Protein Synthesis

One Gene-One Polypeptide Hypothesis

Mendel described that certain "factors" were responsible for the patterns of inheritance that he observed. Today, these factors are known as **genes**, and they direct the production of **proteins** that determine the phenotype – morphology, physical characteristics, body form, etc. of the organism. Genes also direct the production of proteins that function as antibodies or hormones, or proteins that form structural molecules like in muscles, ligaments, hair, nails, etc. The presence or absence of certain proteins may manifest into genetic disorders as well.

Garrod's Hypothesis

- the principle that one gene directs the production of one enzyme was first put forth by Garrod in the early 20th century
- he noticed that a particular inherited trait in humans was the result of an accumulation of a chemical called alkapton – when its levels increase substantially in the blood, it works its way into the urine, and causes the urine to turn black as it makes contact with air
- he hypothesized that a defective enzyme caused an "inborn error in metabolism", resulting in the inability to break down alkapton in the blood
- those without the error were able to make the enzyme necessary to metabolize alkapton
- those with the error were not able to make the enzyme to metabolize alkapton, so it accumulated and showed up in the urine
- Garrod went on further to say that enzymes are under the control of the hereditary material, and thus an error in the hereditary material resulted in an error in an enzyme
- Figure 2, p. 234 illustrates Garrod's hypothesis

Beadle and Tatum's Neurospora crassa

- effectively demonstrated the relationship between genes and enzymes using a red bread mould called *Neurospora crassa*
- they found that one strain of *Neurospora crassa* was able to synthesize all the complex amino acids and vitamins it required for optimum growth if it was given a minimal nutrient medium containing simple inorganic salts, sugar, and one of the B vitamins
- mutant strains of *Neurospora crassa* were created by exposing its spores to X rays or UV radiation
- the mutant strains "stuck out like a soar thumb" because they were not able to grow on a minimal medium because they could no longer manufacture one or more of the complex compounds that they required
- the mutant strains could grow only if a complete nutrient medium was supplied to them

- to determine which amino acid or vitamin that the mutant strains were not able to synthesize, Beadle and Tatum place colonies of the mould in vials, each containing minimal medium plus one additional nutrient
- they discovered four distinct mutant strains, each having a different defective gene:

1. a mutant strain grew only if the minimal medium was supplemented with the amino acid arginine
2. a mutant strain that didn't grow in a minimal medium deficient in citrulline, but grew if argininoscucinate was added
3. a mutant strain that didn't grow in a minimal medium deficient in ornithine, but grew if citrulline was added
4. a mutant strain that didn't grow in a minimal medium deficient in a certain precursor molecule, but grew if ornithine was added

- from the results, they were able to deduce the pathway for the synthesis of arginine – an essential amino acid for *Neurospora* growth
- Figure 3, p. 235 illustrates the pathway
- the results indicate that the lack of a particular enzyme is a result of a defect in the gene that is responsible to make it

Ingram and Sickle Cell Anemia

- genes don't just code for enzymes – they also code for structural proteins
- in quaternary proteins, each tertiary chain is controlled by a different gene
- Vernon Ingram established this relationship by studying the amino acid sequence of hemoglobin from individual with sickle cell anemia
- he discovered that the 6^{th} amino acid in the beta hemoglobin chain of hemoglobin is switched – instead of valine, glutamic acid takes its place
- this base substitution leads to a significant change in the hemoglobin quaternary structure, which in turn, changes the shape the red blood cell (see Figure 5(b), p. 236)
- the ultimate result is that the person acquires debilitating symptoms and quite often dies
- Ingram showed that a gene specifies the kind and location of each amino acid of a given polypeptide chain

Homework: 1-6, p. 236

Protein Synthesis – An Overview

Central Dogma

- the central dogma states that genes are expressed in the phenotype or morphology of an individual via protein action
- the central dogma is split into two parts – **transcription** and **translation**
- Figure 1, p. 237 illustrates that the pathway of protein synthesis is one way – genes to protein
- the information to make a protein comes from a gene in the DNA molecule – referred to as "a template" or "master"
- in order for one gene to direct the production of numerous essential proteins, copies of the "master" must be made
- every copy will eventually amount to a specific polypeptide sequence
- therefore, the central dogma is a sequence of events that begin with a genetic code and end with a protein
- the template region of DNA is first "copied", or **transcribed** into an mRNA, and then "converted" into a polypeptide chain, or **translated**

Ribonucleic Acid (RNA)

- like DNA, RNA is a carrier of genetic information
- RNA differs from DNA in the following ways:
    - RNA contains a ribose sugar, not a deoxyribose sugar (see Figure 2, p. 238)
    - instead of thymine, RNA contains uracil (Figure 3, p. 238), which is still able to pair up with adenine
    - RNA is single-stranded, whereas DNA is double-stranded
- Table 1, p. 238 summarizes the differences between DNA and RNA
- when a gene is transcribed into messenger RNA, only a single-stranded complementary copy is made, where uracil replaces thymine
- the three major classes of RNA molecules are messenger RNA (mRNA) – made by RNA polymerase II, transfer RNA (tRNA) – made by RNA polymerase III, and ribosomal RNA (rRNA) – made by RNA polymerase I
- the mRNA length depends on the length of the gene that is transcribed – the longer the base pair sequence, the longer the mRNA
- tRNAs (about 70-90 nucleotides long) transfer the appropriate amino acid to the ribosome to build a protein, as dictated by the mRNA transcript
- the rRNA is a structural component of a ribosome that, together with proteins, forms the ribosome
- Table 2, p. 239 outlines the characteristics of each type of RNA molecule

Transcription and Translation: An Overview

- the three parts to transcription are initiation, elongation, and termination
- transcription initiation begins with RNA polymerase binding to a promoter region on the DNA template strand, located near the beginning of the gene

- transcription elongation is the process where RNA polymerase puts the appropriate ribonucleotides together, building the mRNA transcript as dictated by the genetic code on the template DNA (see Figure 5, p. 239)
- transcription termination takes place at a location shortly after the end of the coding region of the DNA – the RNA polymerase reaches a signal to stop transcribing
- translation is also subdivided into initiation, elongation, and termination
- translation initiation is when the ribosome recognizes a specific sequence on the mRNA and binds to that site
- translation elongation is when the ribosome moves along the transcript mRNA six nucleotide bases at a time -- the first three bases are translated while the next three are prepared for translation, and each translated triplet codon codes for a specific amino acid that is brought to the ribosome by tRNA and strung together to make the polypeptide chain
- translation termination is when the codon on the mRNA representing a “stop” signal is translated – the ribosome falls off and the polypeptide chain is released
- Figure 6, p. 239 illustrates a summary of protein synthesis

The Genetic Code

- a sequence of three nucleotides are used to code for the 20 amino acids found in proteins
- each triplet of nucleotides is called a **codon**
- the use of three nucleotides results in 4^3 = 64 different possible combinations (see “Figure 7, p. 240)
- more than one codon can code for a single amino acid, which indicates a redundancy in the genetic code
- the only start codon is AUG (methionine), and the three stop codons are UAA, UAG, and UGA
- for an animation of translation in eukaryotes, click on http://bioweb.uwlax.edu/GenWeb/Molecular/Theory/Translation/translation.mov
- for an animation of the complete process of protein synthesis click on http://science.nhmccd.edu/biol/bio1int.htm#dna

Homework: 1-11, p. 241

Transcription

Initiation

- RNA polymerase binds to the segment of DNA that is to be transcribed and opens the double helix
- the binding site for the RNA polymerase is a region just before the coding region of the gene, called a **promoter region**
- the promoter region is high in adenine and thymine bases – the recognition site for RNA polymerase

- A and T bases possess 2 H-bonds between them, while C and G bases possess 3 H-bonds between them
- therefore, less energy is required to unwind and break apart the H-bonds between A & T than would be required to break those between C & G
- Figure 2 (a) and (b), p. 243 illustrate the process of initiation

Elongation

- the mRNA is built in the 5' to 3' direction
- a primer is not required by the RNA polymerase, as was the case for DNA polymerase in DNA replication
- this means that elongation of the mRNA transcript immediately follows the initiation step
- the promoter is not transcribed
- the strand of the DNA that is transcribed is called the **template strand**
- the complimentary strand of the DNA that is not transcribed is called the **coding strand**

Termination

- RNA polymerase recognizes the end of the gene when it comes across a **terminator sequence**
- terminator sequences between prokaryotes and eukaryotes differ
- at the point of termination, the mRNA transcript disassociates with the DNA template strand, and RNA polymerase is free to transcribe another gene
- Figure 2, p. 243 summarizes the entire transcription process

Posttranscriptional Modifications

- in eukaryotic cells, the **primary transcript** needs to undergo **capping**, **tailing** and **base excision**, before it can leave the nucleus
- capping is the addition of a **5' cap** to the start of the transcript, consisting of 7-methyl guanosine, which in turn, forms a modified guanine nucleoside triphosphate
- capping has two functions:
    - to protect the mRNA form digestive nucleases and phosphatases as it exits the nucleus and enters the cytoplasm
    - to initiate translation
- in addition to the cap, a tail, of about 200 adenine ribonucleotides, known as a **poly-A-tail**, is added to the 3' end of the transcript by an enzyme called **poly-A-polymerase**
- the entire mRNA transcript consists of two regions – a coding region called **exons**, and a non-coding region called **introns**
- the introns are interspread among the exons

- the introns are removed so that their translation is prevented
- particles made of RNA and proteins, called **spliceosomes**, remove the introns and join the exons together so that the transcript is one continuous coding gene
- Figure 4, p. 244 illustrates the posttranscriptional modifications that need to take place before a eukaryotic mRNA exits the nucleus
- the introns stay inside the nucleus and are degraded into recycled nucleotides
- the "primed" mRNA transcript then moves out of the nucleus and into the cytoplasm where it will be translated
- there is no "quality control" mechanism for the mRNA strand
- an incorrectly-made mRNA transcript will amount to a defective protein
- however, so long as the original DNA template strand is correct, the multiple copies of mRNA transcripts will more than likely compensate for any mistake in one mRNA strand

Homework: 1-9, p. 248

Translation

The Ribosome

- the assembly of amino acids to the chain occurs on the surface of ribosomes
- ribosomes….
    - hold the mRNA and fRNAs in proper position so that the codons are read accurately
    - catalyze the formation of the peptide bonds between adjacent amino acids
    - determine where translation starts and ends
- all ribosomes are made of two subunits of unequal size
- each subunit contains rRNA and ribosomal proteins
- smaller subunits are flatter, larger subunits are hemispheric or rounded
- ribosomes do not carry genetic information, nor do they work on specific mRNAs – they can bind and translate any mRNA made by the cell
- Figure 2, p. 250 illustrates how the large and small subunit of the ribosome work together to translate a strand of mRNA into protein

The Role of Transfer RNA (tRNA)

- an intermediate “facilitator” that facilitates the match between the mRNA codon and its corresponding amino acid
- there are 64 possible tRNAs – one for each codon
- each tRNA has two important parts to it:
    - a triplet sequence of codons complementing the mRNA codon, called an **anticodon**
    - a binding site for an amino acid
- as the tRNA anticodon pairs up with the mRNA codon, it carries its amino acid
- for example, if the anticodon reads AUA, then the codon on the mRNA is UAU, the code for tyrosine
- tRNAs are recycled once they release their amino acid to the previous amino acid
- the 3’ end of the tRNA is the binding site for the amino acid
- tRNA is a tertiary structure molecule that has an acceptor arm
- all tRNA acceptor arms end with an ACCA – OH 3’ end
- when tRNAs are attached to their amino acids they are charged – when they lack their amino acids they are uncharged
- since all tRNA acceptor arms are all the same, special enzymes, called **aminoacyl-tRNA synthetase** enzymes, are required to link up specific amino acids to specific tRNA
- there are 20 different aminoacyl-tRNA synthetase enzymes – each form an aminoacyl-tRNA complex with the tRNA and its amino acid
- the energy that is accumulated by the complex in forming it, is transferred to the formation of the peptide bond between adjacent amino acids at the transcript

Elongation of the Polypeptide Chain

- the start AUG codon is recognized by the ribosome
- the reason why it can differentiate between a start codon and an AUG codon in the middle of a coding sequence is that the 3’ end of the rRNA of the smaller subunit and the sequence of mRNA situated about 10 nucleotides in front of the initiator codon interact, which helps align the initiator codon of the mRNA with the anticodon of the tRNA initiator
- AUG codes for methionine – the first amino acid in every protein
- the tRNA has two binding sites – the **A (aminoacyl) site** and the **P (peptidyl) site**
- the P site is entered only be an initiator tRNA as an aminoacyl-tRNA complex
- the second aminoacyl-tRNA must then enter the A site as specified by the second codon
- once both tRNAs are bound onto the ribosome, the enzyme peptidyl transferase (located in the large subunit) links the two amino acids together
- the first tRNA releases its amino acid (met), which in turn, binds to the second amino acid (via peptidyl transferase) to become part of a peptidyl-tRNA-complex

- then the initiator tRNA is released from the P site, and the ribosome moves one codon down the mRNA and the peptidyl-tRNA is placed on the P site, leaving the A site empty
- the third codon is then ready to be translated
- the next aminoacyl-tRNA binds to the A site (which is positioned adjacent to the next codon)
- the process is repeated as the ribosome moves along the mRNA in a 5' to 3' direction
- the energy required for both binding each tRNA to the ribosome and for the ribosomal movement is supplied by GTP hydrolysis

Termination of Protein Synthesis

- three mRNA codons stop translation
- there are no tRNAs corresponding to these stop codons, but proteins called **release factors** recognize the stop codons on the mRNA
- the release factors cause the release of the polypeptide chain from its tRNA and from the ribosome
- as the amino acid sequence is released, the ribosome dissociates from the mRNA to become available to translate another mRNA
- in prokaryotes, translation of the mRNA begins while the mRNA is still being synthesized, since there is not "priming" that takes place

- Figure 4, p. 252 illustrates the entire process of protein synthesis

Homework: 1-9, p. 254

An Analogy of Protein Synthesis

Imagine a dream where you find yourself in a building full of answering machines and you notice one with your name on it. The red light on it is flashing, indicating that you have a message waiting to be retrieved. You press the play button, but all you here is a muffled, fragmented sound. You call some audio experts to make a copy of the message for you into a long, written scroll. As clear as the script on this scroll may be, it doesn't help you since it is a language that you cannot understand. You bring the script outside of the building to a translator, where all you see are miles of long winding roads. Along the sides of the road are people in black suits. One of them takes your script and offers to help you translate its message. You look up in the sky, and you notice exactly sixty-four different bird species, flying around aimlessly, and each carrying a coloured envelope in their beak. You notice that there are exactly 20 different colours of envelopes in the sky. As your helper is holding onto your script, the birds in the sky begin to land on her shoulder, one at a time. Each bird offers their letter to her. Another bird lands on her other shoulder with another letter in its beak. The first one then flies away, as the second one offers up its letter. She collects the letters one at a time, and a helper staples them together, in the exact order that she receives them. She continues to do so until she receives no further deliveries from any more birds. Once the stapling is complete, she throws the stapled package into a river where the current brings it to a warehouse consisting of hundreds of conveyor belts with crates on them. Your stapled message is packaged and moved into one of the crates, and you watch it move along the conveyor belt. It eventually gets placed onto a skid, and is shipped out of the warehouse.

Description of Analogy:

- the muffled audio message is the base sequence code on the DNA molecule that codes for a specific protein to be synthesized
- your name on a particular answering machine represents the section of the DNA molecule that codes for the specific protein
- the red light flashing indicates the location on that section where the code for the specific protein begins
- your finger is the messenger protein molecule that enters the nucleus to initiate the "uncoding" of the base sequence
- the audio experts represent special protein molecules that copy, or transcribe, the section of DNA that codes for the desired protein into a complimentary mRNA strand
- the mRNA strand represents the written script or copy of the message
- leaving the building represents the mRNA strand exiting the nucleus and entering the cytoplasm
- the winding roads outside the building represent the endoplasmic reticulum
- the people in black suits represent the ribosomes that help attach the amino acids together to form the polypeptide sequence
- each bird represents a specific tRNA, and the paper letter that each bird is carrying represents a specific amino acid

Control Mechanisms

Types of Control Mechanisms

- The control of gene expression can occur on four levels:

1. Transcription Level – the controlling of the transcription of the template DNA
2. Post-transcription Level – the regulating of the exit of the mRNA transcript out of the nucleus
3. Translation Level – the controlling of the translation of the mRNA
4. Post-translation Level – the regulating of the transformation of the polypeptide into a protein structure

The Operons

Genes are not constantly producing proteins. It would not make sense, from an energy point of view, if transcription and translation were occurring all the time. That is why cells need to control and regulate the expression of genes. Eukaryotic cells have a much more complicated mechanism of gene expression control than prokaryotic cells; however the principle is basically the same. Any cell's environment has a certain degree of change and inconsistency. The environment of the *E. coli* bacterium, for example, is the human colon. This means that its food supply is heavily dependent on the eating habits of its host. If the bacterium is deprived of the amino acid tryptophan, which it needs in order to survive, it responds by activating a metabolic pathway to make its own tryptophan from a precursor molecule. However, if the human host eats a tryptophan-rich meal, the bacterial cell stops producing tryptophan for itself, thus conserving cellular energy, since the desired molecule is available in its environment. Another example of how bacteria tune their metabolism to changing environmental conditions is in the process that produces the beta galactosidase enzyme, the enzyme that breaks down lactose for the bacteria so it can use it as an energy source (alpha glucose and galactose). Only a few molecules of this enzyme are present in an *E. coli* cell that has been growing in absence of lactose in its medium. As soon as lactose is introduced to the medium of an *E coli* cell, it takes only 15 minutes for the number of beta galactosidase molecules to increase to the thousands. Both tryptophan and beta galactosidase production mechanisms can be used to demonstrate the basic process of how gene expression is regulated. These two pathways are called the ***trp* operon** and the ***lac* operon**, respectively. The basic mechanism of both these pathways, described as the operon model, was discovered in 1961 by Franscois Jacob and Jacques Monod, at the Pasteur Institute in Paris.

A. The "Repressible" trp Operon

- *E. coli* synthesizes tryptophan from a precursor molecule in a series of steps, each reaction catalyzed by a specific enzyme (see Figure 18.8, of p. 337 of the handout)
- the five genes that code for the polypeptide chains that make up tryptophan are clustered on the same chromosome
- one promoter for the RNA polymerase serves all five genes
- a single start and stop codons exists for the entire transcription unit, and one "on-off-switch" controls the whole cluster of functionally related genes
- the switch is a segment of DNA called an **operator** – located either within the promoter or between the promoter and the enzyme-coding genes
- together, the promoter region, the operator, and the coding genes make up the *trp* operon (see Figure 18.19 of p. 338 of the handout)
- the RNA polymerase can bind to the promoter and transcribe the genes of the operon so long as the operator is in the "on" mode
- the operator is in the "on" mode if no **repressor** protein is bound to it
- the repressor binds to the operator and blocks attachment of the RNA polymerase to the promoter, preventing transcription of the genes
- repressor proteins are specific – they only bind to operators of a certain operon, and won't turn off any other operon
- the repressor is a product of a gene called a **regulatory gene**, that encodes the *trp* repressor, called *trpR*
- *trpR*s are constantly transcribed and translated into repressor proteins, creating a constant supply of them throughout the cell
- the reason why they don't bind to the operator site is because they are made in the inactive form, and they have to be allosterically activated
- when tryptophan levels increase in the cell, the tryptophan, itself, acts as a corepressor – it binds to the repressor, making it active
- in the active form, the repressor "clicks" into place onto the operator, blocking RNA polymerase from transcribing the coding region, thus turning the operon "off"
- the *trp* operon is said to be a **repressible operon** because its transcription is inhibited when a specific small molecule binds allosterically to a regulatory protein

B. The "Inducible" lac Operon

- the genes lacZ, lacY, and lacA make up an operon called the *lac* operon
- lacZ on the operon codes for beta galactosidase – an enzyme that hydrolyzes lactose into alpha glucose and galactose
- lacY on the operon codes for permease – the membrane protein that transports lactose into the cell
- lacA on the operon codes for transacetylase – an enzyme whose function in lactose metabolism is not known
- the gene for the *lac* repressor is called lacI – unlike the *trpR* regulatory gene of the *trp* operon, this regulatory gene is continuously transcribed and translated to make the repressor protein in the active form
- the result is that the *lac* operon is always "off" – it requires an **inducer** molecule to allosterically bind to the repressor and make it inactive so it can lose its affinity to the operator site

- the inducer for this operon (the molecule that turns the operon "on") is allolactose – an isomer of lactose
- therefore, when lactose (hence allolactose) is present in substantial quantities, it binds to the repressor, makes it inactive, forcing it to "let go" of the operator site, turning the operon "on"
- as the operon gets turned on, the RNA polymerase is not blocked, the genes lacZ, lacY, and lacA are transcribed and translated, the three proteins beta galactosidase, permease, and transacetylase are made, and lactose is metabolized and used as an energy source

C. *Inducible vs. Repressible Operons*

Inducible Operons	**Repressible Operons**
• function in catabolic pathways, which break down a nutrient down to simpler molecules • they produce the appropriate enzymes only when the nutrient is available, so that energy is not wasted • i.e. beta galactosidase is not made if no lactose is present to break down	• function in anabolic pathways, which build large molecules from raw monomer units • they don't make the monomer units of the desired end product if the end product molecule is present • i.e. tryptophan precursors are not made if tryptophan is already present in the cell's medium
• both operons are examples of the **negative control** of genes, because the operons are switched off by the active form of the repressor protein	

D. An Example of Positive Gene Control

- for the enzymes that break down lactose to be synthesized, the another requirement necessary is that the simple sugar glucose be in short supply
- for the *lac* operon, the *E. coli* cell prefers to use more glucose than lactose as a cue to that switches the operon on and off
- when glucose is absent, **cyclic AMP (cAMP)** accumulates
- as cAMP allosterically binds to a **cAMP receptor protein (CRP)**, it changes it from an inactive to an active state
- when CRP is activated, it binds to a specific site next to the *lac* promoter region (see Figure 18.21 of p. 340 of the handout)
- this attachment of CRP makes it easier for the RNA polymerase to bind to the adjacent promoter and start transcription of the operon
- since CRP is a regulatory protein that directly stimulates gene expression, this mechanism is considered to as **positive regulation**

- the conditions are the following:

Condition	first result	second result	third result	Final outcome
lactose present, glucose absent	*lac* repressor protein inactive and cAMP levels increase	operator region is clear and CRP is active	promoter region is not blocked and RNA polymerase binds much better to the promoter region	a lot of *lac* mRNA is synthesized
lactose present, glucose present	*lac* repressor protein inactive and cAMP levels decrease	operator region is clear and CRP is inactive	promoter region is not blocked but RNA polymerase binds less to the promoter region	*lac* mRNA is synthesized – but not a lot
lactose absent, glucose present	*lac* repressor protein is active and cAMP levels decrease	operator region is not clear and CRP is inactive	promoter region is blocked and RNA polymerase binds less to the promoter region	*lac* mRNA is not synthesized at all
lactose absent, glucose absent	*lac* repressor protein is active and cAMP levels increase	operator region is not clear and CRP is active	promoter region is blocked and RNA polymerase binds much better to the promoter region	*lac* mRNA is not synthesizes at all

  - lactose absent, glucose present
  - lactose absent, glucose absent

Homework: 1-6, p. 258

Mutations

What are mutations and what caused them?

- mutations are errors made in the DNA sequence that are inherited
- mutations may be **spontaneous**, caused by errors in the genetic machinery that are most often picked up by DNA polymerase I
- mutations may be **induced** and caused by **mutagenic agents** such as UV radiation, cosmic rays, X-rays, and certain chemicals
- not all mutations are bad -- the effects of mutations can be negative, positive, or can have not effect at all
- for example, cystic fibrosis is a deleterious mutation that is passed on to offspring – it results from a modification in the cystic fibrosis Transmembrane regulator (CFTR) gene
- a result of a series of mutations leading to a positive effect is the increase in the human brain throughout evolution
- eukaryotes have a higher tendency to mask mutations because, for the most part, they are diploid -- diploid organisms have two copies of each gene; hence, if an error is made in one copy, the other copy may compensate for this error
- we may in fact have numerous deleterious mutations in our own genomes, but because we are diploid, they may not surface

- mutations can be **silent mutations** – meaning they have no effect on the operation of the cell because they occur on the noncoding sections of the DNA (the introns), which are taken out of the mRNA by spliceosomes

- another way of silencing mutations is through the redundant nature of the DNA code – more than one triplet code codes for the same amino acid in translation of protein synthesis, which compensates for the incorrect base that may have been strung together during transcription

Types of Mutations

- there are two major types of mutations:

1. Point Mutations – mutations that are specific to one base pair within the DNA molecule

    - *substitutions*

        - when one base pair is exchanged for another, resulting in a different code
        - a change in a base pair during DNA replication may result in coding for a different amino acid during translation can lead to what is known as a **missense mutation** (ex. Sickle cell anemia)
        - if any one of the stop codons replaces a midstream amino acid codon due to a defect in DNA replication, a **nonsense mutation** results, which means that translation will result in a shortened polypeptide chain – only the part of the protein that precedes the "new" stop codon is produced, which will in turn, result in a fragment protein that gets digested by cell proteases

    - *deletions*

        - when one or more nucleotides is/are removed from the DNA sequence during replication
        - this causes a frame shift in the code that leads to a different polypeptide chain assembly, and a codon at the end consisting of two bases, not three
        - the result is often drastic and creates a totally different, undesirable and useless protein
        - this type of mutation is known as a **frameshift mutation**

    - *insertions*

        - when one or more nucleotides is/are inserted into the DNA sequence during replication
        - this causes a frame shift in the code that leads to the same effects as a deletion
        - this is also known as a **frameshift mutation**

2. Chromosomal Mutations - mutations involving entire sections of DNA that are apparent at the chromosomal level

- *translocations*

  - the relocation of groups of base pairs from one part of the genome to another
  - usually occur between two nonhomologous chromosomes
  - a segment of one chromosome breaks and releases a fragment, while the same event takes place at another chromosome
  - basically, two fragments exchange places resulting in unrelated gene sequences being transcribed and translated, making a fusion protein with either a completely altered function or no function at all
  - examples of such mutations are some types of leukemia
  - sometimes fragments of DNA consistently move around, known as "jumping genes" or **transposable elements**
  - not all transposable element events lead to phenotypic changes – it depends on whether or not the sections fall into a coding region of a gene

- *inversion*

  - a reversal in the orientation of a section of chromosome, leading to the backwards transcription and translation

- for an excision repair animation, click on http://www.nature.com/nrc/journal/v1/n1/animation/nrc1001-022a_swf_MEDIA1.html

Homework: 1-7, p. 263

Prokaryotes vs. Eukaryotes

The Major Differences

- protein synthesis in prokaryotes and eukaryotes differs in the following ways:

  - prokaryotes don't possess a nuclear membrane which means that the DNA is transcribed and translated at the same time – a process called **coupled transcription-translation** (see Figure 1, p. 264)
  - prokaryotic genes don't contain any introns

- prokaryotic ribosomes recognize the start of an mRNA transcript by a unique sequence of purine-rich bases known as the Shine-Dalgarno sequence, whereas eukaryotic ribosomes recognize the 5' cap region of the transcript as the starting point
- prokaryotic ribosomes are smaller than eukaryotic ribosomes
- the first tRNA for prokaryotes and eukaryotes both carry methionine as the start amino acid, except for prokaryotes, the methionine in prokaryotes is tagged with a formyl group
- eukaryotic organisms do not possess operons
- a prokaryotic genome is circular chromosome, whereas eukaryotic genomes are organized into many individual chromosomes

Endosymbiotic Relationships Between Organelles and Cells

- this idea is derived from the fact that eukaryotic and prokaryotic cells have very similar protein synthesis mechanisms
- it is believed that eukaryotic cells are a branching off of an ancestor that consisted of a prokaryotic mitochondria and host prokaryotic cell that engulfed it without killing it – the endosymbiotic theory
- mitochondria in eukaryotic cells resemble prokaryotic cells in the following respects:
  - mitochondria have circular genomes that are not contained within a nucleus
  - the sequence of mitochondrial DNA is very similar to the genomes of bacterial cells
  - mitochondria divide by the process of **fission** within a eukaryotic cell, similar to bacteria
  - mitochondria possess their own system of DNA synthesis, transcription, and translation, indicating that mitochondria may once have been free-ling cells
- the genomes of different types of mitochondria from different organisms vary greatly, which suggests that evolutionary change took place over time
- there is in fact a modern day example of the endosymbiotic theory
- a small biflagellate alga, known as a cryptomonad, has taken in a compartmentalized structure that resembles a red alga
- the red alga contains a chloroplast, a plasma membrane, and a small volume of cytoplasm surrounding a small nucleus called a **nucleomorph**
- the nucleomorph contains 551 genes organized into 3 chromosomes
- most of the genes code for DNA replication, and only about 30 of them code for protein synthesis
- the relationship between the cryptomonad and the red alga-like structure is considered a mutualistic endosymbiotic relationship since the red alga-like structure provides food for the cryptomonad by photosynthesizing, and the cryptomonad protects the red alga structure

Summary of Differences

Characteristic	Prokaryotes	Eukaryotes
Genome	• small and circular • all regions are coding, except for promoters and operators • presence of operons	• large and arranged in chromosomes • consists of coding and noncoding regions • absence of operons

Transcription	• coupled with translation • lack of introns means no excision	• occurs in the nucleus • introns excised by spliceosomes and exons joined together
Translation	• commences with formyl-methionine • ribosome recognizes Shine-Dalgarno sequences on mRNA as binding site • ribosomes are smaller than in eukaryotes	• commences with methionine • ribosomes recognizes 5' cap on mRNA as binding site • occurs in the cytoplasm • ribosomes are larger than prokaryotes

Homework: 1-3, p. 265

Gene Organization and Chromosome Structure

- at prophase of mitosis, the cell genome is organized into chromosomes – a complex of DNA chromatin and proteins
- the entire human genome chromatin is approximately 1.8 m in length – a considerable length to be packed within a nucleus that is 1.5 μm long
- a length of chromatin that is 145 to 200 base pair nucleotides long, wraps around 8 positively-charged protein molecules called **histones** – two each of histones H2A, H2B, H3, and H4
- the 145 to 200 nucleotide-long chromatin wraps 1 ¾ times around the 8 histones and is fastened by another histone called H1
- the histones stabilize the DNA chromatin structure (which is negatively-charged)
- the resulting "histone wrap" is called a **nucleosomes**
- the distance between adjacent nucleosomes is approximately 60 nucleotide base pairs long
- the nucleosomes and intermittent regions of DNA chromatin first coil, then the coil undergoes **supercoiling**
- the condensed structure appears as an "X" under a microscope and is referred to as a **chromosome** (see Figure1, p. 267)
- only a fraction of the human genome is known to actually code for specific proteins – about 5% of it!!!
- 95% of the human genome is noncoding!!!
- noncoding regions are filled with **variable number tandem repeats (VNTRs)**, also known as **microsatellites** – sequences of base pairs that repeat over and over again
- microsatellites vary among individuals within the same species, as does their length and their location
- functions of microsatellites:
    - cause deleterious effects such as the case in Huntington's disease
    - some function as a defence mechanism against the shortfalls of DNA replication

- how? -- some microsatellite ends, known as **telomeres**, protect cells from losing valuable genomic material during DNA replication by binding proteins that stop the ends form being degraded and prevent them from "sticking" to other chromosomes

- make up the centromere portion of the chromosomes – maintains the chromosome shape and structure during chromosome movement in mitosis and meiosis

- the human genome also contains **pseudogenes** – a nucleotide sequence that resembles a coding gene but does not get transcribed and translated into a protein
- there are two types of pseudogenes – **long interspersed nuclear elements (LINEs)** and **short interspersed nuclear elements (SINEs)**
- the function of LINEs and SINEs is not completely known – could they possibly be "spare tire" genes?

Homework: 1-5, p. 267