo end-to-end fusion; this leads to chromosomes being torn apart in subsequent cell divisions (catastrophic)
The Cell Nucleus...

• Telomeres and telomerase...functions
  – A. Dermis fibroblast experiment
    • 1. Plate fibroblasts on nutrient media
    • 2. Remove a fraction & replate, they would once again proliferate,
    • 3. After ~50 – 80 population doublings, the cells stop dividing & eventually die
    • 4. One finds a dramatic decrease in telomere length over time in culture
    • 5. Most cells lack telomerase & are unable to prevent the loss of their chromosome ends
    • 6. Shortening is thought to continue to a critical point (a crisis), when cells exhibit extensive chromosome abnormalities & stop dividing
    • 7. A similar decrease in telomere length is found in somatic cells from an elderly adult as compared with telomeres in corresponding cells from an infant or young child
The Cell Nucleus...

• Telomeres and telomerase...functions
  – B. Unlike somatic cells, germ cells of gonads retain telomerase activity
  – C. Telomere shortening plays a key role in protecting humans from cancer

• A. Malignant cells are cells that have escaped the body's normal growth control & keep dividing indefinitely
  – 1. Unlike normal cells that lack detectable telomerase activity, ~90% of human tumors consist of cells that contain an active telomerase enzyme
  – 2. The other 10% or so have an alternate mechanism based on genetic recombination that maintains telomere length in the absence of telomerase
The Cell Nucleus...

- Centromeres - site of indentation on chromosome surface; constriction marks centromere position
  - A. Human centromeres contain a tandemly repeated, 171 bp DNA sequence (α-satellite DNA) that extends for at least 500 kilobases; they are constitutive heterochromatin
The Cell Nucleus…

• Centromeres …
  – B. Centromere DNA associates with specific proteins that distinguish it from other parts of the chromosome
    • 1. Centromeric chromatin contains a special H3 histone variant (CENP-A), which replaces conventional H3 in many of the nucleosomes
    • 2. Centromeric chromatin also binds specific proteins that serve as attachment sites (kinetochores) for the microtubules that separate chromosomes during cell division
    • 3. Chromosomes lacking a centromere fail to assemble a kinetochore & are lost during cell division
The Cell Nucleus...

• Centromeres …
  – C. Unlike telomeres, centromeric DNA exhibits very large differences in nucleotide sequence, even among closely related species
    • 1. They should be conserved because they are responsible for essential cell functions
    • 2. The amino acid sequence of CENP-A, which binds to centromeric DNA – co-evolution between protein and protein target
    • 3. Changes in the DNA sequence lead to the selection of CENP-A sequences that allow the protein to continue to bind at the centromere with high affinity
    • 4. This suggests that the DNA sequence itself may not be that important a determinant of centromere structure & function
The Cell Nucleus…

- Nuclear organization…
  - In EM, eukaryotic nuclei look like no more than scattered clumps of chromatin & a few irregular nucleoli
  - Newer techniques, like fluorescence in situ hybridization (FISH) & imaging of live GFP-labeled cells, make it possible to localize specific gene loci within nuclei and to determine organization within
The Cell Nucleus...

• Nuclear organization…
  – A given interphase chromosome's chromatin fibers are not randomly distributed, but stay in their own specific area
    • 1. Chromatin of human chromosome 18 occupies territory near periphery of nucleus; whereas chromatin of chromosome 19 is localized more centrally within the organelle
The Cell Nucleus…

• Nuclear organization…
  • 2. This difference in nuclear location may be related to the levels of activity of these 2 chromosomes
  • 3. Chromosome 18 happens to be relatively devoid of genes, whereas chromosome 19 happens to be rich in protein-coding sequences, many of which are presumably transcribed in these cells
  – Inactive X chromosome of women found at one edge of nucleus, while active X is situated internally
The Cell Nucleus...

- **Control of Gene Expression: Prokaryotes**
  - Prokaryotes live in constantly changing environment
  - It is advantageous for cells to use available resources in most efficient way so regulate responses
  - Thus, they respond by selective gene expression
The Cell Nucleus…

• Prokaryotic gene expression…
  – Lactose (disaccharide) - made of glucose & galactose
  – Oxidation provides the cell with metabolic intermediates & energy
  – The β-galactoside linkage is broken in the first step of catabolism
The Cell Nucleus...

- **Prokaryotic gene expression...**
  - 1. The β-galactoside linkage between glucose & galactose is broken by *β-galactosidase*
  - 2. If lactose is absent (minimal conditions) —> β-galactosidase not needed & not there (<5 copies of enzyme, 1 of the corresponding mRNA)
  - 3. If lactose is present —> enzyme levels rise ~1000-fold in a few minutes; lactose has induced the synthesis of β-galactosidase
The Cell Nucleus…

• Control of Gene Expression: Prokaryotes
  – Tryptophan - essential amino acid needed for protein synthesis; if it is not in the growth medium, it must be produced by a bacterium
    • 1. In its absence, cells contain enzymes & their mRNAs needed to make tryptophan & they make it
    • 2. If tryptophan is available in medium, bacteria don't need enzymes to make it; the genes of those enzymes are repressed within a few minutes & the production of the enzymes stops
The Cell Nucleus...

- Control of Gene Expression: Prokaryotes
  - The operon - in bacteria, genes for enzymes of metabolic pathway are usually clustered in functional complex under coordinate control
    - 1. Genes - code for operon enzymes; usually adjacent to each other & transcribed onto single piece of mRNA; then translated into individual enzymes; turn on one, turn them all on
    - 2. Promoter - site where RNA polymerase binds to the DNA prior to beginning transcription
    - 3. Operator – typically resides adjacent to or overlapping with the promoter; serves as repressor protein binding site
    - 4. Repressor - gene regulatory protein; binds with high affinity to specific DNA bp sequence; DNA-binding proteins play predominant role in determining if a genome segment is transcribed or not
    - 5. Regulatory gene - encodes repressor protein
The operon's components are shown in green.

Figure 12-28 Cell and Molecular Biology, 4/e © 2005 John Wiley & Sons
The Cell Nucleus…

• The operon…
  – The repressor is key to operon expression; if it binds to operator, it shields promoter from polymerase & prevents transcription of associated genes
    • 1. Repressor binding to operator depends on conformation, which is regulated allosterically by a key compound in the metabolic pathway (lactose or tryptophan)
    • 2. Concentration of key metabolite determines if operon is active or inactive at any given time
The Cell Nucleus...

• The *lac* operon...
  – The *lac* operon is an example of an *inducible* operon – the presence of a key substance induces the transcription of the genes.
  – It regulates production of the enzymes needed to degrade lactose in bacterial cells
  • The genes in the *lac* operon?
    • 1. *z* gene - encodes β-galactosidase
    • 2. *y* gene - encodes galactoside permease; promotes lactose entry into cell
    • 3. *a* gene - encodes thiogalactoside acetyltransferase; its physiological role is unclear
• The *lac* operon...
  
  1. If lactose is present in medium, it enters cell, binds *lac* repressor, changing its shape. Lactose acts as an *inducer*
The Cell Nucleus…

• The *(lac)* operon…
  – 2. Lactose-bound repressor cannot bind operator DNA; structural genes then transcribed, enzymes translated & lactose catabolized
The Cell Nucleus...

• The *lac* operon...
  – 3. If lactose levels fall, it dissociates from repressor, changing repressor back to active shape
  – 4. Repressor binds operator and physically blocks polymerase from reaching structural genes, turns off transcription
  – *Lac* operon movie
The Cell Nucleus…

• The \textit{trp} operon…
  – The \textit{trp} operon is an example of a \textit{repressible} operon – the presence of a key substance represses the transcription of the genes.
  – It regulates production of the enzymes needed to produces tryptophan in bacterial cells
  – Repressor is active only if bound to specific factor which functions as a \textit{co-repressor} (like tryptophan)
The Cell Nucleus…

• The *trp* operon…
  – 1. Without tryptophan, operator site is open to binding by RNA polymerase; polymerase binds & transcribes *trp* operon structural genes
  – 2. Leads to production of enzymes that synthesize tryptophan
  – 3. Tryptophan produced
The Cell Nucleus…

• The *trp* operon…
  – 4. When tryptophan is available, enzymes of tryptophan synthetic pathway are no longer needed
  – 5. Increased tryptophan concentration leads to formation of tryptophan-repressor, an active repressor
  – 6. The newly activated repressor binds DNA at operator, blocks transcription of operon structural genes & production of tryptophan-synthesizing enzymes
Animation at:
http://bcs.whfreeman.com/thelifewire/content/chp13/1302002.html
The Cell Nucleus...

- **Control of Gene Expression: Eukaryotes**
  - 1. Vertebrates have hundreds of different cell types, each far more complex than bacterial cells & each requiring a distinct battery of proteins that allow it to carry out specialized activities
  - 2. Some once thought that cells attained differentiated state by retaining parts of chromosomes needed for functions of that cell type, while eliminating unneeded parts of chromosomes
  - 3. The idea that differentiation was accompanied by a loss of genetic information was finally put to rest in 1950s & 1960s by a series of key experiments in both plants & animals
  - 4. They demonstrated that differentiated cells retain the genes required to become any other cell in that organism
  - 5. For example, Dolly the sheep
Figure 12-31 Cell and Molecular Biology, 4/e (© 2005 John Wiley & Sons)
Control of Gene Expression: Eukaryotes

- All eukaryotic cells have identical gene content; but they regulate which genes are expressed in particular cells
  
  1. Results from wide range of experiments show that cells from most organisms do not lose genetic material as they differentiate, instead they repress genes that confer character of another cell type
  
  2. It is not the presence or absence of genes in a cell that determines properties of that cell, but how those genes are utilized
  
  3. Liver cells express a specific set of liver genes, while nerve cells express a specific set of nerve genes; each cell also represses genes whose products are not involved in their function
  
  4. Regulatory proteins are responsible for turning genes on & off
The Cell Nucleus...

• Control of Gene Expression: Eukaryotes
  – A single human cell contains enough DNA (6 billion bp) to encode several million different polypeptides
    • 1. Most of this DNA does not actually code for proteins, but mammalian genomes are thought to contain between >30,000 protein-coding genes
    • 2. A typical mammalian cell may make ~5,000 different polypeptides at any given time
    • 3. Many of these polypeptides are made by virtually all cells of the organism
    • 4. Cells also make proteins unique to its differentiated state; giving the cell its unique characteristics
    • 5. Regulating eukaryotic gene expression is an extremely complex process, just starting to be understood
The Cell Nucleus...

- Control of Gene Expression: Eukaryotes
  - Example: RBC development in bone marrow of human bone
    - 1. Hemoglobin accounts for >95% of cell's protein, yet genes coding for hemoglobin polypeptides represent <one-millionth of RBC-precursor total DNA
    - 2. RBCs have to find these genes & regulate their expression to a high degree so that production of a few polypeptides becomes the dominant synthetic activity of the cell
The Cell Nucleus...

- Control of Gene Expression: Eukaryotes
  - Three levels of control
    - Transcriptional
    - Processing
    - Translational

Figure 12-32 Cell and Molecular Biology, 4/e © 2005 John Wiley & Sons)
The Cell Nucleus...

- Control of Gene Expression: Eukaryotes
  - Transcriptional control is orchestrated by actions of a large number of proteins called transcription factors (TFs);
  - 1. General TFs - bind at core promoter sites in association with RNA polymerase
  - 2. Sequence-specific TFs - bind to various regulatory sites of particular genes; they either stimulate (transcriptional activators) or inhibit (transcriptional repressors) transcription of adjacent genes
The Cell Nucleus...

- Control of Gene Expression: Eukaryotes
  - Understanding transcription factor function is a complex undertaking
    - A. A single gene may be controlled by many different DNA regulatory sites that bind variety of different TFs
    - B. A single TF may become attached to numerous sites around genome, controlling the expression of a host of different genes
    - C. Each cell type has characteristic pattern of gene transcription, which is determined by the particular complement of TFs contained in that cell
    - D. Control of gene transcription is complex & influenced by various circumstances:
      - 1. Affinity of TFs for particular DNA sequences and
      - 2. Ability of TFs bound at nearby sites on DNA to interact directly with one another
The Cell Nucleus…

- Control of Gene Expression: Eukaryotes
The Cell Nucleus...

• Transcription factors
  – TFs have different domains that mediate different aspects of their function (usually at least 2 domains)
    • A. The **DNA-binding domain** recognizes & binds to specific DNA base pair sequence
    • B. The **activation domain** regulates transcription by interacting with other proteins
      – 1. Many TFs have a surface that promotes their binding with another protein of identical or similar structure to form a dimer
      – 2. Dimer formation is common feature of many different types of TFs & is thought to play an important role in regulating gene expression
The Cell Nucleus…

• Transcription factors
  – DNA-binding domains of most TFs can be grouped into several broad classes whose members possess related structures (motifs) that interact with DNA sequences
  • Most motifs contain a segment (often an α-helix) that is inserted into the major groove of DNA, where it recognizes the sequence of base pairs that line the groove
    – 1. Protein binding to DNA is achieved by a combination of van der Waals forces, ionic bonds & H bonds between amino acid residues & various parts of DNA, including the backbone
The Cell Nucleus...

• Transcription factors
  – Among the most common motifs in DNA-binding proteins are the **zinc finger**, the **helix-loop-helix**, the **leucine zipper** & the **HMG box**
    • 1. Provide a stable framework on which DNA-recognizing surfaces of a protein can be properly positioned to interact with the DNA double helix
    • 2. Such motifs are found in a wide variety of TFs that regulate diverse types of cellular activities in a wide variety of organisms (fungi to animals to plants)
The Cell Nucleus...

• Transcription factors
  – Zinc finger motif
    • A. A zinc ion usually held in place by 2 cysteines & 2 histidines
      – These proteins typically have a number of such fingers acting independently of one another; they are spaced apart so as to project into successive major grooves in target DNA
    • B. The first zinc finger protein discovered was TFIIIA; it has 9 zinc fingers; others include:
      – 1. Egr - helps activate genes needed for cell division
      – 2. GATA - involved in cardiac muscle development
The Cell Nucleus...

- Transcription factors
  - Helix-Loop-Helix (HLH) motif – characterized by 2 α-helical segments separated by an intervening loop
    - HLH domain is often preceded by a stretch of highly basic amino acids whose positively charged side chains contact DNA & determine TF sequence specificity
The Cell Nucleus…

• Transcription factors
  – Proteins with this basic-HLH (bHLH) motif always occur as dimers; the 2 subunits of a dimer are usually encoded by different genes so the protein is a heterodimer
The Cell Nucleus...

- Transcription factors – HLH motif
  - 1. Heterodimerization greatly expands the diversity of regulatory factors that can be generated from a limited number of polypeptides
  - 2. Example: if cell makes 5 different bHLH-containing polypeptides that can form heterodimers with one another in any combination (up to 32 [25] different TFs recognizing 32 different DNA sequences)
  - 3. In reality, combinations between polypeptides are probably restricted
The Cell Nucleus…

• Transcription factors – HLH motif
  – Leucine zipper (LZ) motif - leucines occur every 7th amino acid along an \( \alpha \)-helix of 30 - 40 residues
    A. Since \( \alpha \)-helix repeats every 3.5 residues, all leucines along polypeptide's helical stretch face same direction
• 1. 2 \( \alpha \)-helices of this type can "zip" together along their length to form coiled-coil
• 2. Leucines of one helix are pressed against leucines of the other, so LZ proteins of LZ exist as dimer
The Cell Nucleus...

- Transcription factors – HLH motif
  - HMG-box motif - first discovered in abundant high mobility group (HMG) proteins & named after them
    - A. Consists of 3 α-helices organized into a boomerang-shaped motif capable of binding DNA
    - 1. Called architectural factors; activate transcription by bending DNA, which promotes interaction of other TFs bound at nearby sites
The Cell Nucleus...

- Transcription factors – HLH motif
  - HMG-box motif...
  - Ex.: SRY protein - plays key role in human male sexual differentiation; its gene is on Y chromosome
    - 1. SRY protein activates transcription of genes in pathway leading to testes differentiation
    - 2. Mutations in SRY gene that render protein unable to bind to DNA lead to condition known as sex reversal phenotype (individuals have XY pair of sex chromosomes but develop into females)
  - Ex.: UBF protein - activates rRNA transcription by RNA polymerase I; binds to DNA as dimer whose 2 subunits contain a total of 10 HMG boxes

---

**Figure 12-46b: Cell and Molecular Biology, 4th ed (c) 2005 John Wiley & Sons**
The Cell Nucleus...

- Regulatory Regions
  - The *regulatory region* of a gene can be thought of as a type of integration center for that gene's expression
    - A. Cells exposed to different stimuli respond by synthesizing different TFs, which bind to different sites in DNA
    - B. The extent to which a given gene is transcribed presumably depends upon particular combination of TFs bound to its upstream regulatory elements
      - 1. Roughly 5 – 10% of genome encodes TFs
      - 2. Thus, an nearly unlimited number of possible combinations of interactions among TFs is possible
      - 3. Complexity of interactions is revealed in marked variation in gene expression patterns between cells of different type, different tissue, different developmental stage & different physiological state
The Cell Nucleus...

- Regulatory Regions
  - Promoter elements – regions upstream of a gene that regulate the initiation of transcription.
  - Most eukaryotic promoter elements can be roughly divided into ‘proximal’ and ‘distal’
  - Proximal promoter elements (-50 to -200bp):
    - TATA box –
      - Consensus sequence – TATAAA
      - Usually at ~-30
    - CAAT box –
      - Consensus sequence – CAAT
      - Usually ~-70
    - GC box –
      - Consensus sequence – GGGCGG
      - Often multiple copies within 100 bp upstream of start codon
Figure 12-41 Cell and Molecular Biology, 4/e (© 2005 John Wiley & Sons)
The Cell Nucleus…

• Regulatory Regions
  – Proximal promoter elements (-50 to -200bp):
    • TATA box –
      – Site of assembly of the transcription complex:
      – RNA polymerase II, all necessary transcription factors
    • CAAT box and GC box –
      – Regulate the frequency of transcription via binding of transcription factors
The Cell Nucleus…

• More Regulatory Elements
  – Enhancers
  – Typically are multiple binding sites for sequence-specific transcriptional activators; raise transcription rates above the basal level
    • 1. Have a unique property: they can be moved experimentally from one place to another within a DNA molecule (even be inverted) without affecting the ability of a bound transcription factor to stimulate transcription
    • 2. Deletion of an enhancer can decrease the level of transcription by 100-fold or more
    • 3. Some enhancers are located thousands or even tens of thousands of base pairs upstream or downstream from the gene whose transcription they stimulate
The Cell Nucleus...

- More Regulatory Elements
  - Enhancers
  - Thought to stimulate transcription by influencing events that occur at core promoter
    - A. Enhancers & core promoters can be brought into close proximity because the intervening DNA can form a loop
    - B. What prevents enhancer from binding to inappropriate promoter located even farther downstream?
      - 1. insulators
        - 2. insulator sequences may bind to proteins of nuclear matrix; DNA segments between insulators correspond to looped domains of chromatin
The Cell Nucleus…

• More Regulatory Elements
  – Enhancers and coactivators
  – How do transcriptional activators bound at enhancer stimulate transcription initiation at core promoter?

• **coactivators;** 2 basic types:
  – 1. Those that interact with components of the basal transcription machinery (general TFs & RNA polymerase II) - lead to assembly of preinitiation complex & initiation of RNA synthesis
  – 2. Those that act on chromatin, converting it from a state relatively inaccessible to transcription machinery to a much more transcription-friendly state
Upstream transcription factors and binding sites for the PEPCK gene
The Cell Nucleus…

• Coactivators and chromatin structure
  – Much evidence suggests that packaging DNA into nucleosomes impedes access to DNA & markedly inhibits both the initiation & elongation stages of transcription – how is this overcome?
  – 1. Covalent modifications of core histone tails have big impact on chromatin structure & function
  – 2. *Addition of methyl groups* to core histones can promote *chromatin compaction & transcriptional silencing*, *addition of acetyl groups* to core histones tends to have *opposite effect*
The Cell Nucleus...

- Coactivators and chromatin structure
- Acetyl groups are added to histones via the action of histone acetyltransferases (HAT)
- Many coactivators have HAT activity
Figure 12-46 Cell and Molecular Biology, 4/e (© 2005 John Wiley & Sons)
The Cell Nucleus...

- Transcriptional repression
  - All of the previous slides have been aimed at ways to activate transcription. How is transcription repressed?
  - **Histone deacetylation/histone methylation**
    - Histone deacetylases (HDAC) remove acetyl groups from core histones
    - This is often accompanied by the methylation of a nearby lysine on the histone.
    - These two events compact the nucleosome and restrict access to the DNA
    - Many TFs (corepressors) have HDAC and histone methylation activity
The Cell Nucleus...

- Transcriptional repression
- DNA methylation...
  - As many as 1 in 100 nucleotides in mammals & other vertebrates may have an added methyl group
    - 1. Methyl groups are added to DNA by a small family of enzymes called **DNA methyltransferases**
    - 2. These enzymes are encoded in humans by *DNMT* genes
The Cell Nucleus...

- Transcriptional repression
- DNA methylation...

  - This simple chemical modification (methylation) is thought to serve as a mark that allows identification of certain DNA regions & allows them to be used differently from other regions
    - 1. In mammals, virtually all methylcytosine is part of 5'-CpG-3' dinucleotide within symmetrical sequence; such sequences do not occur randomly in DNA
    - 2. They tend to be concentrated in GC-rich islands that are primarily located in promoter regions that regulate gene expression
The Cell Nucleus…

- Transcriptional repression
- DNA methylation…
  - Methylation of promoter DNA is strongly correlated with gene repression
    1. Majority of added CH₃ groups are part of CpG dinucleotides on transposable elements (thought to keep them in inactive state)
    2. These elements are often demethylated in tumor cells, leading to increased genetic rearrangement
    3. Methylation level in region upstream from γ-globin gene in DNA in fetal liver is greatly decreased compared to the same gene in other fetal tissues
    4. This change correlates with high level of transcription of this gene in liver during fetal development
The Cell Nucleus...

- Transcriptional repression
- DNA methylation...
  - Most evidence suggests that DNA methylation is more important for maintaining a gene in an inactive state than as a mechanism for initial inactivation
    - Inactivation of the genes on one of X chromosomes in female mammals occurs prior to a wave of DNA methylation that is thought to convert the DNA into a more permanently repressed condition
  - Although methylation is thought to inhibit transcription on a rather permanent basis (it is a relatively stable epigenetic mark), it is not irreversible
Figure 12-51 Cell and Molecular Biology, 4/e (© 2005 John Wiley & Sons)
The Cell Nucleus...

• Post-transcriptional control
  – Alternate splicing – a single gene can encode two or more related proteins; multiple processing pathways for the transcript
  • Genes of complex plants & animals have numerous introns & exons —> use a different exon combination, get a different protein
The Cell Nucleus...

- Post-transcriptional control
- Alternate splicing...
  - Usually alternatively spliced proteins are identical along most of their length, but they tend to differ in key regions that may affect cellular location, ligands they can bind, or kinetics of catalytic activity
  - TFs often made from genes that can be alternately spliced, producing variants that can determine pathway of differentiation a cell takes
  - Antibodies may be membrane-bound or soluble depending on which of 2 alternate exons is at 3' end of mRNA
  - Roughly 40 – 60% of human genes are subject to alternate splicing
The Cell Nucleus...

- Post-transcriptional control
- **Translational-level control** encompasses a wide variety of regulatory mechanisms affecting mRNA translation already transported from the nucleus to the cytoplasm;
- 3 aspects of translational-level control
  - A. Localization of mRNAs to certain sites within a cell
  - B. Controlling whether or not an mRNA is translated and, if so, how often
  - C. Controlling the half-life of the mRNA, a property that determines how long the message is translated
- Mechanisms usually work via interactions between mRNAs & cytoplasmic proteins
The Cell Nucleus...

- Post-transcriptional control
  - mRNAs contain noncoding segments, called untranslated regions (UTRs) at both their 5' & 3' ends; these are sites where most translational control is effected
    - 1. 5' UTR extends from methylguanosine cap at start of message to AUG initiation codon
    - 2. 3' UTR extends from termination codon at end of coding region to the end of the poly(A) tail attached to nearly all eukaryotic mRNAs
The Cell Nucleus...

- Translational-level control...
- Cytoplasmic localization of mRNAs
  - Example: the fruit fly, anterior-posterior axis
    - 1. Axis formation is influenced by localization of specific mRNAs along same axis in the oocyte
    - 2. *Bicoid* mRNAs preferentially localized at anterior end; *oskar* mRNAs preferentially localized at opposite end
    - 3. Protein encoded by *bicoid* mRNA is critical for head & thorax development; *oskar* protein is required for formation of germ cells, which develop at posterior end of larva
    - 4. Localizing mRNAs is more efficient than localizing their corresponding proteins, since each mRNA can be translated into large numbers of protein molecules
The Cell Nucleus…

• Translational-level control…
• Cytoplasmic localization of mRNAs
  – 3' UTR governs localization of *bicoid* & *oskar* mRNAs
    • 1. Join foreign gene coding region to DNA sequence encoding 3’ UTR of *oskar* or *bicoid*
    • 2. Place in fruit flies & see what happens when the foreign gene is transcribed during oogenesis —> foreign gene goes to site determined by its 3’ UTR
    • 3. Localization of mRNAs is mediated by specific proteins that recognize mRNA localization sequences (*zipcodes*) in this region of mRNA
The Cell Nucleus…

- Translational-level control…
- **Controlling mRNA translation**
- Example: mRNAs stored in unfertilized egg are templates for proteins synthesized during the early stages of development;
  - rendered inactive by association with inhibitory proteins
  - Activation of these stored mRNAs involves at least two distinct events:
    - 1. Release of bound inhibitory proteins
    - 2. Increase in length of poly(A) tails by action of an enzyme residing in egg cytoplasm
The Cell Nucleus...

- Translational-level control...
- **Controlling mRNA stability**
- The longer an mRNA is present in cell, the more times it can serve as template for polypeptide synthesis
  - *c-fos* mRNA made in response to changes in external conditions in many cells; degraded rapidly in cell (half-life of 10 - 30 min); involved in cell division control
  - In contrast, dominant cell protein mRNAs in a particular cell, like those for hemoglobin, (half-life >24 hours)
The Cell Nucleus...

- Translational-level control...
- **Controlling mRNA stability**
- mRNA longevity is related to length of poly(A) tail
  - 1. Early study - mRNAs lacking poly(A) tails are rapidly degraded after injection into cell, whereas same mRNA with poly(A) tail is relatively stable
  - 2. Typical mRNA has ~200 adenosine residues when it leaves nucleus
  - 3. Gradually reduced in length as it is nibbled away by poly(A) ribonuclease
  - 4. No effect until the tail is reduced to ~30 A residues; once shortened to this length, the mRNA is usually degraded rapidly
The Cell Nucleus…

- Translational-level control…
- **Controlling mRNA stability**
- Tail length not the whole story;
  - mRNAs starting with same size tail have very different half-lives –
  - 3' UTR plays role
  - 3'-UTR of α-globin mRNA contains a number of CCUCC repeats that serve as binding sites for specific proteins that stabilize mRNA; if these sequences are mutated, the mRNA is destabilized
  - Short-lived mRNAs often contain destabilizing sequences (AU-rich elements; AUUUA repeats) in their 3' UTR; thought to bind proteins that destabilize mRNA
The Cell Nucleus...

- Post-translational control...
- **Controlling protein stability**
- Every protein is thought to have characteristic longevity (half-life) or the period of time during which it has a 50% likelihood of being destroyed
  - A. Some enzymes (those of glycolysis or erythrocyte globin molecules) are present for days to weeks
  - B. Other proteins required for a specific, fleeting activity (regulatory proteins that initiate DNA replication or trigger cell division) may survive only a few minutes
  - C. All of the proteins, regardless of expected survival time, are degraded by proteasomes
  - D. Factors controlling a protein's lifetime are not well understood