AP Biology

Unit 6

Student Notes

Unit 6 Student Notes

Table of Contents

A. Mendelian Inheritance —Pages 3-7
B. Important Terms—Pages 3-5
C. Monohybrid Cross—Pages 5-6
D. Dihybrid Cross—Pages 6-7
E. Trihybrid or Larger Cross—Pages 7-8
F. Linked Genes and Dihybrid Crosses—Pages 8-9
G. Non-Mendelian Inheritance Patterns—Pages 9-11
H. Pedigree Charts/Modes of Inheritance—Pages 12-16
I. Important Terms—Page 12
J. Modes of Inheritance--Autosomal Dominant—Page 12
K. Modes of Inheritance--Autosomal Recessive—Page 13
L. Modes of Inheritance--X-linked Dominant—Pages 13-14
M. Modes of Inheritance--X-linked Recessive—Page 14
N. Modes of Inheritance--Y-linked—Page 15
O. Modes of Inheritance--Mitochondrial--Pages 15-16
P. Gene Regulation—Page 16-23
Q. Operons—Pages 16-19
R. Inducible Operons—Pages 17-18
S. Repressible Operons—Pages 18-19
T. Gene Regulation in Eukaryotic Cells—Pages 19-23
U. Epigenetics—Pages 21-22
V. Embryonic Development—Pages 23-25
W. HOX Genes—Pages 25-26
X. DNA Biotechnology—Pages 25-35
Y. Genetic Engineering—Pages 25-35
Z. Possible Applications of Biotechnology and Recombinant DNA—Pages 26-28
AA. DNA Cloning—Pages 27-29
BB. Polymerase Chain Reaction (PCR)—Pages 29-32
CC. Gel Electrophoresis—Pages 32-34
DD. Human Genome Project—Page 35
EE. Gene Editing/CRISPR/CAS9—Page 35

Mendelian Inheritance

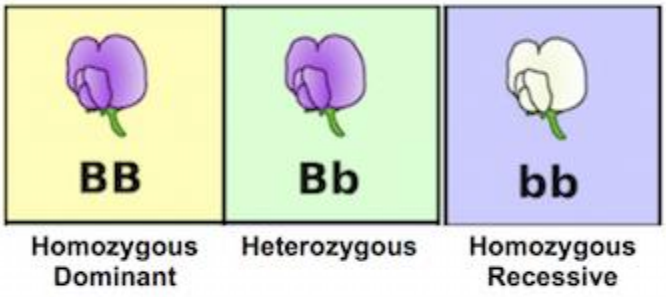
Gregor Mendel—Mendel is considered to be the "Father of Genetics" because of his groundbreaking work with inheritance in pea plants. Mendel conducted his experiments at an Austrian monastery in the late 1800s. He discovered the basic laws of inheritance. He was able to do this before science had an understanding of DNA, chromosomes, or meiosis. One of the characteristics of Mendel's work that made it so successful was that he quantitatively analyzed the results of his pea plant breeding experiments. Most biologists, before Mendel, didn't think that biological processes could be mathematically analyzed. Mendel's main conclusions included:

1. Each trait is controlled by a single pair of alleles. Mendel called them factors.
2. Some alleles are dominant and can mask the appearance of recessive alleles (**Law of Dominance**)
3. The alleles from a pair separate (during meiosis) and only one member of each pair is transmitted to an offspring from each parent (**Law of Segregation**). Realize that Mendel did not know about meiosis, but essentially predicted its existence.
4. The maternal and paternal alleles/factors from each pair assort independently from the alleles in the other pairs. All possible combinations of alleles/factors occur in the gametes. We know today that independent assortment occurs during Metaphase I of Meiosis. (**Law of Independent Assortment**).

Important Terms

Trait—a genetically determined characteristic. Mendel hypothesized that each trait is determined by two alleles (one inherited from each parent). Although this is often true, there are many exceptions to this rule. We will explore some of those exceptions in the Non-Mendelian Inheritance section of the notes (included below).

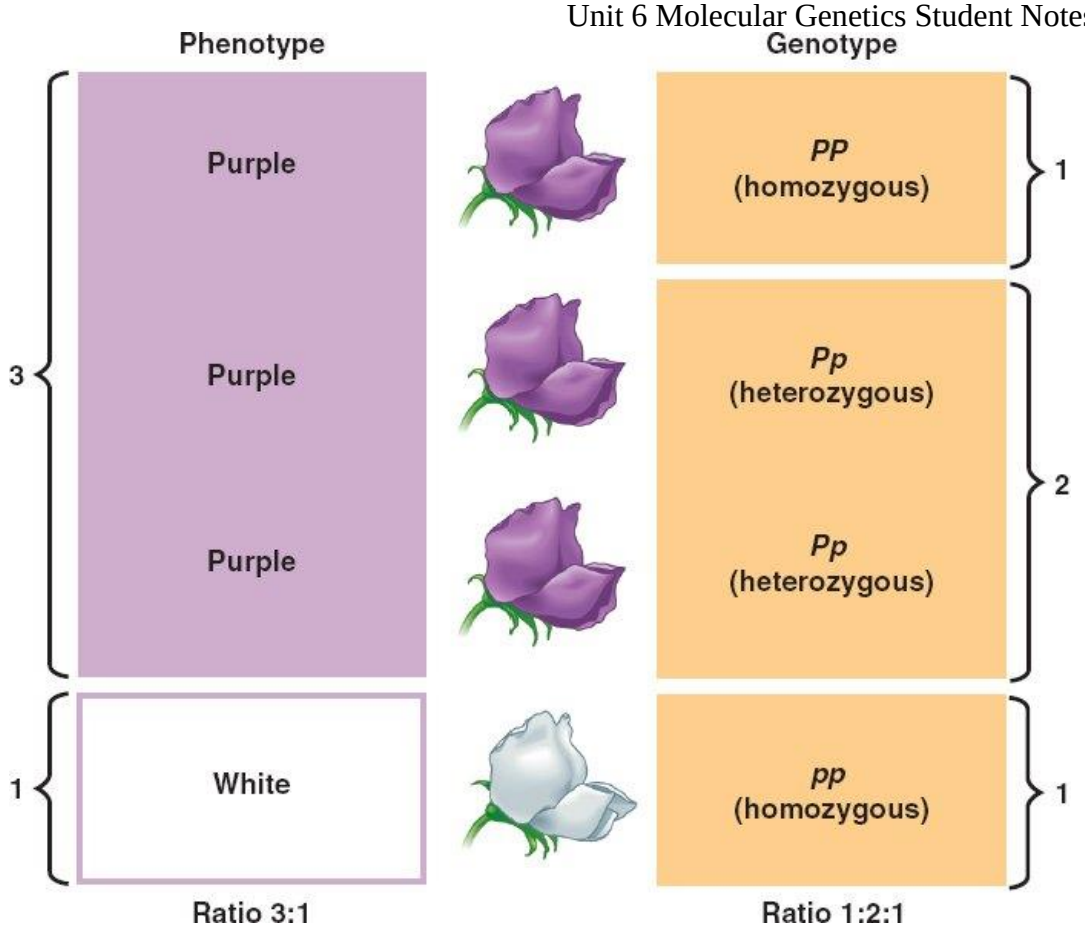
Alleles--This term refers to **different versions of a gene**. For example, a gene might control the color of one's eyes. Different versions (alleles) of that gene might cause the individual to have brown, blue, or green eyes. **Dominant** alleles are represented with capital letters, while r**ecessive** alleles are represented with lower case letters.



Homozygous—An organism is said to be homozygous or **pure bred** for a trait if the two alleles it possesses for the trait are identical. (TT or tt for example).

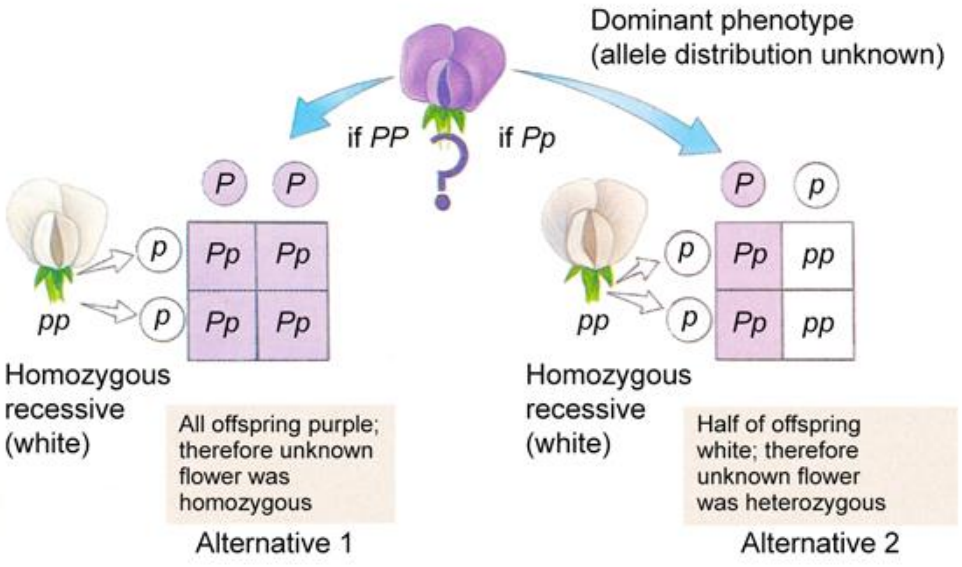
Heterozygous-- An organism is said to be heterozygous or a **hybrid** for a trait if the two alleles it possesses for the trait are different. (Tt for example).

Phenotype--This term refers to the physical traits or appearance of an organism. (Blue eyes or Type A blood, would be examples.)



Genotype--This term refers to an organism's genetic (DNA) make-up for a trait. (Such as TT, Tt or tt.) A genotype of TT might cause a pea plant to have the "tall" phenotype.

If the genotype of an organism is unknown, we can perform a **TESTCROSS** to determine it. To perform this test, you must cross the individual with the unknown genotype with a **homozygous recessive individual**. The phenotypes of the resulting offspring can then be used to deduce the genotype of the parent.

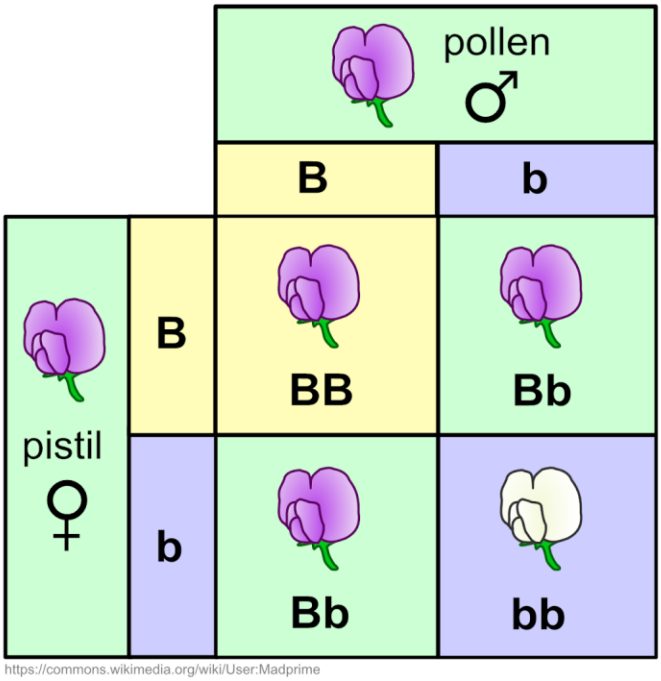


.

Punnett Square--This is a chart which shows the **possible** genotypic outcomes for a mating cross based on the parents' genotypes.

Monohybrid Cross

Monohybrid Cross—This is a cross in which only one trait is analyzed. To complete the cross, you must first determine the possible gamete combinations of the parents. Remember that gametes are haploid. They should only contain one allele for each trait. For example: If we crossed two purple flowered plants, with the genotype Bb, the possible gametes for each plant are "B" or "b". The gametes for one parent are written along the top of the Punnett square, one haploid gamete per column. The gametes of the other parent are written along the left hand side of the Punnett square, one haploid gamete per row. The boxes inside the square are then filled in using the column and row headers as shown below.



The cross illustrated above predicts a **phenotypic ratio** of 3 purple: 1 white and a **genotypic ratio** of 1 BB: 2Bb:1 bb

If you were asked to determine the **probability** (likelihood that an event will occur) that the cross would produce a plant with white flowers, your answer would be ¼ or 25%. If you were asked to determine the probability that the cross would produce a plant with purple flowers, your answer would be ¾ or 75%.

Suppose you were asked to determine the probability that the cross produced two purple-flowered plants in a row?
To calculate this probability, you would need to use the **product rule**. The product rule states that in order to determine the probability of two independent events occurring together, you must multiply the individual probabilities of each event. In this example, we should multiple ¾ X ¾ to get a probability of 9/16 or 56.25%.

The Product Rule

- If two or more events are independent of one another, the likelihood of their simultaneous or consecutive occurrence is the product of their individual probabilities
- This is the **product rule,** also called the **multiplication rule**

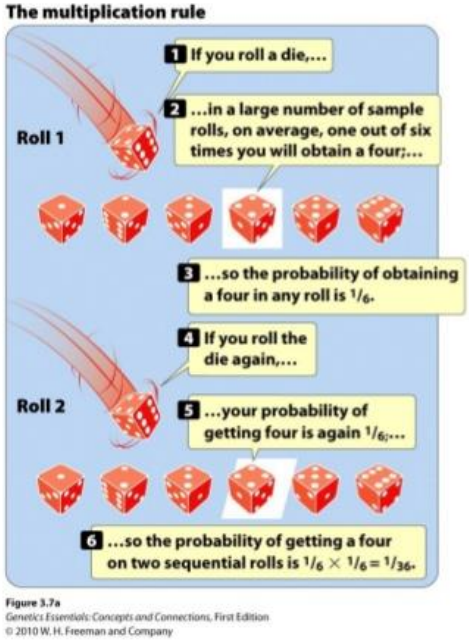


Figure 3.7a
Genetics Essentials: Concepts and Connections, First Edition

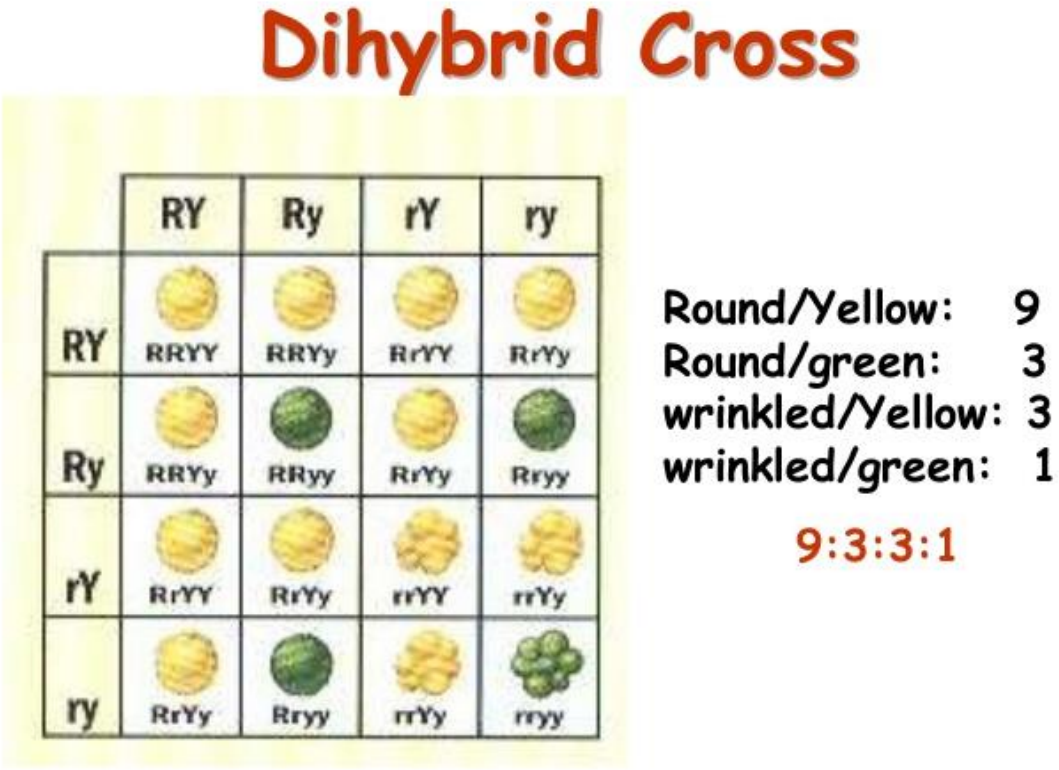
Dihybrid Cross

Dihybrid Cross–A dihybrid cross is a cross in which **the inheritance of two traits is analyzed at the same time.** For example: Suppose that in pea plants "R" is the allele for round seeds, "r" is the allele for wrinkled seeds, "Y" is the allele for yellow seeds, and "y" is the allele for green seeds. The Punnett Square below illustrates the cross between two plants that are heterozygous (RrYy X RrYy) for both traits, **a true dihybrid cross**. To construct the cross, you must first determine the possible gametes that each parent can create**. Each gamete should contain only one allele from each gene.** In this case, each allele contains only one R (either upper or lowercase) and one Y (either upper or lowercase). All possible combinations of the parental alleles are possible. This means that each of the two parents can create the following gamete combinations: RY; Ry; rY; and ry.

To determine the possible number of unique gamete combinations, analyze each gene pair. For the R pair, each parent has a single R and a single r, **2 different alleles.** For the Y pair, each parent also has a single Y and a single y**, 2 different alleles.** To calculate the number of possible unique gametes, multiply the number of unique alleles in each gene pair together. In this case that would mean 2 (from the R pair) X 2 (from the Y pair) =**4 unique gametes per parent.**

Let's say that one of the original parents had the following genotype: **rrYy** To calculate the number of possible alleles from this parent, multiple 1 (from the unique alleles in the r pair) X 2 (from the unique alleles in the Y pair) to get **2 possible unique gamete combinations**. In this case those would be **rY and ry.**

Once the possible gamete combinations are determined, the gametes for one parent are written along the top of the Punnett square, one haploid gamete per column. The gametes of the other parent are written along the left-hand side of the Punnett square, one haploid gamete per row. The boxes inside the square are then filled in using the column and row headers as shown below.



In a true dihybrid cross (one in which both parents are heterozygous for both traits), the phenotypic ratio will always be 9:3:3:1 as shown in the Punnett square above.

Trihybrid or Larger Cross

It is possible to construct Punnett Squares for trihybrid crosses (crosses in which the inheritance of three traits are analyzed at the same time). These squares can be really large (up to 64 squares). You will typically not be asked to draw a trihybrid Punnett Square on the AP exam. You might, however, be asked to answer questions like the following:

How many different possible gametes can an individual with the following genotype produce Aa Bb Dd Ee?

To answer this question, use the same approach as that described above. Multiple the number of unique alleles from each pair together. In this case. 2 unique alleles from the A pair X 2 unique alleles from the B pair X 2 unique alleles from the D pair X 2 unique alleles from the E pair to yield 16 total unique allele combinations/gamete types.

If the genotype of the parent had instead been AA Bb cc, the calculation would have been:
1 unique allele from the A pair X 2 unique alleles from the B pair X 1 unique allele from the c pair to yield only two possible gamete combinations. In this case, those combinations would be ABc and Abc.

What is the probability that the cross between the following genotypes:
Aa BB Dd Ee X Aa Bb Dd ee will produce an offspring with the genotype Aa BB dd Ee?

To answer this question, you could draw a huge Punnett Square, but the easiest and fastest approach is to use the product rule. First, analyze each gene pair separately. Think of doing four separate Punnett squares.

What is the probability that AA (from parent 1) and Aa (from parent 2) will yield (Aa) in the offspring? **The answer is ½.**

What is the probability that BB (from parent 1) and Bb (from parent 2) will yield BB in the offspring? Again, **the answer is ½.**

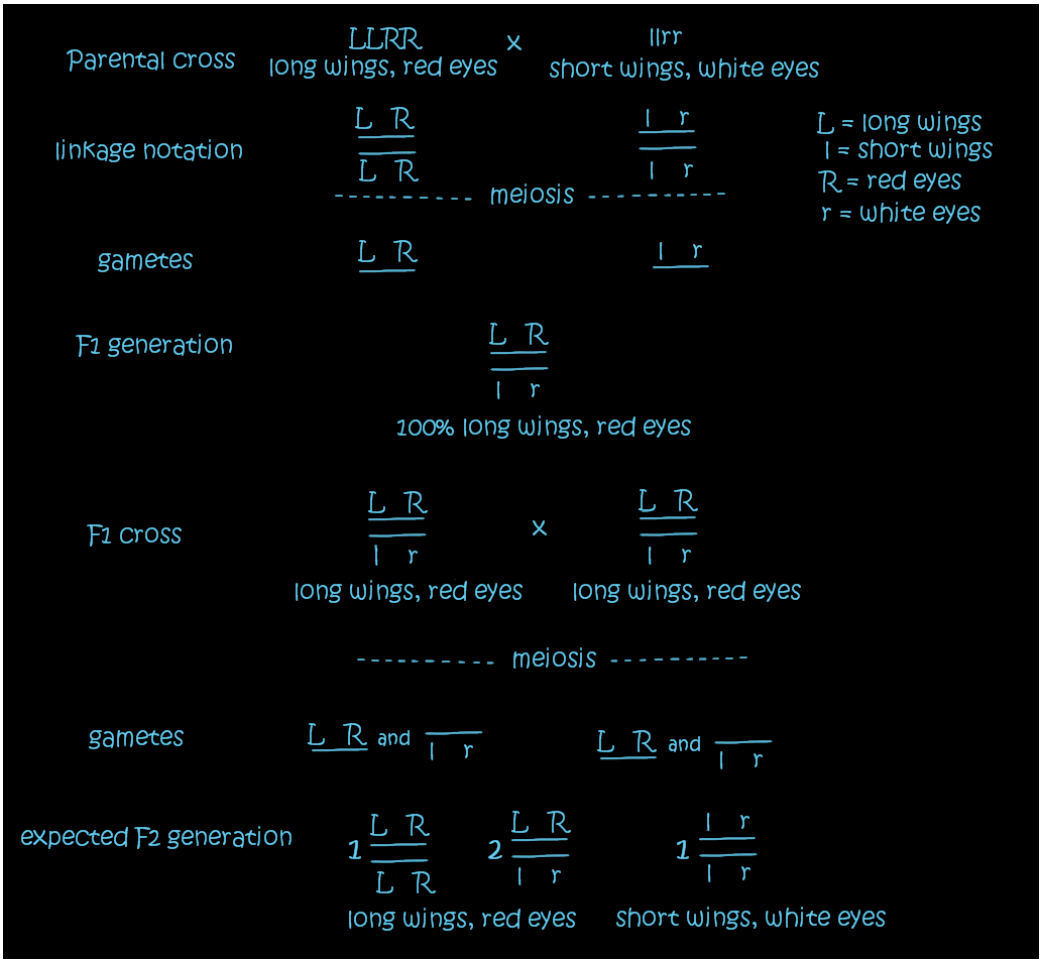
What is the probability that Dd (from parent 1) and Dd (from parent 2) will yield dd in the offspring? This time, **the answer is ¼.**

What is the probability that Ee (from parent 1) and ee (from parent 2) will yield Ee in the offspring? This time **the answer is ½.**

To determine the overall probability that the cross between parents Aa BB Dd Ee X Aa Bb Dd ee will produce an offspring with the genotype Aa BB dd Ee, multiply the probabilities from each gene pair (in bold above). In this case, the calculation is ½ (from the A pair) X ½ (from the B pair) X ¼ (from the C pair) X ½ (from the D pair) **to yield an overall probability of 1/32.**

Linked Genes and Dihybrid Crosses

Linked genes occur on the same chromosome, and therefore, tend to be inherited together (i.e., do not segregate independently). When two heterozygotes are mated in a normal dihybrid cross with independent assortment of alleles, the expected ratio in the offspring is 9:3:3:1. However, as shown in the figure below, in cases of dihybrid crosses involving linkage, the ratio of the offspring produced is 3:1 and only the parental types, with no recombinants, are expected.



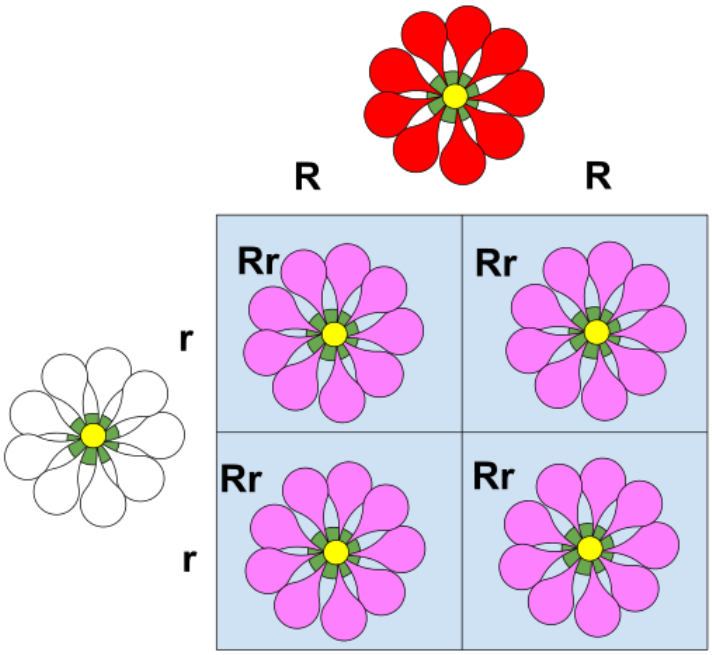
The table below includes the actual observed results of the dihybrid cross involving the two heterozygotes described above. Even though offspring with long wings/white eyes and offspring with short wings/red eyes weren't expected, some were produced. **This often occurs in crosses involving linked genes because of the genetic recombination that occurs during crossing over.**

observed F2 generation	$\frac{L\ R}{l\ r}$	$\frac{L\ r}{l\ r}$	$\frac{l\ R}{l\ r}$	$\frac{l\ r}{l\ r}$
	12	2	4	13
	long wings, red eyes	long wings, white eyes	short wings, red eyes	short wings, white eyes

Non-Mendelian Inheritance

Although Mendel discovered some of the most basic laws of inheritance, scientists have discovered many exceptions to the so-called laws of Mendelian inheritance. The processes/conditions described below are all examples on **non-mendelian inheritance patterns.**

Incomplete Dominance is a form of intermediate inheritance in which one allele for a specific trait is not completely expressed over its paired allele. This results in a third phenotype in which the expressed physical trait is a combination or **blending** of the phenotypes of both alleles. In some plants, there are alleles for red flowers (R) and for white flowers (r). Heterozygous plants (Rr) end up with pink flowers (a blend between red and white).

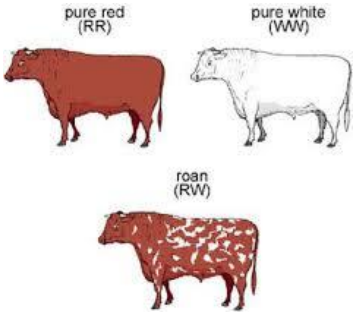


Codominance—This is a condition in which both alleles in a pair are expressed at the same time. They are both equally present in the phenotype (not blended). An example of codominance is human blood type. There are alleles for Type A blood (I^A), Type B (I^B), and type O (i) blood. The alleles for type A and B code for different cell surface proteins that occur on the surface of the red blood cells (RBCs). Individuals with two O alleles (ii) lack these proteins. Individuals with either I^AI^A or I^Ai possess the type A proteins on their RBCs. Individuals with either I^BI^B or I^Bi possess the type B proteins on their RBCs. Individuals with the I^AI^B genotype possess both the A and B proteins on the surface of their RBCs.

Genotype	Red blood cell appearance	Phenotype (blood group)
I^AI^A or $I^A\,i$		A
I^BI^B or $I^B\,i$		B
I^AI^B		AB
ii		O

Another commonly cited example of codominance is the inheritance pattern of coat coloration in some cows.

There are two unique alleles (R) for red coat color and (W) for white color. Cows that inherit one of each allele (RW) have a blotchy phenotype which includes both red and white patches. This type of coloration is known as **roan**.



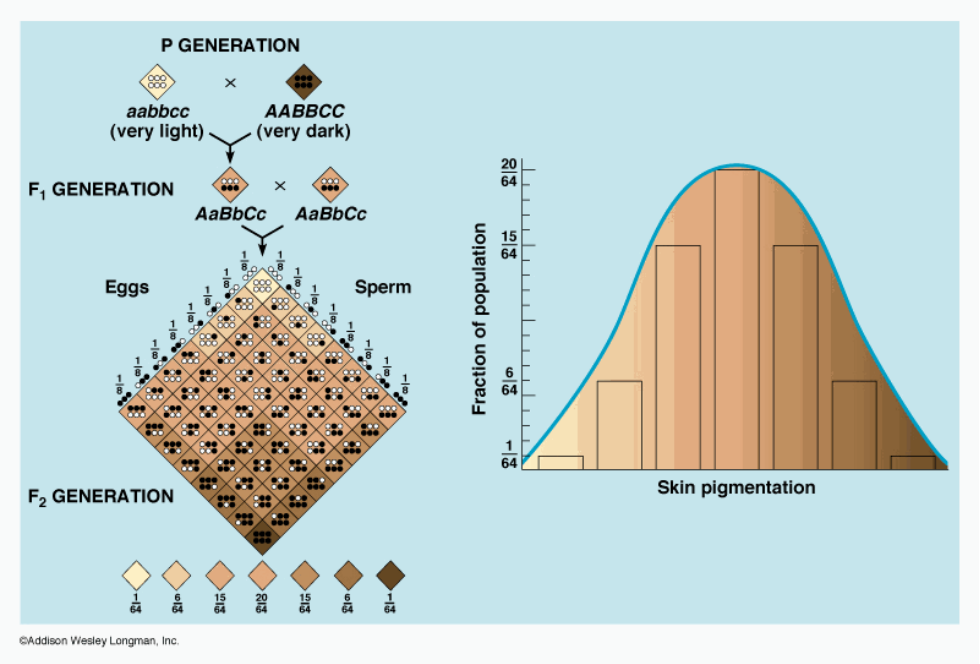
Multiple Alleles—This is when three or more alternative forms of a gene (**alleles**) can occupy the same **locus (location)**. However, only two of the **alleles** can be present in a single organism. For example, the ABO system of blood types is controlled by three **alleles (I^A, I^B, i)**, only two of which are present in an individual.

Pleiotropy is a condition in which one gene affects multiple (seemingly unrelated) characteristics.

Sickle Cell Disease is a great example. The gene AFFECTS the shape of one of the polypeptide chains that make up **hemoglobin** (the molecule that allows red blood cells to transport oxygen). The change in the polypeptide's shape also causes the red blood cells to change shape (from round to sickle shaped). This impedes the flow of blood and leads to multiple symptoms all over the body such as anemia, physical weakness, impaired mental function, lowered disease resistance, and paralysis**. Individuals who are heterozygous for the sickle cell mutation also possess resistance to the protozoan that causes malaria.** This **heterozygote advantage** explains why the normally harmful mutation is so common in human populations where malaria has consistently been a problem.

Epistasis is a condition in which a gene at **one locus** affects a gene at a **second locus.** Albinism in humans is a good example of epistasis. Albinism arises when individuals inherit two defective copies of the allele that normally codes for the enzyme necessary for melanin (pigment) synthesis. Even though alleles at other loci may code for brown eyes, skin, or hair, these individuals are albinos because they lack the ability to produce the pigment (melanin).

Polygenic Inheritance occurs when a trait is governed by two or more sets of alleles. Examples of human traits that are polygenic include height, skin color, and the prevalence of diabetes. Each individual possesses a copy of all the allelic pairs. These may be located on different chromosomes. Each dominant allele has a quantitative effect on the phenotype and the effects are additive. The population typically exhibits **continuous phenotypic variations**. If the frequency of the different phenotypes were graphed, the graph would look like a bell curve. Human skin color is thought to be controlled by 3 pairs of alleles (A, B, C). A person with the alleles *AABBCC* is very dark skinned, a person with *aabbcc* alleles is very light skinned, and a person with *AaBbCc* (or any combo) has an intermediate skin color.



Multifactorial traits are those controlled by multiple genes that are also affected by environmental influences. Hypertension, diabetes, schizophrenia, and allergic conditions are probably all multifactorial traits.

Lethal alleles—There are certain allele combinations which are lethal and prevent the birth of individuals with certain genotypes. For example, **Achondroplasia** is an autosomal dominant form of human dwarfism. Individuals with the genotype Aa are dwarfs. Individuals with the genotype aa are normal. The genotype AA is lethal. Individuals with this genotype die before birth. The Punnett square for the cross between two dwarfs (Aa) would lead one to believe that ¾ of their children should be dwarfs.

	A	a
A	AA	Aa
a	Aa	aa

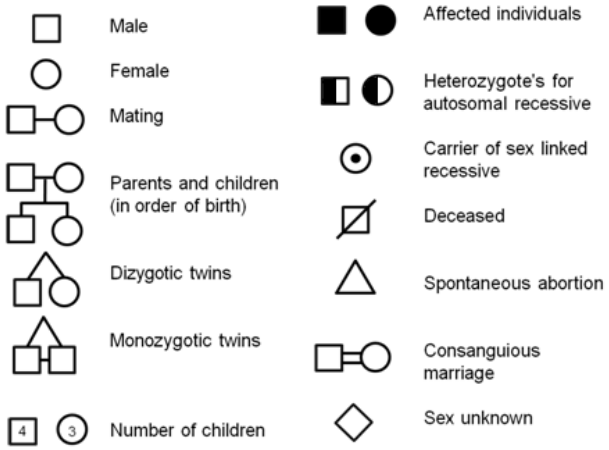
Due to the lethality of the AA genotype. the actual probability that a living dwarf child will be born is 2/3.

Pedigree Charts/Modes of Inheritance

Important Terms

Pedigree Chart--A diagram showing the lineage or genealogy of an individual and all the direct ancestors, usually to analyze or follow the inheritance of trait.
Mode of Inheritance--The manner in which a genetic trait or disorder is passed from one generation to the next. Autosomal dominant, autosomal recessive, X-linked dominant, X-linked recessive, Y-linked, and mitochondrial **inheritance** are examples.
Autosome—A non-sex chromosome. Humans have 44 autosomes per diploid, somatic cell.
X-chromosome—One of the two types of sex chromosomes in humans. Females usually have 2 X chromosomes per somatic cell, while males usually have only 1 X chromosome per cell.
Carrier—Individuals who possess only one copy of a recessive allele. These individuals don't express the trait, but can pass the allele on to their offspring.
Gene Linkage—Genes are linked if they are found together on the same chromosome.

The diagram included below illustrates the symbols commonly used in pedigree charts.

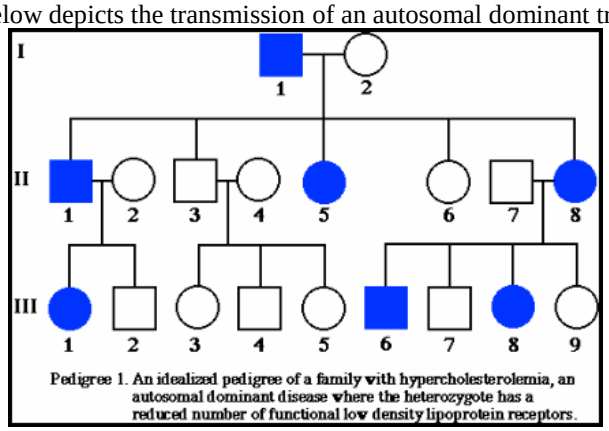


Modes of Inheritance—Autosomal Dominant

The genes for autosomal dominant traits are located on one of the 44 autosomes in a human cell. The genes are dominant, which indicates that individuals with only one copy of the gene are affected. Characteristics of autosomal dominant traits include:

- Males and females are equally likely to have the trait.
- Traits do not skip generations because there are no carriers for the traits.
- The trait is present whenever a single copy of the corresponding allele is present.
- There is male-to-male transmission. This means that fathers can pass to trait to their sons.

Some examples of human genetic conditions that are transmitted via the autosomal dominant mode of inheritance include: Huntington's Disease, Achondroplasia, Polycystic Kidney Disease, and Familial Hypercholesterolemia.
The pedigree chart included below depicts the transmission of an autosomal dominant trait.



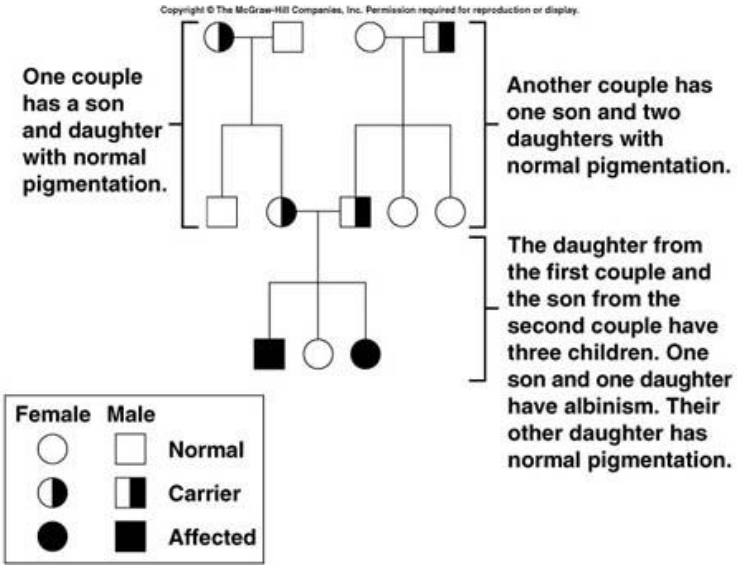
Pedigree 1. An idealized pedigree of a family with hypercholesterolemia, an autosomal dominant disease where the heterozygote has a reduced number of functional low density lipoprotein receptors.

Modes of Inheritance—Autosomal Recessive

The genes for autosomal recessive traits are located on one of the 44 autosomes in a human cell. The genes are recessive, which indicates that individuals must possess two copies of the recessive allele in order to exhibit the trait. Characteristics of autosomal recessive traits include:

- Males and females are equally likely to have the trait.
- Traits often skip generations (due to the possibility of carriers).
- Only homozygous recessive individuals have the trait.
- Traits may appear in siblings without appearing in their parents.
- If a parent has the trait, those offspring who do not have it are heterozygous carriers of the trait.
- Some examples of human genetic conditions that are transmitted via the autosomal recessive mode of inheritance include: Tay Sachs Disease, PKU, and Cystic Fibrosis.

The pedigree chart included below depicts the transmission of an autosomal recessive trait.



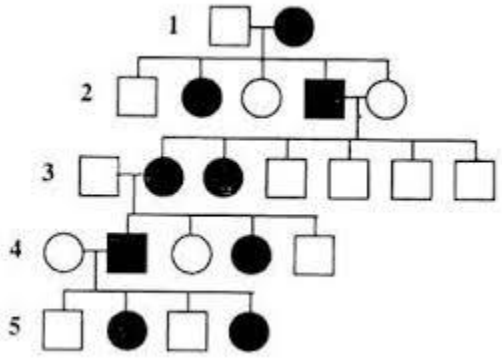
Modes of Inheritance—X-linked Dominant

The genes for X-linked dominant traits are located on the X chromosome in a human cell. The alleles are dominant, so individuals with only one copy of the allele exhibit the trait. Characteristics of X-linked dominant traits include:

- The trait is present whenever the corresponding gene is present.
- There is no male-to-male transmission. Fathers cannot pass the trait to their sons.
- A female who has the trait may or may not pass the allele for the trait to her son or daughter.
- All of the daughters of a male with the trait will inherit the trait.

Some examples of human genetic conditions that are transmitted via the X-linked dominant mode of inheritance include: hypertrichosis, porphyria, and Rett syndrome.

The pedigree chart included below depicts the transmission of an X-linked dominant trait.

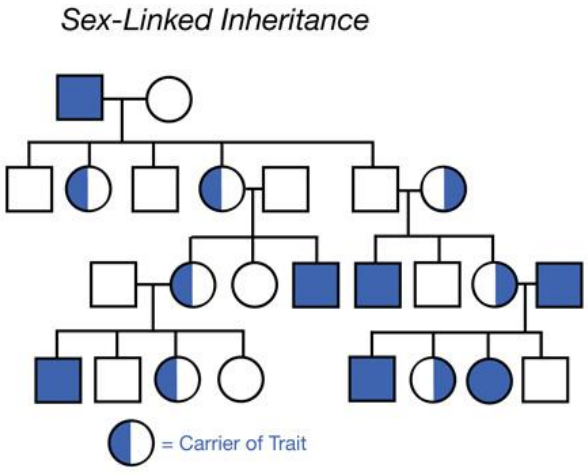


Modes of Inheritance—X-linked Recessive

The alleles for X-linked recessive traits are located on the X chromosome in a human cell. The alleles are recessive, so for a female to exhibit the trait she must possess the allele on both of her X chromosomes. Since a male has only one X chromosome, he is affected if his X chromosome carries the allele. Characteristics of X-linked recessive traits include:

- These traits are far more common in males than in females.
- Traits may skip generations.
- All daughters of a male who has the trait are either affected or are heterozygous carriers.
- The son of a female carrier has a 50 percent chance of having the trait.
- Mothers of males who have the trait are either heterozygous carriers or homozygous and express the trait.
- There is no male-to-male transmission. Fathers cannot pass the trait to their children.
-

Some examples of human genetic conditions that are transmitted via the X-linked recessive mode of inheritance include: hemophilia, red/green colorblindness, Duchenne Muscular Dystrophy, and Lesch-Nyhan syndrome.
The pedigree chart included below depicts the transmission of an X-linked recessive trait.



Inheritance of Red-Green Blindness:
an X-linked Recessive Trait

Modes of Inheritance—Y-linked

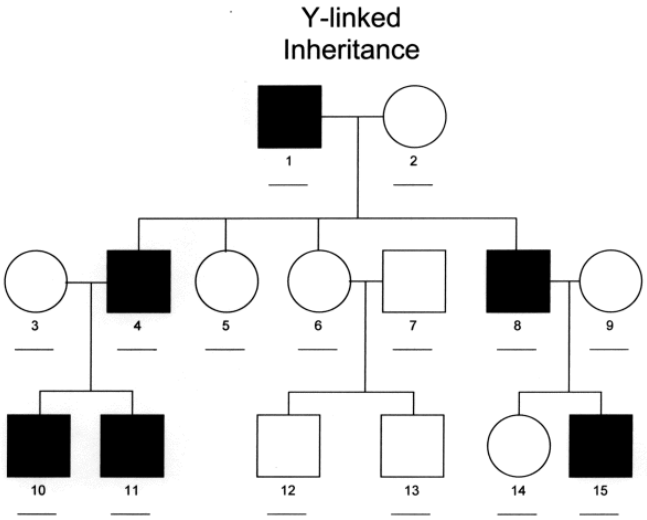
The alleles for Y-linked traits are located on the Y chromosome in a human cell. Since males possess only one Y chromosome, the traits aren't usually referred to as dominant or recessive. A man who possesses a Y-linked allele will exhibit the trait it controls. Since females don't possess a Y chromosome, they are not affected by Y-linked traits.

Characteristics of Y-linked traits include:

- Only males are affected.
- All sons of an affected male will also be affected.
- Females can't possess or pass on the trait.

Some forms of Retinitis pigmentosum (a form of progressive blindness) are transmitted via the Y-linked mode of inheritance

The pedigree chart included below depicts the transmission of an Y-linked trait.



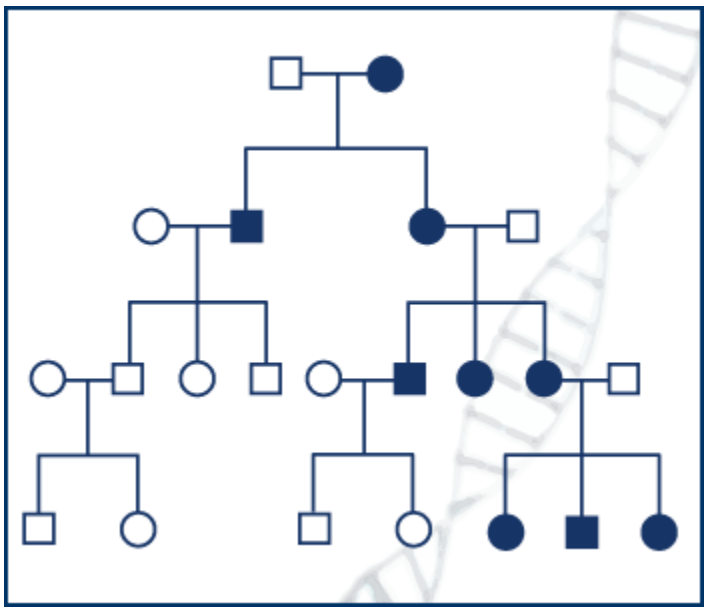
Modes of Inheritance—Mitochondrial

The alleles for mitochondrial traits are located on the small circular chromosome found in the mitochondria. Since children of both sexes inherit their organelles (including the mitochondria) from their mother, mitochondrial traits are always transmitted from mother to child. Since the mitochondria only have one chromosome each, there are no dominant or recessive mitochondrial traits. Individuals who possess a single copy of a mitochondrial allele express the trait. Characteristics of mitochondrial traits include:

- All of the children of an affected female will express the trait.
- Males can express the trait, but are unable to pass it on to their offspring.

The human mitochondrial chromosome is a very small piece of DNA and codes for only 13 proteins which are part of the electron transport chain. Disorders related to mutations in the mitochondrial chromosome usually affect one's ability to make ATP.

The pedigree chart included below depicts the transmission of a mitochondrial trait.



Gene Regulation

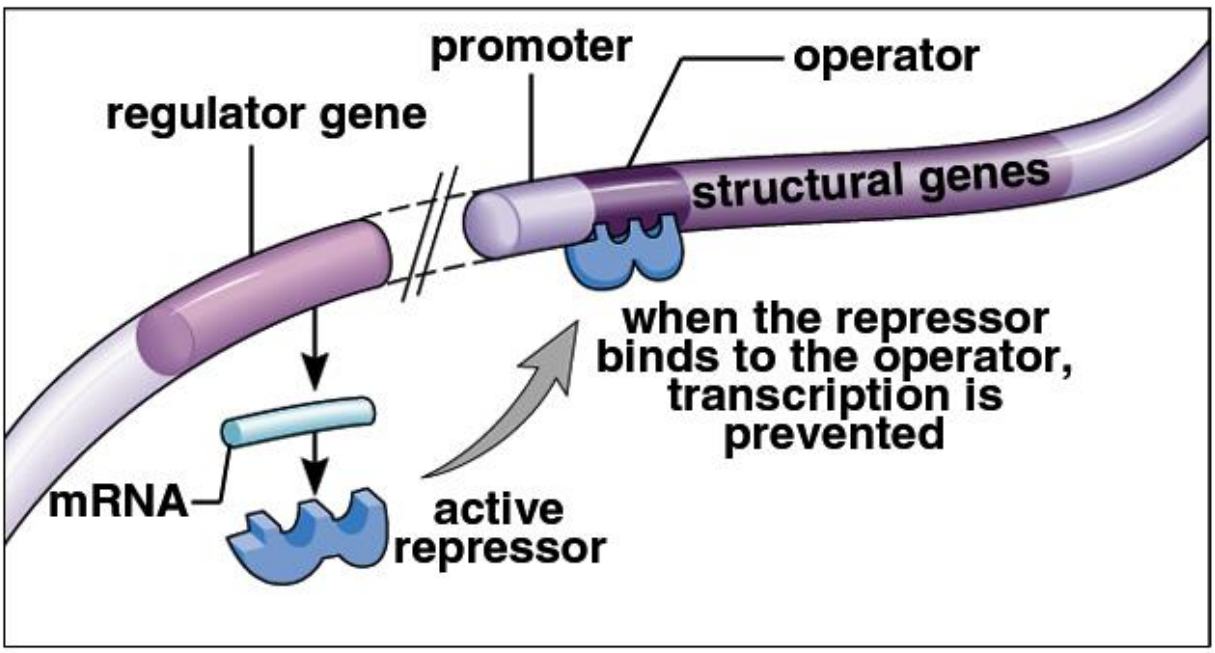
Every cell in a multicellular organism contains the same DNA.
Even though this is true, organisms contain cells that look and function very differently from each other. This is possible because of **selective gene expression**. This essentially means that only certain genes are expressed (used to make proteins) in certain cells. The same genes may be inactive in different cell types within the same organism.

Scientists have worked for years to determine how the processes of selective gene expression and gene regulation are carried out in organisms.
We will first take a look out how genes are regulated in prokaryotes since these cells represent the simplest possible case.

Operons

Prokaryotic genes tend to be organized in related groups. These **structural genes** and the other DNA sequences that help to regulate their expression are known as **operons**.
An operon might be thought of a functional unit of transcription and gene regulation. For a gene to be expressed (on) it must be transcribed and translated.
Operons typically consist of the following parts:
Promoter—This DNA sequence serves as the site of RNA polymerase binding. RNA polymerase is the enzyme that carries out transcription.
Operator—This sequence serves as the site of repressor binding.
Structural genes—The DNA sequences/genes that code for the actual proteins/enzymes of the operon.
Regulatory Genes—These sequences code for repressor proteins.

Repressor Proteins—In certain situations, these proteins bind to the operator and prevent the transcription and expression of the structural genes.

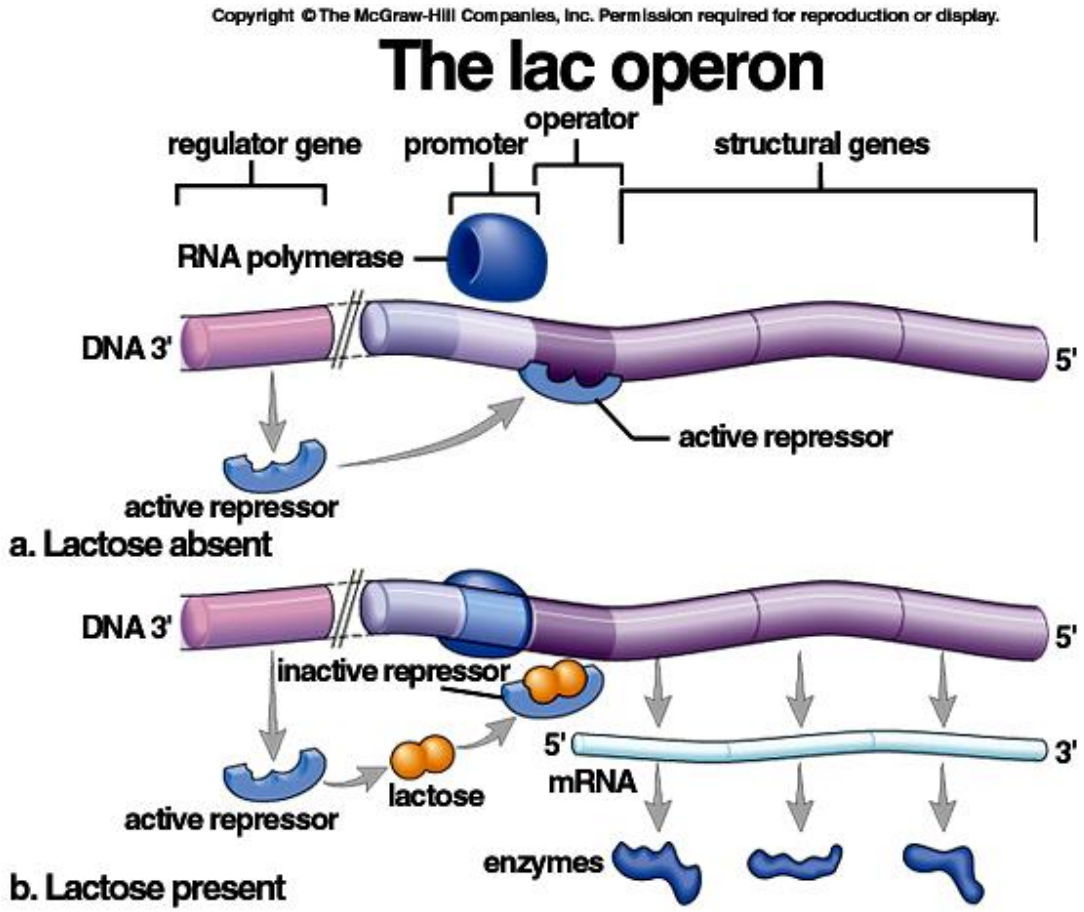


Operons are often classified as either **inducible or repressible**.

Inducible Operons

The structural genes in **inducible operons** are **usually not expressed** (the genes are normally "off"). This means that repressor proteins are normally bound to the operator of the operon. **These proteins block transcription of the structural genes and prevent the expression of the genes.** Under certain environmental or internal conditions, an **inducer** binds to the repressor protein and causes it to change shape and unbind from the operator. Once the repressor protein is released from the operator, RNA polymerase transcribes the genes and they are expressed.

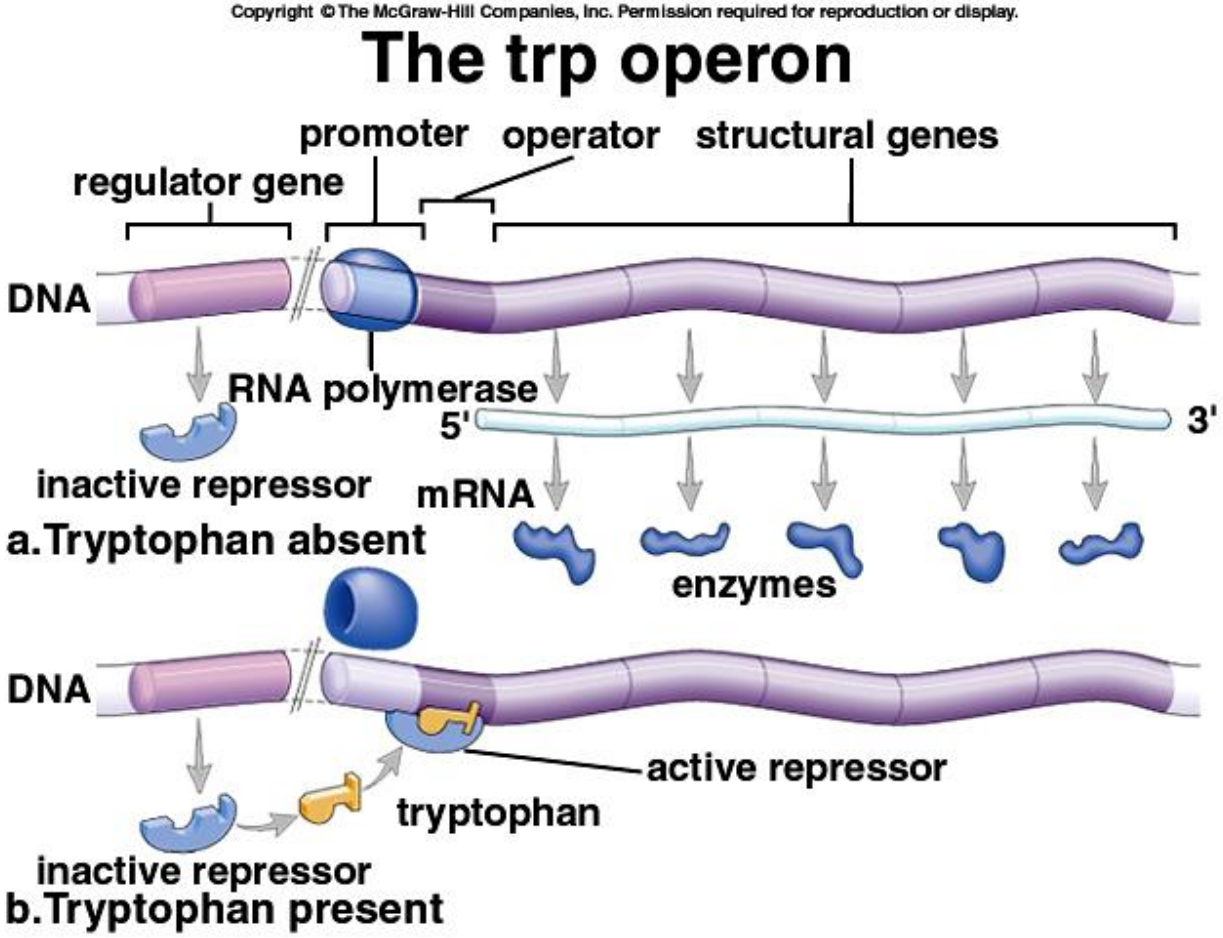
A commonly cited example of an inducible operon is the **lac operon**. The structural genes in this operon code for a group of enzymes (**collectively referred to as lactase**) that break down the carbohydrate lactose.



If bacterial cells don't have access to lactose, it makes no sense for them to manufacture lactase, so in the absence of lactose the structural genes of the lac operon are not expressed.
If, however, the bacterial cells have access to lactose, the **lactose acts as an inducer.** It binds to the repressor proteins and causes them to change shape and unbind from the operator of the lac operon. Once this happens, the genes are **transcribed and translated**. The cell now possesses lactase and can metabolize the lactose and use it as a food source. Once all of the lactose is used up, the repressor proteins once again bind to the operator and the expression of the structural genes ceases.
Operons function much like negative feedback loops and allow prokaryotes to express genes only when they need to be expressed. This helps to save energy and resources for the cells.

Repressible Operons

The structural genes in **repressible operons** are **usually expressed** (the genes are normally "on"). This means that repressor proteins are normally not bound to the operator of the operon. Under certain environmental or internal conditions, a **corepressor** binds to the repressor protein and changes its shape so that it can bind to the operator and then block the transcription and expression of the structural genes.
A commonly cited example of a repressible operon is the **trp operon**. The structural genes in this operon code for a group of enzymes that make the **amino acid tryptophan**. This amino acid is essential for bacteria and **they don't normally get enough of it from their diet**, thus the need from the enzymes that manufacture it.



If bacteria are exposed to large amounts of tryptophan in their environment, it doesn't make sense for them to use energy and resources to manufacture something that is freely available. In this case, the tryptophan acts as an corepressor and binds to the inactive repressor and changes its shape so that it can bind to the operator. Once there, the structural genes are no longer expressed and the cell stops making the enzymes to produce tryptophan. Once the environmental tryptophan is used up, the repressor protein unbinds from the operator and the cell once again expresses the structural genes and produces the enzymes needed to manufacture tryptophan. **Like inducible operons, repressible operons function like negative feedback loops and help bacterial cells to conserve energy and resources.**

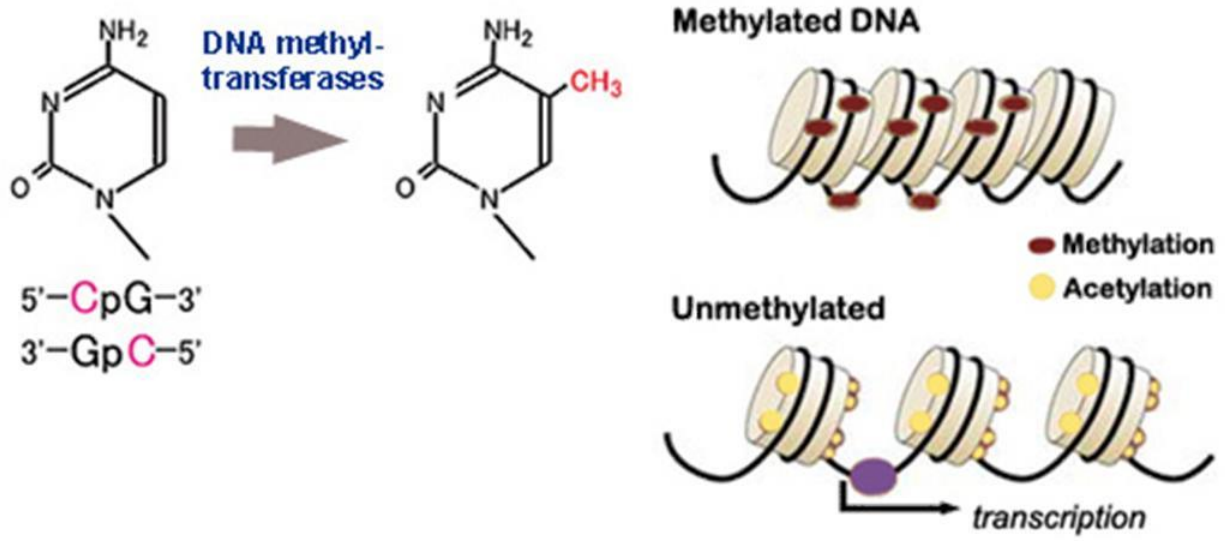
Gene Regulation in Eukaryotic Cells

The expression of eukaryotic genes is controlled on several different levels. These include:

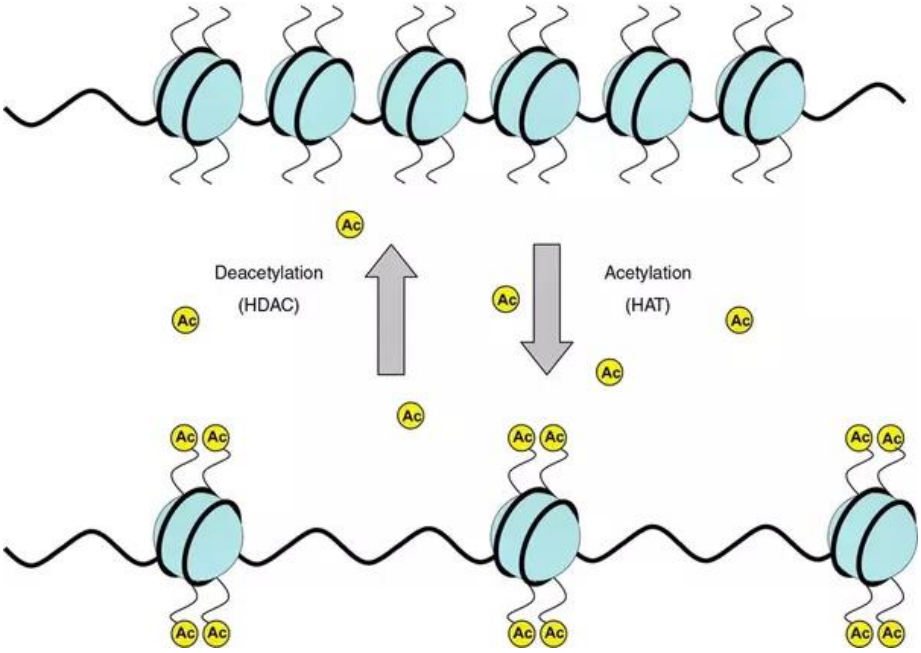
A. **Chromatin Structure**—In Unit 5 you learned that eukaryotic DNA is associated with proteins known as histones. The histones help to organize, compact, and regulate the expression of DNA. DNA is often organized into nucleosomes (a small portion of DNA wrapped around a group of histones). The degree to which the DNA is compacted and coiled affects the accessibility of the

DNA to RNA polymerase and thus affects if the DNA can be transcribed and expressed. Active/expressed genes are typically loosely compacted and coiled and are referred to as **euchromatin**. Inactive genes are tightly compacted/coiled and are referred to as **heterochromatin**. Genes can be inactivated when **methyl groups (-CH_3)** are attached to the cytosines of a gene. The addition of these **methyl tags** causes the DNA to tightly wind around the histones and become inactive and not expressed. Methylation is typically permanent. **Methylation deactivates specific genes in different cells types and plays an important role in cell differentiation.**

An example of the results of methylation is the **Barr body** in mammalian females. Each cell in a female mammal has two "X" chromosomes, but only one active "X" chromosome. The second "X" chromosome is highly methylated and is compacted into a tiny structure known as a Barr Body. The genes of this "X" are completely inactive and do not produce proteins.



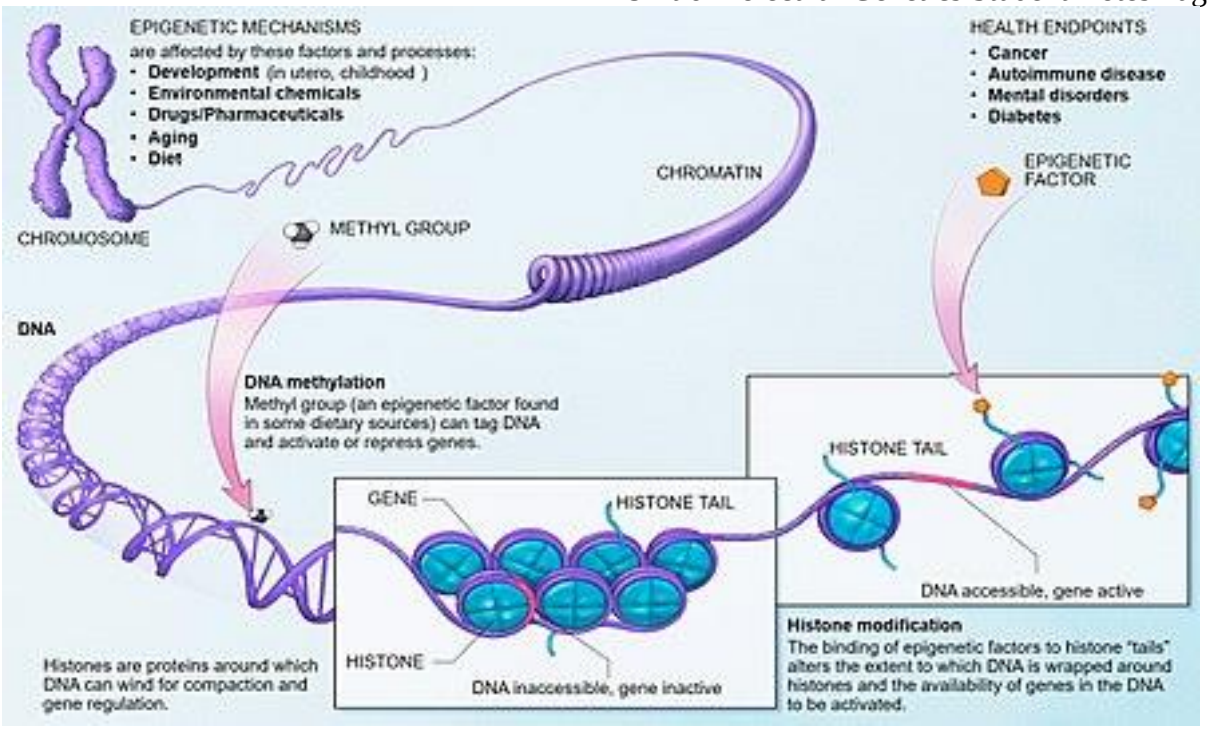
Another **tag** that affects chromatin structure is the **acetyl group (--$COCH_3$).** Acetyl groups are often attached to the histone proteins. This causes the DNA that is wrapped around the histone to uncoil and thus the genes are transcribed and expressed. **Acetylation (the process of adding acetyl groups) plays an important role in allowing cells to respond to changes in their internal and/or external environments. Acetylation is a reversible process.**



Epigenetics

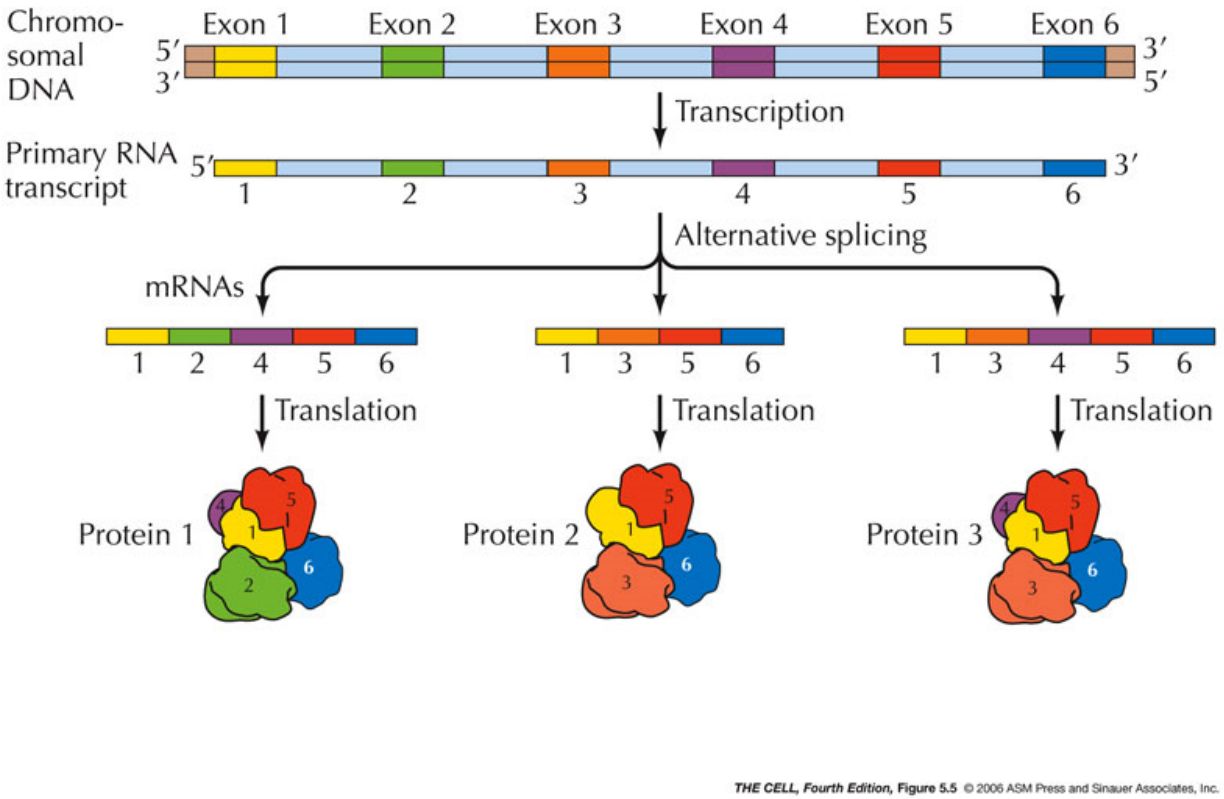
Acetylation and Methylation are often linked to a phenomenon known as **epigenetics.** The term epigenetics refers to changes in the expression of genes that do not depend on the nucleotide sequence itself. Epigenetics is responsible for the changes that occur through the years in identical twins. Even though the twins have identical nucleotides sequences, the patterns of methylation and acetylation in their genomes become more and more different over time due to the twins' different lifestyles and environments.

Methylation and acetylation patterns can also sometimes be inherited from parents. Although in most cases, epigenetic tags are removed from the DNA during either meiosis or events right after fertilization, there are some situations in which the methylation and acetylation patterns are transmitted from parent to child. Although scientists don't yet know how important this **epigenetic form of inheritance** is, it is possible that the diet, lifestyle habits (like smoking, drug use, lack of exercise), and exposure to environmental toxins of an individual's parents may cause epigenetic changes in the DNA of both the parent and child. These changes may then be passed on to future generations without changing the actual nucleotide sequence of the DNA.



B. Transcriptional Control—In eukaryotic cells, proteins known as transcription factors help regulate transcription (and therefore gene expression) by assisting the binding of RNA Polymerase to the promoter. Most eukaryotic genes are also regulated by transcription activators (that enhance transcription) and repressors (which block transcription).

C. Posttranscriptional Control—Once the pre-mRNA or primary transcript is made during transcription, it next has to be processed. One of the most important parts of RNA processing is the removal of the intron sequences and the splicing together of the exon sequences. RNA can be **alternatively processed** and thus one gene can actually be expressed as multiple proteins.



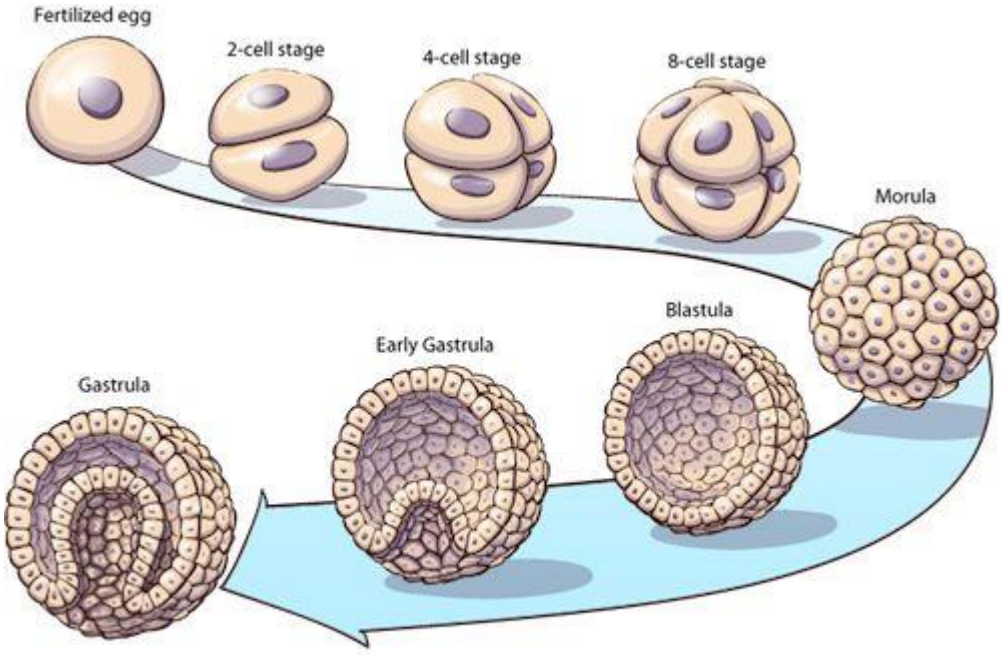
THE CELL, Fourth Edition, Figure 5.5 © 2006 ASM Press and Sinauer Associates, Inc.

D. Translational Control—Small segments of RNA, known as microRNAs, can bind to and disable the translation of mRNA molecules and affect the processes of development and host/pathogen interactions.

E. Posttranslational Control—The expression of a gene can also be controlled by how the protein is folded after translation. Additionally, the lifespan of a protein in a cell, is usually regulated by enzymes known as proteases. These enzymes, usually housed in the lysosomes or proteasomes, break down proteins usually after they have been tagged by a signaling protein such as **ubiquitin.**

Embryonic Development

Embryonic development begins when a haploid egg and a haploid sperm fuse to form the organism's first diploid cell, the **zygote**. This cell goes through a 12 to 24 hour period of rapid cell division (mitosis) known as **cleavage**. During this stage, cell division occurs so rapidly that the cells have little time to grow. When the developing organism reaches the 32 cell stage, known as the **morula**, it is the same size as the original zygote. The morula consists of a solid ball of cells that is still surrounded by the zona pellucida which surrounded and protected the egg.

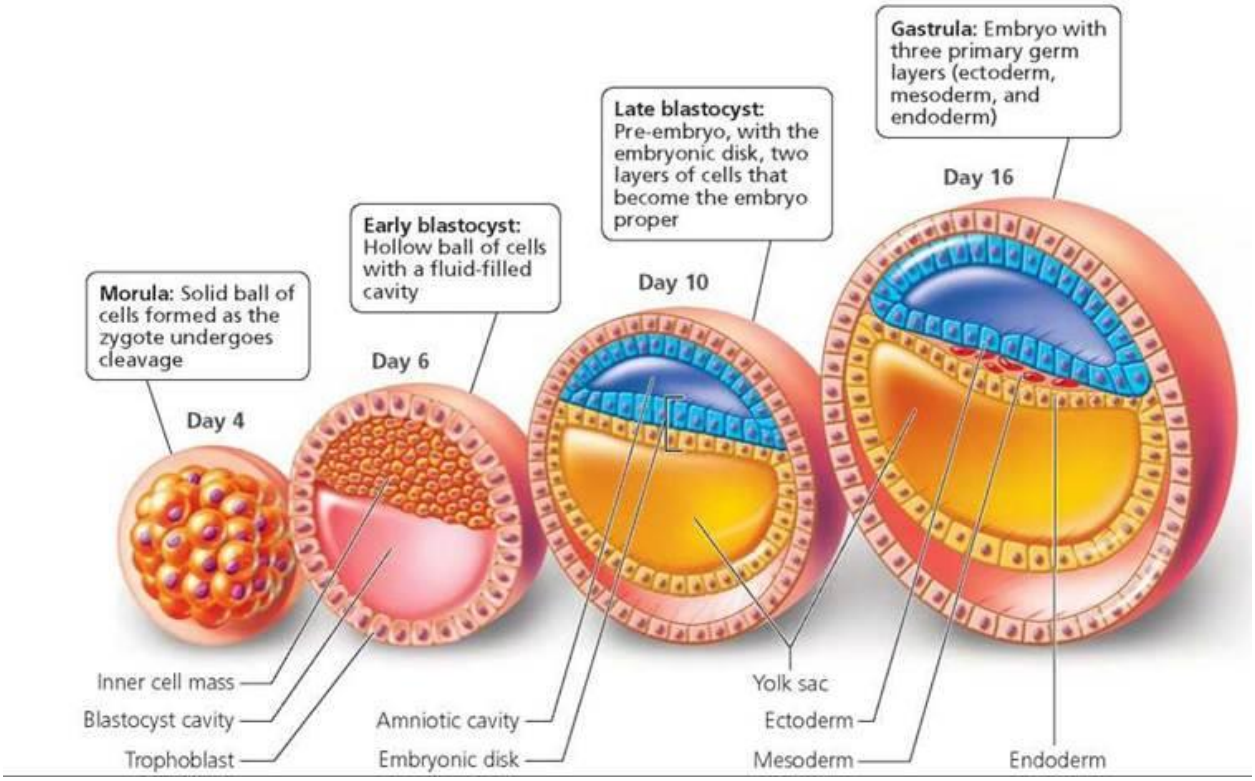


Cell Cleavage
Process by which the number of cells in a developing embryo is multiplied through cell division.

At this point in development all of the cells are **totipotent stem cells**. These cells have the potential to become any type of cell in the embryo or in the extraembryonic tissues (cells that will become the placenta) because essentially none of the genes have been deactivated by methylation.

By day 4 of embryonic development, the cells continue to divide, but they also begin to differentiate. During this stage, known as **blastulation,** two layers develop: an outer shell known as the **trophoblast** (which will eventually form the placenta) and an inner cell mass that will form the embryo. The inner cell mass is pushed off to one side of the structure, while the rest of the inside of the sphere forms a fluid-filled cavity called the **blastocoel**. At this point in development, the entire ball of cells is known as the **blastocyst** (in mammals) or as the **blastula** (in non-mammal animals). During the blastula stage, the cells of the inner cell mass are **pluripotent stem cells**. This means that they have begun the process of differentiation and that some of their genes have been deactivated through methylation. Pluripotent cells are capable of becoming any type of cell in the embryo, but they now lack the ability to become extraembryonic cells.

During Week 3 of embryonic development the cells of the inner cell mass go through the process of **gastrulation**. During gastrulation, the cells organize themselves into three distinct layers known as the germ layers. The cells in the outer layer (**the ectoderm**) will eventually go on to form the epidermis, the hair, the nails, the brain, the spinal cord, and the peripheral nervous system. The cells of the middle germ layer (**the mesoderm**) will go on to form the muscles, the bones, the connective tissues, the notochord, the kidneys, the gonads, and the circulatory system. The cells of the innermost germ layer (**the endoderm**) go on to form the epithelial lining of the digestive tract, the stomach, the colon, the liver, the pancreas, the bladder, and the lungs. Once the cells are organized into the three germ layers, the entire ball of cells is known as the **gastrula.** Also during gastrulation, an invagination forms an opening (**known as the blastopore)** in the gastrula. In protostomes (like worms, mollusks, and arthropods) this opening eventually becomes the mouth. In deuterostomes (like echinoderms and chordates) this opening becomes the anus.



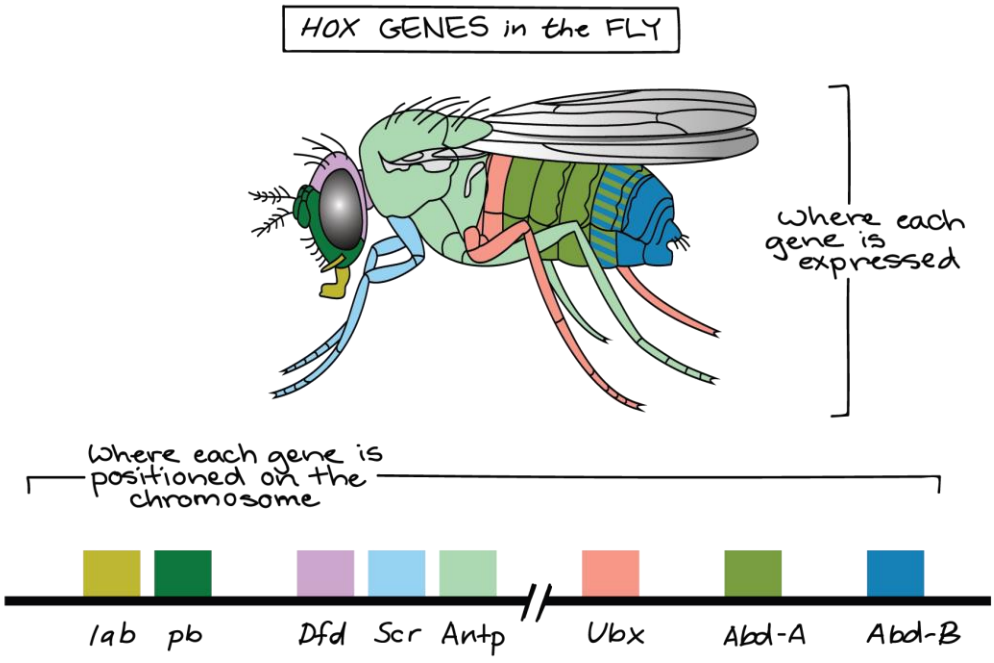
After gastrulation, the cells undergo **neurulation**. During this process, the notochord forms and causes the ectoderm to bend itself into a tube known as the **neural tube**. The neural tube will give rise to the brain and spinal cord. The cells of the mesoderm begin to give rise to the kidneys, gonads, adrenal glands, blood vessels, and muscles of the organs. At the same time, the cells of the endoderm roll themselves into a tube known as the digestive tract. The organs of the gastrointestinal tract start off as outpouchings of the digestive tract. The endoderm also give rise to the lungs.

Hox Genes

Hox genes are a group of related genes that control the body plan of an embryo along the head-tail axis. After the embryonic segments have formed, the Hox proteins determine the type of appendages (e.g. legs, antennae, and wings in fruit flies) or the different types of vertebrae (in humans) that will form on a segment. Hox proteins thus confer segmental identity, but do not form the actual segments themselves.

An analogy for the Hox genes can be made to the role of a play director that calls which scene the actors should carry out next. If the play director calls the scenes in the wrong order, the overall play will be presented in the wrong order. Similarly, mutations in the Hox genes can result in body parts and limbs in the wrong place along the body. Like a play director, the Hox genes do not act in the play or participate in limb formation themselves.

The protein product of each Hox gene is a transcription factor. Each Hox gene contains a well-conserved DNA sequence known as the homeobox, of which the term "Hox" is a contraction. In many animals, the organization of the Hox genes in the chromosome is the same as the order of their expression along the anterior-posterior axis of the developing animal



DNA Biotechnology

Genetic Engineering

Genetic engineering, sometimes called **genetic modification**, is the process of altering the DNA in an organism's genome. This may mean changing one base pair (A-T or C-G), deleting a whole region of DNA, or introducing an additional copy of a gene. It may also mean extracting DNA from another organism's genome and combining it with the DNA of that individual. Genetic engineering is used by scientists to enhance or modify the characteristics of an individual organism. Genetic engineering can be applied to any organism, from a bacterium to a sheep. For example, genetic engineering can be used to produce plants that have a higher nutritional value or that can tolerate exposure to herbicides.

Biotechnology involves the use of living systems, organisms, or parts of organisms to manipulate natural processes in order to develop products, systems, or environments to benefit people. These may be products, such as foods, pharmaceuticals, or compost; systems, such as waste management or water purification; or environments, such as hydroponics. Biotechnology also includes genetic or biomedical engineering.

Recombinant DNA technology is the process of joining together **DNA** molecules from two different species. The recombinant DNA is then inserted into a host organism to produce new gene products that are of value to science, medicine, agriculture, and industry.

Possible Applications of Biotechnology and Recombinant DNA

Gene Therapy-- **Gene therapy** is an experimental technique that uses **genes** to treat or prevent disease. In the future, this technique may allow doctors to treat a disorder by inserting a **gene** into a patient's cells instead of using drugs or surgery.

Pharmaceuticals—Biotechnology is currently used to create medical products such as human insulin, human growth hormone, and vaccines.

Criminal Forensics—Biotechnology is used create DNA fingerprints and identify criminal suspects.

Paternity/Maternity testing—Biotechnology is used to accurately determine the parentage of children.

Agriculture—Biotechnology has been used to create Genetically Modified Foods and Genetically Modified Organisms (GMOs). Some crop plants have been modified to be herbicide resistant. Others have been modified to create larger and hardier fruits. Still others have been modified to be cold weather tolerant and disease resistant.

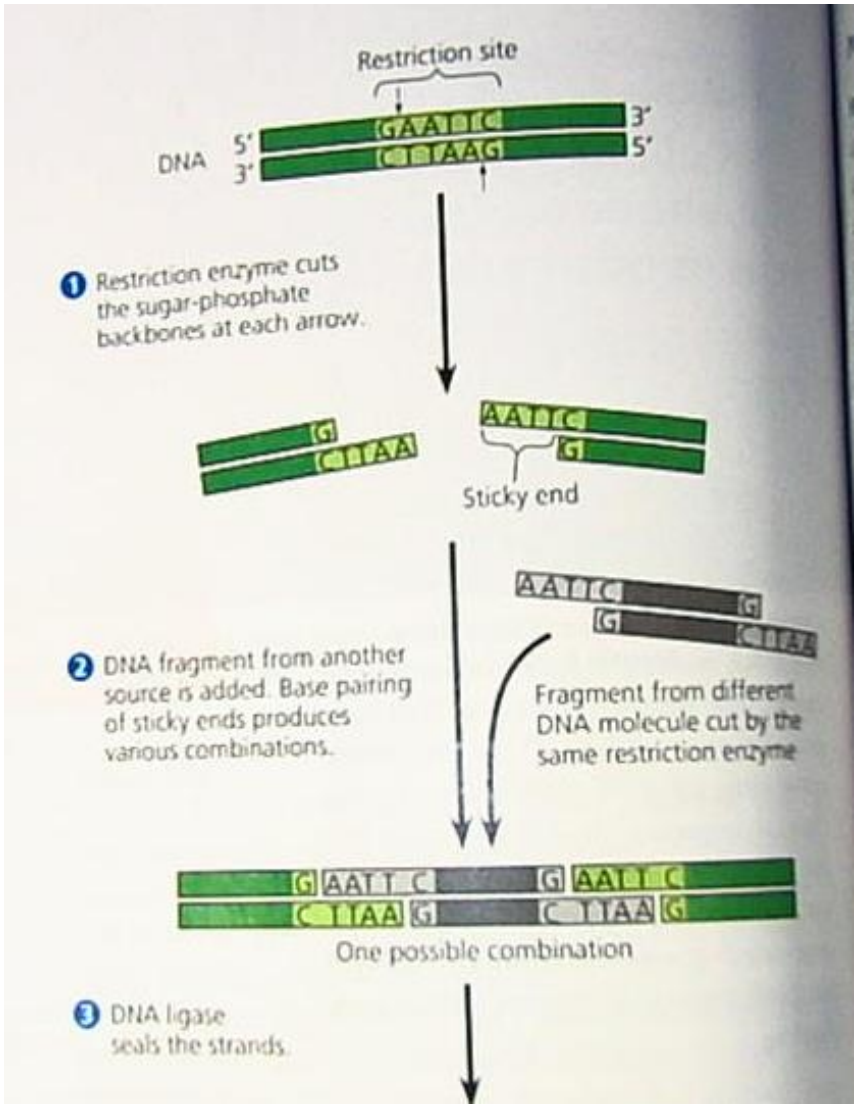
DNA Cloning

DNA cloning is the process of making multiple, identical copies of a particular piece of DNA. In a typical DNA cloning procedure, the gene or other DNA fragment of interest (perhaps a gene for a medically important human protein) is first inserted into a circular piece of bacterial DNA called a **plasmid**. The insertion is done using enzymes that "cut and paste" DNA, and produces a molecule of **recombinant DNA**, or DNA assembled out of fragments from different organisms/sources. In some cases, we need lots of DNA copies to conduct experiments or to build new plasmids. In other cases, the piece of DNA encodes a useful protein, and the bacteria are used as "factories" to make the protein. For instance, the human insulin gene is expressed in *E. coli* bacteria to make insulin for use by diabetics.

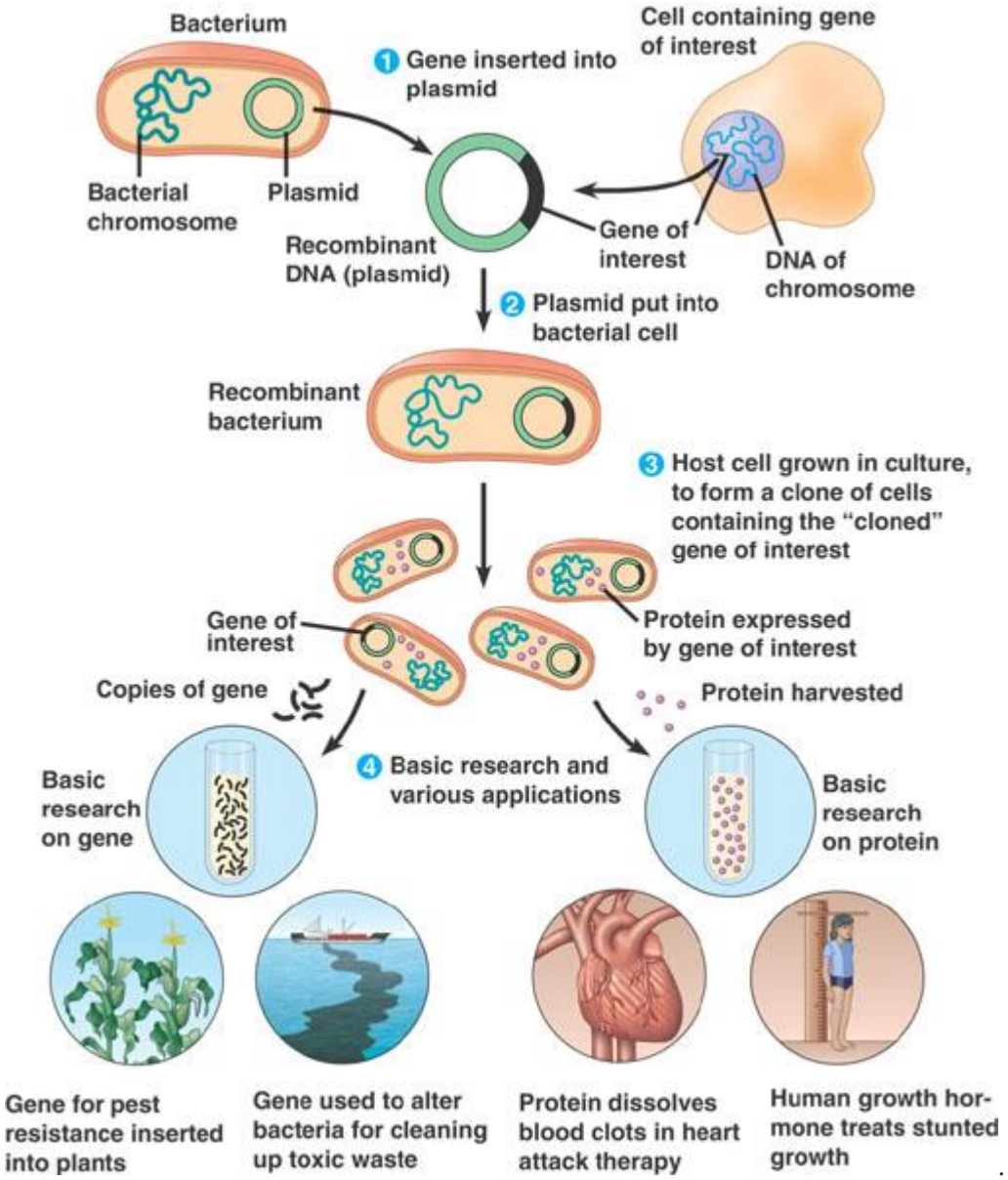
Next, the recombinant plasmid is introduced into and used to **transform** bacteria. Bacteria carrying the plasmid are selected and grown. As they reproduce, they replicate the plasmid and pass it on to their offspring, making copies of the DNA it contains and potentially produce lots of the protein coded for by the inserted gene.

Steps in the DNA Cloning Process

Cut open a bacterial plasmid with **restriction enzymes**. Restriction enzymes are enzymes that were isolated from bacteria. They are used by the bacteria to defend themselves against bacteriophages. They cut DNA after specific nucleotide sequences known as restriction sites. Restriction enzymes have strange names like EcoRI and HinDIII. They are often named after the bacteria they were discovered in. Many restriction enzymes cut the two strands of the DNA molecule at different points and thus leave behind unpaired base sequences known as "sticky ends". If the gene of interest (the one the scientist wants to insert into the plasmid) is cut with the same restriction enzyme, the sticky ends match and bond together. Once bonded, the enzyme ligase is used to permanently bond the strands. At this point, a recombinant DNA plasmid has been created.



Next, the plasmid is inserted into the bacteria. The bacteria are said to be **transgenic** because they contain DNA from another species. A gene that codes for resistance to a particular antibiotic is often also inserted into the plasmid. The insertion of the plasmid into the bacteria is facilitated via a process of chemical treatment and heat shock. Since most of the bacteria are not transformed due to the inefficiency of the process, the bacteria are cultured in a medium containing an antibiotic (the one that the plasmid codes resistance to). This process eliminates the untransformed bacteria and allows only the transformed ones (the ones that took in the recombinant plasmid) to be artificially selected for. The bacteria are then cultured in large vats. Every time one of the bacteria reproduce, they reproduce the plasmid. The bacteria essentially become factories that are used to create the protein coded for by the inserted gene. The proteins are then harvested and purified.



Polymerase Chain Reaction (PCR)

Polymerase chain reaction (**PCR**) is a laboratory technique used to quickly produce many copies (millions or billions!) of a particular region of DNA. This DNA region can be anything the experimenter is interested in. For example, it might be a gene whose function a researcher wants to understand, or a genetic marker used by forensic scientists to match crime scene DNA with the DNA of a suspect or victim.

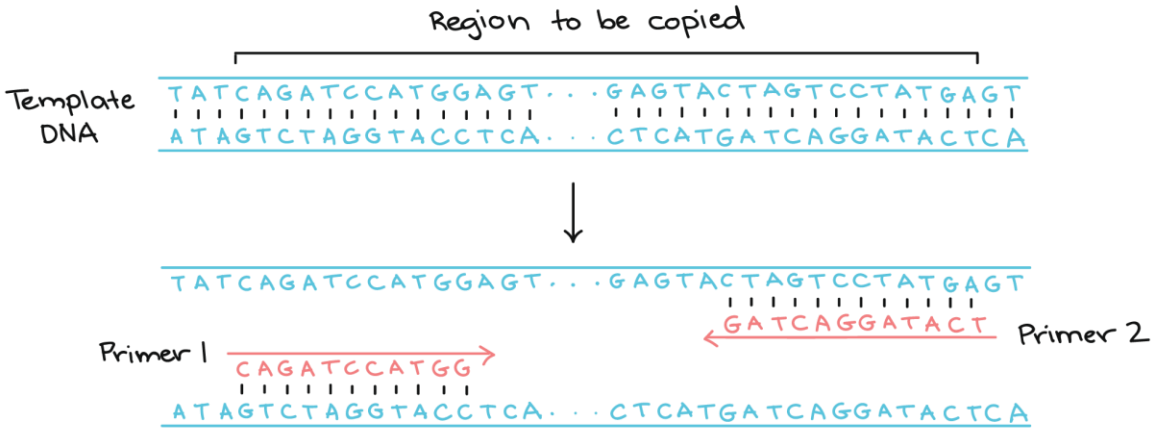
Typically, the goal of PCR is to make enough copies of the target DNA region so that it can be analyzed or used in some other way. For instance, DNA amplified by PCR may be sent for sequencing, visualized by gel electrophoresis, or cloned into a plasmid for further experiments.

PCR is used in many areas of biology and medicine, including molecular biology research, medical diagnostics, and even some branches of ecology.

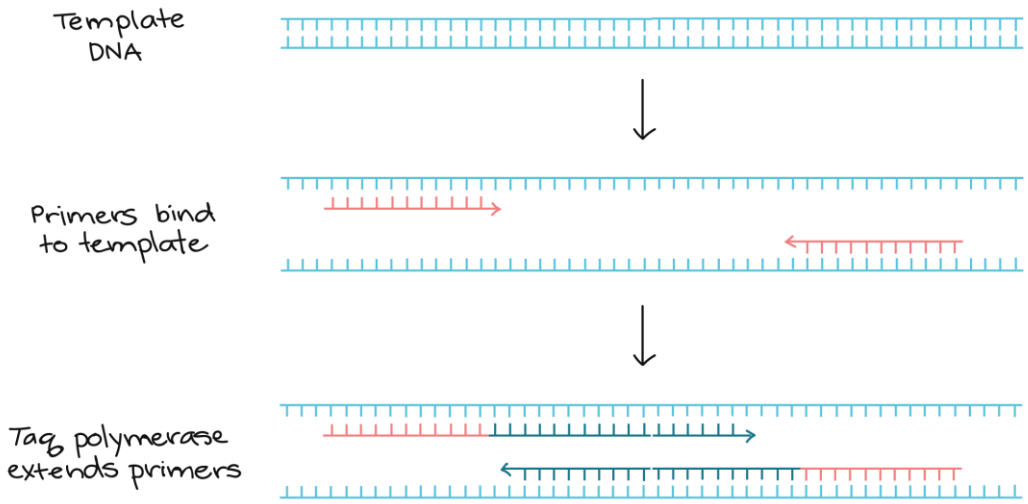
Like DNA replication in an organism, PCR requires a DNA polymerase enzyme that makes new strands of DNA, using existing strands as templates. The DNA polymerase typically used in PCR is called ***Taq* polymerase**, after the heat-tolerant bacterium from which it was isolated (***Thermus aquaticus***).

T. aquaticus lives in hot springs and hydrothermal vents. Its DNA polymerase is very heat-stable and is most active around 70°C (a temperature at which a human or *E. coli* DNA polymerase would be nonfunctional). This heat-stability makes Taq polymerase ideal for PCR. As we'll see, high temperature is used repeatedly in PCR to **denature (separate its strands)** the template DNA.

Like other DNA polymerases, ***Taq* polymerase** can only make DNA if it's given a **primer**, a short sequence of nucleotides that provides a starting point for DNA synthesis. In a PCR reaction, the experimenter determines the region of DNA that will be copied, or amplified, by the primers she or he chooses.
PCR primers are short pieces of single-stranded DNA, usually around 20 nucleotides in length. Two primers are used in each PCR reaction, and they are designed so that they flank the target region (region that should be copied). That is, they are given sequences that will make them bind to opposite strands of the template DNA, just at the edges of the region to be copied. The primers bind to the template by complementary base pairing.



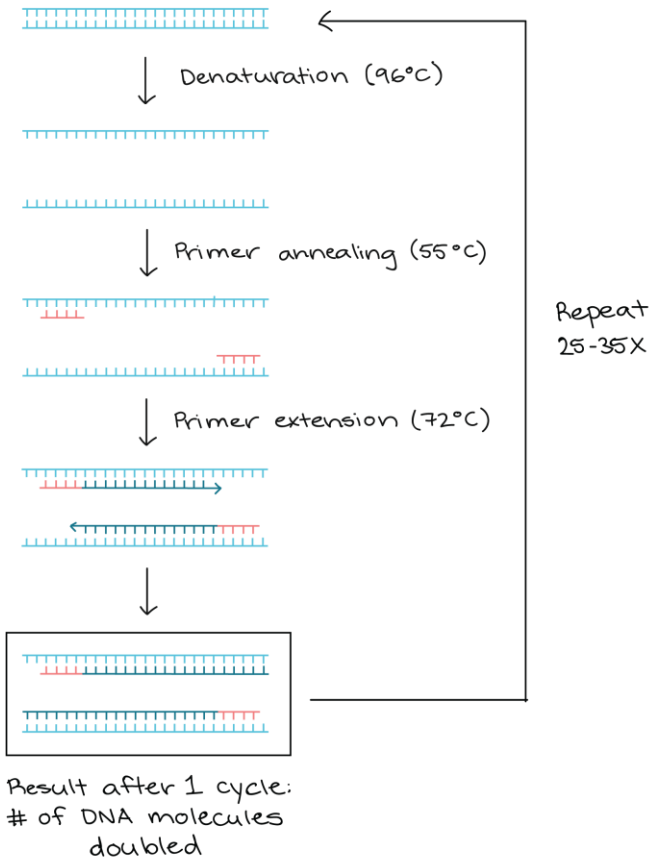
When the primers are bound to the template, they can be extended by the polymerase, and the region that lies between them will get copied.



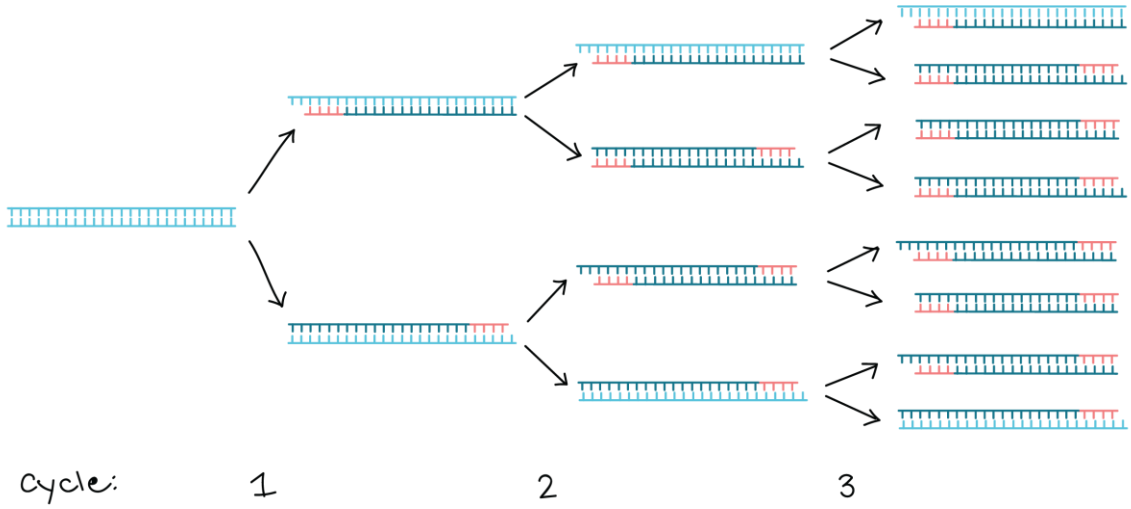
Steps of the PCR Process

1. **Denaturation** (96°C): Heat the reaction strongly to separate, or denature, the DNA strands. This provides single-stranded DNA templates for the next step.
2. **Annealing** (55 - 65°C): Cool the reaction so the primers can bind to their complementary sequences on the single-stranded template DNA.

3. **Extension** (72°C): Raise the reaction temperature so that *Taq* polymerase can extend the primers and synthesize the new strands of DNA.
4. This cycle is then repeated 25 - 35 times in a typical PCR reaction, which generally takes 2 – 4 hours, depending on the length of the DNA region being copied. If the reaction is efficient, billions of the copies of the target region can be made.



The process is able to make so many copies of the original DNA because the new DNA that's made in each round can serve as a template in the next round of DNA synthesis. There are many copies of the primers and many molecules of *Taq* polymerase floating around in the reaction, so the number of DNA molecules can roughly double in each round of cycling. This pattern of exponential growth is shown in the image below.



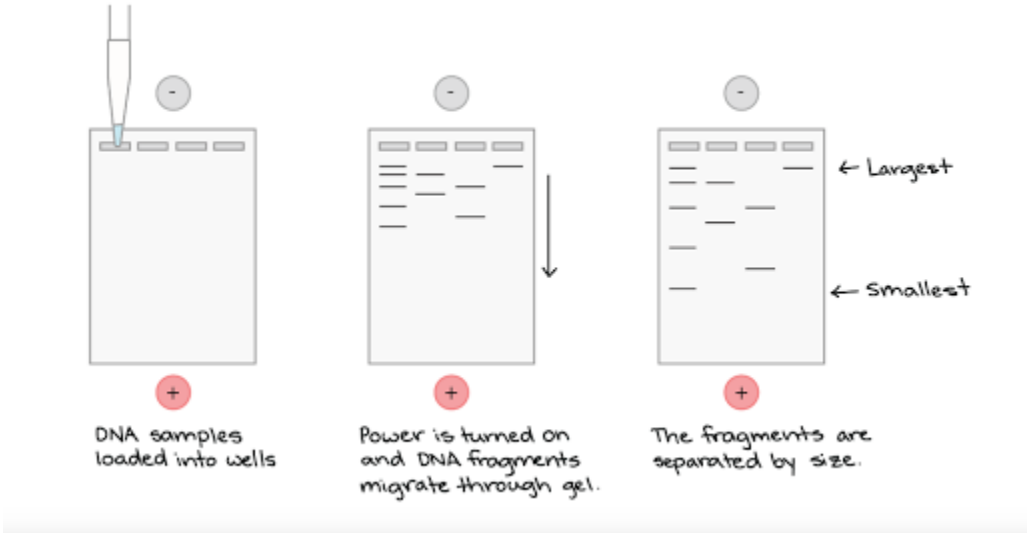
Gel Electrophoresis

Gel electrophoresis is a technique used to separate DNA fragments, RNA fragments, or proteins based on their size and charge. Electrophoresis involves running a current through a gel containing the molecules of interest. Based on their size and charge, the molecules will travel through the gel in different directions or at different speeds, allowing them to be separated from one another.

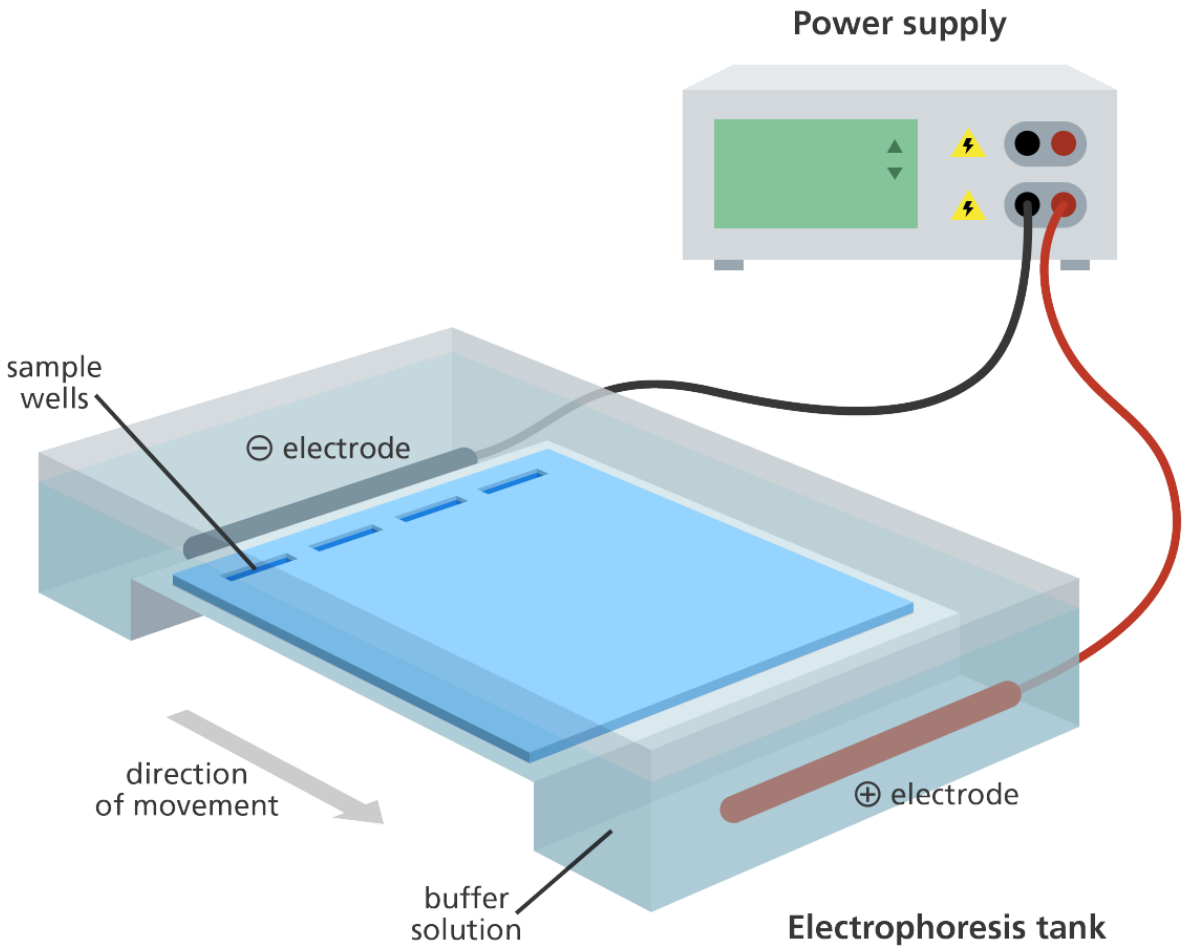
All DNA molecules have the same amount of charge per mass. Because of this, gel electrophoresis of DNA fragments **separates them based on size only**. Using electrophoresis, we can see how many different DNA fragments are present in a sample and how large they are relative to one another. We can also determine the absolute size of a piece of DNA by examining it next to a standard "yardstick" or ladder made up of DNA fragments of known sizes.

As the name suggests, gel electrophoresis involves a gel: a slab of Jello-like material known as **agarose**. The gel contains tiny pores through which the DNA fragments can move.

At one end, the gel has pocket-like indentations called **wells**, which are where the DNA samples are initially loaded into the gel.



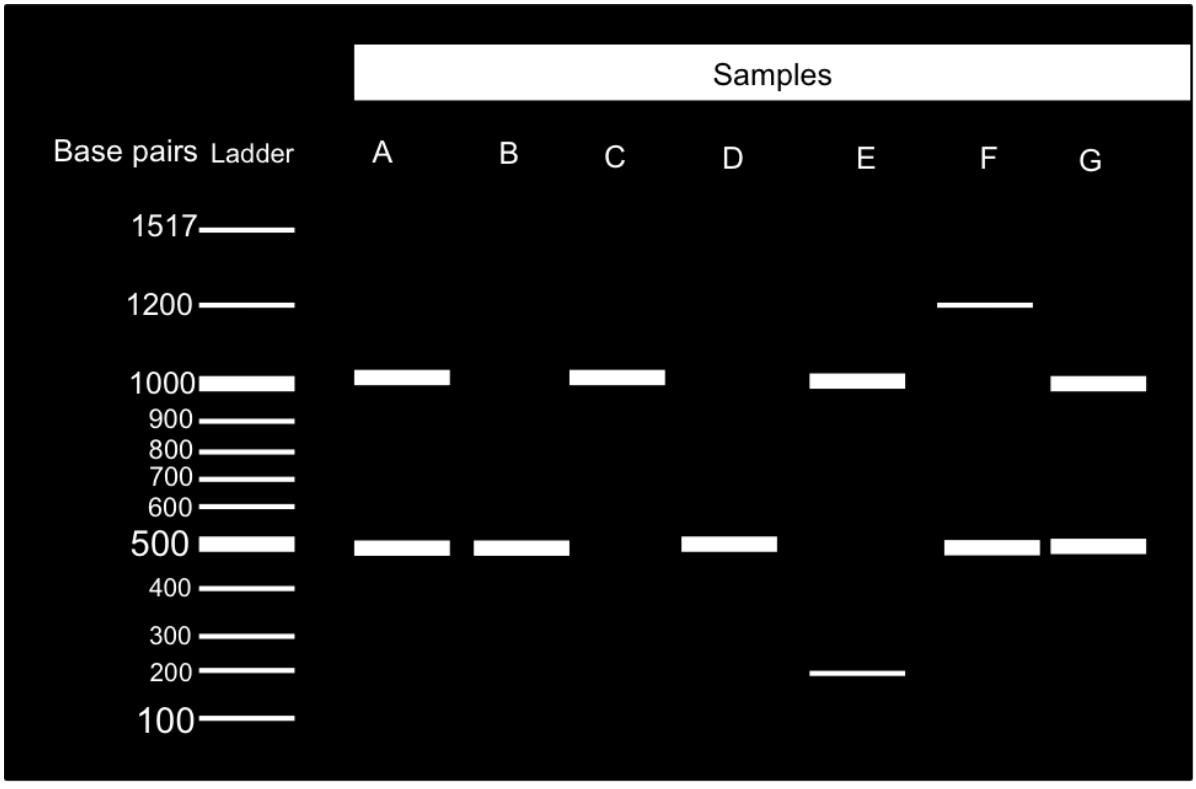
Before the DNA samples are added, the gel must be placed in a **gel box**. One end of the box contains a positive electrode, while the other end is attached to a negative electrode. The main body of the box, where the gel is placed, is filled with a salt-containing buffer solution that can conduct current. The buffer fills the gel box to a level where it just barely covers the gel. The end of the gel with the wells is positioned towards the negative electrode. The end without wells (towards which the DNA fragments will migrate) is positioned towards the positive electrode.



Once the gel is in the box, each of the DNA samples (for instance, each PCR reaction or each restriction-digested plasmid) is carefully transferred into one of the wells using a micropipette.

One well is reserved for a **DNA ladder**, a standard reference that contains DNA fragments of known lengths.
Next, the power to the gel box is turned on, and current begins to flow through the gel. The DNA molecules have a negative charge because of the phosphate groups in their sugar-phosphate backbone, so they start moving through the matrix of the gel towards the positive pole. Due to the size of the pores in the gel, large fragments move through the gel more slowly than do smaller fragments. When the power is turned on and current is passing through the gel, the gel is said to be **running**.

A well-defined "line" of DNA on a gel is called a **band**. Each band contains a large number of DNA fragments of the same size that have all traveled as a group to the same position. A single DNA fragment (or even a small group of DNA fragments) would not be visible by itself on a gel.
By comparing the bands in a sample to the DNA ladder, we can determine their approximate sizes. For instance, the two fragments in lane A are of sizes 1000 and 500 base pairs.



Human Genome Project

The Human Genome Project was an international research effort to determine the nucleotide sequence of the entire human genome and to identify the genes that it contains. The Project was coordinated by the National Institutes of Health and the U.S. Department of Energy. The Human Genome Project formally began in 1990 and was completed in 2003, 2 years ahead of its original schedule. The Human Genome Project's goal was to provide researchers with powerful tools to understand the genetic factors in human disease, paving the way for new strategies for their diagnosis, treatment and prevention. The Human Genome Project has already fueled the discovery of more than 1,800 disease genes.

As a result of the Human Genome Project, today's researchers can find a gene suspected of causing an inherited disease in a matter of days, rather than the years it took before the genome sequence was in hand. There are now more than 2,000 genetic tests for human conditions. These tests enable patients to learn their genetic risks for disease and also help healthcare professionals to diagnose disease. At least 350 biotechnology-based products resulting from the Human Genome Project are currently in clinical trials.

All data generated by the Human Genome Project were made freely and rapidly available on the Internet, serving to accelerate the pace of medical discovery around the globe. The data can now be accessed at:
https://www.ncbi.nlm.nih.gov/projects/genome/guide/human/

Gene Editing/CRISPR/CAS9

The term CRISPR/Cas9 stands for Clustered Regularly Interspaced Short Palindromic Repeats/CRISPR associated protein 9. CRISPR/Cas9 is a system found in bacteria involved in immune defense. Bacteria use CRISPR/Cas9 to cut up the DNA of invading bacterial viruses that might otherwise kill them.

Today, scientists have adapted this molecular machinery for an entirely different purpose – to change the nucleotide sequences of an organism's DNA code.

We might want to correct a disease-causing error that was inherited or crept into our DNA when it replicated. Or, in some cases, we may want to enhance the genetic code of crops, livestock or perhaps even people. CRISPR/Cas9 is already allowing scientists to make these types of genetic edits.