MT HOOD COMMUNITY COLLEGE BIOLOGY 102

MT HOOD COMMUNITY COLLEGE BIOLOGY 102

GENERAL BIOLOGY II: SURVEY OF MOLECULAR LIFE AND GENETICS

LISA BARTEE AND CHRISTINE ANDERSON AND BARTEE, LISA

Open Oregon Educational Resources

Portland, Ore.

Contents

9. Biotechnology

10. Cell Division - Binary Fission and Mitosis

11. Meiosis

12. Patterns of Inheritance

Genetics: Dog Coat Color

Reference Information

The resources provided in this section will enable you to:

Learning Objectives

- Select, evaluate and utilize discipline-specific information and literature to research a biological topic.
  - Differentiate between questions that can and cannot be answered using science.
  - Identify appropriate credible sources of information to research a topic.
  - Evaluate sources of information for their strengths and weaknesses.
- Communicate information using appropriate biological terminology in multiple formats.
  - Use an appropriate written format to present scientific information.
- Use evidence to develop informed opinions on contemporary biological issues while considering cultural and ethical implications.
  - Research current ethical issues in genetics and biotechnology.
  - Form opinions based on published scientific research.

The Process of Science

Learning Objectives

Course Objective for this section: Understand the process of scientific inquiry in order to apply the scientific method to biological questions by designing experiments and using the resulting data to form a conclusion
- Design a controlled experiment to answer a biological question.
- Predict the outcome of an experiment.
- Collect, manipulate, and analyze quantitative and qualitative data
- Answer a biological question using data.

Like geology, physics, and chemistry, **biology** is a science that gathers knowledge about the natural world. Specifically, biology is the study of life. The discoveries of biology are made by a community of researchers who work individually and together using agreed-on methods. In this sense, biology, like all sciences is a social enterprise like politics or the arts. The methods of science include careful observation, record keeping, logical and mathematical reasoning, experimentation, and submitting conclusions to the scrutiny of others. Science also requires considerable imagination and creativity; a well-designed experiment is commonly described as elegant, or beautiful. Like politics, science has considerable practical implications and some science is dedicated to practical applications, such as the prevention of disease (see Figure 1.1). Other science proceeds largely motivated by curiosity. Whatever its goal, there is no doubt that science, including biology, has transformed human existence and will continue to do so.

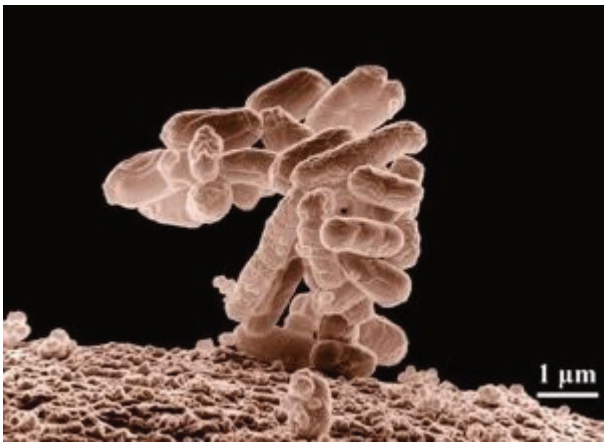


Figure 1.1 Biologists may choose to study Escherichia coli (E. coli), a bacterium that is a normal resident of our digestive tracts but which is also sometimes responsible for disease outbreaks. In this micrograph, the bacterium is visualized using a scanning electron microscope and digital colorization. (credit: Eric Erbe; digital colorization by Christopher Pooley, USDA-ARS)

References

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:RD6ERYiU@5/The-Process-of-Science.

3. Biological Molecules

The large molecules necessary for life that are built from smaller organic molecules are called biological **macromolecules**. There are four major classes of biological macromolecules (carbohydrates, lipids, proteins, and nucleic acids), and each is an important component of the cell and performs a wide array of functions. Combined, these molecules make up the majority of a cell's mass. Biological macromolecules are organic, meaning that they contain carbon atoms. In addition, they may contain atoms of hydrogen, oxygen, nitrogen, phosphorus, sulfur, and additional minor elements.

Figure 1: *The structure of a macromolecule can be compared to a necklace: both are larger structures that are built out of small pieces connected together into a chain. ("Beads on a string" by Daniel is licensed under CC BY-NC-ND 2.0)*

References

Unless otherwise noted, images on this page are licensed under CC-BY 4.0 by OpenStax.

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:QhGQhr4x@6/Biological-Molecules

4. Structure of DNA

Learning Objectives

By the end of this section, you will begin to: Discuss and apply biological theories and concepts of the molecular and genetic of life, including cell division, inheritance, and gene regulation by showing that you can:

- Describe the structure of DNA.

DNA is the genetic material passed from parent to offspring for all life on Earth. The three letters "DNA" have now become associated with crime solving, paternity testing, human identification, and genetic testing. DNA can be retrieved from hair, blood, or saliva. With the exception of identical twins, each person's DNA is unique and it is possible to detect differences between human beings on the basis of their unique DNA sequence.

DNA analysis has many practical applications beyond forensics and paternity testing. DNA testing is used for tracing genealogy and identifying pathogens. In the medical field, DNA is used in diagnostics, new vaccine development, and cancer therapy. It is now possible to determine predisposition to many diseases by analyzing genes.

In the 1950s, Francis Crick and James Watson worked together at the University of Cambridge, England, to determine the structure of DNA. Other scientists, such as Linus Pauling and Maurice Wilkins, were also actively exploring this field. Pauling had discovered the secondary structure of proteins using X-ray crystallography. X-ray crystallography is a method for investigating molecular structure by observing the patterns formed by X-rays shot through a crystal of the substance. The patterns give important information about the structure of the molecule of interest. In Wilkins' lab, researcher Rosalind Franklin was using X-ray crystallography to understand the structure of DNA. Watson and Crick were able to piece together the puzzle of the DNA molecule using Franklin's data (**Figure 9.2**). Watson and Crick also had key pieces of information available from other researchers such as Chargaff's rules. Chargaff had shown that of the four kinds of monomers (nucleotides) present in a DNA molecule, two types were always present in equal amounts and the remaining two types were also always present in equal amounts. This meant they were always paired in some way. In 1962, James Watson, Francis Crick, and Maurice Wilkins were awarded the Nobel Prize in Medicine for their work in determining the structure of DNA.

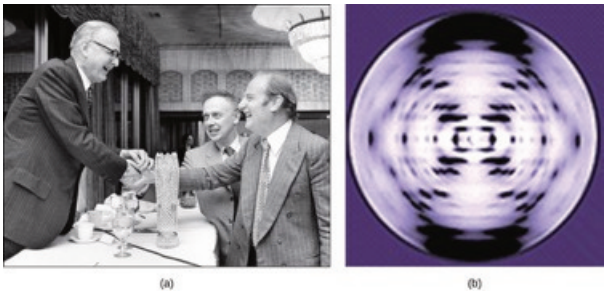


Figure 1: *Pioneering scientists (a) James Watson and Francis Crick are pictured here with American geneticist Maclyn McCarty. Scientist Rosalind Franklin discovered (b) the X-ray diffraction pattern of DNA, which helped to elucidate its double helix structure. (credit a: modification of work by Marjorie McCarty; b: modification of work by NIH)*

References

OpenStax, Concepts of Biology. OpenStax CNX. May 18, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:8v2Xzdco@5/The-Structure-of-DNA

5. DNA Replication

When a cell divides, it is important that each daughter cell receives an identical copy of the DNA. This is accomplished by the process of **DNA replication**. The replication of DNA occurs before the cell begins to divide into two separate cells.

The discovery and characterization of the structure of the double helix provided a hint as to how DNA is copied. Recall that adenine nucleotides pair with thymine nucleotides, and cytosine with guanine. This means that the two strands are **complementary** to each other. For example, a strand of DNA with a nucleotide sequence of AGTCATGA will have a complementary strand with the sequence TCAGTACT (**Figure 9.8**).

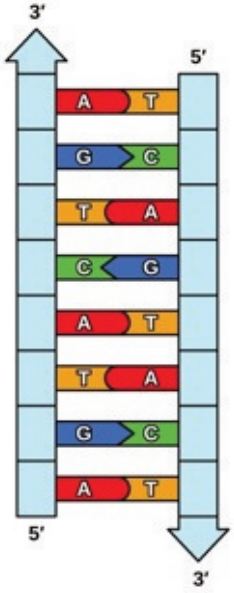


***Figure 1**: The two strands of DNA are complementary, meaning the sequence of bases in one strand can be used to create the correct sequence of bases in the other strand.*

Because of the complementarity of the two strands, having one strand means that it is possible to recreate the other strand. This model for replication suggests that the two strands of the double helix separate during replication, and each strand serves as a template from which the new complementary strand is copied (**Figure 9.9**).

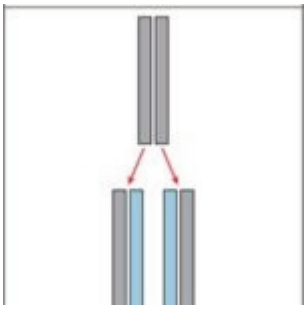


***Figure 2**: The semiconservative model of DNA replication is shown. Gray indicates the original DNA strands, and blue indicates newly synthesized DNA.*

During DNA replication, each of the two strands that make up the double helix serves as a template from which new strands are copied. The new strand will be complementary to the parental or "old" strand. Each new double strand consists of one parental strand and one new daughter strand. This is known as **semiconservative replication**. When two DNA copies are formed, they have an identical sequence of nucleotide bases and are divided equally into two daughter cells.

References

OpenStax, Concepts of Biology. OpenStax CNX. May 18, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:2ousESf0@5/DNA-Replication

6. Protein Synthesis

Learning Objectives

By the end of this section, you will be able to:

- Describe how DNA is used for the process of protein synthesis

References

"What are proteins and what do they do?" by U.S. National Library of Medicine is in the Public Domain

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:TkuNUJis@3/Transcription

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:FUH9XUkW@6/Translation

6.1 What are proteins and what do they do?

In both prokaryotes and eukaryotes, the major purpose of DNA is to provide the information needed to construct the proteins necessary for the cell can perform all of its functions. Proteins are large, complex molecules that play many critical roles in the body. They do most of the work in cells and are required for the structure, function, and regulation of the body's tissues and organs.

Proteins are made up of hundreds or thousands of smaller units called amino acids, which are attached to one another in long chains. There are 20 different types of amino acids that can be combined to make a protein. The sequence of amino acids determines each protein's unique 3-dimensional structure and its specific function.

Proteins can be described according to their large range of functions in the body, listed in alphabetical order:

Function	Description
Antibody	Antibodies bind to specific foreign particles, such as viruses and bacteria, to help protect the body.
Enzyme	Enzymes carry out almost all of the thousands of chemical reactions that take place in cells. They also assist with the formation of new molecules by reading the genetic information stored in DNA.
Messenger	Messenger proteins, such as some types of hormones, transmit signals to coordinate biological processes between different cells, tissues, and organs.
Structural component	These proteins provide structure and support for cells. On a larger scale, they also allow the body to move.
Transport/ storage	These proteins bind and carry atoms and small molecules within cells and throughout the body.

References

6.2 What is a gene?

The information to make proteins is stored in an organism's DNA. Each protein is coded for by a specific section of DNA called a **gene.** A gene is the section of DNA required to produce one protein. Genes are typically hundreds or thousands of base pairs in length because they code for proteins made of hundreds or thousands of amino acids.

Remember that DNA in eukaryotes is found as long linear molecules called **chromosomes.** Chromosomes are millions of base pairs in length and each contain many, many genes. Some example chromosomes are described in the table below.

Chromosome	Size (in base pairs)	# of genes
1	248,956,422	2058
10	133,797,422	733
22	50818468	488

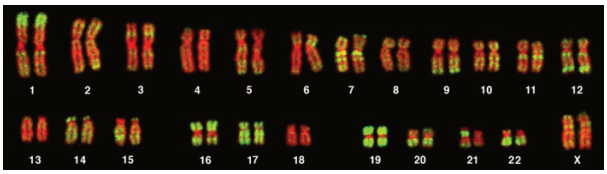


Figure 1: A karyotype showing the sizes of all the human chromosomes. Notice that they decrease in size.

To summarize: many base pairs make up one gene, many genes are found on one chromosome, and many chromosomes can be found in one genome.

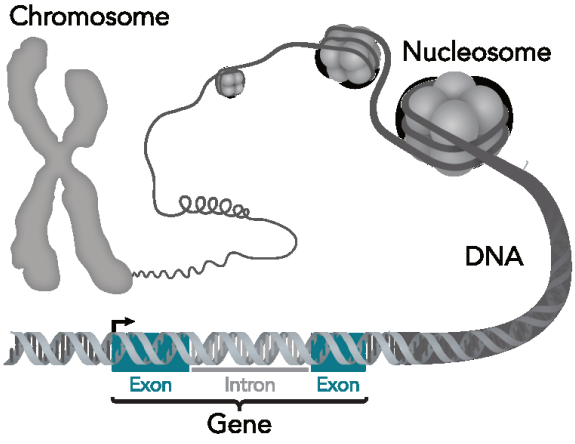


Figure 2: The arrangement of DNA into chromosomes. Photo credit: Thomas Splettstoesser (www.scistyle.com)

6.2 How do genes direct the production of proteins?

The information to make proteins is stored in an organism's DNA. Each protein is coded for by a specific section of DNA called a **gene.** A gene is the section of DNA required to produce one protein. Genes are typically hundreds or thousands of base pairs in length because they code for proteins made of hundreds or thousands of amino acids.

Most genes contain the information needed to make functional molecules called proteins. A few genes produce other molecules that help the cell assemble proteins. The journey from gene to protein is complex and tightly controlled within each cell. It consists of two major steps: transcription and translation. Together, transcription and translation are known as gene expression.

During the process of transcription, the information stored in a gene's DNA is transferred to a similar molecule called RNA (ribonucleic acid) in the cell nucleus. Both RNA and DNA are made up of a chain of nucleotide bases, but they have slightly different chemical properties. The type of RNA that contains the information for making a protein is called messenger RNA (mRNA) because it carries the information, or message, from the DNA out of the nucleus into the cytoplasm.

Translation, the second step in getting from a gene to a protein, takes place in the cytoplasm. The mRNA interacts with a specialized complex called a ribosome, which "reads" the sequence of mRNA bases. Each sequence of three bases, called a codon, usually codes for one particular amino acid. Remember that amino acids are the building blocks of proteins. A type of RNA called transfer RNA (tRNA) assembles the protein, one amino acid at a time. Protein assembly continues until the ribosome encounters a "stop" codon (a sequence of three bases that does not code for an amino acid).

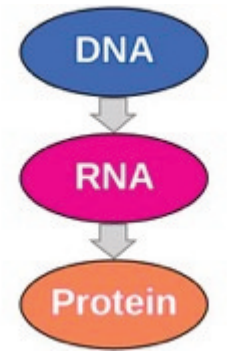


***FIgure 1**: The Central Dogma – DNA is used to make RNA is used to make protein*

The flow of information from DNA to RNA to proteins is one of the fundamental principles of molecular biology. It is so important that it is sometimes called the "central dogma."

References

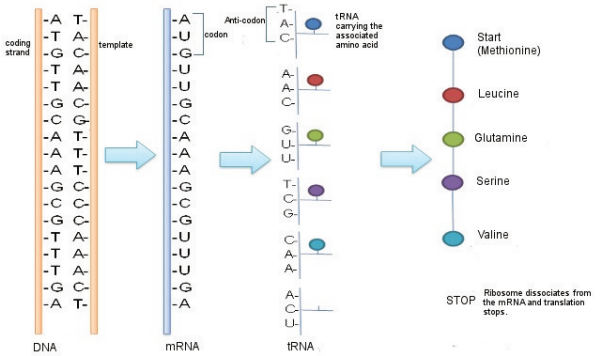


Figure 2: *More detail on the central dogma. ("Overview of Protein Synthesis" by Becky Boone is licensed under CC BY-SA 2.0)*

6.3 Transcription: from DNA to mRNA

Both prokaryotes and eukaryotes perform fundamentally the same process of transcription, with the important difference of the membrane-bound nucleus in eukaryotes. With the genes bound in the nucleus, transcription occurs in the nucleus of the cell and the mRNA transcript must be transported to the cytoplasm. The prokaryotes, which include bacteria and archaea, lack membrane-bound nuclei and other organelles, and transcription occurs in the cytoplasm of the cell.

Transcription requires the DNA double helix to partially unwind in the region of mRNA synthesis. The DNA sequence onto which the proteins and enzymes involved in transcription bind to initiate the process is called a **promoter**. In most cases, promoters exist upstream of the genes they regulate. The specific sequence of a promoter is very important because it determines whether the corresponding gene is transcribed all of the time, some of the time, or hardly at all.

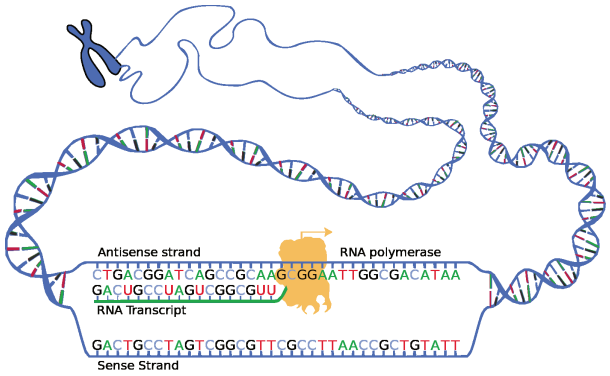


Figure 2: The initiation of transcription begins when DNA is unwound, forming a transcription bubble. Enzymes and other proteins involved in transcription bind at the promoter.

Transcription always proceeds from one of the two DNA strands, which is called the **template strand**. The mRNA product is complementary to the template strand and is almost identical to the other DNA strand, called the **non-template strand**, with the exception that RNA contains a uracil (U) in place of the thymine (T) found in DNA. An enzyme called **RNA polymerase** proceeds along the DNA template adding nucleotides by base pairing with the DNA template in a manner similar to DNA replication.

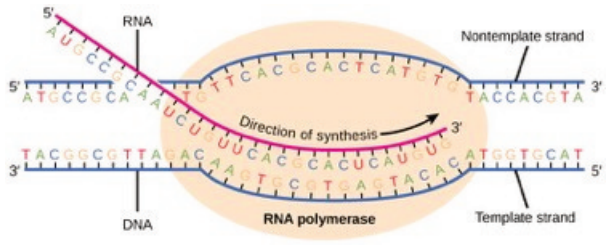


Figure 3: During elongation, RNA polymerase tracks along the DNA template, synthesizes mRNA in the 5′ to 3′ direction, and unwinds then rewinds the DNA as it is read.

Once a gene is transcribed, the RNA polymerase needs to be instructed to dissociate from the DNA template and liberate the newly made mRNA.

In a prokaryotic cell, by the time transcription ends, the transcript would already have been used to begin making copies of the encoded protein because the processes of transcription and translation can occur at the same time since both occur in the cytoplasm (**Figure 4**). In contrast, transcription and translation cannot occur simultaneously in eukaryotic cells since transcription occurs inside the nucleus and translation occurs outside in the cytoplasm.

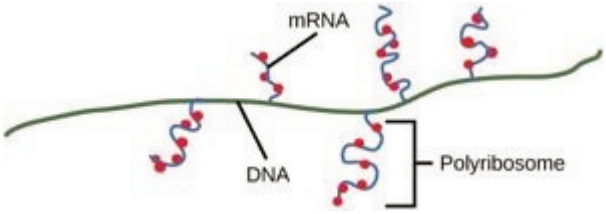


Figure 4: *Multiple polymerases can transcribe a single bacterial gene while numerous ribosomes concurrently translate the mRNA transcripts into polypeptides. In this way, a specific protein can rapidly reach a high concentration in the bacterial cell.*

References

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:TkuNUJis@3/Transcription

6.4 Eukaryotic RNA Processing

Eukaryotic mRNAs must undergo several processing steps before they can be transferred from the nucleus to the cytoplasm and translated into a protein. The additional steps involved in eukaryotic mRNA maturation create a molecule that is much more stable than a prokaryotic mRNA. Eukaryotic mRNAs typically last for several hours, whereas the typical prokaryotic mRNA lasts no more than five seconds.

The mRNA transcript is coated in **RNA-stabilizing proteins** to prevent it from degrading while it is processed and exported out of the nucleus.

A special nucleotide "cap" is added to one end of the growing transcript, which also helps prevent degradation and helps the cell recognize this molecule as an mRNA that should be translated.

Once transcription is complete, an enzyme adds a string of approximately 200 adenine residues to the end of the mRNA, called the **poly-A tail**. This modification further protects the pre-mRNA from degradation and signals to cellular factors that the transcript needs to be exported to the cytoplasm.

Eukaryotic genes are composed of protein-coding sequences called **exons** (*ex*-on signifies that they are *ex*pressed) and *int*ervening sequences called **introns** (*int*-ron denotes their *int*ervening role; you can also think of them as *int*errupting sequences). Introns are removed from the pre-mRNA during processing. Intron sequences in mRNA do not encode amino acids that become part of proteins. It is essential that all of a pre-mRNA's introns be completely and precisely removed before protein synthesis so that the exons join together to code for the correct amino acids. If the process errs by even a single nucleotide, the sequence of the rejoined exons would be shifted, and the resulting protein would be nonfunctional. The process of removing introns and reconnecting exons is called **splicing** (**Figure 5**). Introns are removed and degraded while the pre-mRNA is still in the nucleus.

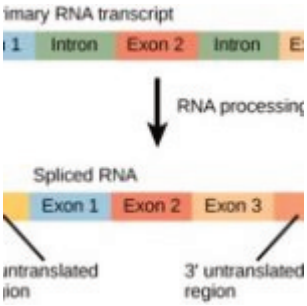


***Figure 5**: Eukaryotic mRNA contains introns that must be spliced out. A 5′ cap and 3′ tail are also added.*

References

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:TkuNUJis@3/Transcription

6.5 Translation

The synthesis of proteins is one of a cell's most energy-consuming metabolic processes. In turn, proteins account for more mass than any other component of living organisms (with the exception of water), and proteins perform a wide variety of the functions of a cell. The process of translation, or protein synthesis, involves decoding an mRNA message into a polypeptide product. Amino acids are covalently strung together in lengths ranging from approximately 50 amino acids to more than 1,000.

The Protein Synthesis Machinery

In addition to the mRNA template, many other molecules contribute to the process of translation. However, the general structures and functions of the protein synthesis machinery are comparable from bacteria to human cells. Translation requires the input of an mRNA template, ribosomes, tRNAs, and various enzymatic factors (**Figure 6**).

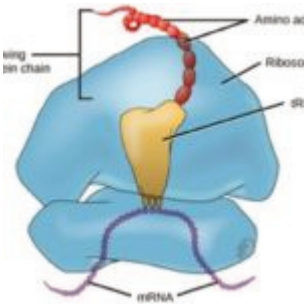


Figure 6: *The protein synthesis machinery includes the large and small subunits of the ribosome, mRNA, and tRNA. (credit: modification of work by NIGMS, NIH)*

Ribosomes are the part of the cell which reads the information in the mRNA molecule and joins amino acids together in the correct order. In *E. coli*, there are 200,000 ribosomes present in every cell at any given time. A ribosome is a very large, complex macromolecule. Ribosomes are located in the cytoplasm in prokaryotes and in the cytoplasm and endoplasmic reticulum of eukaryotes. Ribosomes are made up of two subunits that come together for translation, rather like a hamburger bun comes together around the meat (the mRNA). The small subunit is responsible for binding the mRNA template, whereas the large subunit sequentially binds **tRNAs**, a type of RNA molecule that brings amino acids to the growing chain of the polypeptide. Each mRNA molecule can be simultaneously translated by many ribosomes, all synthesizing protein in the same direction.

Depending on the species, 40 to 60 types of **tRNA** exist in the cytoplasm. Serving as adaptors, specific tRNAs bind to sequences on the mRNA template and add the corresponding amino acid to the polypeptide chain. Therefore, tRNAs are the molecules that actually "translate" the language of RNA into the language of proteins. For each tRNA to function, it must have its specific amino acid bonded to it. In the process of tRNA "charging," each tRNA molecule is bonded to its correct amino acid.

References

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:FUH9XUkW@6/Translation

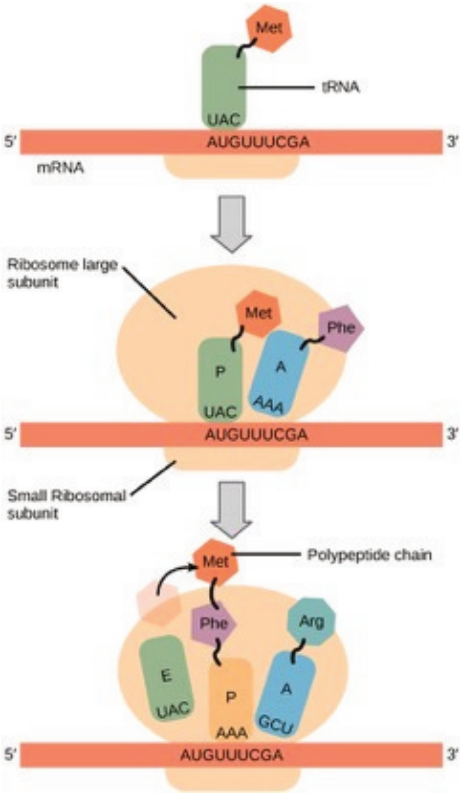


Figure 7: *Translation begins when a tRNA anticodon recognizes a codon on the mRNA. The large ribosomal subunit joins the small subunit, and a second tRNA is recruited. As the mRNA moves relative to the ribosome, the polypeptide chain is formed. Entry of a release factor into the A site terminates translation and the components dissociate.*

6.6 The Genetic Code

To summarize what we know to this point, the cellular process of transcription generates messenger RNA (mRNA), a mobile molecular copy of one or more genes with an alphabet of A, C, G, and uracil (U). Translation of the mRNA template converts nucleotide-based genetic information into a protein product. Protein sequences consist of 20 commonly occurring amino acids; therefore, it can be said that the protein alphabet consists of 20 letters. Each amino acid is defined by a three-nucleotide sequence called the triplet **codon**. The relationship between a nucleotide codon and its corresponding amino acid is called the **genetic code**.

Given the different numbers of "letters" in the mRNA and protein "alphabets," combinations of nucleotides corresponded to single amino acids. Using a three-nucleotide code means that there are a total of 64 (4 × 4 × 4) possible combinations; therefore, a given amino acid is encoded by more than one nucleotide triplet (**Figure 8**).

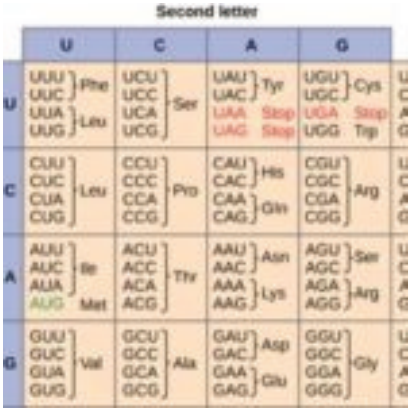


***Figure 8**: This figure shows the genetic code for translating each nucleotide triplet, or codon, in mRNA into an amino acid or a termination signal in a nascent protein. (credit: modification of work by NIH)*

Three of the 64 codons terminate protein synthesis and release the polypeptide from the translation machinery. These triplets are called **stop codons**. Another codon, AUG, also has a special function. In addition to specifying the amino acid methionine, it also serves as the **start codon** to initiate translation. The reading frame for translation is set by the AUG start codon near the 5′ end of the mRNA. The genetic code is universal. With a few exceptions, virtually all species use the same genetic code for protein synthesis, which is powerful evidence that all life on Earth shares a common origin.

References

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:FUH9XUkW@6/Translation

Optional Section – Micropigs

Micropigs are tiny, genetically-edited pigs that have recently been developed by a Chinese genomics institute (Li, 2014). The Chinese scientists used a technique called TALENs to edit the genome of pig cells (Figure 1). Each cell inside a pig contains two copies of the growth hormone receptor gene: one from each of its parents. The TALENs technique was used to delete one of the two copies of this gene.

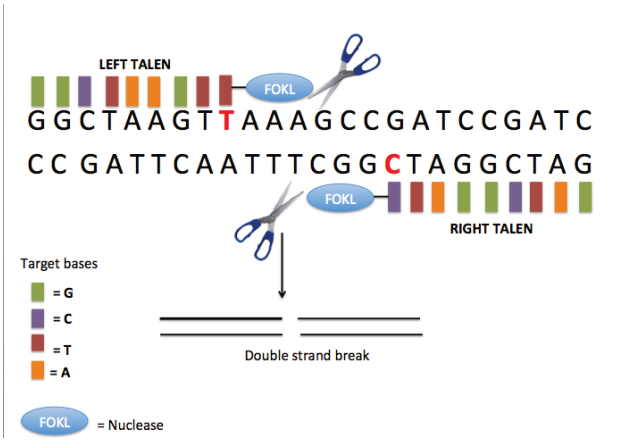


Figure 1 General overview of the TALEN process. The left and right TALEN bind to a specific sequence of genomic DNA inside the nucleus of a cell. When they are correctly bound, nuclease enzymes (represented by scissors) cut the genomic DNA. The TALEN sequences can be edited by scientists to target different DNA sequences in the genome. (Photo credit: Ogletreerd, Wikimedia.

Growth hormone (GH) stimulates the growth of essentially all tissues within the body. GH is a 191 amino acid peptide (protein) hormone which is produced from the GH gene. Cells sense the presence of GH protein hormone with the growth hormone receptor (GHR) protein on the outside of the cell. GHR protein is produced from the GHR gene and is found on the cell membrane on the outside of cells.

The GHR protein has three major parts:

- An extracellular region that sticks out from the outside surface of the cell
- A transmembrane region that goes through the cell membrane and anchors the receptor to the membrane
- An intracellular region on the inside of the cell membrane that transmits signals to the interior of the cell.

The extracellular region binds (attaches) to GH, fitting together like a lock and its key. The binding of growth hormone transmits signals through the cell membrane to the intracellular region of the receptor (Figure 2). These signals "turn on" genes involved in growth and metabolism so that those genes are made into proteins. These proteins stimulate the growth and division of other cells in the organism.

If growth hormone is not present, the organism will not grow to full size. In humans, severe GH deficiency can lead to an adult height of only 4 feet tall. If growth hormone receptor is not present, the "grow" signal from the GH will not be transmitted inside of cells, so growth will not be stimulated.

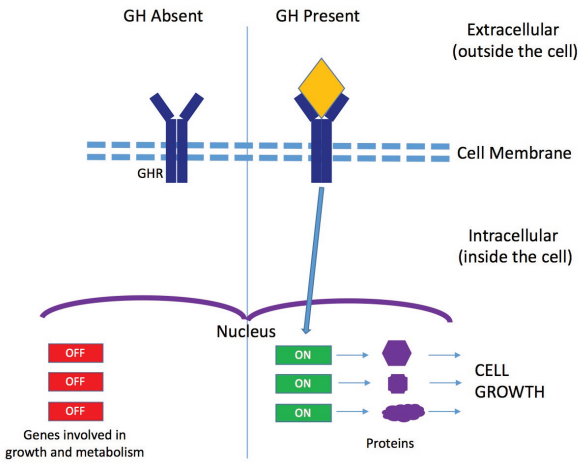


Figure 2 Growth hormone signaling pathway. When GH (growth hormone) binds to GHR (growth hormone receptor), a signal is sent through the cell membrane and into the nucleus of the cell. This signal turns on genes involved in cell metabolism and growth. (Photo credit: Lisa Bartee, 2017)

References

Li f, Li Y, Liu H, Zhang H, Liu C, Zhang X, Dou H, Yang W, Du Y. 2014 Sep. Production of GHR double-allelic knockout Bama pig by TALENs and handmade cloning. Yi Chuan: 36(9):903-11.

7. Mutations

Learning Objectives

By the end of this section, you will be able to:

- Describe how mutations affect protein synthesis and its products.

References

“Mutations and Health” by U.S. National Library of Medicine is in the Public Domain

How Gene Mutations Occur

A gene **mutation** is a permanent alteration in the DNA sequence that makes up a gene, such that the sequence differs from what is found in most people. Mutations range in size; they can affect anywhere from a single DNA building block (base pair) to a large segment of a chromosome that includes multiple genes.

Recall that the DNA sequence found within a gene controls protein synthesis. If the DNA sequence is altered, this can alter the amino acid sequence within a protein.

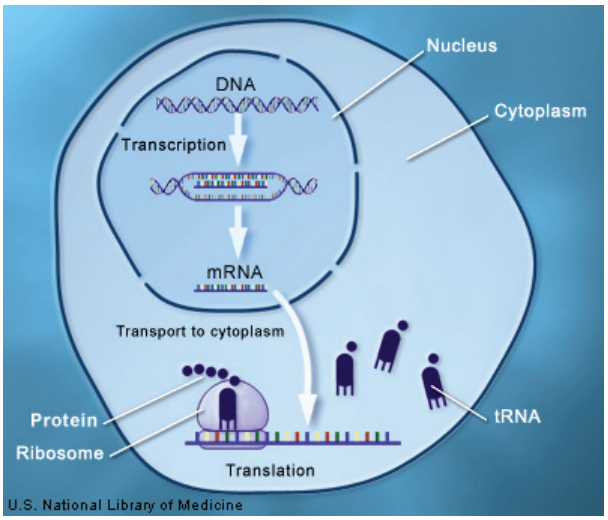


Figure : The process of protein synthesis first creates an mRNA copy of a DNA sequence during the process of transcription. This mRNA is translated into a sequence of amino acids by the ribosome. In this way, the information encoded in the sequence of bases in the DNA making up a gene is used to produce a protein.

Gene mutations can be classified in two major ways:

- **Hereditary mutations** are inherited from a parent and are present throughout a person's life in virtually every cell in the body. These mutations are also called germline mutations because they are present in the parent's egg or sperm cells, which are also called germ cells. When an egg and a sperm cell unite, the resulting fertilized egg cell receives DNA from both parents. If this DNA has a mutation, the child that grows from the fertilized egg will have the mutation in each of his or her cells.
- **Acquired (or somatic) mutations** occur at some time during a person's life and are present only in certain cells, not in every cell in the body. These changes can be caused by environmental factors such as ultraviolet radiation from the sun, or can occur if a mistake is made as DNA copies itself during cell division. Acquired mutations in somatic cells (cells other than sperm and egg cells) cannot be passed on to the next generation.

Genetic changes that are described as ***de novo* (new) mutations** can be either hereditary or somatic. In some cases, the mutation occurs in a person's egg or sperm cell but is not present in any of the person's other cells. In other cases, the mutation occurs in the fertilized egg shortly after the egg and sperm cells unite. It is often impossible to tell exactly when a de novo mutation happened. As the fertilized egg divides, each resulting cell in the growing embryo will have the mutation. *De novo* mutations may explain genetic disorders in which an affected child has a mutation in every cell in the body but the parents do not, and there is no family history of the disorder.

Somatic mutations that happen in a single cell early in embryonic development can lead to a situation called mosaicism. These genetic changes are not present in a parent's egg or sperm cells, or in the fertilized egg, but happen a bit later when the embryo includes several cells. As all the cells divide during growth and development, cells that arise from the cell with

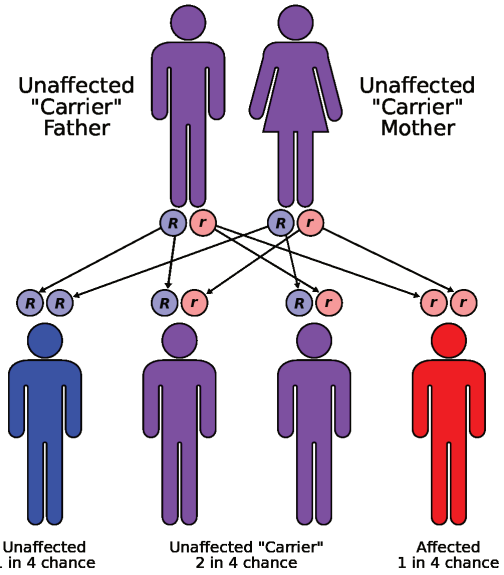


Figure: *The red individual has inherited two mutated alleles of a gene from their parents. This is an example of a hereditary mutation.*

Figure: *The color variation seen in this tulip is caused by a somatic mutation – one which occurred early in the development of this individual flower.*

the altered gene will have the mutation, while other cells will not. Depending on the mutation and how many cells are affected, mosaicism may or may not cause health problems.

Most disease-causing gene mutations are uncommon in the general population. However, other genetic changes occur more frequently. Genetic alterations that occur in more than 1 percent of the population are called **polymorphisms**. They are common enough to be considered a normal variation in the DNA. Polymorphisms are responsible for many of the normal differences between people such as eye color, hair color, and blood type. Although many polymorphisms have no negative effects on a person's health, some of these variations may influence the risk of developing certain disorders.

References

"Mutations and Health" by U.S. National Library of Medicine is in the Public Domain

Intro to Genetic Disorders

To function correctly, each cell depends on thousands of proteins to do their jobs in the right places at the right times. Sometimes, gene mutations prevent one or more of these proteins from working properly. By changing a gene's instructions for making a protein, a mutation can cause the protein to malfunction or to be missing entirely. When a mutation alters a protein that plays a critical role in the body, it can disrupt normal development or cause a medical condition. A condition caused by mutations in one or more genes is called a **genetic disorder**.

In some cases, gene mutations are so severe that they prevent an embryo from surviving until birth. These changes occur in genes that are essential for development, and often disrupt the development of an embryo in its earliest stages. Because these mutations have very serious effects, they are incompatible with life.

It is important to note that genes themselves do not cause disease—genetic disorders are caused by mutations that make a gene function improperly. For example, when people say that someone has "the cystic fibrosis gene," they are usually referring to a mutated version of the *CFTR* gene, which causes the disease. All people, including those without cystic fibrosis, have a version of the *CFTR* gene.

References

"Mutations and Health" by U.S. National Library of Medicine is in the Public Domain

Do all gene affect health and development?

No; only a small percentage of mutations cause genetic disorders—most have no impact on health or development. For example, some mutations alter a gene's DNA sequence but do not change the function of the protein made by the gene.

Often, gene mutations that could cause a genetic disorder are repaired by certain enzymes before the gene is expressed and an altered protein is produced. Each cell has a number of pathways through which enzymes recognize and repair mistakes in DNA. Because DNA can be damaged or mutated in many ways, DNA repair is an important process by which the body protects itself from disease.

A very small percentage of all mutations actually have a positive effect. These mutations lead to new versions of proteins that help an individual better adapt to changes in his or her environment. For example, a beneficial mutation could result in a protein that protects an individual and future generations from a new strain of bacteria.

Because a person's genetic code can have a large number of mutations with no effect on health, diagnosing genetic conditions can be difficult. Sometimes, genes thought to be related to a particular genetic condition have mutations, but whether these changes are involved in development of the condition has not been determined; these genetic changes are known as variants of unknown significance (VOUS). Sometimes, no mutations are found in suspected disease- related genes, but mutations are found in other genes whose relationship to a particular genetic condition is unknown. It is difficult to know whether these variants are involved in the disease.

Figure: *This lobster contains a mutation that causes it to be blue. This is estimated to occur in roughly one in two million lobsters.*

References

"Mutations and Health" by U.S. National Library of Medicine is in the Public Domain

Types of Mutations

The DNA sequence of a gene can be altered in a number of ways. Gene mutations have varying effects on health, depending on where they occur and whether they alter the function of essential proteins. The types of mutations include:

Missense mutation: This type of mutation is a change in one DNA base pair that results in the substitution of one amino acid for another in the protein made by a gene.

Nonsense mutation: A nonsense mutation is also a change in one DNA base pair. Instead of substituting one amino acid for another, however, the altered DNA sequence prematurely signals the cell to stop building a protein. This type of mutation results in a shortened protein that may function improperly or not at all.

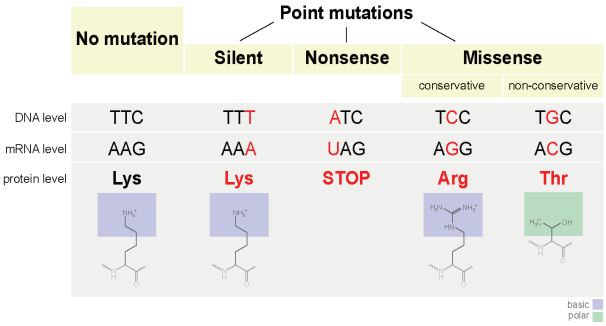


Figure: Some mutations do not change the sequence of amino acids in a protein. Some swap one amino acid for another. Others introduce an early stop codon into the sequence causing the protein to be truncated.

Insertion or Deletion: An insertion changes the number of DNA bases in a gene by adding a piece of DNA. A deletion removes a piece of DNA. Insertions or deletions may be small (one or a few base pairs within a gene) or large (an entire gene, several genes, or a large section of a chromosome). In any of these cases, the protein made by the gene may not function properly.

Duplication: A duplication consists of a piece of DNA that is abnormally copied one or more times. This type of mutation may alter the function of the resulting protein.

Frameshift mutation: This type of mutation occurs when the addition or loss of DNA bases changes a gene's reading frame. A reading frame consists of groups of 3 bases that each code for one amino acid. A frameshift mutation shifts the grouping of these bases and changes the code for amino acids. The resulting protein is usually nonfunctional. Insertions, deletions, and duplications can all be frameshift mutations.

Repeat expansion: Nucleotide repeats are short DNA sequences that are repeated a number of times in a row. For example, a trinucleotide repeat is made up of 3-base- pair sequences, and a tetranucleotide repeat is made up of 4-base-pair sequences. A repeat expansion is a mutation that increases the number of times that the short DNA sequence is repeated. This type of mutation can cause the resulting protein to function.

References

"Mutations and Health" by U.S. National Library of Medicine is in the Public Domain

Changes in Numbers of Genes

People have two copies of most genes, one copy inherited from each parent. In some cases, however, the number of copies varies—meaning that a person can be born with one, three, or more copies of particular genes. Less commonly, one or more genes may be entirely missing. This type of genetic difference is known as copy number variation (CNV).

Copy number variation results from insertions, deletions, and duplications of large segments of DNA. These segments are big enough to include whole genes. Variation in gene copy number can influence the activity of genes and ultimately affect many body functions.

Researchers were surprised to learn that copy number variation accounts for a significant amount of genetic difference between people. More than 10 percent of human DNA appears to contain these differences in gene copy number. While much of this variation does not affect health or development, some differences likely influence a person's risk of disease and response to certain drugs. Future research will focus on the consequences of copy number variation in different parts of the genome and study the contribution of these variations to many types of disease.

References

"Mutations and Health" by U.S. National Library of Medicine is in the Public Domain

Changes in Chromosome Number

Human cells normally contain 23 pairs of chromosomes, for a total of 46 chromosomes in each cell. A change in the number of chromosomes can cause problems with growth, development, and function of the body's systems. These changes can occur during the formation of reproductive cells (eggs and sperm), in early fetal development, or in any cell after birth. A gain or loss of chromosomes from the normal 46 is called **aneuploidy**.

A common form of aneuploidy is **trisomy,** or the presence of an extra chromosome in cells. "Tri-" is Greek for "three"; people with trisomy have three copies of a particular chromosome in cells instead of the normal two copies. Down syndrome is an example of a condition caused by trisomy. People with Down syndrome typically have three copies of chromosome 21 in each cell, for a total of 47 chromosomes per cell.

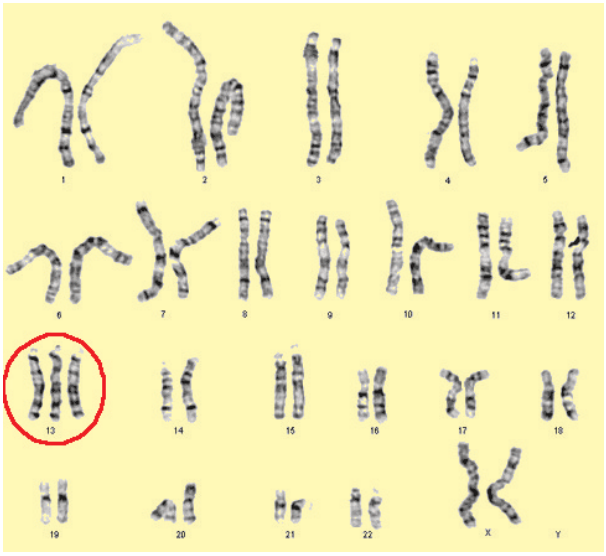


Figure: *This karyotype, which is a picture of all the chromosomes from one individual, is from a person who has Trisomy 13.*

Monosomy, or the loss of one chromosome in cells, is another kind of aneuploidy. "Mono-" is Greek for "one"; people with monosomy have one copy of a particular chromosome in cells instead of the normal two copies. Turner syndrome is a condition caused by monosomy. Women with Turner syndrome usually have only one copy of the X chromosome in every cell, for a total of 45 chromosomes per cell.

Rarely, some cells end up with complete extra sets of chromosomes. Cells with one additional set of chromosomes, for a total of 69 chromosomes, are called **triploid.** Cells with two additional sets of chromosomes, for a total of 92 chromosomes, are called tetraploid. A condition in which every cell in the body has an extra set of chromosomes is not compatible with life.

In some cases, a change in the number of chromosomes occurs only in certain cells. When an individual has two or more cell populations with a different chromosomal makeup, this situation is called **chromosomal mosaicism**. Chromosomal mosaicism occurs from an error in cell division in cells other than eggs and sperm. Most commonly, some cells end up with one extra or missing chromosome (for a total of 45 or 47 chromosomes per cell), while other cells have the usual 46 chromosomes. Mosaic Turner syndrome is one example of chromosomal mosaicism. In females with this condition, some cells have 45 chromosomes because they are missing one copy of the X chromosome, while other cells have the usual number of chromosomes.

Many cancer cells also have changes in their number of chromosomes. These changes are not inherited; they occur in somatic cells (cells other than eggs or sperm) during the formation or progression of a cancerous tumor.

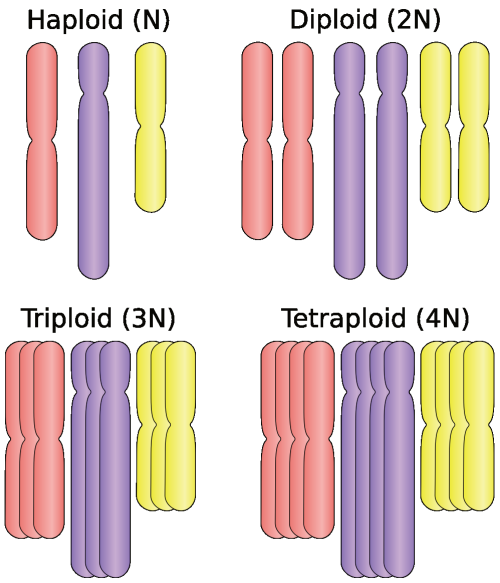


Figure: "Ploid" refers to the number of copies of each chromosome found in a somatic cell.

Figure: Human and other animal cells do not develop if they have an entire extra set of chromosomes. In contrast, plants often have entire copied sets of chromosomes. This strawberry is an example of a plant that is tetraploid.

References

"Mutations and Health" by U.S. National Library of Medicine is in the Public Domain

Complex Multifactorial Disorders

Researchers are learning that nearly all conditions and diseases have a genetic component. Some disorders, such as sickle cell disease and cystic fibrosis, are caused by mutations in a single gene. The causes of many other disorders, however, are much more complex. Common medical problems such as heart disease, diabetes, and obesity do not have a single genetic cause—they are likely associated with the effects of multiple genes in combination with lifestyle and environmental factors. Conditions caused by many contributing factors are called complex or **multifactorial disorders**.

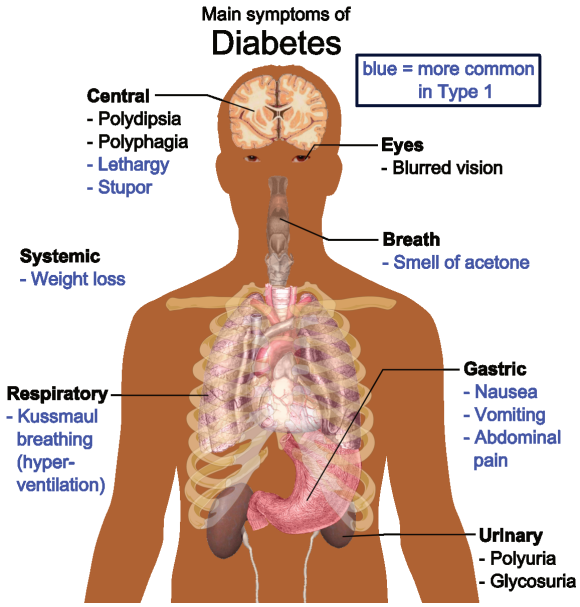


Figure: The main symptoms of diabetes

Although complex disorders often cluster in families, they do not have a clear- cut pattern of inheritance. This makes it difficult to determine a person's risk of inheriting or passing on these disorders. Complex disorders are also difficult to study and treat because the specific factors that cause most of these disorders have not yet been identified. Researchers continue to look for major contributing genes for many common complex disorders.

References

"Mutations and Health" by U.S. National Library of Medicine is in the Public Domain

Genetic Predispositions

A **genetic predisposition** (sometimes also called genetic susceptibility) is an increased likelihood of developing a particular disease based on a person's genetic makeup. A genetic predisposition results from specific genetic variations that are often inherited from a parent. These genetic changes contribute to the development of a disease but do not directly cause it. Some people with a predisposing genetic variation will never get the disease while others will, even within the same family.

Genetic variations can have large or small effects on the likelihood of developing a particular disease. For example, certain mutations in the *BRCA1* or *BRCA2* genes greatly increase a person's risk of developing breast cancer and ovarian cancer. Variations in other genes, such as *BARD1* and *BRIP1*, also increase breast cancer risk, but the contribution of these genetic changes to a person's overall risk appears to be much smaller.

Current research is focused on identifying genetic changes that have a small effect on disease risk but are common in the general population. Although each of these variations only slightly increases a person's risk, having changes in several different genes may combine to increase disease risk significantly. Changes in many genes, each with a small effect, may underlie susceptibility to many common diseases, including cancer, obesity, diabetes, heart disease, and mental illness.

In people with a genetic predisposition, the risk of disease can depend on multiple factors in addition to an identified genetic change. These include
other genetic factors (sometimes called modifiers) as well as lifestyle and environmental factors. Although a person's genetic makeup cannot be altered, some lifestyle and environmental modifications (such as having more frequent disease screenings and maintaining a healthy weight) may be able to reduce disease risk in people with a genetic predisposition.

References

"Mutations and Health" by U.S. National Library of Medicine is in the Public Domain

Genetics and Statistics

Statistical data can provide general information about how common a condition is, how many people have the condition, or how likely it is that a person will develop the condition. Statistics are not personalized, however—they offer estimates based on groups of people. By taking into account a person's family history, medical history, and other factors, a genetics professional can help interpret what statistics mean for a particular patient.

Some statistical terms are commonly used when describing genetic conditions and other disorders. These terms include:

Statistical term	Description
Incidence	The incidence of a gene mutation or a genetic disorder is the number of people who are born with the mutation or disorder in a specified group per year. Incidence is often written in the form "1 in [a number]" or as a total number of live births.
Prevalence	The prevalence of a gene mutation or a genetic disorder is the total number of people in a specified group at a given time who have the mutation or disorder. This term includes both newly diagnosed and pre-existing cases in people of any age. Prevalence is often written in the form "1 in [a number]" or as a total number of people who have a condition.
Mortality	Mortality is the number of deaths from a particular disorder occurring in a specified group per year. Mortality is usually expressed as a total number of deaths.
Lifetime risk	Lifetime risk is the average risk of developing a particular disorder at some point during a lifetime. Lifetime risk is often written as a percentage or as "1 in [a number]." It is important to remember that the risk per year or per decade is much lower than the lifetime risk. In addition, other factors may increase or decrease a person's risk as compared with the average.

Use of Statistics Terms

- About 1 in 200,000 people in the United States are born with syndrome A each year.
- An estimated 15,000 infants with syndrome B were born last year worldwide.
- Approximately 1 in 100,000 people in the United States have syndrome A at the present time.
- About 100,000 children worldwide currently have syndrome B.
- An estimated 12,000 people worldwide died from syndrome C in 2002.
- Approximately 1 percent of people in the United States develop disorder D during their lifetimes.
- The lifetime risk of developing disorder D is 1 in 100.

References

"Mutations and Health" by U.S. National Library of Medicine is in the Public Domain

Gene Regulation

Learning Objectives

By the end of this section, you will be able to:

- Discuss the purpose and mechanisms of gene regulation.

Each cell expresses, or turns on, only a fraction of its genes. The rest of the genes are repressed, or turned off. The process of turning genes on and off is known as **gene regulation**. Gene regulation is an important part of normal development. Genes are turned on and off in different patterns during development to make a brain cell look and act different from a liver cell or a muscle cell, for example. Gene regulation also allows cells to react quickly to changes in their environments. Although we know that the regulation of genes is critical for life, this complex process is not yet fully understood.

For a cell to function properly, necessary proteins must be synthesized at the proper time. All organisms and cells control or regulate the transcription and translation of their DNA into protein. The process of turning on a gene to produce RNA and protein is called **gene expression**. Whether in a simple unicellular organism or in a complex multicellular organism, each cell controls when and how its genes are expressed. For this to occur, there must be a mechanism to control when a gene is expressed to make RNA and protein, how much of the protein is made, and when it is time to stop making that protein because it is no longer needed.

Cells in multicellular organisms are specialized; cells in different tissues look very different and perform different functions. For example, a muscle cell is very different from a liver cell, which is very different from a skin cell. These differences are a consequence of the expression of different sets of genes in each of these cells. All cells have certain basic functions they must perform for themselves, such as converting the energy in sugar molecules into energy in ATP. Each cell also has many genes that are not expressed, and expresses many that are not expressed by other cells, such that it can carry out its specialized functions. In addition, cells will turn on or off certain genes at different times in response to changes in the environment or at different times during the development of the organism. Unicellular organisms, both eukaryotic and prokaryotic, also turn on and off genes in response to the demands of their environment so that they can respond to special conditions.

The control of gene expression is extremely complex. Malfunctions in this process are detrimental to the cell and can lead to the development of many diseases, including cancer.

References

OpenStax, Concepts of Biology. OpenStax CNX. May 18, 2016 http://cnx.org/contents/b3c1e1d2-839c-42b0-a314-e119a8aafbdd@9.10

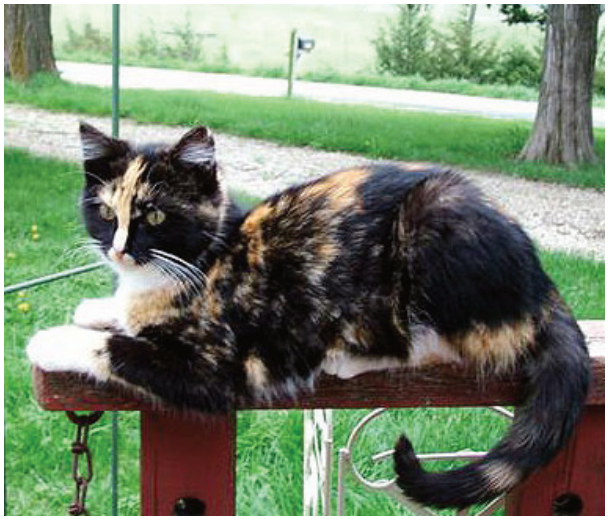


Figure 1: *The unique color pattern of this cat's fur is caused by either the orange or the black allele of a gene being randomly silenced (turned off).*

8.1 Prokaryotic versus Eukaryotic Gene Expression

To understand how gene expression is regulated, we must first understand how a gene becomes a functional protein in a cell. The process occurs in both prokaryotic and eukaryotic cells, just in slightly different fashions.

Because prokaryotic organisms lack a cell nucleus, the processes of transcription and translation occur almost simultaneously. When the protein is no longer needed, transcription stops. As a result, the primary method to control what type and how much protein is expressed in a prokaryotic cell is through the regulation of DNA transcription into RNA. All the subsequent steps happen automatically. When more protein is required, more transcription occurs. Therefore, in prokaryotic cells, the control of gene expression is almost entirely at the transcriptional level.

Eukaryotic cells, in contrast, have intracellular organelles and are much more complex. Recall that in eukaryotic cells, the DNA is contained inside the cell's nucleus and it is transcribed into mRNA there. The newly synthesized mRNA is then transported out of the nucleus into the cytoplasm, where ribosomes translate the mRNA into protein. The processes of transcription and translation are physically separated by the nuclear membrane; transcription occurs only within the nucleus, and translation only occurs outside the nucleus in the cytoplasm. The regulation of gene expression can occur at all stages of the process (**Figure 2**):

- Epigenetic level: regulates how tightly the DNA is wound around histone proteins to package it into chromosomes
- Transcriptional level: regulates how much transcription takes place
- Post-transcriptional level: regulates aspects of RNA processing (such as splicing) and transport out of the nucleus
- Translational level: regulates how much of the RNA is translated into protein
- Post-translational level: regulates how long the protein lasts after it has been made and whether the protein is processed into an active form

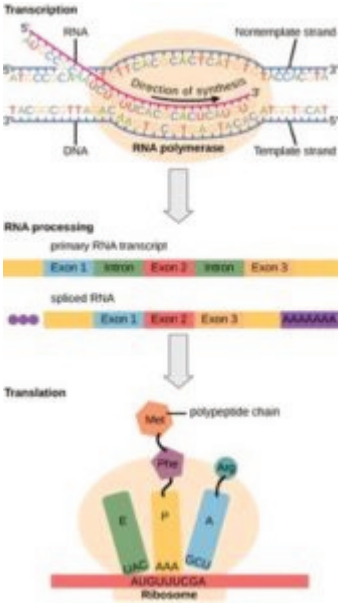


Figure 2: *Eukaryotic gene expression is regulated during transcription and RNA processing, which take place in the nucleus, as well as during protein translation, which takes place in the cytoplasm. Further regulation may occur through post-translational modifications of proteins.*

The differences in the regulation of gene expression between prokaryotes and eukaryotes are summarized in **Table 1**.

Table 1: Differences in the Regulation of Gene Expression of Prokaryotic and Eukaryotic Organisms

Prokaryotic organisms	**Eukaryotic organisms**
Lack nucleus	Contain nucleus
RNA transcription and protein translation occur almost simultaneously	RNA transcription occurs prior to protein translation, and it takes place in the nucleus. RNA translation to protein occurs in the cytoplasm. RNA post-processing includes addition of a 5′ cap, poly-A tail, and excision of introns and splicing of exons.
Gene expression is regulated primarily at the transcriptional level	Gene expression is regulated at many levels (epigenetic, transcriptional, post-transcriptional, translational, and posttranslational)

References

OpenStax, Concepts of Biology. OpenStax CNX. May 18, 2016 http://cnx.org/contents/b3c1e1d2-839c-42b0-a314-e119a8aafbdd@9.10

8.2 What is the epigenome?

DNA modifications that do not change the DNA sequence can affect gene activity. Chemical compounds that are added to single genes can regulate their activity; these modifications are known as epigenetic changes. The epigenome comprises all of the chemical compounds that have been added to the entirety of one's DNA (genome) as a way to regulate the activity (expression) of all the genes within the genome. The chemical compounds of the epigenome are not part of the DNA sequence, but are on or attached to DNA ("epi-" means above in Greek). Epigenomic modifications remain as cells divide and in some cases can be inherited through the generations. Environmental influences, such as a person's diet and exposure to pollutants, can also impact the epigenome.

Epigenetic changes can help determine whether genes are turned on or off and can influence the production of proteins in certain cells, ensuring that only necessary proteins are produced. For example, proteins that promote bone growth are not produced in muscle cells. Patterns of epigenome modification vary among individuals, different tissues within an individual, and even different cells.

A common type of epigenomic modification is called methylation. Methylation involves attaching small molecules called methyl groups, each consisting of one carbon atom and three hydrogen atoms, to segments of DNA. When methyl groups are added to a particular gene, that gene is turned off or silenced, and no protein is produced from that gene.

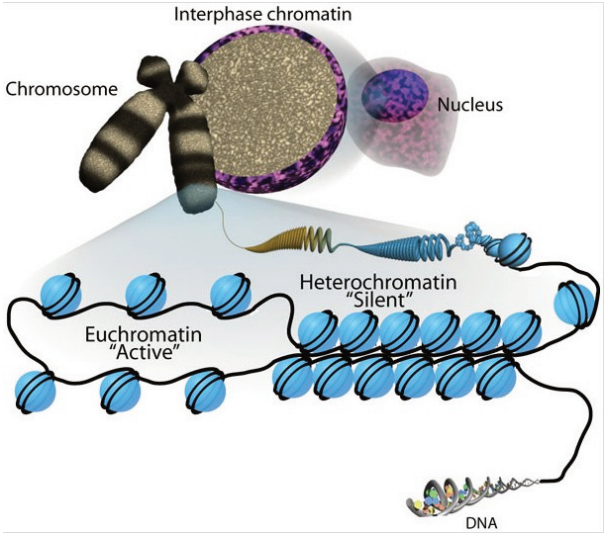


***Figure 3**: The difference in chromatin packaging between an active (euchromatic) and inactive (heterochromatic) region of DNA.*

Because errors in the epigenetic process, such as modifying the wrong gene or failing to add a compound to a gene, can lead to abnormal gene activity or inactivity, they can cause genetic disorders. Conditions including cancers, metabolic disorders, and degenerative disorders have all been found to be related to epigenetic errors.

Scientists continue to explore the relationship between the genome and the chemical compounds that modify it. In particular, they are studying what effect the modifications have on gene function, protein production, and human health.

References

OpenStax, Concepts of Biology. OpenStax CNX. May 18, 2016 http://cnx.org/contents/b3c1e1d2-839c-42b0-a314-e119a8aafbdd@9.10

8.3 Alternative RNA splicing

In the 1970s, genes were first observed that exhibited **alternative RNA splicing**. Alternative RNA splicing is a mechanism that allows different protein products to be produced from one gene when different combinations of introns (and sometimes exons) are removed from the transcript (**Figure 9.23**). This alternative splicing can be haphazard, but more often it is controlled and acts as a mechanism of gene regulation, with the frequency of different splicing alternatives controlled by the cell as a way to control the production of different protein products in different cells, or at different stages of development. Alternative splicing is now understood to be a common mechanism of gene regulation in eukaryotes; according to one estimate, 70% of genes in humans are expressed as multiple proteins through alternative splicing.

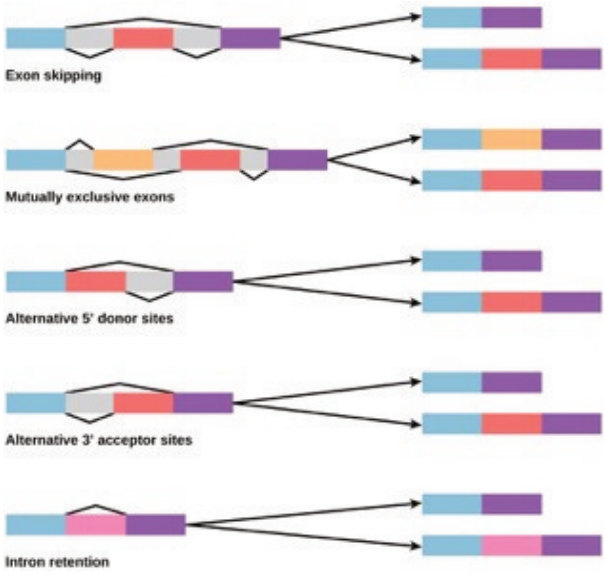


***Figure 4**: There are five basic modes of alternative splicing. Segments of pre-mRNA with exons shown in blue, red, orange, and pink can be spliced to produce a variety of new mature mRNA segments.*

How could alternative splicing evolve? Introns have a beginning and ending recognition sequence, and it is easy to imagine the failure of the splicing mechanism to identify the end of an intron and find the end of the next intron, thus removing two introns and the intervening exon. In fact, there are mechanisms in place to prevent such exon skipping, but mutations are likely to lead to their failure. Such "mistakes" would more than likely produce a nonfunctional protein. Indeed, the cause of many genetic diseases is alternative splicing rather than mutations in a sequence. However, alternative splicing would create a protein variant without the loss of the original protein, opening up possibilities for adaptation of the new variant to new functions. Gene duplication has played an important role in the evolution of new functions in a similar way—by providing genes that may evolve without eliminating the original functional protein.

References

OpenStax, Concepts of Biology. OpenStax CNX. May 18, 2016 http://cnx.org/contents/b3c1e1d2-839c-42b0-a314-e119a8aafbdd@9.10

9. Biotechnology

The latter half of the twentieth century began with the discovery of the structure of DNA, then progressed to the development of the basic tools used to study and manipulate DNA. These advances, as well as advances in our understanding of and ability to manipulate cells, have led some to refer to the twenty-first century as the biotechnology century. The rate of discovery and of the development of new applications in medicine, agriculture, and energy is expected to accelerate, bringing huge benefits to humankind and perhaps also significant risks. Many of these developments are expected to raise significant ethical and social questions that human societies have not yet had to consider.

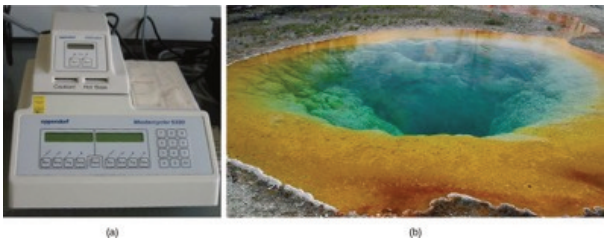


Figure 1: *(a) A thermal cycler, such as the one shown here, is a basic tool used to study DNA in a process called the polymerase chain reaction (PCR). The polymerase enzyme most often used with PCR comes from a strain of bacteria that lives in (b) the hot springs of Yellowstone National Park. (credit a: modification of work by Magnus Manske; credit b: modification of work by Jon Sullivan)*

References

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:Pdu1uR8Y@2/Introduction

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:8CA_YwJq@3/Cloning-and-Genetic-Engineerin

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:rytx-nDe@3/Biotechnology-in-Medicine-and-

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:TE1njgbY@4/Genomics-and-Proteomics

9.1 Manipulating Genetic Material

Biotechnology is the use of artificial methods to modify the genetic material of living organisms or cells to produce novel compounds or to perform new functions. Biotechnology has been used for improving livestock and crops since the beginning of agriculture through selective breeding. Since the discovery of the structure of DNA in 1953, and particularly since the development of tools and methods to manipulate DNA in the 1970s, biotechnology has become synonymous with the manipulation of organisms' DNA at the molecular level. The primary applications of this technology are in medicine (for the production of vaccines and antibiotics) and in agriculture (for the genetic modification of crops). Biotechnology also has many industrial applications, such as fermentation, the treatment of oil spills, and the production of biofuels, as well as many household applications such as the use of enzymes in laundry detergent.

To accomplish the applications described above, biotechnologists must be able to extract, manipulate, and analyze nucleic acids.

Review of Nucleic Acid Structure

To understand the basic techniques used to work with nucleic acids, remember that nucleic acids are macromolecules made of nucleotides (a sugar, a phosphate, and a nitrogenous base). The phosphate groups on these molecules each have a net negative charge. An entire set of DNA molecules in the nucleus of eukaryotic organisms is called the genome. DNA has two complementary strands linked by hydrogen bonds between the paired bases.

Unlike DNA in eukaryotic cells, RNA molecules leave the nucleus. Messenger RNA (mRNA) is analyzed most frequently because it represents the protein-coding genes that are being expressed in the cell.

Isolation of Nucleic Acids

To study or manipulate nucleic acids, the DNA must first be extracted from cells. Various techniques are used to extract different types of DNA (**Figure 10.2**). Most nucleic acid extraction techniques involve steps to break open the cell, and then the use of enzymatic reactions to destroy all undesired macromolecules. Cells are broken open using a detergent solution containing buffering compounds. To prevent degradation and contamination, macromolecules such as proteins and RNA are inactivated using enzymes. The DNA is then brought out of solution using alcohol. The resulting DNA, because it is made up of long polymers, forms a gelatinous mass.

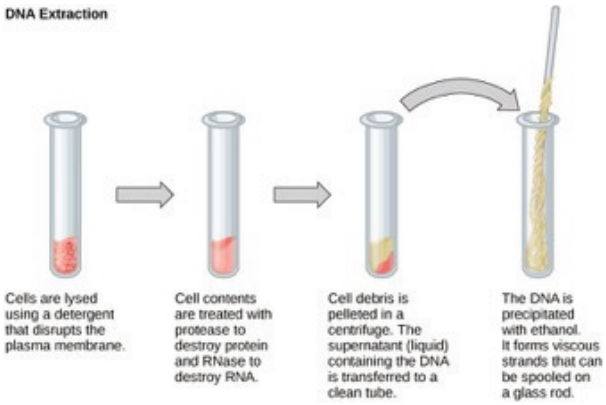


***Figure 2**: This diagram shows the basic method used for the extraction of DNA.*

RNA is studied to understand gene expression patterns in cells. RNA is naturally very unstable because enzymes that break down RNA are commonly present in nature. Some are even secreted by our own skin and are very difficult to

inactivate. Similar to DNA extraction, RNA extraction involves the use of various buffers and enzymes to inactivate other macromolecules and preserve only the RNA.

Gel Electrophoresis

Because nucleic acids are negatively charged ions at neutral or alkaline pH in an aqueous environment, they can be moved by an electric field. **Gel electrophoresis** is a technique used to separate charged molecules on the basis of size and charge. The nucleic acids can be separated as whole chromosomes or as fragments. The nucleic acids are loaded into a slot at one end of a gel matrix, an electric current is applied, and negatively charged molecules are pulled toward the opposite end of the gel (the end with the positive electrode). Smaller molecules move through the pores in the gel faster than larger molecules; this difference in the rate of migration separates the fragments on the basis of size. The nucleic acids in a gel matrix are invisible until they are stained with a compound that allows them to be seen, such as a dye. Distinct fragments of nucleic acids appear as bands at specific distances from the top of the gel (the negative electrode end) that are based on their size (**Figure 10.3**). A mixture of many fragments of varying sizes appear as a long smear, whereas uncut genomic DNA is usually too large to run through the gel and forms a single large band at the top of the gel.

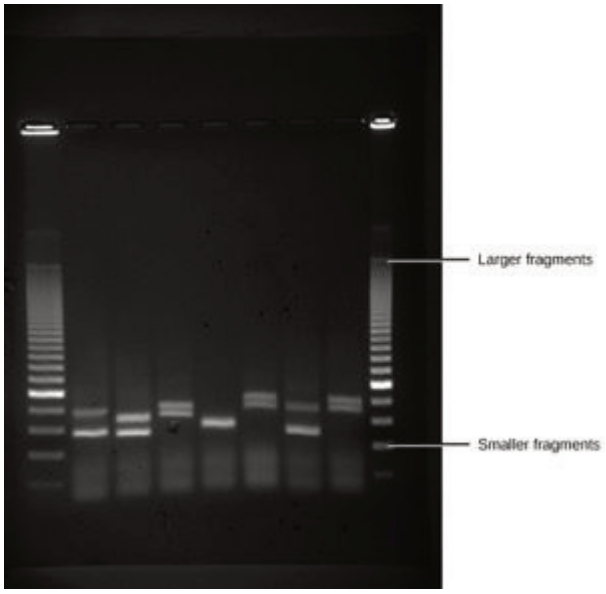


Figure 3: *Shown are DNA fragments from six samples run on a gel, stained with a fluorescent dye and viewed under UV light. (credit: modification of work by James Jacob, Tompkins Cortland Community College)*

Polymerase Chain Reaction

DNA analysis often requires focusing on one or more specific regions of the genome. It also frequently involves situations in which only one or a few copies of a DNA molecule are available for further analysis. These amounts are insufficient for most procedures, such as gel electrophoresis. **Polymerase chain reaction (PCR)** is a technique used to rapidly increase the number of copies of specific regions of DNA for further analyses (**Figure 10.4**). PCR uses a special form of DNA polymerase, the enzyme that replicates DNA, and other short nucleotide sequences called primers that base pair to a specific portion of the DNA being replicated. PCR is used for many purposes in laboratories. These include: 1) the identification of the owner of a DNA sample left at a crime scene; 2) paternity analysis; 3) the comparison of small amounts of ancient DNA with modern organisms; and 4) determining the sequence of nucleotides in a specific region.

References

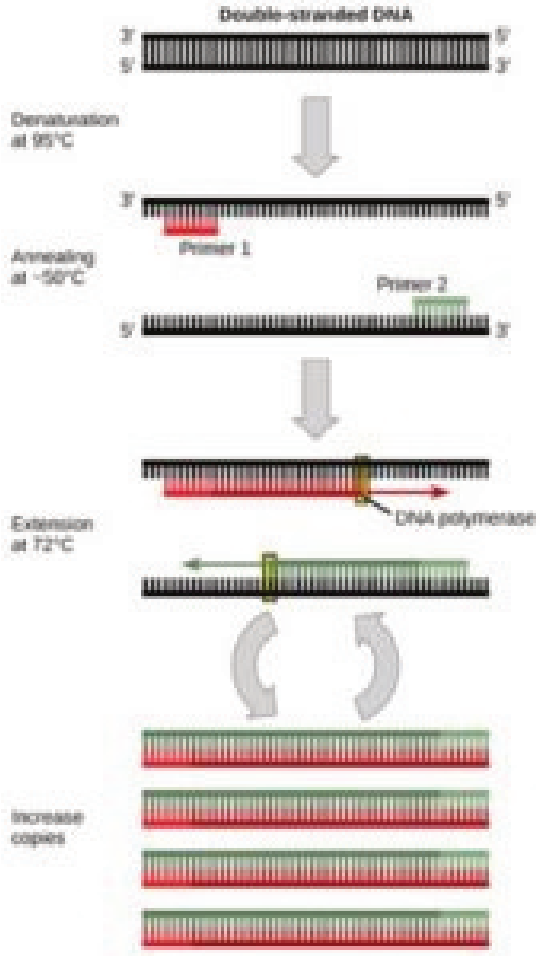


Figure 4: *Polymerase chain reaction, or PCR, is used to produce many copies of a specific sequence of DNA using a special form of DNA polymerase.*

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:8CA_YwJq@3/Cloning-and-Genetic-Engineerin

9.2 Cloning

In general, **cloning** means the creation of a perfect replica. Typically, the word is used to describe the creation of a genetically identical copy. In biology, the re-creation of a whole organism is referred to as "reproductive cloning." Long before attempts were made to clone an entire organism, researchers learned how to copy short stretches of DNA—a process that is referred to as molecular cloning.

Molecular Cloning

Cloning allows for the creation of multiple copies of genes, expression of genes, and study of specific genes. To get the DNA fragment into a bacterial cell in a form that will be copied or expressed, the fragment is first inserted into a plasmid. A **plasmid** (also called a vector in this context) is a small circular DNA molecule that replicates independently of the chromosomal DNA in bacteria. In cloning, the plasmid molecules can be used to provide a "vehicle" in which to insert a desired DNA fragment. Modified plasmids are usually reintroduced into a bacterial host for replication. As the bacteria divide, they copy their own DNA (including the plasmids). The inserted DNA fragment is copied along with the rest of the bacterial DNA. In a bacterial cell, the fragment of DNA from the human genome (or another organism that is being studied) is referred to as foreign DNA to differentiate it from the DNA of the bacterium (the host DNA).

Plasmids occur naturally in bacterial populations (such as *Escherichia coli*) and have genes that can contribute favorable traits to the organism, such as antibiotic resistance (the ability to be unaffected by antibiotics). Plasmids have been highly engineered as vectors for molecular cloning and for the subsequent large-scale production of important molecules, such as insulin. A valuable characteristic of plasmid vectors is the ease with which a foreign DNA fragment can be introduced. These plasmid vectors contain many short DNA sequences that can be cut with different commonly available **restriction enzymes**. Restriction enzymes (also called restriction endonucleases) recognize specific DNA sequences and cut them in a predictable manner; they are naturally produced by bacteria as a defense mechanism against foreign DNA. Many restriction enzymes make staggered cuts in the two strands of DNA, such that the cut ends have a 2- to 4-nucleotide single-stranded overhang. The sequence that is recognized by the restriction enzyme is a four- to eight-nucleotide sequence that is a palindrome. Like with a word palindrome, this means the sequence reads the same forward and backward. In most cases, the sequence reads the same forward on one strand and backward on the complementary strand. When a staggered cut is made in a sequence like this, the overhangs are complementary (**Figure 10.5**).

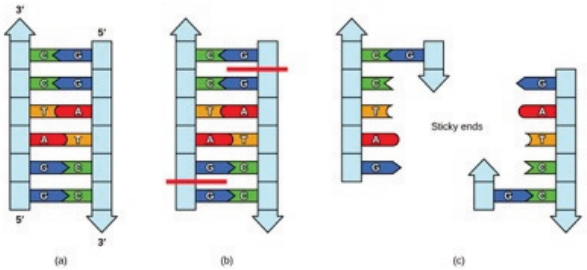


Figure 5: *In this (a) six-nucleotide restriction enzyme recognition site, notice that the sequence of six nucleotides reads the same in the 5′ to 3′ direction on one strand as it does in the 5′ to 3′ direction on the complementary strand. This is known as a palindrome. (b) The restriction enzyme makes breaks in the DNA strands, and (c) the cut in the DNA results in "sticky ends". Another piece of DNA cut on either end by the same restriction enzyme could attach to these sticky ends and be inserted into the gap made by this cut.*

Because these overhangs are capable of coming back together by hydrogen bonding with complementary overhangs on a piece of DNA cut with the same restriction enzyme, these are called "sticky ends." The process of forming hydrogen

bonds between complementary sequences on single strands to form doublestranded DNA is called **annealing**. Addition of an enzyme called DNA ligase, which takes part in DNA replication in cells, permanently joins the DNA fragments when the sticky ends come together. In this way, any DNA fragment can be spliced between the two ends of a plasmid DNA that has been cut with the same restriction enzyme (**Figure 10.6**).

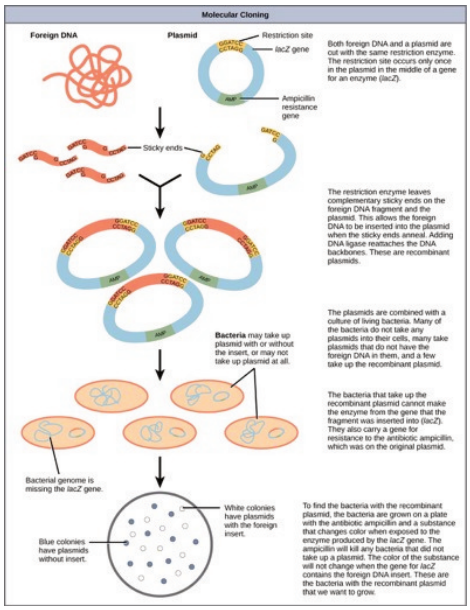


***Figure 6**: This diagram shows the steps involved in molecular cloning.*

Plasmids with foreign DNA inserted into them are called **recombinant DNA** molecules because they contain new combinations of genetic material. Proteins that are produced from recombinant DNA molecules are called **recombinant proteins**. Not all recombinant plasmids are capable of expressing genes. Plasmids may also be engineered to express proteins only when stimulated by certain environmental factors, so that scientists can control the expression of the recombinant proteins.

Reproductive Cloning

Reproductive cloning is a method used to make a clone or an identical copy of an entire multicellular organism. Most multicellular organisms undergo reproduction by sexual means, which involves the contribution of DNA from two individuals (parents), making it impossible to generate an identical copy or a clone of either parent. Recent advances in biotechnology have made it possible to reproductively clone mammals in the laboratory.

Natural sexual reproduction involves the union, during fertilization, of a sperm and an egg. Each of these gametes is haploid, meaning they contain one set of chromosomes in their nuclei. The resulting cell, or zygote, is then diploid and contains two sets of chromosomes. This cell divides mitotically to produce a multicellular organism. However, the union of just any two cells cannot produce a viable zygote; there are components in the cytoplasm of the egg cell that are essential for the early development of the embryo during its first few cell divisions. Without these provisions, there would be no subsequent development. Therefore, to produce a new individual, both a diploid genetic complement and an egg cytoplasm are required. The approach to producing an artificially cloned individual is to take the egg cell of one individual and to remove the haploid nucleus. Then a diploid nucleus from a body cell of a second individual, the donor, is put into the egg cell. The egg is then stimulated to divide so that development proceeds. This sounds simple, but in fact it takes many attempts before each of the steps is completed successfully.

The first cloned agricultural animal was Dolly, a sheep who was born in 1996. The success rate of reproductive cloning at the time was very low. Dolly lived for six years and died of a lung tumor (**Figure 10.7**). There was speculation that

because the cell DNA that gave rise to Dolly came from an older individual, the age of the DNA may have affected her life expectancy. Since Dolly, several species of animals (such as horses, bulls, and goats) have been successfully cloned.

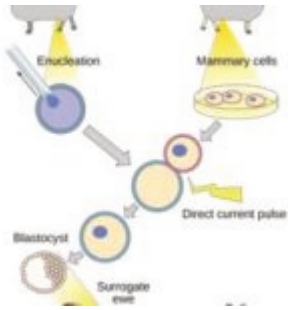


***Figure 7**: The creation of Dolly the sheep*

There have been attempts at producing cloned human embryos as sources of embryonic stem cells. In the procedure, the DNA from an adult human is introduced into a human egg cell, which is then stimulated to divide. The technology is similar to the technology that was used to produce Dolly, but the embryo is never implanted into a surrogate mother. The cells produced are called embryonic stem cells because they have the capacity to develop into many different kinds of cells, such as muscle or nerve cells. The stem cells could be used to research and ultimately provide therapeutic applications, such as replacing damaged tissues. The benefit of cloning in this instance is that the cells used to regenerate new tissues would be a perfect match to the donor of the original DNA. For example, a leukemia patient would not require a sibling with a tissue match for a bone-marrow transplant.

References

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:8CA_YwJq@3/Cloning-and-Genetic-Engineerin

9.3 Genetic Engineering

Using recombinant DNA technology to modify an organism's DNA to achieve desirable traits is called **genetic engineering**. Addition of foreign DNA in the form of recombinant DNA vectors that are generated by molecular cloning is the most common method of genetic engineering. An organism that receives the recombinant DNA is called a **genetically modified organism** (GMO). If the foreign DNA that is introduced comes from a different species, the host organism is called **transgenic**. Bacteria, plants, and animals have been genetically modified since the early 1970s for academic, medical, agricultural, and industrial purposes. These applications will be examined in more detail in the next module.

Although the classic methods of studying the function of genes began with a given phenotype and determined the genetic basis of that phenotype, modern techniques allow researchers to start at the DNA sequence level and ask: "What does this gene or DNA element do?" This technique, called **reverse genetics**, has resulted in reversing the classical genetic methodology. One example of this method is analogous to damaging a body part to determine its function. An insect that loses a wing cannot fly, which means that the wing's function is flight. The classic genetic method compares insects that cannot fly with insects that can fly, and observes that the non-flying insects have lost wings. Similarly in a reverse genetics approach, mutating or deleting genes provides researchers with clues about gene function. Alternately, reverse genetics can be used to cause a gene to overexpress itself to determine what phenotypic effects may occur.

References

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:8CA_YwJq@3/Cloning-and-Genetic-Engineerin

9.4 Biotechnology in Medicine and Agriculture

It is easy to see how biotechnology can be used for medicinal purposes. Knowledge of the genetic makeup of our species, the genetic basis of heritable diseases, and the invention of technology to manipulate and fix mutant genes provides methods to treat diseases. Biotechnology in agriculture can enhance resistance to disease, pests, and environmental stress to improve both crop yield and quality.

Genetic Diagnosis and Gene Therapy

The process of testing for suspected genetic defects before administering treatment is called genetic diagnosis by genetic testing. In some cases in which a genetic disease is present in an individual's family, family members may be advised to undergo genetic testing. For example, mutations in the *BRCA* genes may increase the likelihood of developing breast and ovarian cancers in women and some other cancers in women and men. A woman with breast cancer can be screened for these mutations. If one of the highrisk mutations is found, her female relatives may also wish to be screened for that particular mutation, or simply be more vigilant for the occurrence of cancers. Genetic testing is also offered for fetuses (or embryos with in vitro fertilization) to determine the presence or absence of disease-causing genes in families with specific debilitating diseases.

Gene therapy is a genetic engineering technique that may one day be used to cure certain genetic diseases. In its simplest form, it involves the introduction of a non-mutated gene at a random location in the genome to cure a disease by replacing a protein that may be absent in these individuals because of a genetic mutation. The non-mutated gene is usually introduced into diseased cells as part of a vector transmitted by a virus, such as an adenovirus, that can infect the host cell and deliver the foreign DNA into the genome of the targeted cell (**Figure 10.8**). To date, gene therapies have been primarily experimental procedures in humans. A few of these experimental treatments have been successful, but the methods may be important in the future as the factors limiting its success are resolved.

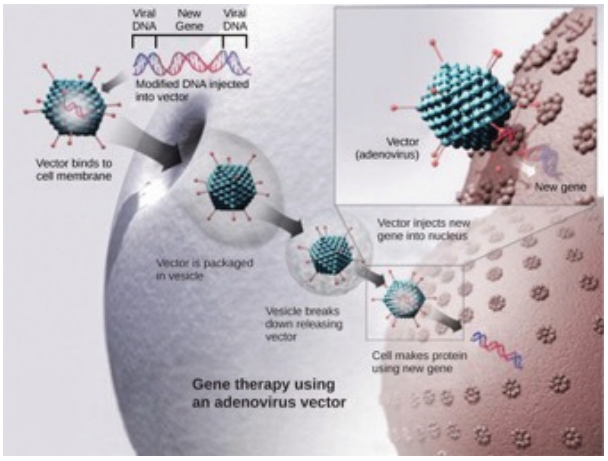


***Figure 8**: This diagram shows the steps involved in curing disease with gene therapy using an adenovirus vector. (credit: modification of work by NIH)*

Production of Vaccines, Antibiotics, and Hormones

Traditional vaccination strategies use weakened or inactive forms of microorganisms or viruses to stimulate the immune system. Modern techniques use specific genes of microorganisms cloned into vectors and mass-produced in bacteria to make large quantities of specific substances to stimulate the immune system. The substance is then used as a vaccine. In some cases, such as the H1N1 flu vaccine, genes cloned from the virus have been used to combat the constantly changing strains of this virus.

Antibiotics kill bacteria and are naturally produced by microorganisms such as fungi; penicillin is perhaps the most well-known example. Antibiotics are produced on a large scale by cultivating and manipulating fungal cells. The fungal cells have typically been genetically modified to improve the yields of the antibiotic compound.

Recombinant DNA technology was used to produce large-scale quantities of the human hormone insulin in *E. coli* as early as 1978. Previously, it was only possible to treat diabetes with pig insulin, which caused allergic reactions in many humans because of differences in the insulin molecule. In addition, human growth hormone (HGH) is used to treat growth disorders in children. The HGH gene was cloned from a cDNA (complementary DNA) library and inserted into *E. coli* cells by cloning it into a bacterial vector.

Transgenic Animals

Although several recombinant proteins used in medicine are successfully produced in bacteria, some proteins need a eukaryotic animal host for proper processing. For this reason, genes have been cloned and expressed in animals such as sheep, goats, chickens, and mice. Animals that have been modified to express recombinant DNA are called transgenic animals (**Figure 10.9**).

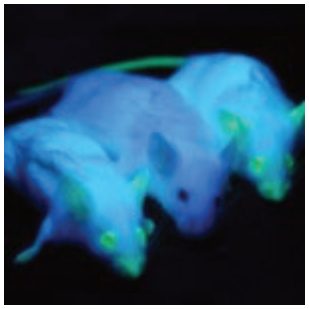


Figure 9: *It can be seen that two of these mice are transgenic because they have a gene that causes them to fluoresce under a UV light. The non-transgenic mouse does not have the gene that causes fluorescence. (credit: Ingrid Moen et al.)*

Several human proteins are expressed in the milk of transgenic sheep and goats. In one commercial example, the FDA has approved a blood anticoagulant protein that is produced in the milk of transgenic goats for use in humans. Mice have been used extensively for expressing and studying the effects of recombinant genes and mutations.

Transgenic Plants

Manipulating the DNA of plants (creating genetically modified organisms, or GMOs) has helped to create desirable traits such as disease resistance, herbicide, and pest resistance, better nutritional value, and better shelf life (**Figure 10.10**). Plants are the most important source of food for the human population. Farmers developed ways to select for plant varieties with desirable traits long before modernday biotechnology practices were established.

Transgenic plants have received DNA from other species. Because they contain unique combinations of genes and are not restricted to the laboratory, transgenic plants and other GMOs are closely monitored by government agencies to ensure that they are fit for human consumption and do not endanger other plant and animal life. Because foreign genes can spread to other species in the environment, particularly in the pollen and seeds of plants, extensive testing is required to ensure ecological stability. Staples like corn, potatoes, and tomatoes were the first crop plants to be genetically engineered.

Figure 10: *Corn, a major agricultural crop used to create products for a variety of industries, is often modified through plant biotechnology. (credit: Keith Weller, USDA)*

Transformation of Plants Using Agrobacterium tumefaciens

In plants, tumors caused by the bacterium *Agrobacterium tumefaciens* occur by transfer of DNA from the bacterium to the plant. The artificial introduction of DNA into plant cells is more challenging than in animal cells because of the thick plant cell wall. Researchers used the natural transfer of DNA from *Agrobacterium* to a plant host to introduce DNA fragments of their choice into plant hosts. In nature, the disease-causing *A. tumefaciens* have a set of plasmids that contain genes that integrate into the infected plant cell's genome. Researchers manipulate the plasmids to carry the desired DNA fragment and insert it into the plant genome.

The Organic Insecticide Bacillus thuringiensis

Bacillus thuringiensis (Bt) is a bacterium that produces protein crystals that are toxic to many insect species that feed on plants. Insects that have eaten Bt toxin stop feeding on the plants within a few hours. After the toxin is activated in the intestines of the insects, death occurs within a couple of days. The crystal toxin genes have been cloned from the bacterium and introduced into plants, therefore allowing plants to produce their own crystal Bt toxin that acts against insects. Bt toxin is safe for the environment and non-toxic to mammals (including humans). As a result, it has been approved for use by organic farmers as a natural insecticide. There is some concern, however, that insects may evolve resistance to the Bt toxin in the same way that bacteria evolve resistance to antibiotics.

FlavrSavr Tomato

The first GM crop to be introduced into the market was the FlavrSavr Tomato produced in 1994. Molecular genetic technology was used to slow down the process of softening and rotting caused by fungal infections, which led to increased shelf life of the GM tomatoes. Additional genetic modification improved the flavor of this tomato. The FlavrSavr tomato did not successfully stay in the market because of problems maintaining and shipping the crop.

References

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:rytx-nDe@3/Biotechnology-in-Medicine-and-

9.5 Genomics and Proteomics

The study of nucleic acids began with the discovery of DNA, progressed to the study of genes and small fragments, and has now exploded to the field of **genomics**. Genomics is the study of entire genomes, including the complete set of genes, their nucleotide sequence and organization, and their interactions within a species and with other species. The advances in genomics have been made possible by DNA sequencing technology. Just as information technology has led to Google Maps that enable us to get detailed information about locations around the globe, genomic information is used to create similar maps of the DNA of different organisms.

Mapping Genomes

Genome mapping is the process of finding the location of genes on each chromosome. The maps that are created are comparable to the maps that we use to navigate streets. A **genetic map** is an illustration that lists genes and their location on a chromosome. Genetic maps provide the big picture (similar to a map of interstate highways) and use genetic markers (similar to landmarks). A genetic marker is a gene or sequence on a chromosome that shows genetic linkage with a trait of interest. The genetic marker tends to be inherited with the gene of interest, and one measure of distance between them is the recombination frequency during meiosis. Early geneticists called this linkage analysis.

Physical maps get into the intimate details of smaller regions of the chromosomes (similar to a detailed road map) (**Figure 10.11**). A physical map is a representation of the physical distance, in nucleotides, between genes or genetic markers. Both genetic linkage maps and physical maps are required to build a complete picture of the genome. Having a complete map of the genome makes it easier for researchers to study individual genes. Human genome maps help researchers in their efforts to identify human disease-causing genes related to illnesses such as cancer, heart disease, and cystic fibrosis, to name a few. In addition, genome mapping can be used to help identify organisms with beneficial traits, such as microbes with the ability to clean up pollutants or even prevent pollution. Research involving plant genome mapping may lead to methods that produce higher crop yields or to the development of plants that adapt better to climate change.

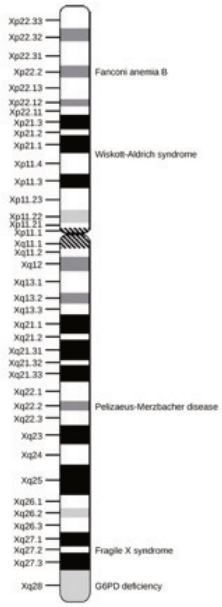


***Figure 11**: This is a physical map of the human X chromosome. (credit: modification of work by NCBI, NIH)*

Genetic maps provide the outline, and physical maps provide the details. It is easy to understand why both types of genome-mapping techniques are important to show the big picture. Information obtained from each technique is used in combination to study the genome. Genomic mapping is used with different model organisms that are used for research. Genome mapping is still an ongoing process, and as more advanced techniques are developed, more advances are expected. Genome mapping is similar to completing a complicated puzzle using every piece of available data. Mapping information generated in laboratories all over the world is entered into central databases, such as the National Center for Biotechnology Information (NCBI). Efforts are made to make the information more easily accessible to researchers and the general public. Just as we use global positioning systems instead of paper maps to navigate through roadways, NCBI allows us to use a genome viewer tool to simplify the data mining process.

Whole Genome Sequencing

Although there have been significant advances in the medical sciences in recent years, doctors are still confounded by many diseases and researchers are using whole genome sequencing to get to the bottom of the problem. **Whole genome sequencing** is a process that determines the DNA sequence of an entire genome. Whole genome sequencing is a brute-force approach to problem solving when there is a genetic basis at the core of a disease. Several laboratories now provide services to sequence, analyze, and interpret entire genomes.

In 2010, whole genome sequencing was used to save a young boy whose intestines had multiple mysterious abscesses. The child had several colon operations with no relief. Finally, a whole genome sequence revealed a defect in a pathway that controls apoptosis (programmed cell death). A bone marrow transplant was used to overcome this genetic disorder, leading to a cure for the boy. He was the first person to be successfully diagnosed using whole genome sequencing.

The first genomes to be sequenced, such as those belonging to viruses, bacteria, and yeast, were smaller in terms of the number of nucleotides than the genomes of multicellular organisms. The genomes of other model organisms, such as the mouse (*Mus musculus*), the fruit fly (*Drosophila melanogaster*), and the nematode (*Caenorhabditis elegans*) are now known. A great deal of basic research is performed in **model organisms** because the information can be applied to other organisms. A model organism is a species that is studied as a model to understand the biological processes in other species that can be represented by the model organism. For example, fruit flies are able to metabolize alcohol like humans, so the genes affecting sensitivity to alcohol have been studied in fruit flies in an effort to understand the variation in sensitivity to alcohol in humans. Having entire genomes sequenced helps with the research efforts in these model organisms (**Figure 10.12**).

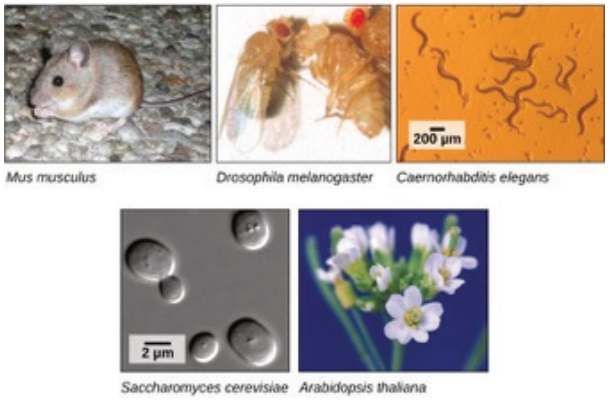


Figure 12: *Much basic research is done with model organisms, such as the mouse, Mus musculus; the fruit fly, Drosophila melanogaster; the nematode Caenorhabditis elegans; the yeast Saccharomyces cerevisiae; and the common weed, Arabidopsis thaliana. (credit "mouse": modification of work by Florean Fortescue; credit "nematodes": modification of work by "snickclunk"/Flickr; credit "common weed": modification of work by Peggy Greb, USDA; scale-bar data from Matt Russell)*

The first human genome sequence was published in 2003. The number of whole genomes that have been sequenced steadily increases and now includes hundreds of species and thousands of individual human genomes.

References

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:TE1njgbY@4/Genomics-and-Proteomics

9.6 Applying Genomics

The introduction of DNA sequencing and whole genome sequencing projects, particularly the Human Genome Project, has expanded the applicability of DNA sequence information. Genomics is now being used in a wide variety of fields, such as metagenomics, pharmacogenomics, and mitochondrial genomics. The most commonly known application of genomics is to understand and find cures for diseases.

Predicting Disease Risk at the Individual Level

Predicting the risk of disease involves screening and identifying currently healthy individuals by genome analysis at the individual level. Intervention with lifestyle changes and drugs can be recommended before disease onset. However, this approach is most applicable when the problem arises from a single gene mutation. Such defects only account for about 5 percent of diseases found in developed countries. Most of the common diseases, such as heart disease, are multifactorial or polygenic, which refers to a phenotypic characteristic that is determined by two or more genes, and also environmental factors such as diet. In April 2010, scientists at Stanford University published the genome analysis of a healthy individual (Stephen Quake, a scientist at Stanford University, who had his genome sequenced); the analysis predicted his propensity to acquire various diseases. A risk assessment was done to analyze Quake's percentage of risk for 55 different medical conditions. A rare genetic mutation was found that showed him to be at risk for sudden heart attack. He was also predicted to have a 23 percent risk of developing prostate cancer and a 1.4 percent risk of developing Alzheimer's disease. The scientists used databases and several publications to analyze the genomic data. Even though genomic sequencing is becoming more affordable and analytical tools are becoming more reliable, ethical issues surrounding genomic analysis at a population level remain to be addressed. For example, could such data be legitimately used to charge more or less for insurance or to affect credit ratings?

Genome-wide Association Studies

Since 2005, it has been possible to conduct a type of study called a genome-wide association study, or GWAS. A GWAS is a method that identifies differences between individuals in single nucleotide polymorphisms (SNPs) that may be involved in causing diseases. The method is particularly suited to diseases that may be affected by one or many genetic changes throughout the genome. It is very difficult to identify the genes involved in such a disease using family history information. The GWAS method relies on a genetic database that has been in development since 2002 called the International HapMap Project. The HapMap Project sequenced the genomes of several hundred individuals from around the world and identified groups of SNPs. The groups include SNPs that are located near to each other on chromosomes so they tend to stay together through recombination. The fact that the group stays together means that identifying one marker SNP is all that is needed to identify all the SNPs in the group. There are several million SNPs identified, but identifying them in other individuals who have not had their complete genome sequenced is much easier because only the marker SNPs need to be identified.

In a common design for a GWAS, two groups of individuals are chosen; one group has the disease, and the other group does not. The individuals in each group are matched in other characteristics to reduce the effect of confounding variables causing differences between the two groups. For example, the genotypes may differ because the two groups are mostly taken from different parts of the world. Once the individuals are chosen, and typically their numbers are a thousand or more for the study to work, samples of their DNA are obtained. The DNA is analyzed using automated systems to identify large differences in the percentage of particular SNPs between the two groups. Often the study examines a million or more SNPs in the DNA. The results of GWAS can be used in two ways: the genetic differences may be used as markers for susceptibility to the disease in undiagnosed individuals, and the particular genes identified can be targets for research into the molecular pathway of the disease and potential therapies. An offshoot of the discovery of gene associations with disease has been the formation of companies that provide socalled "personal genomics" that will identify risk levels for various diseases based on an individual's SNP complement. The science behind these services is controversial.

Because GWAS looks for associations between genes and disease, these studies provide data for other research into causes, rather than answering specific questions themselves. An association between a gene difference and a disease does

not necessarily mean there is a cause-and-effect relationship. However, some studies have provided useful information about the genetic causes of diseases. For example, three different studies in 2005 identified a gene for a protein involved in regulating inflammation in the body that is associated with a disease-causing blindness called age-related macular degeneration. This opened up new possibilities for research into the cause of this disease. A large number of genes have been identified to be associated with Crohn's disease using GWAS, and some of these have suggested new hypothetical mechanisms for the cause of the disease.

Pharmacogenomics

Pharmacogenomics involves evaluating the effectiveness and safety of drugs on the basis of information from an individual's genomic sequence. Personal genome sequence information can be used to prescribe medications that will be most effective and least toxic on the basis of the individual patient's genotype. Studying changes in gene expression could provide information about the gene transcription profile in the presence of the drug, which can be used as an early indicator of the potential for toxic effects. For example, genes involved in cellular growth and controlled cell death, when disturbed, could lead to the growth of cancerous cells. Genome-wide studies can also help to find new genes involved in drug toxicity. The gene signatures may not be completely accurate, but can be tested further before pathologic symptoms arise.

Metagenomics

Traditionally, microbiology has been taught with the view that microorganisms are best studied under pure culture conditions, which involves isolating a single type of cell and culturing it in the laboratory. Because microorganisms can go through several generations in a matter of hours, their gene expression profiles adapt to the new laboratory environment very quickly. On the other hand, many species resist being cultured in isolation. Most microorganisms do not live as isolated entities, but in microbial communities known as biofilms. For all of these reasons, pure culture is not always the best way to study microorganisms. **Metagenomics** is the study of the collective genomes of multiple species that grow and interact in an environmental niche. Metagenomics can be used to identify new species more rapidly and to analyze the effect of pollutants on the environment (**Figure 10.13**). Metagenomics techniques can now also be applied to communities of higher eukaryotes, such as fish.

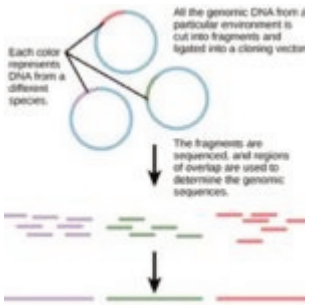


Figure 13: *Metagenomics involves isolating DNA from multiple species within an environmental niche. The DNA is cut up and sequenced, allowing entire genome sequences of multiple species to be reconstructed from the sequences of overlapping pieces.*

Creation of New Biofuels

Knowledge of the genomics of microorganisms is being used to find better ways to harness biofuels from algae and cyanobacteria. The primary sources of fuel today are coal, oil, wood, and other plant products such as ethanol. Although plants are renewable resources, there is still a need to find more alternative renewable sources of energy to meet our

population's energy demands. The microbial world is one of the largest resources for genes that encode new enzymes and produce new organic compounds, and it remains largely untapped. This vast genetic resource holds the potential to provide new sources of biofuels (**Figure 10.14**).

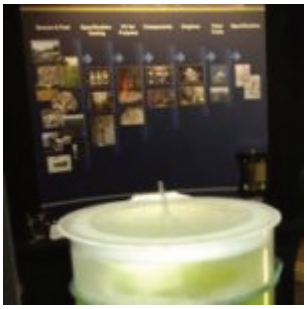


***Figure 14**: Renewable fuels were tested in Navy ships and aircraft at the first Naval Energy Forum. (credit: modification of work by John F Williams, US Navy)*

Mitochondrial Genomics

Mitochondria are intracellular organelles that contain their own DNA. Mitochondrial DNA mutates at a rapid rate and is often used to study evolutionary relationships. Another feature that makes studying the mitochondrial genome interesting is that in most multicellular organisms, the mitochondrial DNA is passed on from the mother during the process of fertilization. For this reason, mitochondrial genomics is often used to trace genealogy.

Genomics in Forensic Analysis

Information and clues obtained from DNA samples found at crime scenes have been used as evidence in court cases, and genetic markers have been used in forensic analysis. Genomic analysis has also become useful in this field. In 2001, the first use of genomics in forensics was published. It was a collaborative effort between academic research institutions and the FBI to solve the mysterious cases of anthrax (**Figure 10.15**) that was transported by the US Postal Service. Anthrax bacteria were made into an infectious powder and mailed to news media and two U.S. Senators. The powder infected the administrative staff and postal workers who opened or handled the letters. Five people died, and 17 were sickened from the bacteria. Using microbial genomics, researchers determined that a specific strain of anthrax was used in all the mailings; eventually, the source was traced to a scientist at a national biodefense laboratory in Maryland.

Genomics in Agriculture

Genomics can reduce the trials and failures involved in scientific research to a certain extent, which could improve the quality and quantity of crop yields in agriculture (**Figure 10.16**). Linking traits to genes or gene signatures helps to improve crop breeding to generate hybrids with the most desirable qualities. Scientists use genomic data to identify desirable traits, and then transfer those traits to a different organism to create a new genetically modified organism, as described in the previous module. Scientists are discovering how genomics can improve the quality and quantity of agricultural production. For example, scientists could use desirable traits to create a useful product or enhance an existing product, such as making a drought-sensitive crop more tolerant of the dry season.

References

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:TE1njgbY@4/Genomics-and-Proteomics

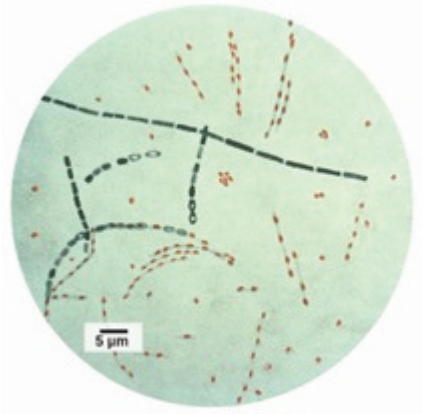


Figure 15: *Bacillus anthracis is the organism that causes anthrax. (credit: modification of work by CDC; scale-bar data from Matt Russell)*

Figure 16: *Transgenic agricultural plants can be made to resist disease. These transgenic plums are resistant to the plum pox virus. (credit: Scott Bauer, USDA ARS)*

9.7 Proteomics

Proteins are the final products of genes that perform the function encoded by the gene. Proteins are composed of amino acids and play important roles in the cell. All enzymes (except ribozymes) are proteins and act as catalysts that affect the rate of reactions. Proteins are also regulatory molecules, and some are hormones. Transport proteins, such as hemoglobin, help transport oxygen to various organs. Antibodies that defend against foreign particles are also proteins. In the diseased state, protein function can be impaired because of changes at the genetic level or because of direct impact on a specific protein.

A proteome is the entire set of proteins produced by a cell type. Proteomes can be studied using the knowledge of genomes because genes code for mRNAs, and the mRNAs encode proteins. The study of the function of proteomes is called **proteomics**. Proteomics complements genomics and is useful when scientists want to test their hypotheses that were based on genes. Even though all cells in a multicellular organism have the same set of genes, the set of proteins produced in different tissues is different and dependent on gene expression. Thus, the genome is constant, but the proteome varies and is dynamic within an organism. In addition, RNAs can be alternatively spliced (cut and pasted to create novel combinations and novel proteins), and many proteins are modified after translation. Although the genome provides a blueprint, the final architecture depends on several factors that can change the progression of events that generate the proteome.

Genomes and proteomes of patients suffering from specific diseases are being studied to understand the genetic basis of the disease. The most prominent disease being studied with proteomic approaches is cancer (**Figure 10.17**). Proteomic approaches are being used to improve the screening and early detection of cancer; this is achieved by identifying proteins whose expression is affected by the disease process. An individual protein is called a **biomarker**, whereas a set of proteins with altered expression levels is called a **protein signature**. For a biomarker or protein signature to be useful as a candidate for early screening and detection of a cancer, it must be secreted in body fluids such as sweat, blood, or urine, so that large-scale screenings can be performed in a noninvasive fashion. The current problem with using biomarkers for the early detection of cancer is the high rate of false-negative results. A false-negative result is a negative test result that should have been positive. In other words, many cases of cancer go undetected, which makes biomarkers unreliable. Some examples of protein biomarkers used in cancer detection are CA-125 for ovarian cancer and PSA for prostate cancer. Protein signatures may be more reliable than biomarkers to detect cancer cells. Proteomics is also being used to develop individualized treatment plans, which involves the prediction of whether or not an individual will respond to specific drugs and the side effects that the individual may have. Proteomics is also being used to predict the possibility of disease recurrence.

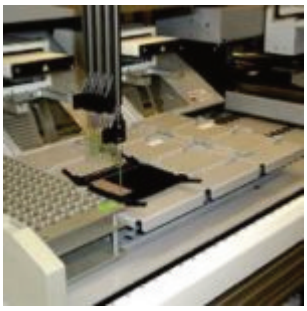


***Figure 17**: This machine is preparing to do a proteomic pattern analysis to identify specific cancers so that an accurate cancer prognosis can be made. (credit: Dorie Hightower, NCI, NIH)*

The National Cancer Institute has developed programs to improve the detection and treatment of cancer. The Clinical Proteomic Technologies for Cancer and the Early Detection Research Network are efforts to identify protein signatures specific to different types of cancers. The Biomedical Proteomics Program is designed to identify protein signatures and design effective therapies for cancer patients.

References

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:TE1njgbY@4/Genomics-and-Proteomics

10. Cell Division – Binary Fission and Mitosis

Learning Objectives

By the end of this section, you will be able to:

- Describe the process and consequences of binary fission and mitosis.

References

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:Vbi92lHB@9/The-Cell-Cycle

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:LlKfCy5H@4/Prokaryotic-Cell-Division

10.1 Prokaryotic Cell Division

The cell division process of prokaryotes (such as E. coli bacteria) is called **binary fission**. For unicellular organisms, cell division is the only method to produce new individuals. The outcome of this type of cell reproduction is a pair of daughter cells that are genetically identical to the parent cell. In unicellular organisms, daughter cells are whole individual organisms. This is a less complicated and much quicker process than cell division in eukaryotes. Because of the speed of bacterial cell division, populations of bacteria can grow very rapidly.

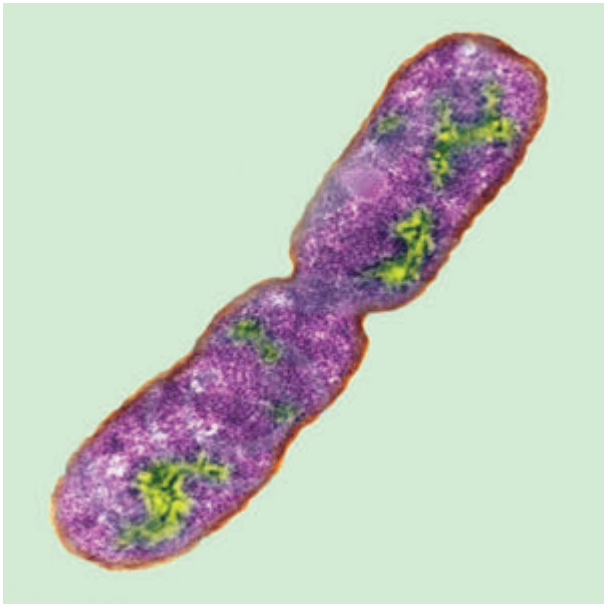


***Figure 1**: An E. coli bacteria dividing into two identical daughter cells*

To achieve the outcome of identical daughter cells, there are some essential steps. The genomic DNA must be replicated (using DNA replication) and then one copy must be moved into each of the daughter cells. The cytoplasmic contents must also be divided to give both new cells the machinery to sustain life. In bacterial cells, the genome consists of a single, circular DNA chromosome; therefore, the process of cell division is very simple.

References

Unless otherwise noted, images on this page are licensed under CC-BY 4.0 by OpenStax.

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:Vbi92lHB@9/The-Cell-Cycle

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:LlKfCy5H@4/Prokaryotic-Cell-Division

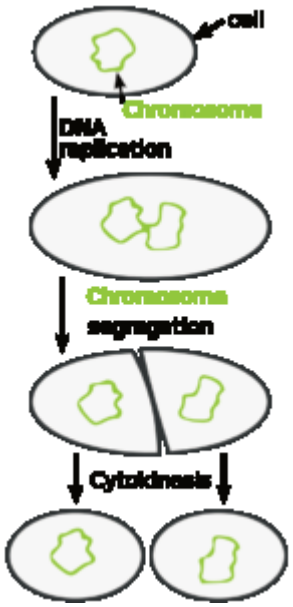


Figure 2: *Prokaryotic cell division occurs via a process called binary fission.*

10.2 Eukaryotic Cell Division

Eukaryotes have two major types of cell division: mitosis and meiosis. Mitosis is used to produce new body cells for growth and healing, while meiosis is used to produce sex cells (eggs and sperm). Meiosis will be discussed in a later chapter.

The **cell cycle** is an ordered series of events involving cell growth and cell division that produces two new daughter cells via mitosis. The length of the cell cycle is highly variable even within the cells of an individual organism. In humans, the frequency of cell turnover ranges from a few hours in early embryonic development to an average of two to five days for epithelial cells, or to an entire human lifetime spent without dividing in specialized cells such as cortical neurons or cardiac muscle cells. There is also variation in the time that a cell spends in each phase of the cell cycle. When fast-dividing mammalian cells are grown in culture (outside the body under optimal growing conditions), the length of the cycle is approximately 24 hours. The timing of events in the cell cycle is controlled by mechanisms that are both internal and external to the cell.

Cells on the path to cell division proceed through a series of precisely timed and carefully regulated stages of growth, DNA replication, and division that produce two genetically identical cells. The cell cycle has two major phases: interphase and the mitotic phase (**Figure 6.3**). During **interphase**, the cell grows and DNA is replicated. During the **mitotic phase**, the replicated DNA and cytoplasmic contents are separated and the cell divides.

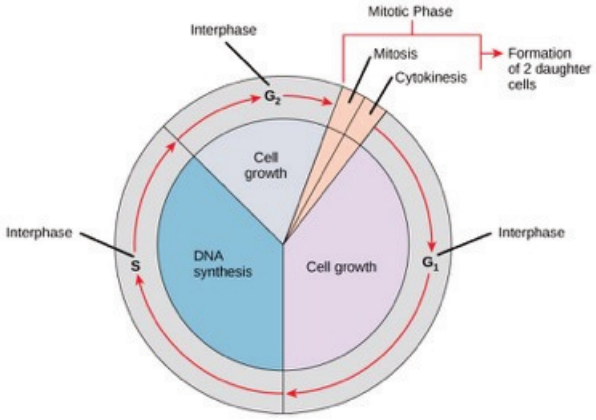


Figure 3: *A cell moves through a series of phases in an orderly manner. During interphase, G1 involves cell growth and protein synthesis, the S phase involves DNA replication and the replication of the centrosome, and G2 involves further growth and protein synthesis. The mitotic phase follows interphase. Mitosis is nuclear division during which duplicated chromosomes are segregated and distributed into daughter nuclei. Usually the cell will divide after mitosis in a process called cytokinesis in which the cytoplasm is divided and two daughter cells are formed.*

Interphase

During interphase, the cell undergoes normal processes while also preparing for cell division. For a cell to move from interphase to the mitotic phase, many internal and external conditions must be met.

The Mitotic Phase

To make two daughter cells, the contents of the nucleus and the cytoplasm must be divided. The mitotic phase is a multistep process during which the duplicated chromosomes are aligned, separated, and moved to opposite poles of the cell, and then the cell is divided into two new identical daughter cells. The first portion of the mitotic phase, **mitosis**, is

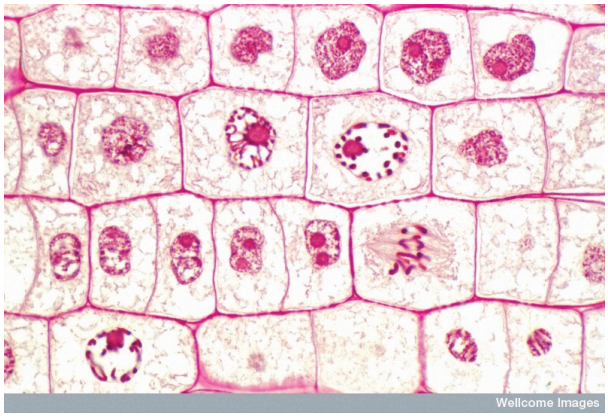


Figure 4: Mitosis in onion root cells. The cells in this image are in various stages of mitosis. (Credit: Spike Walker. Wellcome Images images@wellcome.ac.uk http://images.wellcome.ac.uk)

composed of five stages, which accomplish nuclear division. The second portion of the mitotic phase, called cytokinesis, is the physical separation of the cytoplasmic components into two daughter cells. Although the stages of mitosis are similar for most eukaryotes, the process of cytokinesis is quite different for eukaryotes that have cell walls, such as plant cells.

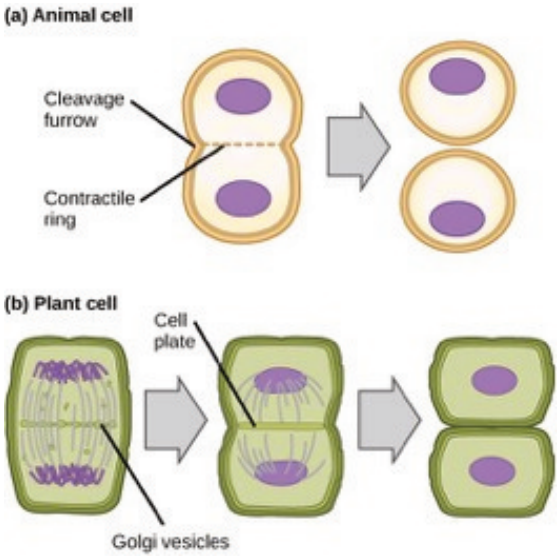


Figure 5: In part (a), a cleavage furrow forms at the former metaphase plate in the animal cell. The plasma membrane is drawn in by a ring of actin fibers contracting just inside the membrane. The cleavage furrow deepens until the cells are pinched in two. In part (b), Golgi vesicles coalesce at the former metaphase plate in a plant cell. The vesicles fuse and form the cell plate. The cell plate grows from the center toward the cell walls. New cell walls are made from the vesicle contents.

G0 Phase

Not all cells adhere to the classic cell-cycle pattern in which a newly formed daughter cell immediately enters interphase, closely followed by the mitotic phase. Cells in the **G0 phase** are not actively preparing to divide. The cell is in a quiescent (inactive) stage, having exited the cell cycle. Some cells enter G0 temporarily until an external signal triggers the onset of G1. Other cells that never or rarely divide, such as mature cardiac muscle and nerve cells, remain in G0 permanently (**Figure 6.6**).

Figure 6.6 Cells that are not actively preparing to divide enter an alternate phase called G0. In some cases, this is a temporary condition until triggered to enter G1. In other cases, the cell will remain in G0 permanently.

References

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:Vbi92lHB@9/The-Cell-Cycle

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:LlKfCy5H@4/Prokaryotic-Cell-Division

10.3 Control of the Cell Cycle

It is essential that daughter cells be exact duplicates of the parent cell. Mistakes in the duplication or distribution of the chromosomes lead to mutations that may be passed forward to every new cell produced from the abnormal cell. To prevent a compromised cell from continuing to divide, there are internal control mechanisms that operate at three main **cell cycle checkpoints** at which the cell cycle can be stopped until conditions are favorable.

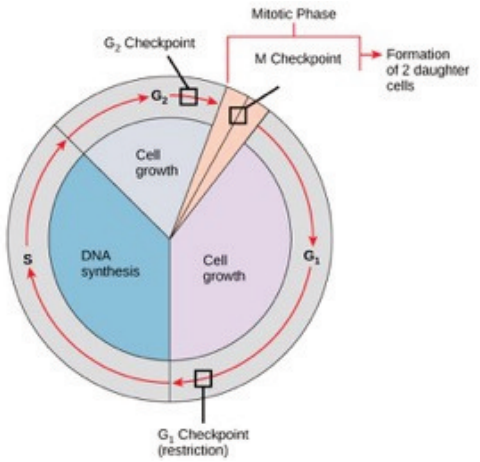


Figure 6: *The cell cycle is controlled at three checkpoints. Integrity of the DNA is assessed at the G1 checkpoint. Proper chromosome duplication is assessed at the G2 checkpoint. Attachment of each kinetochore to a spindle fiber is assessed at the M checkpoint.*

The first checkpoint determines whether all conditions are favorable for cell division to proceed. This checkpoint is the point at which the cell irreversibly commits to the cell-division process. In addition to adequate reserves and cell size, there is a check for damage to the genomic DNA. A cell that does not meet all the requirements will not be released into the S phase.

The second checkpoint bars the entry to the mitotic phase if certain conditions are not met. The most important role of this checkpoint is to ensure that all of the chromosomes have been replicated and that the replicated DNA is not damaged.

The final checkpoint occurs in the middle of mitosis. This checkpoint determines if all of the copied chromosomes are arranged appropriately to be separated to opposite sides of the cell. If this doesn't happen correctly, incorrect numbers of chromosomes can be partitioned into each of the daughter cells, which would likely cause them to die.

References

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:Vbi92lHB@9/The-Cell-Cycle

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:LlKfCy5H@4/Prokaryotic-Cell-Division

10.4 Cancer and the Cell Cycle

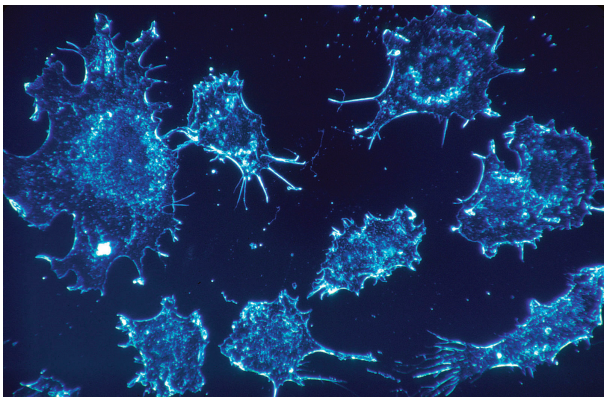


***Figure 7**: Cancer cells in culture from human connective tissue, illuminated by darkfield amplified contrast, at a magnification of 500x.*

Cancer is a collective name for many different diseases caused by a common mechanism: uncontrolled cell division. Despite the redundancy and overlapping levels of cell-cycle control, errors occur. One of the critical processes monitored by the cell-cycle checkpoint surveillance mechanism is the proper replication of DNA. Even when all of the cell-cycle controls are fully functional, a small percentage of replication errors (mutations) will be passed on to the daughter cells. If one of these changes to the DNA nucleotide sequence occurs within a gene, a gene mutation results. All cancers begin when a gene mutation gives rise to a faulty protein that participates in the process of cell reproduction. The change in the cell that results from the malformed protein may be minor. Even minor mistakes, however, may allow subsequent mistakes to occur more readily. Over and over, small, uncorrected errors are passed from parent cell to daughter cells and accumulate as each generation of cells produces more non-functional proteins from uncorrected DNA damage. Eventually, the pace of the cell cycle speeds up as the effectiveness of the control and repair mechanisms decreases. Uncontrolled growth of the mutated cells outpaces the growth of normal cells in the area, and a tumor can result.

References

Unless otherwise noted, images on this page are licensed under CC-BY 4.0 by OpenStax.

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:Vbi92lHB@9/The-Cell-Cycle

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:LlKfCy5H@4/Prokaryotic-Cell-Division

11. Meiosis

Figure 1: *Each of us, like these other large multicellular organisms, begins life as a fertilized egg. After trillions of cell divisions, each of us develops into a complex, multicellular organism. (credit a: modification of work by Frank Wouters; credit b: modification of work by Ken Cole, USGS; credit c: modification of work by Martin Pettitt)*

The ability to reproduce *in kind* is a basic characteristic of all living things. *In kind* means that the offspring of any organism closely resembles its parent or parents. Hippopotamuses give birth to hippopotamus calves; Monterey pine trees produce seeds from which Monterey pine seedlings emerge; and adult flamingos lay eggs that hatch into flamingo chicks. *In kind* does not generally mean *exactly the same*. While many single-celled organisms and a few multicellular organisms can produce genetically identical clones of themselves through mitotic cell division, many single-celled organisms and most multicellular organisms reproduce regularly using another method.

Sexual reproduction is the production by parents of sex cells and the fusion of two sex cells to form a single, unique cell. In multicellular organisms, this new cell will then undergo mitotic cell divisions to develop into an adult organism. A type of cell division called meiosis leads to the cells that are part of the sexual reproductive cycle. Sexual reproduction, specifically meiosis and fertilization, introduces variation into offspring that may account for the evolutionary success of sexual reproduction. The vast majority of eukaryotic organisms can or must employ some form of meiosis and fertilization to reproduce.

References

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:qOUtHXNY@3/Sexual-Reproduction

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:1Q8z96mT@4/Meiosis

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:6-3MVU-j@4/Errors-in-Meiosis

11.1 Sexual Reproduction

Sexual reproduction was an early evolutionary innovation after the appearance of eukaryotic cells. The fact that most eukaryotes reproduce sexually is evidence of its evolutionary success. In many animals, it is the only mode of reproduction. And yet, scientists recognize some real disadvantages to sexual reproduction. On the surface, offspring that are genetically identical to the parent may appear to be more advantageous. If the parent organism is successfully occupying a habitat, offspring with the same traits would be similarly successful. There is also the obvious benefit to an organism that can produce offspring by asexual budding, fragmentation, or asexual eggs. These methods of reproduction do not require another organism of the opposite sex. There is no need to expend energy finding or attracting a mate. That energy can be spent on producing more offspring. Indeed, some organisms that lead a solitary lifestyle have retained the ability to reproduce asexually. In addition, asexual populations only have female individuals, so every individual is capable of reproduction. In contrast, the males in sexual populations (half the population) are not producing offspring themselves. Because of this, an asexual population can grow twice as fast as a sexual population in theory. This means that in competition, the asexual population would have the advantage. All of these advantages to asexual reproduction, which are also disadvantages to sexual reproduction, should mean that the number of species with asexual reproduction should be more common.

However, multicellular organisms that exclusively depend on asexual reproduction are exceedingly rare. Why is sexual reproduction so common? This is one of the important questions in biology and has been the focus of much research from the latter half of the twentieth century until now. A likely explanation is that the variation that sexual reproduction creates among offspring is very important to the survival and reproduction of those offspring. The only source of variation in asexual organisms is mutation. This is the ultimate source of variation in sexual organisms. In addition, those different mutations are continually reshuffled from one generation to the next when different parents combine their unique genomes, and the genes are mixed into different combinations by the process of **meiosis**. Meiosis is the division of the contents of the nucleus that divides the chromosomes among gametes. Variation is introduced during meiosis, as well as when the gametes combine in fertilization.

References

OpenStax, Biology. OpenStax CNX. May 27, 2016http://cnx.org/contents/s8Hh0oOc@9.10:qOUtHXNY@3/Sexual-Reproduction

11.2 Overview of Meiosis

Sexual reproduction requires **fertilization**, a union of two cells from two individual organisms. If those two cells each contain one set of chromosomes, then the resulting cell contains two sets of chromosomes. The number of sets of chromosomes in a cell is called its ploidy level. **Haploid** cells contain one set of chromosomes. Cells containing two sets of chromosomes are called **diploid**. If the reproductive cycle is to continue, the diploid cell must somehow reduce its number of chromosome sets before fertilization can occur again, or there will be a continual doubling in the number of chromosome sets in every generation. So, in addition to fertilization, sexual reproduction includes a nuclear division, known as meiosis, that reduces the number of chromosome sets.

Most animals and plants are diploid, containing two sets of chromosomes; in each **somatic cell** (the non-reproductive cells of a multicellular organism), the nucleus contains two copies of each chromosome that are referred to as homologous chromosomes. Somatic cells are sometimes referred to as "body" cells. **Homologous chromosomes** are matched pairs containing genes for the same traits in identical locations along their length. Diploid organisms inherit one copy of each homologous chromosome from each parent; all together, they are considered a full set of chromosomes. In animals, haploid cells containing a single copy of each homologous chromosome are found only within gametes. Gametes fuse with another haploid gamete to produce a diploid cell.

The nuclear division that forms haploid cells, which is called meiosis, is related to mitosis. As you have learned, mitosis is part of a cell reproduction cycle that results in identical daughter nuclei that are also genetically identical to the original parent nucleus. In mitosis, both the parent and the daughter nuclei contain the same number of chromosome sets—diploid for most plants and animals. Meiosis employs many of the same mechanisms as mitosis. However, the starting nucleus is always diploid and the nuclei that result at the end of a meiotic cell division are haploid. To achieve the reduction in chromosome number, meiosis consists of one round of chromosome duplication and two rounds of nuclear division. Because the events that occur during each of the division stages are analogous to the events of mitosis, the same stage names are assigned. However, because there are two rounds of division, the stages are designated with a "I" or "II." Thus, **meiosis I** is the first round of meiotic division and reduces the number of chromosome sets from two to one. The genetic information is also mixed during this division to create unique recombinant chromosomes. **Meiosis II**, in which the second round of meiotic division takes place in a way that is similar to mitosis, separates the sister chromatids (the identical copies of each chromosome produced during DNA replication that are attached at the centromere).

References

OpenStax, Biology. OpenStax CNX. May 27, 2016http://cnx.org/contents/s8Hh0oOc@9.10:1Q8z96mT@4/Meiosis

11.3 Interphase

Meiosis is preceded by an interphase which is nearly identical to the interphase preceding mitosis. During interphase, the DNA of the chromosomes is replicated. After DNA replication, each chromosome becomes composed of two identical copies (called sister chromatids) that are held together at the centromere until they are pulled apart during meiosis II.

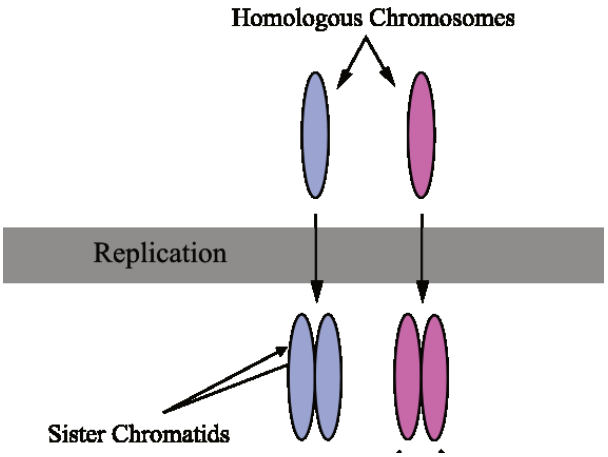


Figure 2: *Sister chromatids are identical copies of a chromosome that are held together at the centromere. They are produced during DNA replication. (Credit: User:SyntaxError55, from Wikimedia)*

References

OpenStax, Biology. OpenStax CNX. May 27, 2016http://cnx.org/contents/s8Hh0oOc@9.10:1Q8z96mT@4/Meiosis

11.4 Meiosis I

Early in meiosis I, the chromosomes can be seen clearly microscopically. As the nuclear envelope begins to break down, the proteins associated with homologous chromosomes bring the pair close to each other. The tight pairing of the homologous chromosomes is called **synapsis**. In synapsis, the genes on the chromatids of the homologous chromosomes are precisely aligned with each other. An exchange of chromosome segments between non-sister homologous chromatids occurs and is called **crossing over**. The crossover events are the first source of genetic variation produced by meiosis. A single crossover event between homologous non-sister chromatids leads to a reciprocal exchange of equivalent DNA between a maternal chromosome and a paternal chromosome. Now, when that sister chromatid is moved into a gamete, it will carry some DNA from one parent of the individual and some DNA from the other parent. The **recombinant** sister chromatid has a combination of maternal and paternal genes that did not exist before the crossover.

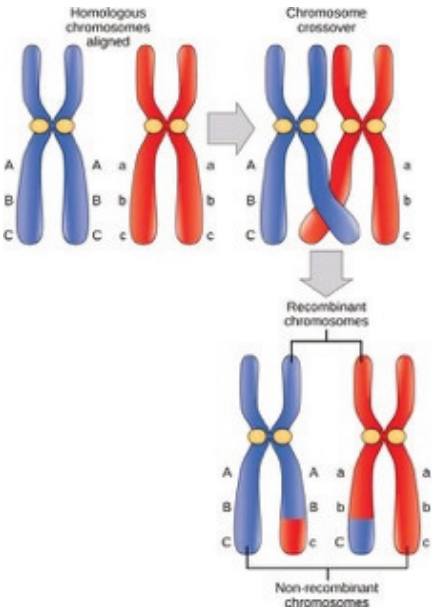


Figure 3: *In this illustration of the effects of crossing over, the blue chromosome came from the individual's father and the red chromosome came from the individual's mother. Crossover occurs between non-sister chromatids of homologous chromosomes. The result is an exchange of genetic material between homologous chromosomes. The chromosomes that have a mixture of maternal and paternal sequence are called recombinant and the chromosomes that are completely paternal or maternal are called non-recombinant.*

During meiosis I, the homologous chromosomes are arranged in the center of the cell with the kinetochores facing opposite poles. The orientation of each pair of homologous chromosomes at the center of the cell is random. This randomness, called independent assortment, is the physical basis for the generation of the second form of genetic variation in offspring. Consider that the homologous chromosomes of a sexually reproducing organism are originally inherited as two separate sets, one from each parent. Using humans as an example, one set of 23 chromosomes is present in the egg donated by the mother. The father provides the other set of 23 chromosomes in the sperm that fertilizes the egg. In metaphase I, these pairs line up at the midway point between the two poles of the cell. Because there is an equal chance

that a microtubule fiber will encounter a maternally or paternally inherited chromosome, the arrangement of the tetrads at the metaphase plate is random. Any maternally inherited chromosome may face either pole. Any paternally inherited chromosome may also face either pole. The orientation of each tetrad is independent of the orientation of the other 22 tetrads.

In each cell that undergoes meiosis, the arrangement of the tetrads is different. The number of variations depends on the number of chromosomes making up a set. There are two possibilities for orientation (for each tetrad); thus, the possible number of alignments equals 2^n where n is the number of chromosomes per set. Humans have 23 chromosome pairs, which results in over eight million (223) possibilities. This number does not include the variability previously created in the sister chromatids by crossover. Given these two mechanisms, it is highly unlikely that any two haploid cells resulting from meiosis will have the same genetic composition (**Figure 4**).

To summarize the genetic consequences of meiosis I: the maternal and paternal genes are recombined by crossover events occurring on each homologous pair during prophase I; in addition, the random assortment of tetrads at metaphase produces a unique combination of maternal and paternal chromosomes that will make their way into the gametes.

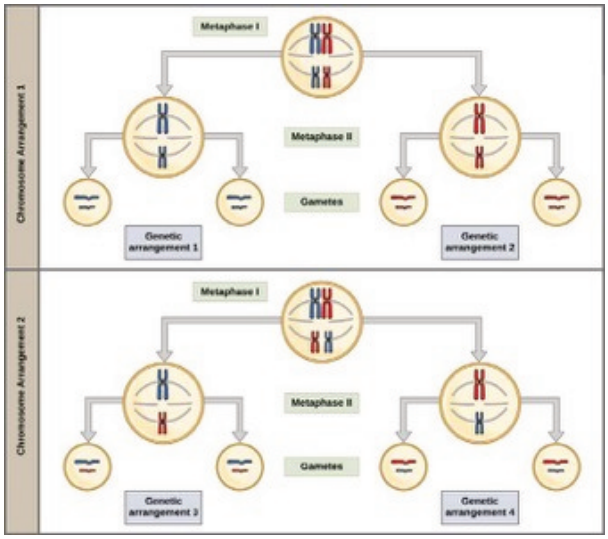


***Figure 4**: To demonstrate random, independent assortment at metaphase I, consider a cell with n = 2. In this case, there are two possible arrangements at the equatorial plane in metaphase I, as shown in the upper cell of each panel. These two possible orientations lead to the production of genetically different gametes. With more chromosomes, the number of possible arrangements increases dramatically.*

Cytokinesis, the physical separation of the cytoplasmic components into two daughter cells, occurs without reformation of the nuclei in other organisms. In nearly all species, cytokinesis separates the cell contents by either a cleavage furrow (in animals and some fungi), or a cell plate that will ultimately lead to formation of cell walls that separate the two daughter cells (in plants). At each pole, there is just one member of each pair of the homologous chromosomes, so only one full set of the chromosomes is present. This is why the cells are considered haploid—there is only one chromosome set, even though there are duplicate copies of the set because each homolog still consists of two sister chromatids that are still attached to each other. However, although the sister chromatids were once duplicates of the same chromosome, they are no longer identical at this stage because of crossovers.

References

OpenStax, Biology. OpenStax CNX. May 27, 2016http://cnx.org/contents/s8Hh0oOc@9.10:1Q8z96mT@4/Meiosis

11.5 Meiosis II

In meiosis II, the connected sister chromatids remaining in the haploid cells from meiosis I will be split to form four haploid cells. The two cells produced in meiosis I go through the events of meiosis II in synchrony. Overall, meiosis II resembles the mitotic division of a haploid cell. During meiosis II, the sister chromatids are pulled apart by the spindle fibers and move toward opposite poles.

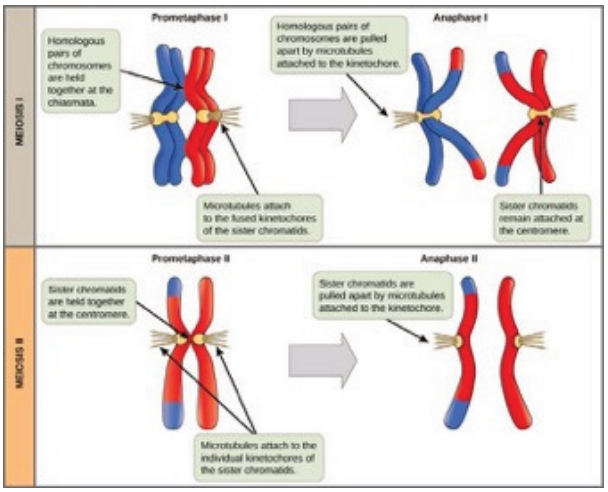


***Figure 5**: In prometaphase I, microtubules attach to the fused kinetochores of homologous chromosomes. In anaphase I, the homologous chromosomes are separated. In prometaphase II, microtubules attach to individual kinetochores of sister chromatids. In anaphase II, the sister chromatids are separated.*

Nuclear envelopes form around the chromosomes. Cytokinesis separates the two cells into four genetically unique haploid cells. At this point, the nuclei in the newly produced cells are both haploid and have only one copy of the single set of chromosomes. The cells produced are genetically unique because of the random assortment of paternal and maternal homologs and because of the recombination of maternal and paternal segments of chromosomes—with their sets of genes—that occurs during crossover.

References

OpenStax, Biology. OpenStax CNX. May 27, 2016http://cnx.org/contents/s8Hh0oOc@9.10:1Q8z96mT@4/Meiosis

11.6 Comparing Meiosis and Mitosis

Mitosis and meiosis, which are both forms of division of the nucleus in eukaryotic cells, share some similarities, but also exhibit distinct differences that lead to their very different outcomes. Mitosis is a single nuclear division that results in two nuclei, usually partitioned into two new cells. The nuclei resulting from a mitotic division are genetically identical to the original. They have the same number of sets of chromosomes: one in the case of haploid cells, and two in the case of diploid cells. On the other hand, meiosis is two nuclear divisions that result in four nuclei, usually partitioned into four new cells. The nuclei resulting from meiosis are never genetically identical, and they contain one chromosome set only—this is half the number of the original cell, which was diploid (**Figure 6**).

The differences in the outcomes of meiosis and mitosis occur because of differences in the behavior of the chromosomes during each process. Most of these differences in the processes occur in meiosis I, which is a very different nuclear division than mitosis. In meiosis I, the homologous chromosome pairs become associated with each other, are bound together, experience chiasmata and crossover between sister chromatids, and line up along the metaphase plate in tetrads with spindle fibers from opposite spindle poles attached to each kinetochore of a homolog in a tetrad. All of these events occur only in meiosis I, never in mitosis.

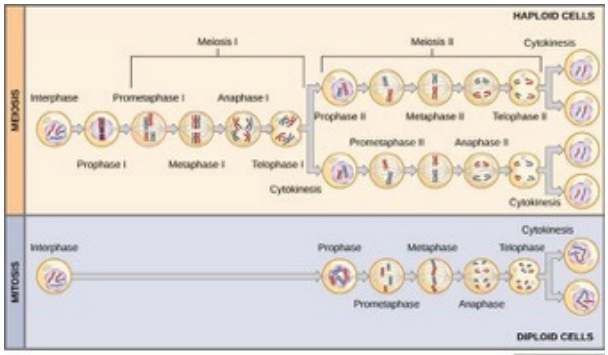
Homologous chromosomes move to opposite poles during meiosis I so the number of sets of chromosomes in each nucleus-to-be is reduced from two to one. For this reason, meiosis I is referred to as a **reduction division**. There is no such reduction in ploidy level in mitosis.

Meiosis II is much more analogous to a mitotic division. In this case, duplicated chromosomes (only one set of them) line up at the center of the cell with divided kinetochores attached to spindle fibers from opposite poles. During anaphase II, as in mitotic anaphase, the kinetochores divide and one sister chromatid is pulled to one pole and the other sister chromatid is pulled to the other pole. If it were not for the fact that there had been crossovers, the two products of each meiosis II division would be identical as in mitosis; instead, they are different because there has always been at least one crossover per chromosome. Meiosis II is not a reduction division because, although there are fewer copies of the genome in the resulting cells, there is still one set of chromosomes, as there was at the end of meiosis I.

Cells produced by mitosis will function in different parts of the body as a part of growth or replacing dead or damaged cells. They may even be involved in asexual reproduction in some organisms. Cells produced by meiosis in a diploid-dominant organism such as an animal will only participate in sexual reproduction.

References

OpenStax, Biology. OpenStax CNX. May 27, 2016http://cnx.org/contents/s8Hh0oOc@9.10:1Q8z96mT@4/Meiosis



		OUTCOME				
PROCESS	DNA synthesis	Synapsis of homologous chromosomes	Crossover	Homologous chromosomes line up at metaphase plate	Sister chromatids line up at metaphase plate	Number and genetic composition of daughter cells
MEIOSIS	Occurs in S phase of interphase	During prophase I	During prophase I	During metaphase I	During metaphase II	Four haploid cells at the end of meiosis II
MITOSIS	Occurs in S phase of interphase	Does not occur in mitosis	Does not occur in mitosis	Does not occur in mitosis	During metaphase	Two diploid cells at the end of mitosis

Figure 6: *Meiosis and mitosis are both preceded by one round of DNA replication; however, meiosis includes two nuclear divisions. The four daughter cells resulting from meiosis are haploid and genetically distinct. The daughter cells resulting from mitosis are diploid and identical to the parent cell.*

11.7 Errors in Meiosis

Inherited disorders can arise when chromosomes behave abnormally during meiosis. Chromosome disorders can be divided into two categories: abnormalities in chromosome number and chromosome structural rearrangements. Because even small segments of chromosomes can span many genes, chromosomal disorders are characteristically dramatic and often fatal.

Disorders in Chromosome Number

The isolation and microscopic observation of chromosomes forms the basis of cytogenetics and is the primary method by which clinicians detect chromosomal abnormalities in humans. A **karyotype** is the number and appearance of chromosomes, including their length, banding pattern, and centromere position. To obtain a view of an individual's karyotype, cytologists photograph the chromosomes and then cut and paste each chromosome into a chart, or **karyogram** (**Figure 7**).

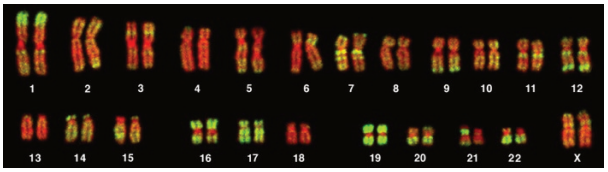


***Figure 7**: This karyogram shows the chromosomes of a female human immune cell during mitosis. (Credit: Andreas Bolzer, et al)*

By observing a karyogram, geneticists can actually visualize the chromosomal composition of an individual to confirm or predict genetic abnormalities in offspring even before birth.

Nondisjunctions, Duplications, and Deletions

Of all the chromosomal disorders, abnormalities in chromosome number are the most easily identifiable from a karyogram. Disorders of chromosome number include the duplication or loss of entire chromosomes, as well as changes in the number of complete sets of chromosomes. They are caused by **nondisjunction**, which occurs when pairs of homologous chromosomes or sister chromatids fail to separate during meiosis. The risk of nondisjunction increases with the age of the parents.

Nondisjunction can occur during either meiosis I or II, with different results (**Figure 7.8**). If homologous chromosomes fail to separate during meiosis I, the result is two gametes that lack that chromosome and two gametes with two copies of the chromosome. If sister chromatids fail to separate during meiosis II, the result is one gamete that lacks that chromosome, two normal gametes with one copy of the chromosome, and one gamete with two copies of the chromosome.

An individual with the appropriate number of chromosomes for their species is called **euploid**; in humans, euploidy corresponds to 22 pairs of **autosomes** and one pair of sex chromosomes. An individual with an error in chromosome number is described as **aneuploid**, a term that includes **monosomy** (loss of one chromosome) or **trisomy** (gain of an extraneous chromosome). Monosomic human zygotes missing any one copy of an autosome invariably fail to develop to birth because they have only one copy of essential genes. Most autosomal trisomies also fail to develop to birth; however, duplications of some of the smaller chromosomes (13, 15, 18, 21, or 22) can result in offspring that survive for several weeks to many years. Trisomic individuals suffer from a different type of genetic imbalance: an excess in gene dose. Cell functions are calibrated to the amount of gene product produced by two copies (doses) of each gene; adding a third copy (dose) disrupts this balance. The most common trisomy is that of chromosome 21, which leads to Down syndrome. Individuals with this inherited disorder have characteristic physical features and developmental delays in growth and cognition. The incidence of Down syndrome is correlated with maternal age, such that older women are more likely to give birth to children with Down syndrome (**Figure 7.9**).

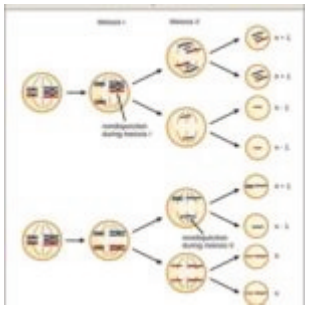


Figure 8: Following meiosis, each gamete has one copy of each chromosome. Nondisjunction occurs when homologous chromosomes (meiosis I) or sister chromatids (meiosis II) fail to separate during meiosis.

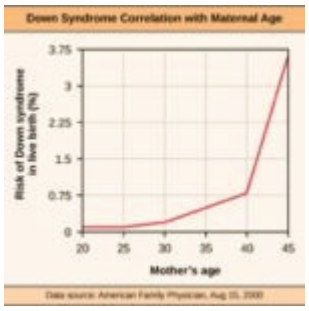


Figure 9: The incidence of having a fetus with trisomy 21 increases dramatically with maternal age.

Humans display dramatic deleterious effects with autosomal trisomies and monosomies. Therefore, it may seem counterintuitive that human females and males can function normally, despite carrying different numbers of the X chromosome. In part, this occurs because of a process called **X inactivation**. Early in development, when female mammalian embryos consist of just a few thousand cells, one X chromosome in each cell inactivates by condensing into a structure called a Barr body. The genes on the inactive X chromosome are not expressed. The particular X chromosome (maternally or paternally derived) that is inactivated in each cell is random, but once the inactivation occurs, all cells descended from that cell will have the same inactive X chromosome. By this process, females compensate for their double genetic dose of X chromosome.

In so-called "tortoiseshell" cats, X inactivation is observed as coat-color variegation (**Figure 10**). Females heterozygous for an X-linked coat color gene will express one of two different coat colors over different regions of their body, corresponding to whichever X chromosome is inactivated in the embryonic cell progenitor of that region. When you see a tortoiseshell cat, you will know that it has to be a female.

In an individual carrying an abnormal number of X chromosomes, cellular mechanisms will inactivate all but one X in each of her cells. As a result, X-chromosomal abnormalities are typically associated with mild mental and physical defects, as well as sterility. If the X chromosome is absent altogether, the individual will not develop.

Several errors in sex chromosome number have been characterized. Individuals with three X chromosomes, called triplo-X, appear female but express developmental delays and reduced fertility. The XXY chromosome complement, corresponding to one type of Klinefelter syndrome, corresponds to male individuals with small testes, enlarged breasts, and reduced body hair. The extra X chromosome undergoes inactivation to compensate for the excess genetic dosage. Turner

Figure 10: Embryonic inactivation of one of two different X chromosomes encoding different coat colors gives rise to the tortoiseshell phenotype in cats. (credit: Michael Bodega)

syndrome, characterized as an X0 chromosome complement (i.e., only a single sex chromosome), corresponds to a female individual with short stature, webbed skin in the neck region, hearing and cardiac impairments, and sterility.

An individual with more than the correct number of chromosome sets (two for diploid species) is called **polyploid**. For instance, fertilization of an abnormal diploid egg with a normal haploid sperm would yield a triploid zygote. Polyploid animals are extremely rare, with only a few examples among the flatworms, crustaceans, amphibians, fish, and lizards. Triploid animals are sterile because meiosis cannot proceed normally with an odd number of chromosome sets. In contrast, polyploidy is very common in the plant kingdom, and polyploid plants tend to be larger and more robust than euploids of their species.

Chromosome Structural Rearrangements

Cytologists have characterized numerous structural rearrangements in chromosomes, including partial duplications, deletions, inversions, and translocations. Duplications and deletions often produce offspring that survive but exhibit physical and mental abnormalities. Cri-du-chat (from the French for "cry of the cat") is a syndrome associated with nervous system abnormalities and identifiable physical features that results from a deletion of most of the small arm of chromosome 5 (**Figure 7.11**). Infants with this genotype emit a characteristic high-pitched cry upon which the disorder's name is based.

Chromosome inversions and translocations can be identified by observing cells during meiosis because homologous chromosomes with a rearrangement in one of the pair must contort to maintain appropriate gene alignment and pair effectively during prophase I.

A **chromosome inversion** is the detachment, 180° rotation, and reinsertion of part of a chromosome (**Figure 7.12**). Unless they disrupt a gene sequence, inversions only change the orientation of genes and are likely to have more mild effects than aneuploid errors.

A translocation occurs when a segment of a chromosome dissociates and reattaches to a different, nonhomologous chromosome. Translocations can be benign or have devastating effects, depending on how the positions of genes are altered with respect to regulatory sequences. Notably, specific translocations have been associated with several cancers and with schizophrenia. Reciprocal translocations result from the exchange of chromosome segments between two nonhomologous chromosomes such that there is no gain or loss of genetic information (**Figure 12**).

References

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:6-3MVU-j@4/Errors-in-Meiosis

Figure 11: *This individual with cri-du-chat syndrome is shown at various ages: (A) age two, (B) age four, (C) age nine, and (D) age 12. (credit: Paola Cerruti Mainardi)*

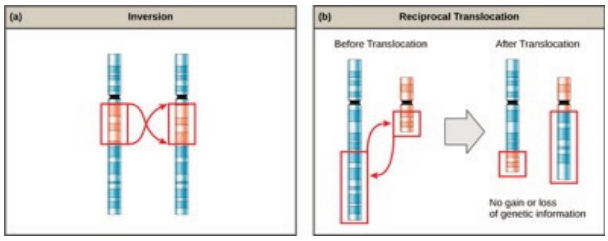


Figure 12: *An (a) inversion occurs when a chromosome segment breaks from the chromosome, reverses its orientation, and then reattaches in the original position. A (b) reciprocal translocation occurs between two nonhomologous chromosomes and does not cause any genetic information to be lost or duplicated. (credit: modification of work by National Human Genome Research Institute (USA)*

12. Patterns of Inheritance

Learning Objectives

By the end of this section, you will be able to:

- Describe the molecular basis of inheritance.
- Determine the outcome in crosses involving various types of inheritance.
- Present and decipher information about trait inheritance using a pedigree.

References

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/GFy_h8cu@10.57:O2lXSTlf@2/Introduction

12.1 Mendelian Genetics

Figure 1: Experimenting with thousands of garden peas, Mendel uncovered the fundamentals of genetics. (credit: modification of work by Jerry Kirkhart)

Genetics is the study of heredity. Johann Gregor Mendel set the framework for genetics long before chromosomes or genes had been identified, at a time when meiosis was not well understood. Mendel selected a simple biological system and conducted methodical, quantitative analyses using large sample sizes. Because of Mendel's work, the fundamental principles of heredity were revealed. We now know that genes, carried on chromosomes, are the basic functional units of heredity with the ability to be replicated, expressed, or mutated. Today, the postulates put forth by Mendel form the basis of classical, or Mendelian, genetics. Not all genes are transmitted from parents to offspring according to Mendelian genetics, but Mendel's experiments serve as an excellent starting point for thinking about inheritance.

Mendel's Experiments

***Figure 2**: Johann Gregor Mendel set the framework for the study of genetics.*

Johann Gregor Mendel (1822–1884) (**Figure 2**) was a lifelong learner, teacher, scientist, and man of faith. As a young adult, he joined the Augustinian Abbey of St. Thomas in Brno in what is now the Czech Republic. Supported by the monastery, he taught physics, botany, and natural science courses at the secondary and university levels. In 1856, he began a decade-long research pursuit involving inheritance patterns in honeybees and plants, ultimately settling on pea plants as his primary **model system** (a system with convenient characteristics that is used to study a specific biological phenomenon to gain understanding to be applied to other systems). In 1865, Mendel presented the results of his experiments with nearly 30,000 pea plants to the local natural history society. He demonstrated that traits are transmitted faithfully from parents to

offspring in specific patterns. In 1866, he published his work, *Experiments in Plant Hybridization,* in the proceedings of the Natural History Society of Brünn.

Mendel's work went virtually unnoticed by the scientific community, which incorrectly believed that the process of inheritance involved a blending of parental traits that produced an intermediate physical appearance in offspring. This hypothetical process appeared to be correct because of what we know now as continuous variation. **Continuous variation** is the range of small differences we see among individuals in a characteristic like human height. It does appear that offspring are a "blend" of their parents' traits when we look at characteristics that exhibit continuous variation. Mendel worked instead with traits that show **discontinuous variation**. Discontinuous variation is the variation seen among individuals when each individual shows one of two—or a very few—easily distinguishable traits, such as violet or white flowers. Mendel's choice of these kinds of traits allowed him to see experimentally that the traits were not blended in the offspring as would have been expected at the time, but that they were inherited as distinct traits. In 1868, Mendel became abbot of the monastery and exchanged his scientific pursuits for his pastoral duties. He was not recognized for his extraordinary scientific contributions during his lifetime; in fact, it was not until 1900 that his work was rediscovered, reproduced, and revitalized by scientists on the brink of discovering the chromosomal basis of heredity.

Mendel's Crosses

Mendel's seminal work was accomplished using the garden pea, *Pisum sativum*, to study inheritance.

This species naturally self-fertilizes, meaning that pollen encounters ova within the same flower. The flower petals remain sealed tightly until pollination is completed to prevent the pollination of other plants. The result is highly inbred, or "true-breeding," pea plants. These are plants that always produce offspring that look like the parent. By experimenting with true-breeding pea plants, Mendel avoided the appearance of unexpected traits in offspring that might occur if the plants were not true breeding. The garden pea also grows to maturity within one season, meaning that several generations could be evaluated over a relatively short time. Finally, large quantities of garden peas could be cultivated simultaneously, allowing Mendel to conclude that his results did not come about simply by chance.

Mendel performed **hybridizations**, which involve mating two true-breeding individuals that have different traits. In the pea, which is naturally self-pollinating, this is done by manually transferring pollen from the anther of a mature pea plant of one variety to the stigma of a separate mature pea plant of the second variety.

Plants used in first-generation crosses were called **P**, or parental generation, plants (**Figure 8.3**). Mendel collected the seeds produced by the P plants that resulted from each cross and grew them the following season. These offspring were called the **F1**, or the first filial (filial = daughter or son), generation. Once Mendel examined the characteristics in the F1 generation of plants, he allowed them to self-fertilize naturally. He then collected and grew the seeds from the F1 plants to produce the **F2**, or second filial, generation. Mendel's experiments extended beyond the F2 generation to the F3 generation, F4 generation, and so on, but it was the ratio of characteristics in the P, F1, and F2 generations that were the most intriguing and became the basis of Mendel's postulates.

References

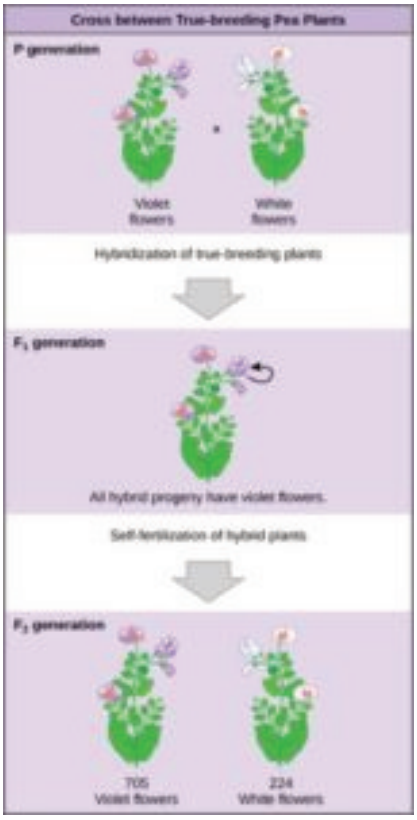


Figure 3: *Mendel's process for performing crosses included examining flower color.*

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/GFy_h8cu@10.57:x4Wo0QTK@7/Mendels-Experiments-and-the-La

12.2 Garden Pea Characteristics Revealed the Basics of Heredity

In his 1865 publication, Mendel reported the results of his crosses involving seven different characteristics, each with two contrasting traits. A **trait** is defined as a variation in the physical appearance of a heritable characteristic. The characteristics included plant height, seed texture, seed color, flower color, pea-pod size, pea-pod color, and flower position. For the characteristic of flower color, for example, the two contrasting traits were white versus violet. To fully examine each characteristic, Mendel generated large numbers of F1 and F2 plants and reported results from thousands of F2 plants.

What results did Mendel find in his crosses for flower color? First, Mendel confirmed that he was using plants that bred true for white or violet flower color. Irrespective of the number of generations that Mendel examined, all self-crossed offspring of parents with white flowers had white flowers, and all self-crossed offspring of parents with violet flowers had violet flowers. In addition, Mendel confirmed that, other than flower color, the pea plants were physically identical. This was an important check to make sure that the two varieties of pea plants only differed with respect to one trait, flower color.

Once these validations were complete, Mendel applied the pollen from a plant with violet flowers to the stigma of a plant with white flowers. After gathering and sowing the seeds that resulted from this cross, Mendel found that 100 percent of the F1 hybrid generation had violet flowers. Conventional wisdom at that time would have predicted the hybrid flowers to be pale violet or for hybrid plants to have equal numbers of white and violet flowers. In other words, the contrasting parental traits were expected to blend in the offspring. Instead, Mendel's results demonstrated that the white flower trait had completely disappeared in the F1 generation.

Importantly, Mendel did not stop his experimentation there. He allowed the F1 plants to self-fertilize and found that 705 plants in the F2 generation had violet flowers and 224 had white flowers. This was a ratio of 3.15 violet flowers to one white flower, or approximately 3:1. When Mendel transferred pollen from a plant with violet flowers to the stigma of a plant with white flowers and vice versa, he obtained approximately the same ratio irrespective of which parent—male or female—contributed which trait. This is called a **reciprocal cross**—a paired cross in which the respective traits of the male and female in one cross become the respective traits of the female and male in the other cross. For the other six characteristics that Mendel examined, the F1 and F2 generations behaved in the same way that they behaved for flower color. One of the two traits would disappear completely from the F1 generation, only to reappear in the F2 generation at a ratio of roughly 3:1 (**Figure 4**).

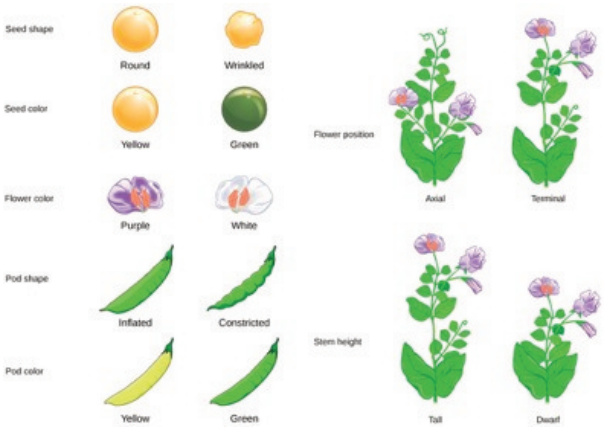


Figure 4: Mendel identified seven pea plant characteristics.

Upon compiling his results for many thousands of plants, Mendel concluded that the characteristics could be divided into expressed and latent traits. He called these dominant and recessive traits, respectively. **Dominant** traits are those that are inherited unchanged in a hybridization. **Recessive** traits become latent, or disappear in the offspring of a hybridization. The recessive trait does, however, reappear in the progeny of the hybrid offspring. An example of a dominant trait is the violet-colored flower trait. For this same characteristic (flower color), white-colored flowers are a recessive trait. The fact

that the recessive trait reappeared in the F2 generation meant that the traits remained separate (and were not blended) in the plants of the F1 generation. Mendel proposed that this was because the plants possessed two copies of the trait for the flower-color characteristic, and that each parent transmitted one of their two copies to their offspring, where they came together. Moreover, the physical observation of a dominant trait could mean that the genetic composition of the organism included two dominant versions of the characteristic, or that it included one dominant and one recessive version. Conversely, the observation of a recessive trait meant that the organism lacked any dominant versions of this characteristic.

Laws of Inheritance

The seven characteristics that Mendel evaluated in his pea plants were each expressed as one of two versions, or traits. Mendel deduced from his results that each individual had two discrete copies of the characteristic that are passed individually to offspring. We now call those two copies genes, which are carried on chromosomes. The reason we have two copies of each gene is that we inherit one from each parent. In fact, it is the chromosomes we inherit and the two copies of each gene are located on paired chromosomes. Recall that in meiosis these chromosomes are separated out into haploid gametes. This separation, or segregation, of the homologous chromosomes means also that only one of the copies of the gene gets moved into a gamete. The offspring are formed when that gamete unites with one from another parent and the two copies of each gene (and chromosome) are restored.

For cases in which a single gene controls a single characteristic, a diploid organism has two genetic copies that may or may not encode the same version of that characteristic. For example, one individual may carry a gene that determines white flower color and a gene that determines violet flower color. Gene variants that arise by mutation and exist at the same relative locations on homologous chromosomes are called **alleles**. Mendel examined the inheritance of genes with just two allele forms, but it is common to encounter more than two alleles for any given gene in a natural population.

References

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/GFy_h8cu@10.57:x4Wo0QTK@7/Mendels-Experiments-and-the-La

12.3 Phenotypes and Genotypes

Two alleles for a given gene in a diploid organism are expressed and interact to produce physical characteristics. The observable traits expressed by an organism are referred to as its **phenotype**. An organism's underlying genetic makeup, consisting of both the physically visible and the non-expressed alleles, is called its **genotype**. Mendel's hybridization experiments demonstrate the difference between phenotype and genotype. For example, the phenotypes that Mendel observed in his crosses between pea plants with differing traits are connected to the diploid genotypes of the plants in the P, F1, and F2 generations. We will use a second trait that Mendel investigated, seed color, as an example. Seed color is governed by a single gene with two alleles. The yellow-seed allele is dominant and the green-seed allele is recessive. When true-breeding plants were cross-fertilized, in which one parent had yellow seeds and one had green seeds, all of the F1 hybrid offspring had yellow seeds. That is, the hybrid offspring were phenotypically identical to the true-breeding parent with yellow seeds. However, we know that the allele donated by the parent with green seeds was not simply lost because it reappeared in some of the F2 offspring (**Figure 5**). Therefore, the F1 plants must have been genotypically different from the parent with yellow seeds.

The P plants that Mendel used in his experiments were each homozygous for the trait he was studying. Diploid organisms that are **homozygous** for a gene have two identical alleles, one on each of their homologous chromosomes. The genotype is often written as *YY* or *yy*, for which each letter represents one of the two alleles in the genotype. The dominant allele is capitalized and the recessive allele is lower case. The letter used for the gene (seed color in this case) is usually related to the dominant trait (yellow allele, in this case, or "*Y*"). Mendel's parental pea plants always bred true because both produced gametes carried the same allele. When P plants with contrasting traits were cross-fertilized, all of the offspring were **heterozygous** for the contrasting trait, meaning their genotype had different alleles for the gene being examined. For example, the F1 yellow plants that received a *Y* allele from their yellow parent and a *y* allele from their green parent had the genotype *Yy*.

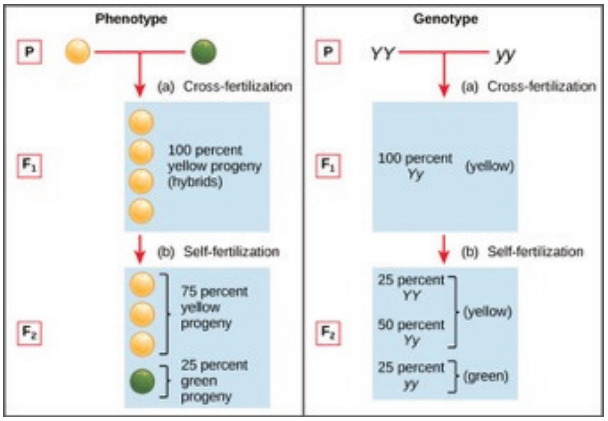


Figure 5: *Phenotypes are physical expressions of traits that are transmitted by alleles. Capital letters represent dominant alleles and lowercase letters represent recessive alleles. The phenotypic ratios are the ratios of visible characteristics. The genotypic ratios are the ratios of gene combinations in the offspring, and these are not always distinguishable in the phenotypes.*

Law of Dominance

Our discussion of homozygous and heterozygous organisms brings us to why the F1 heterozygous offspring were identical to one of the parents, rather than expressing both alleles. In all seven peaplant characteristics, one of the two contrasting alleles was dominant, and the other was recessive. Mendel called the dominant allele the expressed unit factor; the recessive allele was referred to as the latent unit factor. We now know that these so-called unit factors are actually genes on

homologous chromosomes. For a gene that is expressed in a dominant and recessive pattern, homozygous dominant and heterozygous organisms will look identical (that is, they will have different genotypes but the same phenotype), and the recessive allele will only be observed in homozygous recessive individuals (**Table 1**).

Table 1: Correspondence between Genotype and Phenotype for a Dominant Recessive Characteristic.

	Homozygous	**Heterozygous**	**Homozygous**
Genotype	*YY*	*Yy*	*yy*
Phenotype	yellow	yellow	green

Mendel's **law of dominance** states that in a heterozygote, one trait will conceal the presence of another trait for the same characteristic. For example, when crossing true-breeding violet-flowered plants with true-breeding white-flowered plants, all of the offspring were violet-flowered, even though they all had one allele for violet and one allele for white. Rather than both alleles contributing to a phenotype, the dominant allele will be expressed exclusively. The recessive allele will remain latent, but will be transmitted to offspring in the same manner as that by which the dominant allele is transmitted. The recessive trait will only be expressed by offspring that have two copies of this allele (**Figure 6**), and these offspring will breed true when self-crossed.

Figure 6: *The allele for albinism, expressed here in humans, is recessive. Both of this child's parents carried the recessive allele.*

References

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/GFy_h8cu@10.57:4qg08nt-@8/Characteristics-and-Traits

12.4 Monohybrid Cross and the Punnett Square

When fertilization occurs between two true-breeding parents that differ by only the characteristic being studied, the process is called a **monohybrid** cross, and the resulting offspring are called monohybrids. Mendel performed seven types of monohybrid crosses, each involving contrasting traits for different characteristics. Out of these crosses, all of the F1 offspring had the phenotype of one parent, and the F2 offspring had a 3:1 phenotypic ratio. On the basis of these results, Mendel postulated that each parent in the monohybrid cross contributed one of two paired unit factors to each offspring, and every possible combination of unit factors was equally likely.

The results of Mendel's research can be explained in terms of probabilities, which are mathematical measures of likelihood. The probability of an event is calculated by the number of times the event occurs divided by the total number of opportunities for the event to occur. A probability of one (100 percent) for some event indicates that it is guaranteed to occur, whereas a probability of zero (0 percent) indicates that it is guaranteed to not occur, and a probability of 0.5 (50 percent) means it has an equal chance of occurring or not occurring.

To demonstrate this with a monohybrid cross, consider the case of true-breeding pea plants with yellow versus green seeds. The dominant seed color is yellow; therefore, the parental genotypes were *YY* for the plants with yellow seeds and *yy* for the plants with green seeds. A **Punnett square**, devised by the British geneticist Reginald Punnett, is useful for determining probabilities because it is drawn to predict all possible outcomes of all possible random fertilization events and their expected frequencies. **Figure 7** shows a Punnett square for a cross between a plant with yellow peas and one with green peas. To prepare a Punnett square, all possible combinations of the parental alleles (the genotypes of the gametes) are listed along the top (for one parent) and side (for the other parent) of a grid. The combinations of egg and sperm gametes are then made in the boxes in the table on the basis of which alleles are combining. Each box then represents the diploid genotype of a zygote, or fertilized egg. Because each possibility is equally likely, genotypic ratios can be determined from a Punnett square. If the pattern of inheritance (dominant and recessive) is known, the phenotypic ratios can be inferred as well. For a monohybrid cross of two true-breeding parents, each parent contributes one type of allele. In this case, only one genotype is possible in the F1 offspring. All offspring are *Yy* and have yellow seeds.

When the F1 offspring are crossed with each other, each has an equal probability of contributing either a *Y* or a *y* to the F2 offspring. The result is a 1 in 4 (25 percent) probability of both parents contributing a *Y*, resulting in an offspring with a yellow phenotype; a 25 percent probability of parent A contributing a *Y* and parent B a *y*, resulting in offspring with a yellow phenotype; a 25 percent probability of parent A contributing a *y* and parent B a *Y*, also resulting in a yellow phenotype; and a (25 percent) probability of both parents contributing a *y*, resulting in a green phenotype. When counting all four possible outcomes, there is a 3 in 4 probability of offspring having the yellow phenotype and a 1 in 4 probability of offspring having the green phenotype. This explains why the results of Mendel's F2 generation occurred in a 3:1 phenotypic ratio. Using large numbers of crosses, Mendel was able to calculate probabilities, found that they fit the model of inheritance, and use these to predict the outcomes of other crosses.

References

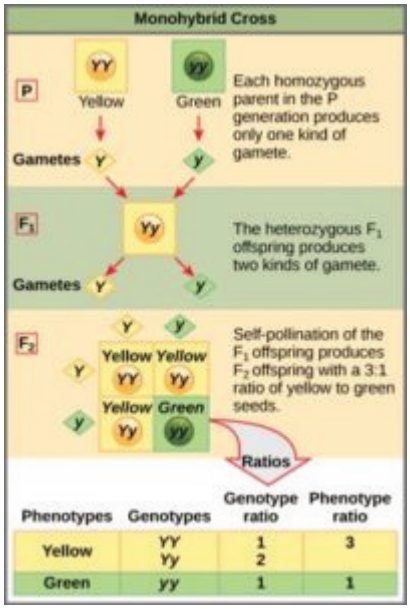


***Figure 7**: This Punnett square shows the cross between plants with yellow seeds and green seeds. The cross between the true-breeding P plants produces F1 heterozygotes that can be self-fertilized. The self-cross of the F1 generation can be analyzed with a Punnett square to predict the genotypes of the F2 generation. Given an inheritance pattern of dominant-recessive, the genotypic and phenotypic ratios can then be determined.*

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/GFy_h8cu@10.57:4qg08nt-@8/Characteristics-and-Traits

12.5 Laws of Inheritance

Law of Segregation

Observing that true-breeding pea plants with contrasting traits gave rise to F1 generations that all expressed the dominant trait and F2 generations that expressed the dominant and recessive traits in a 3:1 ratio, Mendel proposed the **law of segregation**. This law states that paired unit factors (genes) must segregate equally into gametes such that offspring have an equal likelihood of inheriting either factor. For the F2 generation of a monohybrid cross, the following three possible combinations of genotypes result: homozygous dominant, heterozygous, or homozygous recessive. Because heterozygotes could arise from two different pathways (receiving one dominant and one recessive allele from either parent), and because heterozygotes and homozygous dominant individuals are phenotypically identical, the law supports Mendel's observed 3:1 phenotypic ratio. The equal segregation of alleles is the reason we can apply the Punnett square to accurately predict the offspring of parents with known genotypes. The physical basis of Mendel's law of segregation is the first division of meiosis in which the homologous chromosomes with their different versions of each gene are segregated into daughter nuclei. This process was not understood by the scientific community during Mendel's lifetime (**Figure 8**).

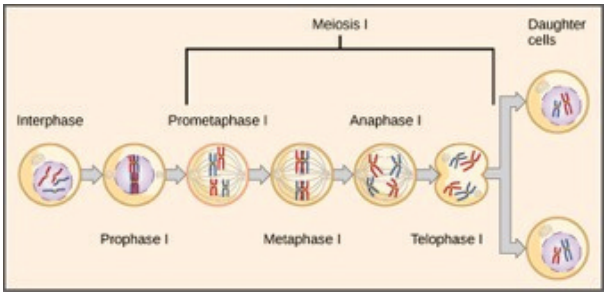


***Figure 8**: The first division in meiosis is shown.*

Test Cross

Beyond predicting the offspring of a cross between known homozygous or heterozygous parents, Mendel also developed a way to determine whether an organism that expressed a dominant trait was a heterozygote or a homozygote. Called the **test cross**, this technique is still used by plant and animal breeders. In a test cross, the dominant-expressing organism is crossed with an organism that is homozygous recessive for the same characteristic. If the dominant-expressing organism is a homozygote, then all F1 offspring will be heterozygotes expressing the dominant trait (**Figure 9**). Alternatively, if the dominant-expressing organism is a heterozygote, the F1 offspring will exhibit a 1:1 ratio of heterozygotes and recessive homozygotes (**Figure 9**). The test cross further validates Mendel's postulate that pairs of unit factors segregate equally.

Law of Independent Assortment

Mendel's **law of independent assortment** states that genes do not influence each other with regard to the sorting of alleles into gametes, and every possible combination of alleles for every gene is equally likely to occur. Independent assortment of genes can be illustrated by the **dihybrid** cross, a cross between two true-breeding parents that express different traits for two characteristics. Consider the characteristics of seed color and seed texture for two pea plants, one that has wrinkled, green seeds (*rryy*) and another that has round, yellow seeds (*RRYY*). Because each parent is homozygous, the law of segregation indicates that the gametes for the wrinkled–green plant all are *ry*, and the gametes for the round–yellow plant are all *RY*. Therefore, the F1 generation of offspring all are *RrYy* (**Figure 10**).

The gametes produced by the F1 individuals must have one allele from each of the two genes. For example, a gamete could get an *R* allele for the seed shape gene and either a *Y* or a *y* allele for the seed color gene. It cannot get both an *R* and an *r* allele; each gamete can have only one allele per gene. The law of independent assortment states that a gamete into which

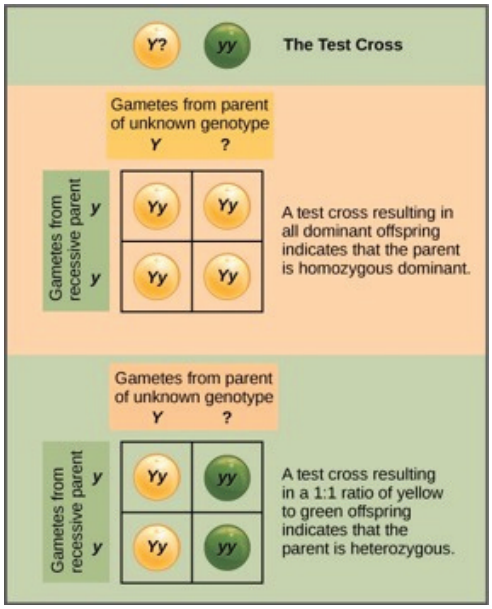


Figure 9: *A test cross can be performed to determine whether an organism expressing a dominant trait is a homozygote or a heterozygote.*

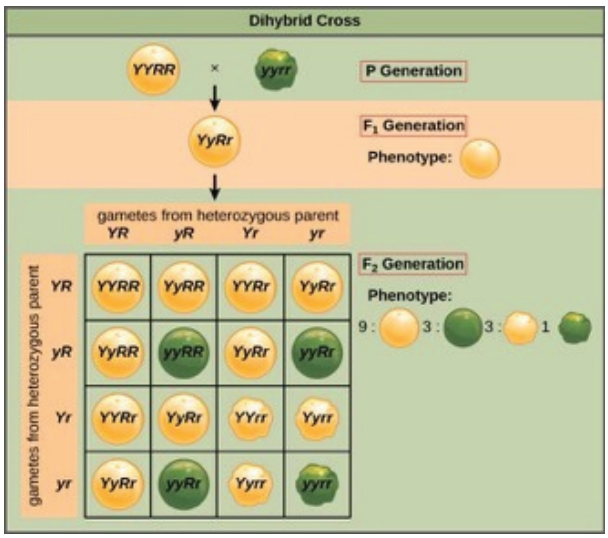


Figure 10: *A dihybrid cross in pea plants involves the genes for seed color and texture. The P cross produces F1 offspring that are all heterozygous for both characteristics. The resulting 9:3:3:1 F2 phenotypic ratio is obtained using a Punnett square.*

an *r* allele is sorted would be equally likely to contain either a *Y* or a *y* allele. Thus, there are four equally likely gametes that can be formed when the *RrYy* heterozygote is self-crossed, as follows: *RY*, *rY*, *Ry*, and *ry*. Arranging these gametes along the top and left of a 4 × 4 Punnett square (**Figure 10**) gives us 16 equally likely genotypic combinations. From these genotypes, we find a phenotypic ratio of 9 round–yellow:3 round–green:3 wrinkled–yellow:1 wrinkled–green (**Figure 10**). These are the offspring ratios we would expect, assuming we performed the crosses with a large enough sample size.

The physical basis for the law of independent assortment also lies in meiosis I, in which the different homologous pairs line up in random orientations. Each gamete can contain any combination of paternal and maternal chromosomes (and therefore the genes on them) because the orientation of tetrads on the metaphase plane is random (**Figure 11**).

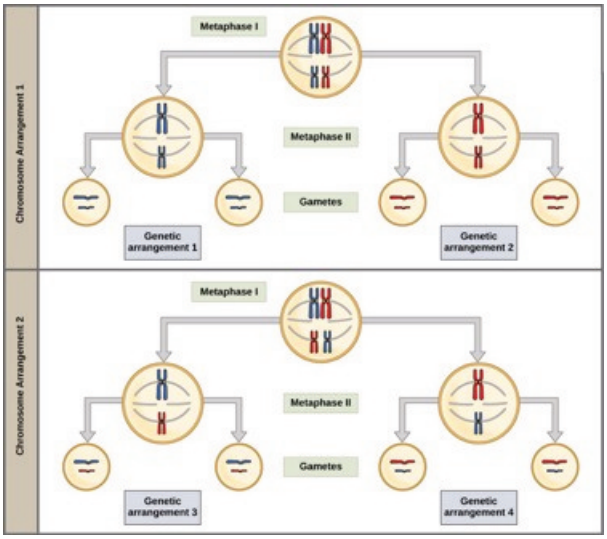


Figure 11: *The random segregation into daughter nuclei that happens during the first division in meiosis can lead to a variety of possible genetic arrangements.*

References

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/GFy_h8cu@10.57:cUeevuaC@3/Laws-of-Inheritance

12.6 Extensions of the Laws of Inheritance

Mendel studied traits with only one mode of inheritance in pea plants. The inheritance of the traits he studied all followed the relatively simple pattern of dominant and recessive alleles for a single characteristic. There are several important modes of inheritance, discovered after Mendel's work, that do not follow the dominant and recessive, single-gene model.

Alternatives to Dominance and Recessiveness

Mendel's experiments with pea plants suggested that: 1) two types of "units" or alleles exist for every gene; 2) alleles maintain their integrity in each generation (no blending); and 3) in the presence of the dominant allele, the recessive allele is hidden, with no contribution to the phenotype. Therefore, recessive alleles can be "carried" and not expressed by individuals. Such heterozygous individuals are sometimes referred to as "carriers." Since then, genetic studies in other organisms have shown that much more complexity exists, but that the fundamental principles of Mendelian genetics still hold true. In the sections to follow, we consider some of the extensions of Mendelism.

Incomplete Dominance

Mendel's results, demonstrating that traits are inherited as dominant and recessive pairs, contradicted the view at that time that offspring exhibited a blend of their parents' traits. However, the heterozygote phenotype occasionally does appear to be intermediate between the two parents. For example, in the snapdragon, *Antirrhinum majus* (**Figure 12**), a cross between a homozygous parent with white flowers (*CWCW*) and a homozygous parent with red flowers (*CRCR*) will produce offspring with pink flowers (*CRCW*). (Note that different genotypic abbreviations are used for Mendelian extensions to distinguish these patterns from simple dominance and recessiveness.) This pattern of inheritance is described as **incomplete dominance**, meaning that one of the alleles appears in the phenotype in the heterozygote, but not to the exclusion of the other, which can also be seen. The allele for red flowers is incompletely dominant over the allele for white flowers. However, the results of a heterozygote self-cross can still be predicted, just as with Mendelian dominant and recessive crosses. In this case, the genotypic ratio would be 1 *CRCR*:2 *CRCW*:1 *CWCW*, and the phenotypic ratio would be 1:2:1 for red:pink:white. The basis for the intermediate color in the heterozygote is simply that the pigment produced by the red allele (anthocyanin) is diluted in the heterozygote and therefore appears pink because of the white background of the flower petals.

Figure 12: *These pink flowers of a heterozygote snapdragon result from incomplete dominance. (credit: "storebukkebruse"/Flickr)*

Codominance

A variation on incomplete dominance is **codominance**, in which both alleles for the same characteristic are simultaneously expressed in the heterozygote. An example of codominance occurs in the ABO blood groups of humans. The A and B alleles are expressed in the form of A or B molecules present on the surface of red blood cells. Homozygotes (*IAIA* and *IBIB*)

express either the A or the B phenotype, and heterozygotes (*IAIB*) express both phenotypes equally. The *IAIB* individual has blood type AB. In a selfcross between heterozygotes expressing a codominant trait, the three possible offspring genotypes are phenotypically distinct. However, the 1:2:1 genotypic ratio characteristic of a Mendelian monohybrid cross still applies (**Figure 13**).

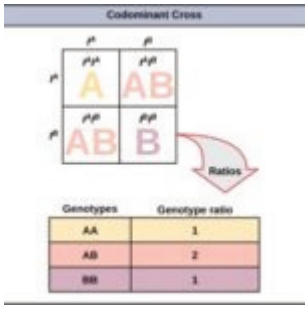


***Figure 13**: This Punnet square shows an AB/AB blood type cross*

References

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/GFy_h8cu@10.57:4qg08nt-@8/Characteristics-and-Traits

12.7 Multiple Alleles

Mendel implied that only two alleles, one dominant and one recessive, could exist for a given gene. We now know that this is an oversimplification. Although individual humans (and all diploid organisms) can only have two alleles for a given gene, multiple alleles may exist at the population level, such that many combinations of two alleles are observed. Note that when many alleles exist for the same gene, the convention is to denote the most common phenotype or genotype in the natural population as the **wild type** (often abbreviated "+"). All other phenotypes or genotypes are considered variants (mutants) of this typical form, meaning they deviate from the wild type. The variant may be recessive or dominant to the wild-type allele.

An example of multiple alleles is the ABO blood-type system in humans. In this case, there are three alleles circulating in the population. The *IA* allele codes for A molecules on the red blood cells, the *IB* allele codes for B molecules on the surface of red blood cells, and the *i* allele codes for no molecules on the red blood cells. In this case, the *IA* and *IB* alleles are codominant with each other and are both dominant over the *i* allele. Although there are three alleles present in a population, each individual only gets two of the alleles from their parents. This produces the genotypes and phenotypes shown in **Figure 14**. Notice that instead of three genotypes, there are six different genotypes when there are three alleles. The number of possible phenotypes depends on the dominance relationships between the three alleles.

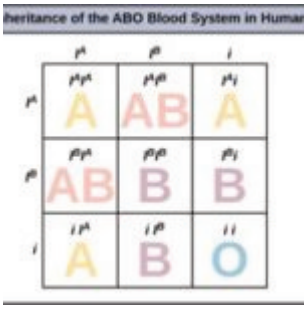


***Figure 14**: Inheritance of the ABO blood system in humans is shown.*

Multiple Alleles Confer Drug Resistance in the Malaria Parasite

Malaria is a parasitic disease in humans that is transmitted by infected female mosquitoes, including *Anopheles gambiae*, and is characteried by cyclic high fevers, chills, flu-like symptoms, and severe anemia. *Plasmodium falciparum* and *P. vivax* are the most common causative agents of malaria, and *P. falciparum* is the most deadly. When promptly and correctly treated, *P. falciparum* malaria has a mortality rate of 0.1 percent. However, in some parts of the world, the parasite has evolved resistance to commonly used malaria treatments, so the most effective malarial treatments can vary by geographic region.

References

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/GFy_h8cu@10.57:4qg08nt-@8/Characteristics-and-Traits

12.8 Sex-Linked Traits

In humans, as well as in many other animals and some plants, the sex of the individual is determined by sex chromosomes—one pair of non-homologous chromosomes. Until now, we have only considered inheritance patterns among non-sex chromosomes, or autosomes. In addition to 22 homologous pairs of autosomes, human females have a homologous pair of X chromosomes, whereas human males have an XY chromosome pair. Although the Y chromosome contains a small region of similarity to the X chromosome so that they can pair during meiosis, the Y chromosome is much shorter and contains fewer genes. When a gene being examined is present on the X, but not the Y, chromosome, it is **X-linked**.

Eye color in *Drosophila*, the common fruit fly, was the first X-linked trait to be identified. Thomas Hunt Morgan mapped this trait to the X chromosome in 1910. Like humans, *Drosophila* males have an XY chromosome pair, and females are XX. In flies the wild-type eye color is red (X*W*) and is dominant to white eye color (X*w*) (**Figure 15**). Because of the location of the eye-color gene, reciprocal crosses do not produce the same offspring ratios. Males are said to be **hemizygous**, in that they have only one allele for any X-linked characteristic. Hemizygosity makes descriptions of dominance and recessiveness irrelevant for XY males. *Drosophila* males lack the white gene on the Y chromosome; that is, their genotype can only be XWY or X*w*Y. In contrast, females have two allele copies of this gene and can be X*W*X*W*, X*W*X*w*, or X*w*X*w*.

Figure 15: *In Drosophila, the gene for eye color is located on the X chromosome. Red eye color is wild-type and is dominant to white eye color.*

In an X-linked cross, the genotypes of F1 and F2 offspring depend on whether the recessive trait was expressed by the male or the female in the P generation. With respect to *Drosophila* eye color, when the P male expresses the white-eye phenotype and the female is homozygously red-eyed, all members of the F1 generation exhibit red eyes (**Figure 16**). The F1 females are heterozygous (X*W*X*w*), and the males are all X*W*Y, having received their X chromosome from the homozygous dominant P female and their Y chromosome from the P male. A subsequent cross between the X*W*X*w* female and the X*W*Y male would produce only red-eyed females (with X*W*X*W* or X*W*X*w* genotypes) and both red- and white-eyed males (with X*W*Y or XwY genotypes). Now, consider a cross between a homozygous white-eyed female and a male with red eyes. The F1 generation would exhibit only heterozygous red-eyed females (XWXw) and only white-eyed males (XwY). Half of the F2 females would be red-eyed (XWXw) and half would be white-eyed (XwXw). Similarly, half of the F2 males would be red-eyed (XWY) and half would be whiteeyed (XwY).

Discoveries in fruit fly genetics can be applied to human genetics. When a female parent is homozygous for a recessive X-linked trait, she will pass the trait on to 100 percent of her male offspring, because the males will receive the Y chromosome from the male parent. In humans, the alleles for certain conditions (some color-blindness, hemophilia, and muscular dystrophy) are X-linked. Females who are heterozygous for these diseases are said to be carriers and may not

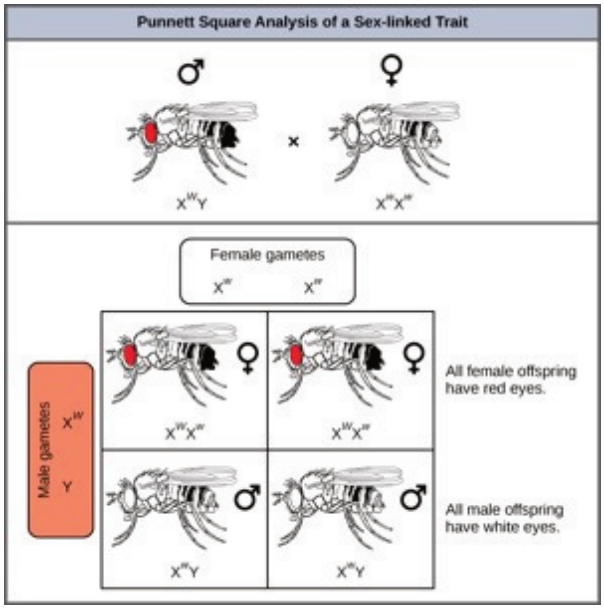


Figure 16: *Crosses involving sex-linked traits often give rise to different phenotypes for the different sexes of offspring, as in the case for this cross involving red and white eye color in Drosophila. In the diagram, w is the white-eye mutant allele and W is the wild-type, red-eye allele.*

exhibit any phenotypic effects. These females will pass the disease to half of their sons and will pass carrier status to half of their daughters; therefore, X-linked traits appear more frequently in males than females.

In some groups of organisms with sex chromosomes, the sex with the non-homologous sex chromosomes is the female rather than the male. This is the case for all birds. In this case, sex-linked traits will be more likely to appear in the female, in whom they are hemizygous.

References

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/GFy_h8cu@10.57:4qg08nt-@8/Characteristics-and-Traits

12.9 Linked Genes Violate the Law of Independent Assortment

Although all of Mendel's pea plant characteristics behaved according to the law of independent assortment, we now know that some allele combinations are not inherited independently of each other. Genes that are located on separate, non-homologous chromosomes will always sort independently. However, each chromosome contains hundreds or thousands of genes, organized linearly on chromosomes like beads on a string. The segregation of alleles into gametes can be influenced by **linkage**, in which genes that are located physically close to each other on the same chromosome are more likely to be inherited as a pair. However, because of the process of recombination, or "crossover," it is possible for two genes on the same chromosome to behave independently, or as if they are not linked. To understand this, let us consider the biological basis of gene linkage and recombination.

Homologous chromosomes possess the same genes in the same order, though the specific alleles of the gene can be different on each of the two chromosomes. Recall that during interphase and prophase I of meiosis, homologous chromosomes first replicate and then synapse, with like genes on the homologs aligning with each other. At this stage, segments of homologous chromosomes exchange linear segments of genetic material (**Figure 17**). This process is called **recombination**, or crossover, and it is a common genetic process. Because the genes are aligned during recombination, the gene order is not altered. Instead, the result of recombination is that maternal and paternal alleles are combined onto the same chromosome. Across a given chromosome, several recombination events may occur, causing extensive shuffling of alleles.

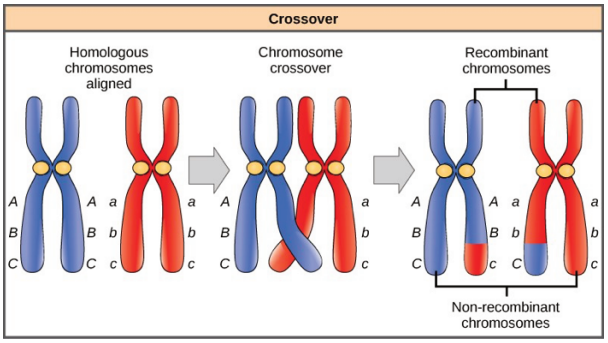


***Figure 17**: The process of crossover, or recombination, occurs when two homologous chromosomes align and exchange a segment of genetic material.*

When two genes are located on the same chromosome, they are considered linked, and their alleles tend to be transmitted through meiosis together. To exemplify this, imagine a dihybrid cross involving flower color and plant height in which the genes are next to each other on the chromosome. If one homologous chromosome has alleles for tall plants and red flowers, and the other chromosome has genes for short plants and yellow flowers, then when the gametes are formed, the tall and red alleles will tend to go together into a gamete and the short and yellow alleles will go into other gametes. These are called the parental genotypes because they have been inherited intact from the parents of the individual producing gametes. But unlike if the genes were on different chromosomes, there will be no gametes with tall and yellow alleles and no gametes with short and red alleles. If you create a Punnett square with these gametes, you will see that the classical Mendelian prediction of a 9:3:3:1 outcome of a dihybrid cross would not apply. As the distance between two genes increases, the probability of one or more crossovers between them increases and the genes behave more like they are on separate chromosomes. Geneticists have used the proportion of recombinant gametes (the ones not like the parents) as a measure of how far apart genes are on a chromosome. Using this information, they have constructed linkage maps of genes on chromosomes for well-studied organisms, including humans.

Mendel's seminal publication makes no mention of linkage, and many researchers have questioned whether he encountered linkage but chose not to publish those crosses out of concern that they would invalidate his independent assortment postulate. The garden pea has seven chromosomes, and some have suggested that his choice of seven

characteristics was not a coincidence. However, even if the genes he examined were not located on separate chromosomes, it is possible that he simply did not observe linkage because of the extensive shuffling effects of recombination.

References

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/GFy_h8cu@10.57:cUeevuaC@3/Laws-of-Inheritance

12.10 Epistasis

Mendel's studies in pea plants implied that the sum of an individual's phenotype was controlled by genes (or as he called them, unit factors), such that every characteristic was distinctly and completely controlled by a single gene. In fact, single observable characteristics are almost always under the influence of multiple genes (each with two or more alleles) acting in unison. For example, at least eight genes contribute to eye color in humans.

In some cases, several genes can contribute to aspects of a common phenotype without their gene products ever directly interacting. In the case of organ development, for instance, genes may be expressed sequentially, with each gene adding to the complexity and specificity of the organ. Genes may function in complementary or synergistic fashions, such that two or more genes expressed simultaneously affect a phenotype. An apparent example of this occurs with human skin color, which appears to involve the action of at least three (and probably more) genes. Cases in which inheritance for a characteristic like skin color or human height depend on the combined effects of numerous genes are called polygenic inheritance.

Genes may also oppose each other, with one gene suppressing the expression of another. In **epistasis**, the interaction between genes is antagonistic, such that one gene masks or interferes with the expression of another. "Epistasis" is a word composed of Greek roots meaning "standing upon." The alleles that are being masked or silenced are said to be hypostatic to the epistatic alleles that are doing the masking. Often the biochemical basis of epistasis is a gene pathway in which expression of one gene is dependent on the function of a gene that precedes or follows it in the pathway.

An example of epistasis is pigmentation in mice. The wild-type coat color, agouti (AA) is dominant to solid-colored fur (aa). However, a separate gene C, when present as the recessive homozygote (cc), negates any expression of pigment from the A gene and results in an albino mouse (**Figure 18**). Therefore, the genotypes *AAcc*, *Aacc*, and *aacc* all produce the same albino phenotype. A cross between heterozygotes for both genes (*AaCc* x *AaCc*) would generate offspring with a phenotypic ratio of 9 agouti:3 black:4 albino (**Figure 18**). In this case, the *C* gene is epistatic to the *A* gene.

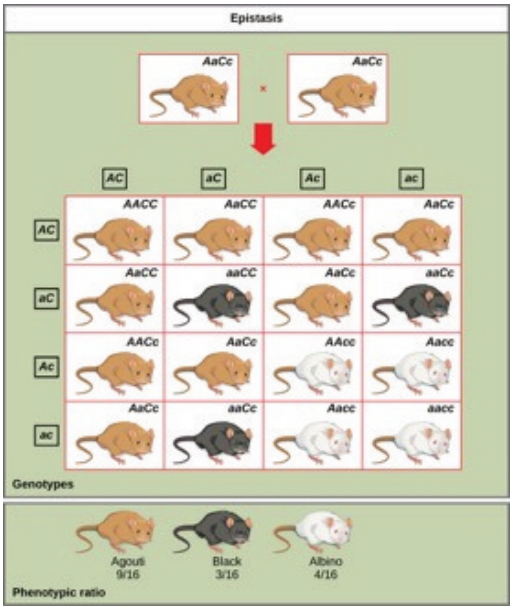


Figure 18: *In this example of epistasis, one gene (C) masks the expression of another (A) for coat color. When the C allele is present, coat color is expressed; when it is absent (cc), no coat color is expressed. Coat color depends on the A gene, which shows dominance, with the recessive homozygote showing a different phenotype than the heterozygote or dominant homozygote.*

References

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/GFy_h8cu@10.57:cUeevuaC@3/Laws-of-Inheritance

Genetics: Dog Coat Color

Learning Objectives

By the end of this section, you will be able to:

- Describe the molecular basis of inheritance.
- Determine the outcome in crosses involving complete dominance.
- Present and decipher information about inheritance using a pedigree.

Figure 1: *Experimenting with thousands of garden peas, Johann Gregor Mendel uncovered the fundamentals of genetics. (credit: modification of work by Jerry Kirkhart)*

Remember that a trait is an aspect of the physical appearance of an organism that can vary. Organisms get their traits from proteins; proteins are produced using the information found in the organism's DNA. Variation in the DNA between different organisms causes the production of proteins that contain differing orders of amino acids. These proteins can have different shapes and therefore different functions. When proteins function differently, this leads to differences in traits.

Recall that diploid organisms have two copies of each chromosome: a pair of homologous chromosomes. The reason that they have two copies is because they inherited one copy of each chromosome from each parent. Each parent donates one haploid gamete (egg or sperm) to the reproductive process. A haploid gamete contains one copy of each chromosome because during meiosis the number of chromosomes is cut in half: the DNA is copied once and then divided twice. This separation of the homologous chromosomes means that only one of the copies of the gene gets moved into a gamete. The offspring are formed when that gamete unites with one from another parent and the two copies of each gene (and chromosome) are restored.

A diploid organism has two copies of a given gene. The two copies may or may not encode the same version of that characteristic. For example, one individual pea plant (such as those studied by Mendel) would have two copies of the gene that controls flower color. That individual could carry one version of the gene that leads to white flower color and a second different version of that same gene that leads to violet flower color. The interaction between these two different versions of the same gene will lead to the visible flower color in the pea plant. Gene variations that arise by mutation and exist at the

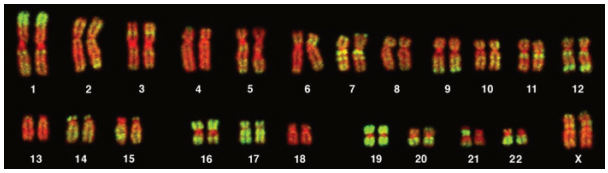


Figure 2: *A karyogram is a picture of all the chromosomes in a cell, organized into homologous pairs. This is a human karyogram which shows the 46 chromosomes present in diploid human somatic cells.*

same relative locations on homologous chromosomes are called **alleles**. Mendel examined the inheritance of genes with just two allele forms, but it is common to encounter more than two alleles for many genes in a natural population.

Each individual (assuming it is a diploid organism) will have two alleles for a specific gene: one from each of its two parents. These two alleles are expressed and interact to produce physical characteristics. The observable traits expressed by an organism are referred to as its **phenotype**. An organism's underlying genetic makeup, consisting of both the physically visible and the non-expressed alleles, is called its **genotype**.

Diploid organisms that are **homozygous** for a gene have two identical alleles, one on each of their homologous chromosomes. If the organism has two different alleles, this is referred to as **heterozygous.**

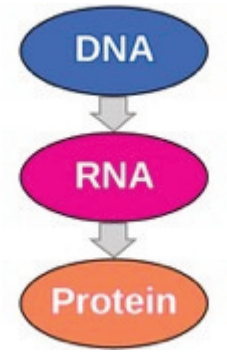
This chapter will address a simple type of inheritance: complete dominance. In this type of inheritance, there are two alleles: dominant and recessive. A **dominant** allele will completely cover up a recessive allele. This means that if one dominant allele is present, the organism will have the trait conferred by that allele. In order for the recessive phenotype to be seen, the organism must have two **recessive** alleles. Just because an allele is dominant does not automatically make it better than a recessive trait. It also does not make it more common than the recessive trait. All it means for an allele to be dominant is that it is able to cover up the recessive allele.

We typically abbreviate the genotype of an organism by using single letters. The letter chosen is often the first letter of the dominant trait. A homozygous dominant genotype would be written AA, a heterozygous genotype as Aa, and a homozygous recessive genotype as aa.

Introduction to Genetics

"Genetics" is the study of how traits are inherited.A **trait** is defined as a variation in the physical appearance of a heritable characteristic. It seeks to understand how traits are passed from generation to generation. Before you start learning about the details of inheritance, let's review some topics that are important in order to understand genetics.

Recall that genes are segments of DNA that are typically several hundred or thousand bases long. Each gene directs the production of a protein through the process of protein synthesis: DNA gets transcribed to produce an mRNA; mRNA provides to code for a ribosome to produce a chain of amino acids. Read this section of the book if you need to review this topic: How do genes direct the production of proteins?

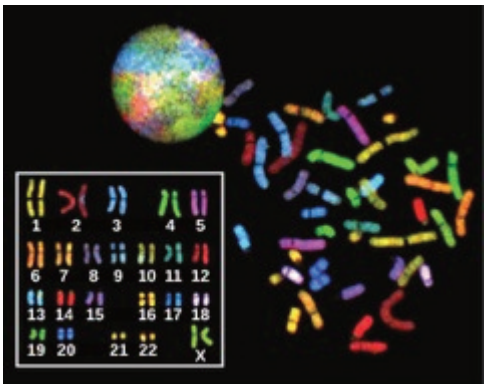


The Central Dogma – DNA is used to make RNA is used to make protein

Recall that eukaryotic genes are found on chromosomes and that each eukaryotic chromosome typically contains hundreds or thousands of genes. In most eukaryotes, including humans and other animals, each cell contains two copies of each chromosome. The reason we have two copies of each gene is that we inherit one from each parent.

In contrast to eukaryotes, prokaryotes have one circular chromosome. This means they have one copy of each gene.

Read this section of the book if you need to review this topic: How DNA is arranged in the cell



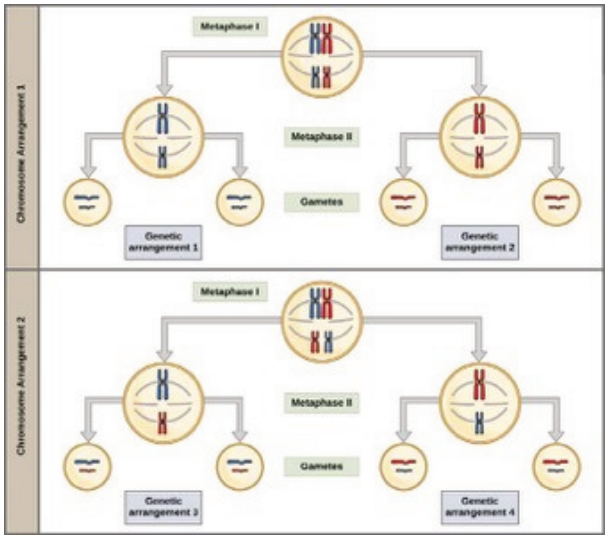
There are 23 pairs of chromosomes in a female human body cell. These chromosomes are viewed within the nucleus (top), removed from a cell during cell division (right), and arranged according to length (left) in an arrangement called a karyotype. In this image, the chromosomes were exposed to fluorescent stains to distinguish them. (credit: "718 Bot"/Wikimedia Commons, National Human Genome Research)

Chromosomes are inherited by the offspring from the parents via the egg or sperm. Inside one egg or one sperm is one

copy of each chronometer (so 23 total in humans). When an egg is fertilized by a sperm, the resulting **zygote** (fertilized egg) will contain two copies of each chromosome, just like each of its parents.

Meiosis is the process that produces eggs and sperm. Eggs and sperm are also known as **gametes**. During meiosis, one copy of each paired chromosome is moved into the gamete. Cells with one copy of each chromosome are known as "**haploid**". This separation, or segregation, of the **homologous** (paired) chromosomes means also that only one of the copies of the gene gets moved into a gamete.

The offspring are formed when that gamete unites with one from another parent and the two copies of each gene (and chromosome) are restored. Read this section of the book if you need more information on this topic: Overview of Meiosis



During meiosis, the DNA is copied once, then the cell divides twice. This produces cells with half as much genetic information as the original cell (1 copy of each chromosome). These cells become the sex cells (eggs or sperm). When two sex cells unite during fertilization, the original number of chromosomes (2 copies of each one) is restored.

The offspring will receive two copies of each gene (one from each parent), but the copies are not necessarily identical. You already knew this – you don't get identical information from your mother and your father because they have different DNA (which gives them different traits). The different versions of one specific gene are known as **alleles.** As you learn about genetics, you will learn about how the information from both alleles of a specific gene interact to give an individual their trait. The genetic information that an individual has is called their **genotype.** The genotype of an individual produces the individual's **phenotype,** or physical traits.

References

OpenStax, Concepts of Biology. OpenStax CNX. May 18, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:8v2Xzdco@5/The-Structure-of-DNA

Pedigrees and Punnett Squares

Pedigrees

Inheritance of a trait through generations can be shown visually using a pedigree, such as is pictured in **Figure 3**. Square shapes represent males; circles represent females. Filled-in shapes are individuals that have whatever trait is being shown in the pedigree. Two individuals connected together with a horizontal line between them are the parents of the individuals that are connected by vertical lines below them. Siblings are typically shown in birth order with the oldest sibling to the left.

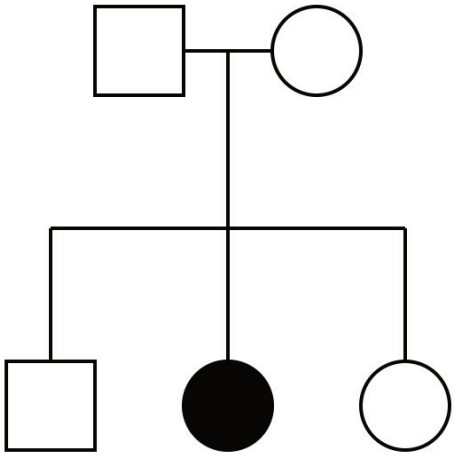


Figure 3: *A simple pedigree. In this pedigree, the parents (at the top) have produced three children: a male and two females. The first female has the condition being shown in the pedigree.*

Punnett Squares

As discussed above, diploid individuals have two copies of each chromosome: one from their male parent, one from their female parent. This means they have two copies of each gene. They can have two of the same alleles (homozygous) or two different alleles (heterozygous). Regardless of their genotype, they will pass one copy of each chromosome to their offspring. This is because meiosis produces haploid gametes that contain one copy of each chromosome. Since genes are present on chromosomes, this means they will pass one copy of each gene to their offspring. That means that an offspring inherits one allele of each gene from each of its two parents. This is illustrated in Figure 4.

An easy, organized way of illustrating the offspring that can result from two specific parents is to use a Punnett square. The gametes that can be generated by each parent are represented above the rows and next to the columns of the square. Each gamete is haploid for the "A gene", meaning it only contains one copy of that gene. In the Punnett square seen in Figure 5, haploid eggs are above each column and haploid sperm are next to each row. When a haploid sperm and a haploid egg (each with 1 copy of the "A gene") combine during the process of fertilization, a diploid offspring (with 2 copies of the A gene) is the result.

A Punnett square shows the probability of an offspring with a given genotype resulting from a cross. It does not show actual offspring. For example, the Punnett square in Figure 5 shows that there is a 25% chance that a homozygous recessive offspring will result from the cross Aa x Aa. It does *not* mean that these parents must have 4 offspring and that they will

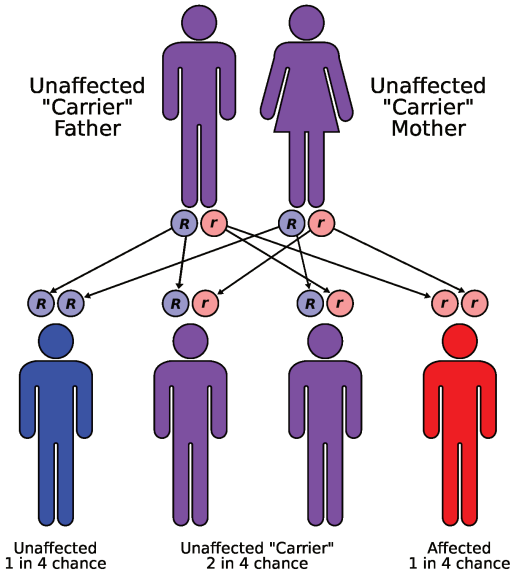


Figure 4: *Two parents who are heterozygous each pass one chromosome / gene / allele to each offspring. Each resulting offspring has two of each chromosome / gene. The individual can have two of the same or two different alleles.*

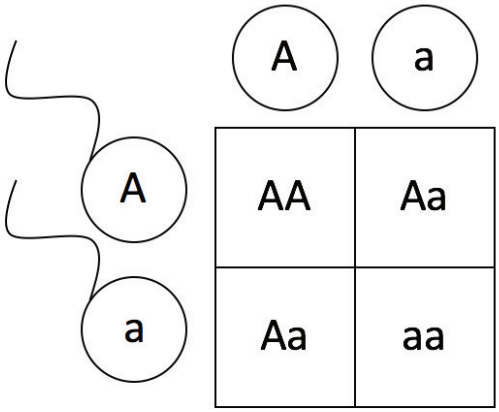


Figure 5: *A Punnett square showing a cross between two individuals who are both heterozygous for A.*

have the ratio 1 AA : 2 Aa : 1 aa. It's just like flipping a coin: you expect 50% heads, but you wouldn't be too surprised to see 7 heads out of 10 coin flips. Additionally, the probability does not change for successive offspring. The probability that the first offspring will have the genotype "aa" is 25% and the probability of the second offspring having the genotype "aa" is still 25%. Again, it's just like flipping a coin: if you flip heads the first time, that doesn't change the probability of getting heads on the next flip.

Organisms don't just inherit one trait at a time, though. They inherit all their traits at once. Sometimes, we want to determine the probability of an individual inheriting two different traits. The easiest way to do this is to determine the probability of the individual inheriting each trait separately, then multiply those probabilities together. An example of this can be seen in **Figure 6**.

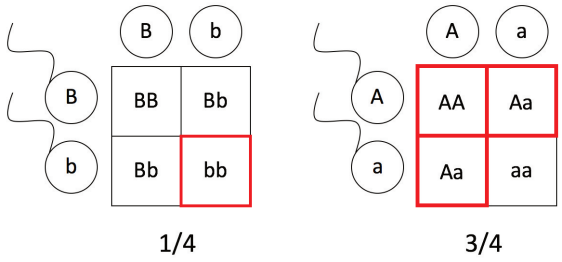


Figure 6: *These two Punnett square show the cross between two individuals who are both heterozygous for two different genes: BbAa x BbAa. We can determine the probability of an offspring having the recessive trait for "B" and the dominant trait for "A". The probability of the offspring having the recessive phenotype for "B" is 1/4. The probability of the offspring having the dominant phenotype for "A" is 3/4. 1/4 x 3/4 = 3/16.*

Another way of determining the probability of getting two different traits is to use a dihybrid Punnett square. **Figure 7** shows three generations of the inheritance of pea seed color and shape. Peas can be either yellow or green, and they can be either round or wrinkled. These are two of the traits that Mendel studied in his work with peas. In the first generation (the "P" generation), two true-breeding (homozygous) individuals are crossed. Their offspring will get one allele of the Y gene and one allele of the R gene from each parent. This means that all their offspring (the "F1" generation) will be heterozygous for both genes. The results (the "F2" generation) from crossing two heterozygous individuals can be seen in the 4×4 Punnett square in **Figure 7**.

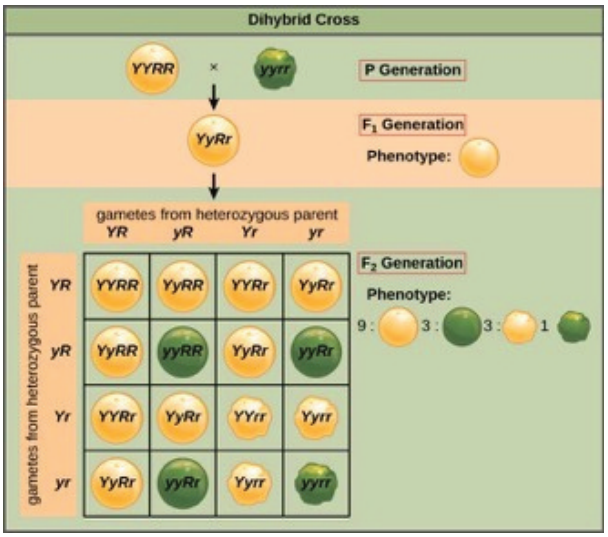


Figure 7: *This dihybrid cross shows the expected offspring from the F2 generation after crossing YYRR x yyrr. Compare the results from this Punnett square to the results seen in the previous figure. They match!*

The gametes produced by the F1 individuals must have one allele from each of the two genes. For example, a gamete could get an *R* allele for the seed shape gene and either a *Y* or a *y* allele for the seed color gene. It cannot get both an *R* and an *r* allele; each gamete can have only one allele per gene. The law of independent assortment states that a gamete into which an *r* allele is sorted would be equally likely to contain either a *Y* or a *y* allele. Thus, there are four equally likely gametes that can be formed when the *RrYy* heterozygote is self-crossed, as follows: *RY*, *rY*, *Ry*, and *ry*. Arranging these gametes along the top and left of a 4 × 4 Punnett square (**Figure 7**) gives us 16 equally likely genotypic combinations. From these

genotypes, we find a phenotypic ratio of 9 round–yellow:3 round–green:3 wrinkled–yellow:1 wrinkled–green (**Figure 7**). These are the offspring ratios we would expect, assuming we performed the crosses with a large enough sample size.

We can look for individuals who have the recessive phenotype for Y and the dominant phenotype for R. These individuals must have two little y's and at least one big R. The possible genotypes are yyRR or yyRr. Examining the Punnett square in **Figure 7**, we can find 3 individuals with these genotypes (they are round and green). If you compare the results from **Figure 6** and **Figure 7**, you'll see that we have arrived at the same value: 3/16!

Black fur color: a dominant trait

Black fur color is dominant over brown

Figure 8: This chocolate lab has two recessive alleles of the TYRP1 gene. (Credit: Rob Hanson; photo from Wikimedia.)

Most of us are familiar with the labrador retriever dog breed, such as the chocolate lab seen in **Figure 8**. But have you ever thought about what makes this dog brown? The difference between brown and black coat color in dogs is caused by a mutation in the TYRP1 gene. The TYRP1 gene provides instructions for making an enzyme called tyrosinase-related protein 1. This enzyme is required to produce a pigment called eumelanin. Eumelanin is a dark colored pigment.

The TYRP1 gene is located on chromosome 11 in dogs (Parker, 2001).

A group of scientists who were interested in determining what caused the difference between black and brown coats sequenced the DNA within the protein-coding region of the TYRP1 gene (Schmutz, 2002). They identified three variations in the DNA making up the TYRP1 gene between brown dogs and black dogs. These variations in DNA sequence are examples of different **alleles** of the TYRP1 gene.

Table 1: Variations in the TYRP1 allele that lead to brown color in dogs. Data from Schmutz, 2002.

Location	Black DNA sequence	Brown DNA sequence	Effect on protein
exon 2	TGT	CGT	changes a cysteine amino acid to a serine
exon 5	CAG	TAG	introduces a premature stop codon which results in 330 amino acids instead of 512 amino acids in the protein
exon 5	CCT	— (deleted)	deletion of a proline amino acid

All of these variations in the DNA sequence are predicted to cause a change in the amino acid sequence of the TYRP1 protein. These changes affect the production of eumelanin pigment, which is black in color. When eumelanin is not being produced correctly, the dog appears brown instead of black.

Like other diploid organisms, dogs all have two copies of the TYRP1 gene (one from their male parent, one from their female parent). Dogs that are homozygous for the black allele (dogs that have two copies of the black allele) are obviously going to be black in color. Dogs that are homozygous for the brown allele are obviously going to be brown. Dogs that are heterozygous (dogs that have one black allele and one brown allele) appear black. The black and brown colors do not blend together: the black allele covers up the brown allele. This means that the black allele is **dominant** over the brown allele. Remember that dominant alleles cover up **recessive** alleles. If there is one dominant allele present, the dog will appear

black. The brown allele is recessive to the black allele. There must be two copies of the recessive brown allele present in order for the dog to appear brown.

***Figure 9**: Black and brown phenotypes in labrador retrievers. (Credit: demealiffe; from Wikimedia)*

Remember that genotypes can be abbreviated with a single letter and that the letter which is chosen is typically the first letter of the dominant trait. In this case, the letter "B" is used to represent the dominant black allele, while "b" represents a recessive brown allele.

The reason that the black allele is dominant over the brown allele in this specific situation is because the black allele produces functional TYRP1 protein, while the brown allele does not. The presence of one functional allele produces enough TYRP1 protein allows the cells to produce eumelanin and appear black.

Remember: dominant does not mean "better" or "more normal". Black color does not confer any special advantages on dogs compared to brown color. It's just a difference.

***Figure 10**: What alleles of TYRP1 does this black lab puppy have? We can't tell by looking at it. The puppy could be homozygous (BB) or heterozygous (Bb). Since black is completely dominant over brown, both options would be black. (Credit: Alice Birkin)*

Let's visualize the inheritance of black and brown using a pedigree. The pedigree in **Figure 11** shows a litter of puppies. The shaded symbol shows a brown puppy, while open symbols are black individuals.

To interpret this pedigree, let's start with information that we already know:

- Brown is recessive, which means brown individuals must have the genotype bb. In this pedigree, brown individuals are filled in.

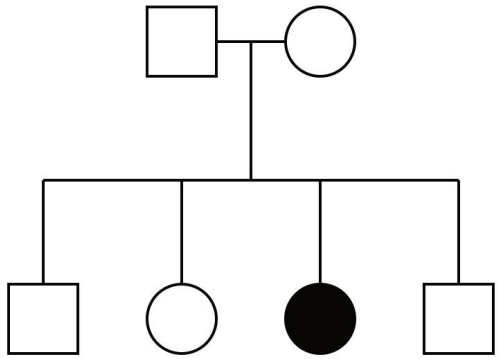


Figure 11: *An example litter of puppies. The filled-in symbol shows a brown individual.*

- Black is dominant, which means black individuals must have at least one B allele. Their genotype could be either BB or Bb. In this pedigree, black individuals are not filled in.

Figure 12 shows the same pedigree, but with information about the individual's genotypes filled in.

1. The shaded individual, who is a brown female puppy, must have the genotype bb. If she had any B alleles, she would be black because the black allele is dominant over the brown allele.
2. In order for the brown puppy to have the genotype bb, she must have gotten two "b" alleles: one from each of her parents. We know that her parents are both black (because they are unshaded), which means they must have a least one "B" allele. This means that both parents must be heterozygous: Bb.
3. The three black puppies must have at least one "B" allele in order for them to be black in color. However, we can't tell whether they are homozygous dominant (BB) or heterozygous (Bb) since both of those genotypes would result in black color. One way to represent this on a pedigree is B-, meaning that the second allele could be either B or b.

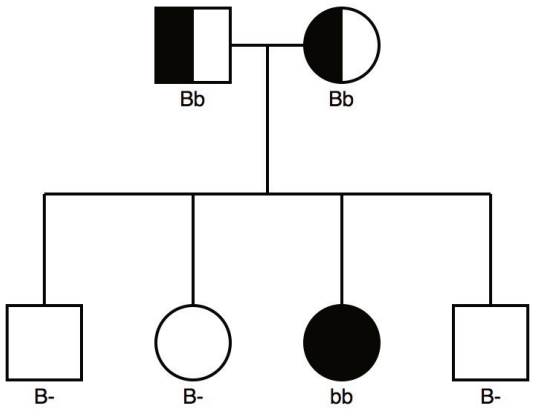


Figure 12: *Genotypes of the individuals in this pedigree.*

We can also show the cross between these parents as a Punnett square (**Figure 13**). We would expect 1/4 of the offspring to have the genotype bb, and that is what we see in the pedigree above.

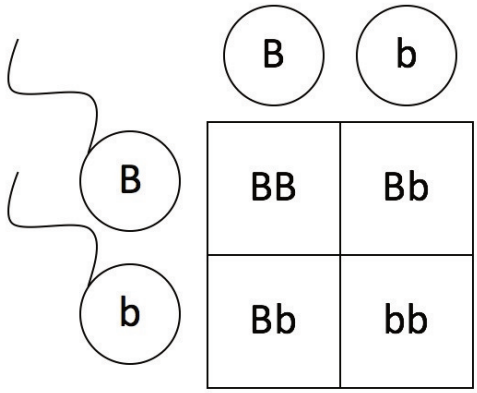


***Figure 13**: The information from the pedigree shown in Figure can also be shown as a Punnett square.*

Human Connection

A small number of mutations in the TYRP1 gene have been found to cause oculocutaneous albinism type 3. This condition includes a form of albinism called rufous oculocutaneous albinism, which has been described primarily in dark-skinned people from southern Africa. Affected individuals have reddish-brown skin, ginger or red hair, and hazel or brown irises. Two TYRP1 mutations are known to cause this form of albinism in individuals from Africa. One mutation replaces a protein building block (amino acid) in tyrosine-related protein 1 with a signal that prematurely stops protein production. This mutation, written as Ser166Ter or S166X, affects the amino acid serine at protein position 166. The other mutation, written as 368delA, deletes a single DNA building block from the TYRP1gene. Other alterations in this gene have been reported in a few affected people of non-African heritage. Most TYRP1 mutations lead to the production of an abnormally short, nonfunctional version of tyrosinase-related protein 1. Because this enzyme plays a role in normal pigmentation, its loss leads to the changes in skin, hair, and eye coloration that are characteristic of oculocutaneous albinism.

Photo credit: Muntuwandi; from Wikipedia.

References

Parker HG, Yuhua X, Mellersh CS, Khan S, Shibuya H, Johnson GS, Ostrander EA. Sept 2001. Meiotic linkage mapping of 52 genes onto the canine map does not identify significant levels of microrearrangement. Mamm Genome. 12(9):713–8.

Schmutz SM, Berryere TG, Goldfinch AD. 2002. TYRP1 and MC1R genotypes and their effects on coat color in dogs. Mammalian Genome 13, 380–387.

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:FtsD6vMd@3/Mendels-Experiments

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:sbdXt0s3@4/Laws-of-Inheritance

Information about TYRP1 and oculocutaneous albinism type 3: "Tyrp1" by Genetics Home Reference: Your Guide to Understanding Genetic Conditions, National Institutes of Health: U.S> National Library of Medicine is in the Public Domain

Yellow fur color: a recessive trait

Yellow color in dogs

Labrador retrievers don't only come in brown and black, they also come in yellow. Yellow color in labs is caused by variations in a different gene: MC1R. This gene controls the production of the melanocortin 1 receptor protein. MC1R is located on chromosome 5 in dogs (Schmutz, 2001).

***Figure 14**: This yellow lab is producing light-colored pheomelanin instead of dark-colored eumelanin. (Credit: Djmirko; from Wikimedia)*

Melanocytes make two forms of melanin, eumelanin and pheomelanin. The relative amounts of these two pigments help determine the color of an individual's hair and skin. Individuals who produce mostly eumelanin tend to have brown or black hair and dark skin that tans easily (in humans). Eumelanin also protects skin from damage caused by ultraviolet (UV) radiation in sunlight. Individuals who produce mostly pheomelanin tend to have red or blond hair, freckles, and light-colored skin that tans poorly. Because pheomelanin does not protect skin from UV radiation, people with more pheomelanin have an increased risk of skin damage caused by sun exposure.

The melanocortin 1 receptor controls which type of melanin is produced by melanocytes. When the receptor is activated, it triggers a series of chemical reactions inside melanocytes that stimulate these cells to make eumelanin. If the receptor is not activated or is blocked, melanocytes make pheomelanin instead of eumelanin. This means that if the receptor is working correctly and is turned on, dark pigment will be produced. If the receptor is not functional or is not turned on, light pigment will be produced.

Schmutz et. al. (2002) determined the DNA sequence for the MC1R gene from dogs of various colors. They determined that black and brown dogs all have one allele of MC1R, while yellow and red dogs have a different allele. The allele that leads to yellow or red color has a premature stop codon which results in a shorter-than-normal protein. This protein would be predicted to not function correctly. Remember that when the melanocortin 1 receptor is not functioning correctly, light pheomelanin pigment is produced and not dark eumelanin.

Dogs that are homozygous for the functioning allele of MC1R (which would cause eumelanin to be produced) are dark in color. Dogs that are homozygous for the non-functioning allele (which would cause pheomelanin to be produced) are light in color. Dogs that are heterozygous are dark in color. What does this tell you about which allele is dominant? If you said "the dark allele is dominant because it covers up the light allele", you're correct. We will use "E" to represent the genotype at MC1R because the dominant phenotype in this case is the production of eumelanin. Dogs that have the

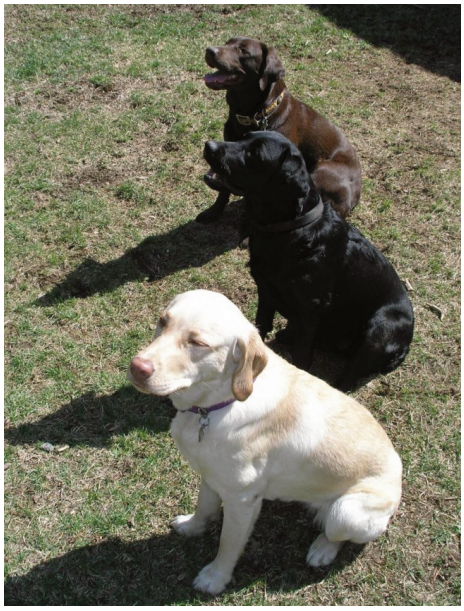


***Figure 14**: The three recognized colors of labs are due to black eumelanin, brown eumelanin, or pheomelanin. (Credit: Erikeltic, from Wikimedia)*

genotype EE or Ee will produce eumelanin and be dark. Dogs that have the genotype "ee" will produce pheomelanin and be light.

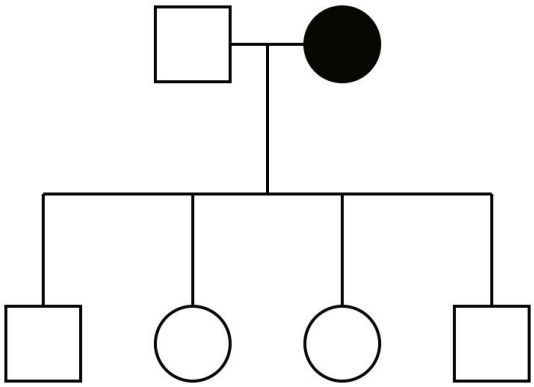


***Figure 15**: In this pedigree, the shaded individual is yellow. She therefore has the genotype ee and produces pheomelanin. We can't tell the genotype of her mate by looking (he could be Ee or EE), but since all of their puppies were dark in color, we would predict that his genotype was EE. In this cross: EE x ee, 100% of the puppies would have the genotype Ee, so 100% of the puppies would produce eumelanin instead of pheomelanin.*

The cross shown in **Figure 15** can also be shown as a Punnett square. Since we are unsure whether the male dog has the genotype "EE" or "Ee", we have to make two Punnett squares. Since all of the puppies resulting from this cross were black, we would predict that the first Punnett square shows the cross. However, it is possible that the second Punnett square is correct. There are only 4 puppies, so it's not hard to imagine that they could all be black even though the Punnett square predicts only 50% black. It would be comparable to flipping a coin 4 times and getting 4 heads in a row. Getting 4 heads in a row is less likely, but definitely possible.

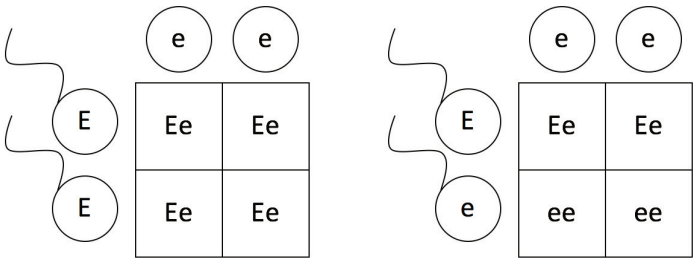


Figure 16: *Cross from* ***Figure 15*** *shown as Punnett squares*

It is very important to note here that yellow dogs still have the TYRP1 gene, even though they are not black or brown!

Human Connection

Common variations (polymorphisms) in the MC1R gene are associated with normal differences in skin and hair color. Certain genetic variations are most common in people with red hair, fair skin, freckles, and an increased sensitivity to sun exposure. These MC1R polymorphisms reduce the ability of the melanocortin 1 receptor to stimulate eumelanin production, causing melanocytes to make mostly pheomelanin. Although MC1R is a key gene in normal human pigmentation, researchers believe that the effects of other genes also contribute to a person's hair and skin coloring.

The melanocortin 1 receptor is also active in cells other than melanocytes, including cells involved in the body's immune and inflammatory responses. The receptor's function in these cells is unknown.

Photo credit: dusdin on flickr; from Wikipedia.

Resources

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:FtsD6vMd@3/Mendels-Experiments

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:sbdXt0s3@4/Laws-of-Inheritance

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:zLLYW2hj@5/Extensions-of-the-Laws-of-Inhe

Human Connection – information about MC1R: "MC1R" by Genetics Home Reference: Your Guide to Understanding Genetic Conditions, National Institutes of Health: U.S> National Library of Medicine is in the Public Domain

Schmutz SM, Moker JS, Berryere TG, Christison KM, Dolf G. 2001. An SNP is used to map MC1R to dog chromosome 5. Anim Genet. 32(1):43-4.

Schmutz SM, Berryere TG, Goldfinch AD. 2002. TYRP1 and MC1R genotypes and their effects on coat color in dogs. Mammalian Genome 13, 380-387.

Epistasis: the relationship between black, brown, and yellow fur

Epistasis

Dogs don't have either the TYRP1 gene *or* the MC1R gene – they have both. In fact, every dog will have two copies of the TYRP1 gene and two copies of the MC1R gene. Since both genes control aspects of coat color, it makes sense that they interact. In fact, TYRP1 and MC1R have what is called an epistatic relationship: the action of one gene controls the expression of a second gene. Another way to phrase this relationship is that the effect of one gene is dependent on another gene.

Remember that TYRP1 is required for the production of eumelanin. The dominant allele of TYRP1 (B) produces black eumelanin, while the recessive allele (b) produces brown eumelanin. However, if a dog is homozygous recessive for MC1R (ee), they lack the ability to produce eumelanin at all. If no eumelanin is being produced, it doesn't matter whether it would have been black or brown: there is none. This means that any dog that is homozygous recessive for MC1R will appear yellow regardless of its genotype at TYRP1. These two genes are epistatic: the action of MC1R controls the expression of TYRP1. The effect of TYRP1 is dependent on MC1R.

If a dog has at least one dominant functioning allele of MC1R, then its genotype at TYRP1 can be seen. If the dog has at least one dominant allele of TYRP1, it will appear black. If it has two recessive alleles, it will appear brown.

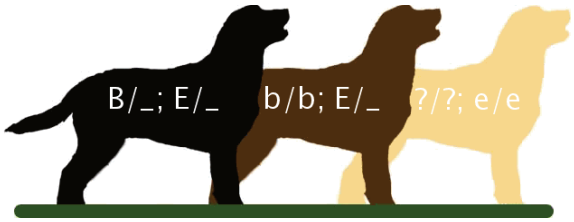


Figure 17: *Genotypes for TYRP1 (B) and MC1R (E) that lead to the three recognized colors of labs. (Credit EArellano, from Wikimedia)*

A pedigree can be used to show the inheritance of two different genes such as TYRP1 and MC1R.

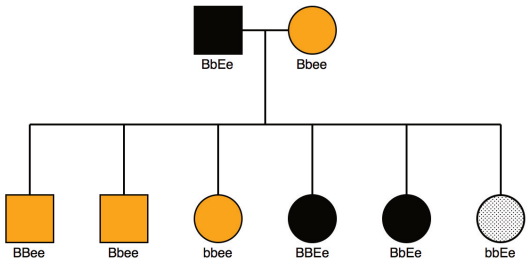


Figure 18: *In this pedigree, a cross between an individual who is heterozygous for both MC1R and TYRP1 and an individual who has the genotype "Bbee"is shown. Black individuals are shaded black, yellow individuals are shaded yellow, and brown individuals are shaded grey. The 6 different possible genotypes are each shown as one offspring. This does not give you any information about the probability of getting a certain genotype of offspring – it gives you the actual number of offspring observed and their traits.*

Punnett squares can also be used to show this cross. If the probability of inheriting one trait is multiplied by the probability of inheriting the second trait, the overall probability of getting any given offspring can be determined.

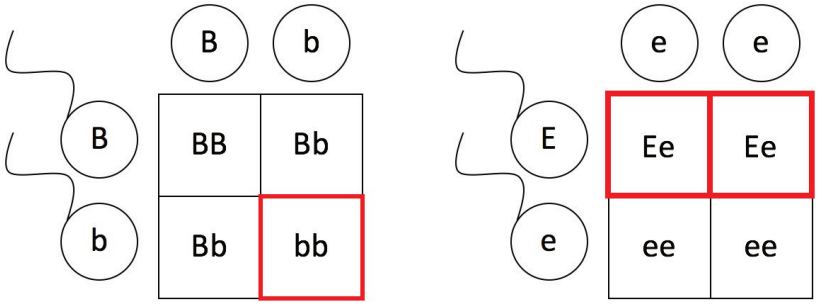


Figure 19: *These two Punnett squares can be used to determine the results of a cross between these individuals: Bbee x BbEe. If you wanted to determine the probability of getting a brown dog, you would multiply the probability of getting bb by the probability of having at least one dominant E. That would equal 1/4 x 1/2 = 1/8. This gives you the probability of getting a brown dog, but doesn't tell you anything about the number of brown dogs actually observed.*

Human Connection

Individuals who have albinism lack the ability to produce any pigment. If no pigment is being produced, the color that the pigment would have been is unimportant. The effect of the pigment genes is controlled by the gene that allows pigment to be produced. This is an example of epistasis.

Albinism can occur in humans (see the section on TYRP1) as well as other animals, such as the squirrel seen below.

Photo credit: Stephenkniatt from Wikipedia.

References:

Unless otherwise noted, images on this page are licensed under CC-BY 4.0 by OpenStax.

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:zLLYW2hj@5/Extensions-of-the-Laws-of-Inhe

Schmutz SM, Berryere TG, Goldfinch AD. 2002. TYRP1 and MC1R genotypes and their effects on coat color in dogs. Mammalian Genome 13, 380-387.

Brindle color: partial dominance and epistasis

Brindle coloration is caused by different alleles at the "K locus", which is probably a gene called ASIP that controls pigment switching (Ciampolini, 2013). There are three alleles of the K locus: K^B, k^{br}, and k^y (Kerns, 2007). The K^B allele is dominant over the other two alleles and produces solid black color. k^{br} produces the brindle color pattern and is dominant over the k^y allele. This means that dogs with the genotype $k^{br}k^{br}$ or $k^{br}k^y$ will have the brindle color pattern. Dogs with the genotype k^yk^y are yellow in color.

***Figure** : This boxer shows the brindle color pattern, which looks sort of like tiger stripes. (Credit: Steve Henderson Location: Memphis, TN)*

The K locus and MC1R (which controls the difference between dark eumelanin and light pheomelanin production) have an epistatic relationship. If a dog has two recessive alleles for MC1R and is therefore unable to make eumelanin, the dog will appear yellow regardless of its genotype at the K locus.

References

Ciampolini R, Cecchi F, Spaterna A, Bramante A, Bardet SM, Oulmouden A. 2013. Characterization of different 5′-untranslated exons of the ASIP gene in black-and-tan Doberman Pinscher and **brindle** Boxer **dogs**. Anim Genet. 44(1):114-7.

Kerns JA, Cargill EJ, Clark LA, Candille SI, Berryere TG, Olivier M, Lust G, Todhunter RJ, Schmutz SM, Murphy KE, Barsh GS. 2007. Linkage and segregation analysis of black and **brindle** coat color in domestic **dogs**. Genetics. 176(3):1679-89.

Incomplete dominance: when traits blend

Flower color in snapdragons

Mendel's results in crossing peas, black vs brown fur color, and eumelanin production vs pheomelanin production all demonstrate traits are inherited as dominant and recessive. This contradicts the historical view that offspring always exhibited a blend of their parents' traits. However, sometimes heterozygote phenotype is intermediate between the two parents. For example, in the snapdragon, *Antirrhinum majus* (**Figure 20**), a cross between a homozygous parent with white flowers (C^WC^W) and a homozygous parent with red flowers (C^RC^R) will produce offspring with pink flowers (C^RC^W) (**Figure 21**).

Figure 20: *These pink flowers of a heterozygote snapdragon result from incomplete dominance. (credit: "storebukkebruse"/Flickr)*

Note that different genotypic abbreviations are used to distinguish these patterns from simple dominance and recessiveness. The abbreviation C^W can be read as "at the flower color gene (C), the white allele is present."

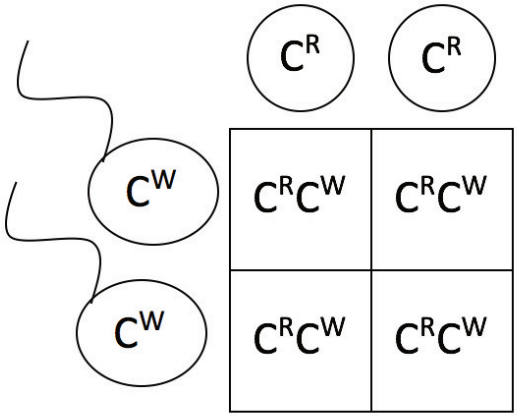


Figure 21: *A cross between a red and white snapdragon will yield 100% pink offspring.*

This pattern of inheritance is described as **incomplete dominance**, meaning that neither of the alleles is completely dominant over the other: both alleles can be seen at the same time. The allele for red flowers is incompletely dominant over the allele for white flowers. Red + white = pink. The results of a cross where the alleles are incompletely dominant can still be predicted, just as with complete dominant and recessive crosses. **Figure 22** shows the results from a cross between two heterozygous individuals: C^RC^W x C^RC^W . The expected offspring would have the genotypic ratio 1 C^RC^R:2 C^RC^W:1 C^WC^W, and the phenotypic ratio would be 1:2:1 for red:pink:white. The basis for the intermediate color in the heterozygote is simply that the pigment produced by the red allele (anthocyanin) is diluted in the heterozygote and therefore appears pink because of the white background of the flower petals.

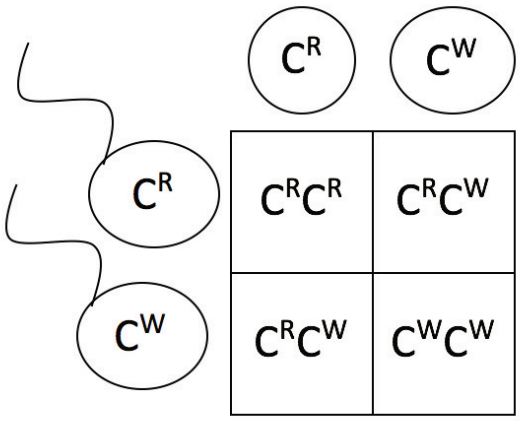


***Figure 22**: The results of crossing two pink snapdragons.*

Straight, curly, and wavy hair in dogs

***Figure 23**: The wavy hair on this labradoodle is caused by incomplete dominance. (Credit: Localpups, Flickr)*

Another example of incomplete dominance is the inheritance of straight, wavy, and curly hair in dogs. The KRT71 gene is used to synthesize the keratin 71 protein. Genes in the KRT family provide instructions for making proteins called keratins. Keratins are a group of tough, fibrous proteins that form the structural framework of epithelial cells, which are cells that line the surfaces and cavities of the body. Epithelial cells make up tissues such as the hair, skin, and nails. These cells also line the internal organs and are an important part of many glands.

Keratins are best known for providing strength and resilience to cells that form the hair, skin, and nails. These proteins allow tissues to resist damage from friction and minor trauma, such as rubbing and scratching. Keratins are also involved in several other critical cell functions, including cell movement (migration), regulation of cell size, cell growth and division

(proliferation), wound healing, and transport of materials within cells. Different combinations of keratin proteins are found in different tissues.

The mutation which causes curly hair in dogs, such as the labradoodle seen in Figure 23, is in exon 2 of the gene and is predicted to substantially disrupt the structure of the keratin 71 protein (Cadieu, 2009). This change in protein shape prevents the keratin proteins from interacting together correctly within the hair, altering the structure of the hair and resulting in a curly coat (Runkel, 2006).

When a dog has two curly alleles (K^CK^C), it has a very curly coat, such as on the poodle in **Figure 24**. A dog with two straight alleles (K^+K^+) has a straight coat. Dogs that are heterozygous (K^+K^C) have an intermediate or wavy coat like the labradoodle in **Figure 23**.

Figure 24: *This poodle has two copies of the curly allele of the KRT71 gene (K^CK^C). Compare his curly hair to the wavy hair of the labradoodle in Figure 23. The labradoodle is heterozygous (K^+K^C). (Credit B. Schoener; From Wikimedia)*

Human Connection – Blood Type

Blood is classified into different groups according to the presence or absence of molecules called antigens on the surface of every red blood cell in a person's body. Antigens determine blood type and can either be proteins or complexes of sugar molecules (polysaccharides). The genes in the blood group antigen family provide instructions for making antigen proteins. Blood group antigen proteins serve a variety of functions within the cell membrane of red blood cells. These protein functions include transporting other proteins and molecules into and out of the cell, maintaining cell structure, attaching to other cells and molecules, and participating in chemical reactions.

There are 29 recognized blood groups, most involving only one gene. Variations (polymorphisms) within the genes that determine blood group give rise to the different antigens for a particular blood group protein. For example, changes in a few DNA building blocks (nucleotides) in the ABO gene give rise to the A, B, and O blood types of the ABO blood group. The changes that occur in the genes that determine blood group typically affect only blood type and are not associated with adverse health conditions, although exceptions do occur.

The A and B alleles are codominant, which is similar to incomplete dominance in that heterozygotes have an

intermediate phenotype. If both the A and B alleles are present, both will be seen in the phenotype. The O allele is recessive to both A and B.

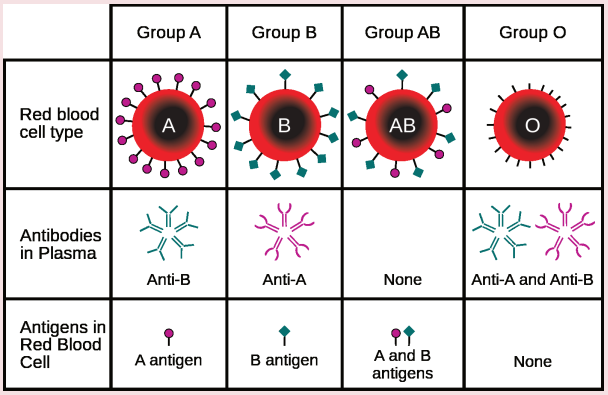


Photo credit: InvictaHOG, from Wikipedia.

References

Unless otherwise noted, images on this page are licensed under CC-BY 4.0 by OpenStax.

Cadieu E, Neff MW, Quignon P, Walsh K, Chase K, Parker HG, Vonholdt BM, Rhue A, Boyko A, Byers A, Wong A, Mosher DS, Elkahloun AG, Spady TC, André C, Lark KG, Cargill M, Bustamante CD, Wayne RK, Ostrander EA. 2009. Coat variation in the domestic **dog** is governed by variants in three genes. Science. 326(5949):150-3.

Runkel F, Klaften M, Koch K, Böhnert V, Büssow H, Fuchs H, Franz T, Hrabé de Angelis M. 2006. Morphologic and molecular characterization of two novel Krt71 (Krt2-6g) mutations: Krt71rco12 and Krt71rco13. Mamm Genome. 17(12):1172-82.

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:zLLYW2hj@5/Extensions-of-the-Laws-of-Inhe

"Blood Group Antigens" by Genetics Home Reference: Your Guide to Understanding Genetic Conditions, National Institutes of Health: U.S> National Library of Medicine is in the Public Domain

"Keratins" by Genetics Home Reference: Your Guide to Understanding Genetic Conditions, National Institutes of Health: U.S> National Library of Medicine is in the Public Domain

White spotting: When there's more than two alleles

So far, we have discussed genes which have only two alleles. However, that is not always the case: there can be more than two alleles for a given gene. One example is the MITF gene, which is the major gene that controls white spotting in dogs. This protein is required for the migration and survival of melanocytes into the skin during development. If it is not functional, it impairs the ability of the skin to make pigment, thus "covering up" the effect of other color genes. There are thought to be at least four alleles that can contribute (Karlsson, 2007). Depending on which alleles are present in a dog, the amount of white can vary from none (a solid-colored dog) to mostly white (**Table 2** and **Figure 24**).

Table 2: Combinations of different alleles for MITF result in different amounts of white present in the coat.

Alleles	Amount of white
SS	None (solid colored)
Ss^{i}	Small amounts of white possible on chin, chest, feet, and tail tip
Ss^{p}	Pied markings where the coat is more than 50% colored, with white on the face, chest, feet, collar, underbelly, and tail tip
$s^{i}s^{p}$	Approximately even amounts of color and white
$s^{i}s^{e}$	More than 50% white with irregular splashes of color
$s^{e}s^{e}$	Mostly white with only minimal areas of color, perhaps on one or both ears, an eye patch, or a spot near the tail

Figure 24: *These dogs have different combinations of alleles of the MITF gene. The first dog probably has the genotype "SS"; the dog in the center is likely "Ss^{p}"; the dog on the right is likely "$s^{e}s^{e}$". (Credits: Funny black dog by X posid from Publicdomainpictures. A black and white dog by Petr Kratochvil from Free stock photos. White dog with black ears by RetyiRetyi from Pixabay.)*

Human Connection – Blood Type

Human blood type was discussed in the previous section. You may remember that there are three alleles for the ABO gene: A, B, and O. A and B are codominant, meaning that if both alleles are present, both will be seen in the phenotype. A person with type AB blood has one A allele and one B allele.

O is recessive to A and B. A person with the genotype AO will have Type A blood. A person with the genotype BO will have type B blood. Type O blood results from two O alleles.

	A	B	O
A	AA	AB	AO
B	AB	BB	BO
O	AO	BO	OO

Photo credit: Kalaiarasy, from Wikipedia.

References

Karlsson EK, Baranowska I, Wade CM, Salmon Hillbertz NH, Zody MC, Anderson N, Biagi TM, Patterson N, Pielberg GR, Kulbokas EJ 3rd, Comstock KE, Keller ET, Mesirov JP, von Euler H, Kämpe O, Hedhammar A, Lander ES, Andersson G, Andersson L, Lindblad-Toh K. 2007. Efficient mapping of mendelian traits in **dogs** through genome-wide association. Nat Genet. 39(11):1321-8.

Hemophilia: a sex-linked disorder

So far, all the genes we have discussed have had two copies present in all individuals. This is because the individual inherited one from the male parent's haploid gamete and one from the female parent's haploid gamete. The two gametes came together during fertilization to produce a diploid individual. There is, however, one exception to this: genes which are present on the sex chromosomes.

In humans, as well as in many other animals and some plants, the sex of the individual is determined by **sex chromosomes –** one pair of non-homologous chromosomes. Until now, we have only considered inheritance patterns among non-sex chromosomes, or **autosomes.** In addition to 22 homologous pairs of autosomes, human females have a homologous pair of X chromosomes, whereas human males have an XY chromosome pair. Although the Y chromosome contains a small region of similarity to the X chromosome so that they can pair during meiosis, the Y chromosome is much shorter and contains fewer genes. When a gene being examined is present on the X, but not the Y, chromosome, it is **X-linked**.

The X chromosome is one of two sex chromosomes. Humans and most mammals have two sex chromosomes, the X and Y. Females have two X chromosomes in their cells, while males have X and Y chromosomes in their cells. Egg cells all contain an X chromosome, while sperm cells contain an X or a Y chromosome. This arrangement means that during fertilization, it is the male that determines the sex of the offspring since the female can only give an X chromosome to the offspring.

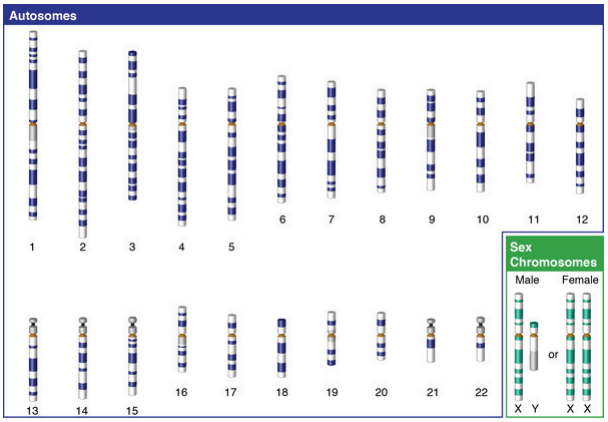


Figure 24: A diagram showing the autosomal and sex chromosomes. Remember that in a diploid cell, there would be two copies of each autosomal chromosome present. (Credit: Darryl Lega, NHGRI)

Most sex-linked genes are present on the X chromosome simply because it is much larger than the Y chromosome. The X chromosome spans about 155 million DNA base pairs and represents approximately 5 percent of the total DNA in cells. The X chromosome likely contains 800 to 900 genes. In contrast, the Y chromosome has approximately 59 million base pairs and only 50-60 genes. Sex is determined by the SRY gene, which is located on the Y chromosome and is responsible for the development of a fetus into a male. This means that the presence of a Y chromosome is what causes a fetus to develop as male. Other genes on the Y chromosome are important for male fertility.

Hemophilia is a bleeding disorder that slows the blood clotting process. People with this condition experience prolonged bleeding or oozing following an injury, surgery, or having a tooth pulled. In severe cases of hemophilia, continuous bleeding occurs after minor trauma or even in the absence of injury (spontaneous bleeding). Serious complications can result from bleeding into the joints, muscles, brain, or other internal organs. Milder forms of hemophilia do not necessarily involve spontaneous bleeding, and the condition may not become apparent until abnormal bleeding occurs following surgery or a serious injury.

The major types of this condition are hemophilia A (also known as classic hemophilia or factor VIII deficiency)

and hemophilia B (also known as Christmas disease or factor IX deficiency). Although the two types have very similar signs and symptoms, they are caused by mutations in different genes.

Hemophilia A and hemophilia B are inherited in an X-linked recessive pattern. The genes associated with these conditions are located on the X chromosome, which is one of the two sex chromosomes. In males (who have only one X chromosome), one altered copy of the gene in each cell is sufficient to cause the condition. In females (who have two X chromosomes), a mutation would have to occur in both copies of the gene to cause the disorder. Because it is unlikely that females will have two altered copies of this gene, it is very rare for females to have hemophilia. A characteristic of X-linked inheritance is that fathers cannot pass X-linked traits to their sons.

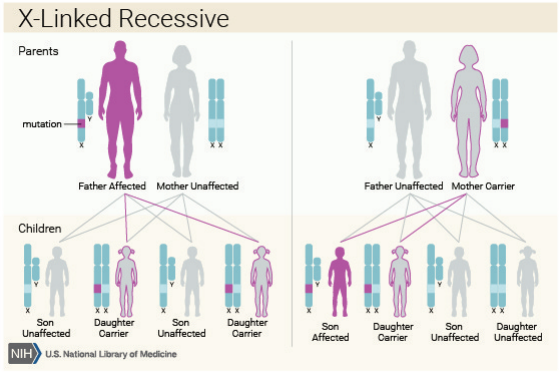


Figure 25: X-linked recessive inheritance. (Credit: U.S. National Library of Medicine)

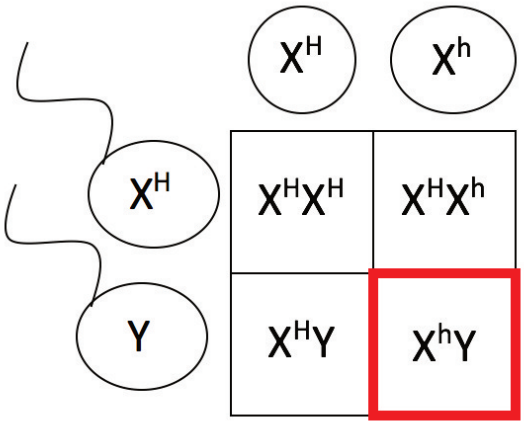


Figure 25: If a carrier female and a normal male produce offspring, there is a 25% total chance that they will have a child with hemophilia. None of their daughters will have the disease (although all will be carriers). Half their sons will be hemophiliacs.

In X-linked recessive inheritance, a female with one altered copy of the gene in each cell is called a carrier. Carrier females have about half the usual amount of coagulation factor VIII or coagulation factor IX, which is generally enough for normal blood clotting. However, about 10 percent of carrier females have less than half the normal amount of one of these coagulation factors; these individuals are at risk for abnormal bleeding, particularly after an injury, surgery, or tooth extraction.

Colorblindness is another example of a sex-linked trait in humans. The genes that produce the photopigments necessary for color vision are located on the X chromosome. If one of these genes is not functional because it contains a harmful

mutation, the individual will be colorblind. Men are much more likely than women to be colorblind: up to 100 times more men than women have various types of colorblindness (http://www.colour-blindness.com/general/prevalence/).

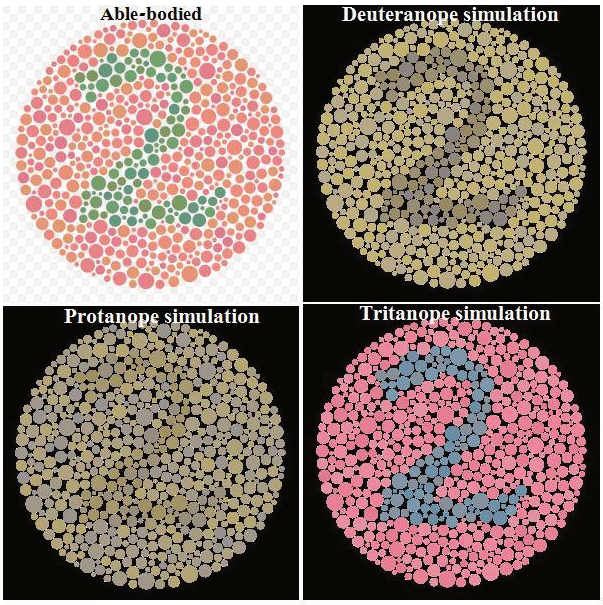


Figure 26: *A test image for color-blindness as seen by someone with normal color vision and several types of colorblindness. (Credit: Sakurambo)*

References:

OpenStax, Biology. OpenStax CNX. May 27, 2016 http://cnx.org/contents/s8Hh0oOc@9.10:zLLYW2hj@5/Extensions-of-the-Laws-of-Inhe

"X chromosome" by Genetics Home Reference: Your Guide to Understanding Genetic Conditions, National Institutes of Health: U.S> National Library of Medicine is in the Public Domain

"Y chromosome" by Genetics Home Reference: Your Guide to Understanding Genetic Conditions, National Institutes of Health: U.S> National Library of Medicine is in the Public Domain

"Hemophilia" by Genetics Home Reference: Your Guide to Understanding Genetic Conditions, National Institutes of Health: U.S> National Library of Medicine is in the Public Domain

Overall phenotypes: putting it all together

None of the genes discussed in these sections occur in isolation: one individual dog would have all the genes for color, hair structure, and hemoglobin (dogs can get hemophilia too). Genes interact together to produce the overall phenotype of the individual.

Example 1: Sugar

For example, look at Sugar in **Figure 26**. She has short hair that is mostly white. The colored portion of her hair is the tiger-striped pattern termed "brindle."

Figure 26: *Sugar has short hair with colored spots. (Credit: Lisa Bartee)*

The difference between short and long hair in dogs is caused by different alleles of a gene called FGF5. This gene produces a protein that is important in regulating the hair growth cycle. When the protein doesn't function correctly, the growth phase of the hair cycle is longer, resulting in long hair. Short hair is the dominant trait. Since Sugar has short hair, we know she has at least one dominant allele of FGF5. We can use the letter "S" for short hair. Sugar's genotype for FGF5 is therefore "S-", meaning she has one dominant allele and we can't tell by looking at her what her second allele is.

Sugar's hair is also straight, which means that she has two straight alleles of KRT71. Her genotype would be $K^{+}K^{+}$.

Sugar is more than 50% white with irregular splashes of color, which means that her genotype for MITF (the gene that controls white spotting) is $s^{i}s^{e}$.

The brindle pattern is caused by the k^{br} allele at the K locus. Sugar can't have the K^{B} allele or she would have solid color instead of the brindle pattern because K^{B} is dominant over k^{br} and k^{y}. She could have either the genotype $k^{br}k^{br}$ or $k^{br}k^{y}$, since the k^{br} allele is dominant over the yellow allele (k^{y}).

Sugar has black eumelanin pigment in her hair and nose. This means she has the dominant phenotype for TYPR1, so her genotype would be "B-". Because she has eumelanin and not pheomelanin in her coat, she has the dominant phenotype for MC1R, so her genotype would be "E-".

Sugar is a female dog who does not have hemophilia. This means that her genotype would be either $X^{H}X^{H}$ or $X^{H}X^{h}$.

Putting all these together, we could say that Sugar's overall coat genotype is S- $K^{+}K^{+}$ $s^{i}s^{e}$ B- E- $X^{H}X$–

We could potentially determine some of the unknown alleles in her genotype if we knew anything about her parents,

but Sugar was adopted from the Multnomah County Animal Shelter after being picked up as a stray. Therefore, her ancestry is unknown.

Example 2: Rags

***Figure 27**: Rags is similar in color to Sugar, but has a very different fur type. (Credit: Lisa Bartee)*

Rags has "furnishings", a term used to describe his beard and mustache. Furnishings are caused by a mutation in the RSPO2 gene. This gene produces a protein that is involved in establishing hair follicles. The allele that leads to furnishings is dominant over the allele for no furnishings. Rags must therefore have the genotype "F-" at RSPO2. This allele also causes the long-ish hair on his legs and tail.

Gene	Genotype	Phenotype
RSPO2	FF or Ff	has furnishings
FGF5	SS or Ss	short fur (his longer fur is caused by the furnishings allele)
KRT71	$K^{+}K^{+}$	straight fur
MITF	$s^{i}s^{e}$	more than 50% white
K locus	$K^{B}K^{B}$, $K^{B}k^{br}$, or $K^{B}k^{y}$	Solid color, not brindle or yellow.
TYRP1	BB or Bb	Produces black eumelanin, not brown
MC1R	EE or Ee	Produces eumelanin instead of pheomelanin
F8	$X^{H}Y$	Male, no hemophilia

Example 3: Black poodle

Figure : Black poodle. (Credit: B. Schoener from Wikimedia)

Gene	Genotype	Phenotype
RSPO2	ff	no furnishings
FGF5	ss	long fur
KRT71	K^CK^C	curly fur
MITF	SS	entirely solid color
K locus	K^BK^B, K^Bk^{br}, or K^Bk^y	Solid color, not brindle or yellow.
TYRP1	BB or Bb	Produces black eumelanin, not brown
MC1R	EE or Ee	Produces eumelanin instead of pheomelanin

Example 4: Golden Retriever

Figure : *Golden Retriever. (Credit: Dirk Vorderstraße)*

Gene	Genotype	Phenotype
RSPO2	ff	no furnishings
FGF5	ss	long fur
KRT71	$K^{+}K^{+}$	straight fur
MITF	SS	entirely solid color
K locus	$K^{B}K^{B}$, $K^{B}k^{br}$, or $K^{B}k^{y}$	Solid color, not brindle or yellow.
TYRP1	BB or Bb	Produces black eumelanin, not brown (seen in the nose)
MC1R	ee	Produces pheomelanin instead of eumelanin, so appears yellow

References

Additional complexity

Figure 29: An English Cocker Spaniel. (Credit eNil)

We haven't exhaustively discussed all the genes that can affect dog appearance. For example, what gene (or genes) causes the English Springer Spaniel in **Figure 29** to be red? What gene(s) cause it to be speckled on it's back? Or lead to its freckles? There are estimated to be about 19,000 genes in the dog genome (Ostrander, 2005). The interactions of all these genes together lead to the overall phenotype of one individual dog.

If you're interested in learning more about the genes that are involved in the appearance of dogs, check out the Dog Coat Color Genetics website at http://www.doggenetics.co.uk.

References

Ostrander EA, Wayne RK. 2005. The Canine Genome. Genome Res. 15: 1706-1716.

It's not all in the genes

Not all traits are directly caused by DNA alone. The environment also plays a large role in shaping an individual's traits. Some examples can be seen below.

- **Height and weight:** A number of genes interact to determine the general height and weight that a person will have. But the environment has a major influence as well. If an individual is malnourished, their growth may be slowed and they may be smaller than they would have been if they had gotten enough food. In contrast, if a person consumes more calories than they need, their weight will likely increase regardless of their genetics.

- **Fingerprints:** the general characteristics of a person's fingerprints are determined by genetics, but the specific pattern is generated randomly during development. Identical twins typically have fingerprints that are similar, but not identical.

- **Intelligence:** Like most aspects of human behavior and cognition, intelligence is a complex trait that is influenced by both genetic and environmental factors. Roughly 50% of a person's IQ appears to be determined by genetic factors. Factors related to a child's home environment and parenting, education and availability of learning resources, and nutrition, among others, also contribute to intelligence. A person's environment and genes influence each other, and it can be challenging to tease apart the effects of the environment from those of genetics. For example, if a child's IQ is similar to that of his or her parents, is that similarity due to genetic factors passed down from parent to child, to shared environmental factors, or (most likely) to a combination of both? It is clear that both environmental and genetic factors play a part in determining intelligence.

- **Cancer Risk:** For example, a person could inherit a mutation in the BRCA1 gene, which increases the risk of developing breast or ovarian cancer. Researchers have identified more than 1,800 mutations in the BRCA1 gene. Most BRCA1 gene mutations lead to the production of an abnormally short version of the BRCA1 protein or prevent any protein from being made from one copy of the gene. As a result, less of this protein is available to help repair damaged DNA or fix mutations that occur in other genes. As these defects accumulate, they can trigger cells to grow and divide uncontrollably to form a tumor. These mutations are present in every cell in the body and can be passed from one generation to the next. As a result, they are associated with cancers that cluster in families. However, not everyone who inherits a mutation in the BRCA1 gene will develop cancer. Other genetic, environmental, and lifestyle factors also contribute to a person's cancer risk.
- In contrast, cancer can be caused by purely environmental factors. According to the CDC, cigarette smoking is the number one risk factor for lung cancer. In the United States, cigarette smoking is linked to about 90% of lung cancers and people who smoke are 15 to 30 times more likely to get lung cancer or die from lung cancer than people who do not smoke. Radon exposure also increases the likelihood that a person will develop lung cancer.

References

Figure 30: *The colors on the poodle seen in this figure have no relationship to his DNA: he was dyed for a parade. (Credit: skeeze)*

"Is intelligence determined by genetics?" by Genetics Home Reference: Your Guide to Understanding Genetic Conditions, National Institutes of Health: U.S> National Library of Medicine is in the Public Domain