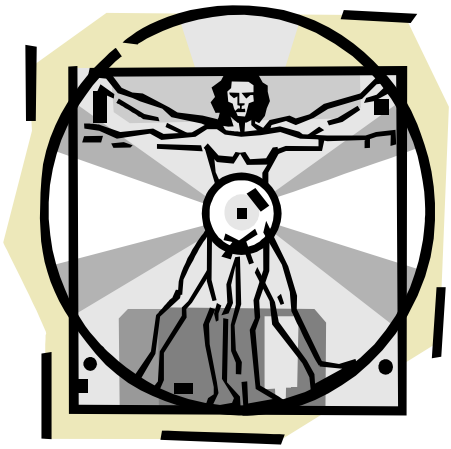
Unit 8—Genetics

Introduction

There are some advantages of having a large population. For example, if one person in a billion is a truly exceptional scientific genius, it means that there are six of them in a world of 6 billion people and somewhere out there another one is being born every twelve to fifteen years. These people are the ones that have great insights that change the world for everyone else and now there are enough of them alive at any one time to work together or to compete for new discoveries. As the population of the world increases, so does the number of exceptionally brilliant scientists and so does the pace of discovery.



One of the areas that is benefiting from collaboration and new scientific discoveries is our understanding of our own biology. If you are a college freshman, more has been discovered about how our bodies work since you were born than was discovered in all of previous history and we are just getting started. In this unit, you will learn about these discoveries and how they are turning into technologies that affect your life and your future.

Cells and Chromosomes

This topic can be complicated and understanding can be lost in a flurry of scientific terms. The purpose of this unit is to provide a framework of understanding. The student is encouraged to further their understanding of this topic by taking classes in biology, computer science, and ***bioinformatics***—a new field that combines biology and computer science.

Cells, the Nucleus and Chromosomes

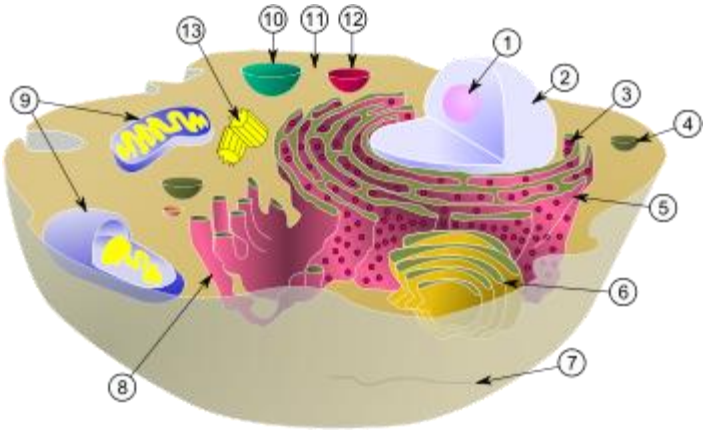
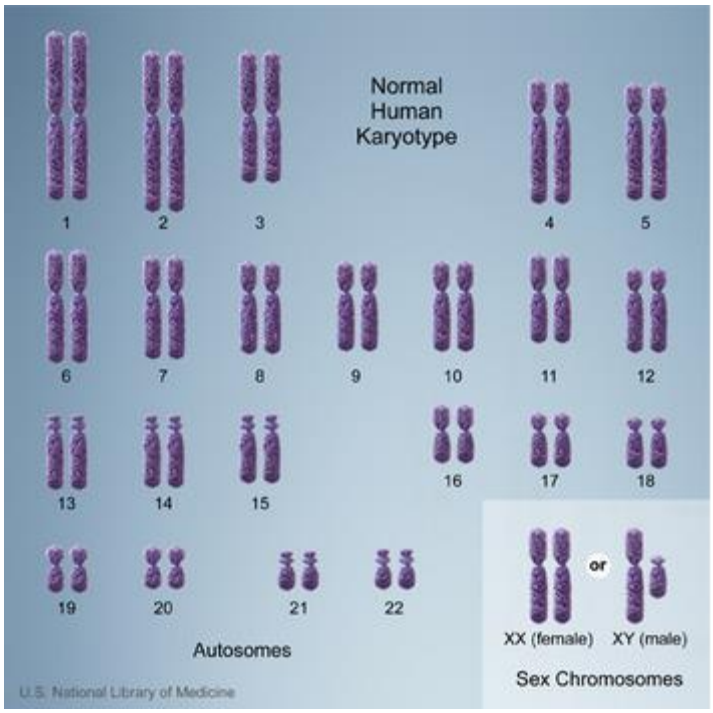


Figure 8.1 Schematic of typical animal cell, showing subcellular components. Organelles: (1) nucleolus (2) nucleus (3) ribosome (4) vesicle (5) rough endoplasmic reticulum (ER) (6) Golgi apparatus (7) Cytoskeleton (8) smooth ER (9) mitochondria (10) vacuole (11) c

The invention of the microscope and the use of dyes allowed biologists to see the fine details of the structure of our bodies. They found that we are comprised of small units called ***cells*** that have an outer wall and various internal parts that include a nucleus. (Cell nucleus, 2008) <Link> The cell nucleus contains the master set of instructions for creating the cell and the chemicals it needs to function. In this discussion, we are only interested in the contents of the nucleus. Inside the nucleus is a wad of material that is normally intertwined like a ball of rubber bands. Scientists noticed that when the cell was about to reproduce this group of material separated into individually identifiable pieces. They learned what chemicals to use to get the nucleus to do this on demand so they could study these pieces and they used dyes to make them show up better in their microscopes. Because of the use of colored dyes used to identify them, the bodies were named ***chromosomes*** which mean colored (chroma) bodies (somes) and were first observe in the mid-1800s. Chromosomes contain the instructions for making new cells and for making the chemicals each cell needs to function.

Upon closer examination, a chromosome is a single long molecule that is tightly bunched. A section near the middle forms a waist which is called a ***centromere***. Chromosomes occur in pairs where one of the chromosomes in each pair comes from the mother and the other comes from the father (for animals or plants that reproduce sexually). The two chromosomes are usually the same and they act like backups of each other. If one of them is defective, the other can still work normally.



8.2 Pairs of chromosomes

Study of chromosomes produced insights regarding the differences and similarities between humans and other animals. Humans have 46 chromosomes that are arranged in 23 pairs. Other animals and plants have widely different numbers of chromosomes. However, the animals that most closely resemble humans—chimpanzees and gorillas—have a similar number.

Species	*Chromosomes*
Common fruit fly	8
Dove	16
Earthworm	36
Tibetan fox	36
Domestic cat	38
Domestic pig	38
Lab mouse	40
Lab rat	42
Bread wheat	42
Rabbit	44
Human	46
Gorillas, Chimpanzees	48
Cultivated tobacco	48
Garden snail	54
Domestic sheep	54
Elephants	56
Silkworm	56

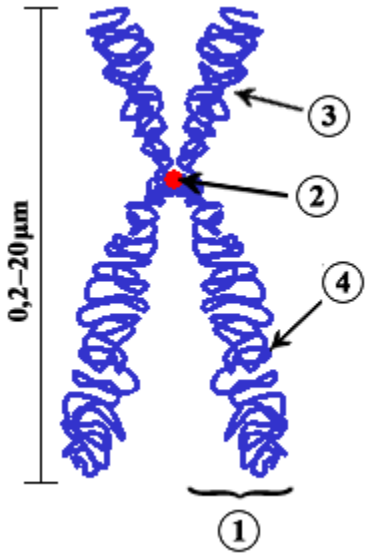
Cow	60
Horse	64
Dog	78

8.3 Number of Chromosomes per animal

Cell Duplication

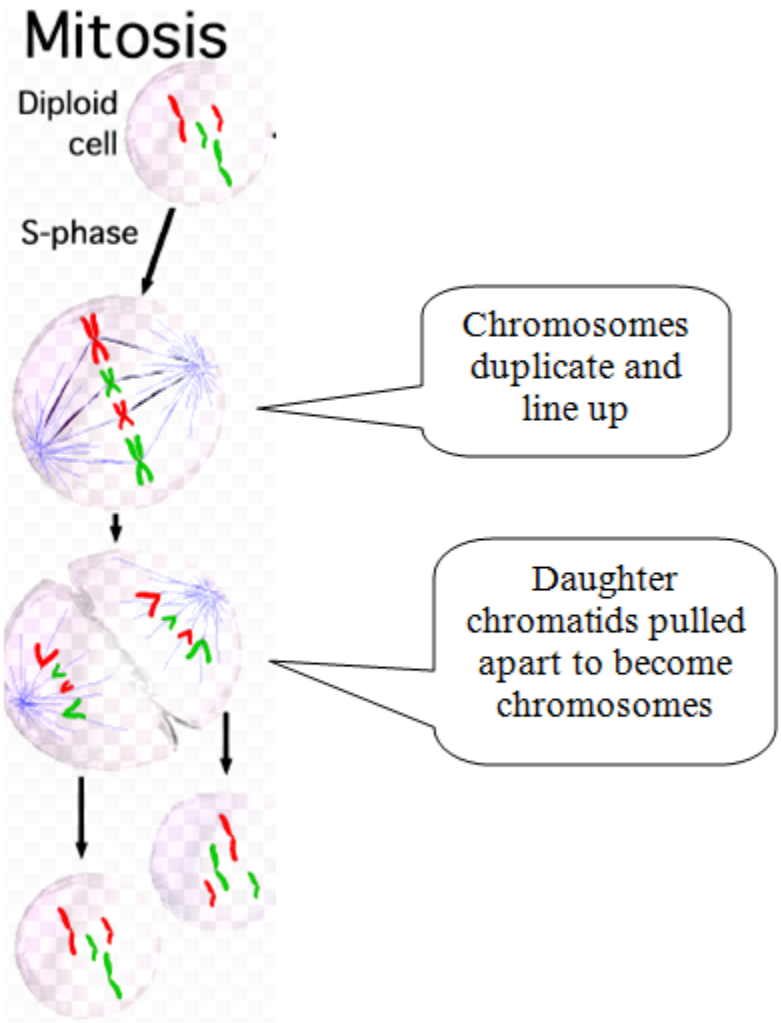
When you work out and build more muscles, you end up with more muscle cells that are duplicates. Even if you are not growing, you have to replace the cells that die from various causes with new cells. To meet both of these needs, an individual cell can divide into two new cells, each of which is identical to the original. They are called ***daughter cells***. This type of duplication is called ***Mitosis***. To accomplish this feat, the chromosomes must be duplicated so they can continue to create the proteins that are necessary in each cell.

During mitosis, each chromosome reproduces itself to form two daughter chromatids that are attached at the centromere.



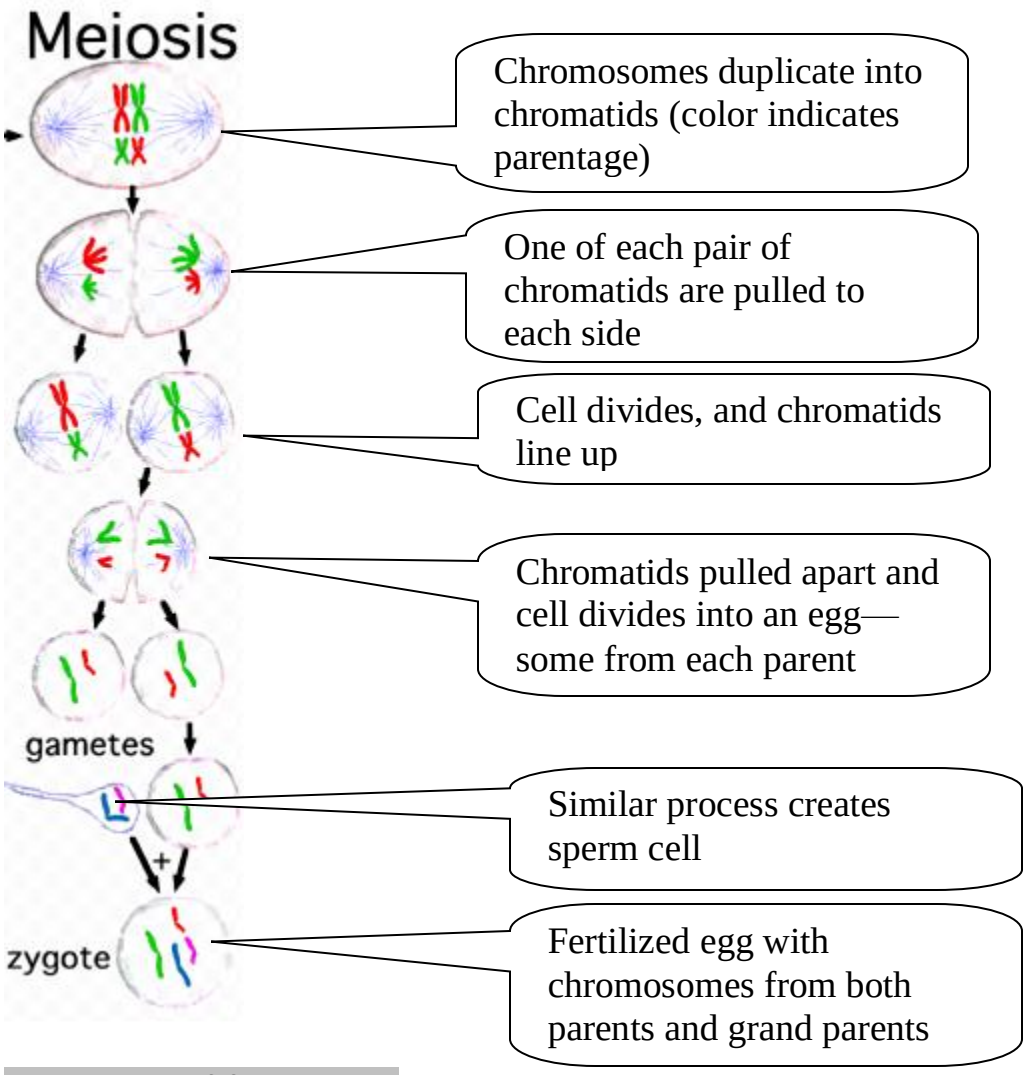
8.4 Chromosome after duplication

Figure 8.5 shows the process of mitosis for two pairs of chromosomes out of the 23 pairs for simplicity. The colors indicate the source of each chromosome—red for the person's father and green for their mother.



8.5 Process of cell duplication

When the human body prepares an egg or sperm, each egg or sperm gets half of the chromosomes—one from each pair. This type of cell division is called ***meiosis*** as shown in Figure 8.6. The egg or sperm has one complete set of 23 chromosomes. When the egg and sperm unite to form a fertilized egg, the resulting embryo has 23 pairs.

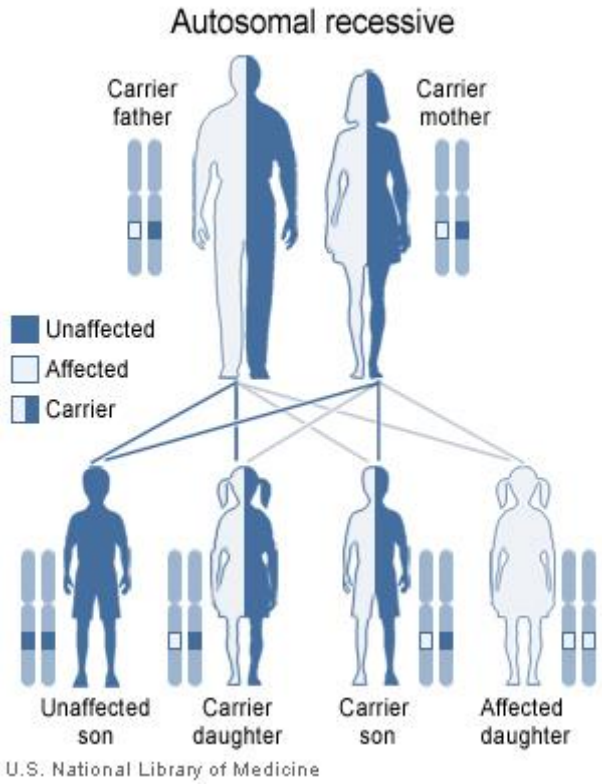


8.6 Process of fertilization

Mutations

Mitosis and meiosis are very accurate methods of making copies and errors are rare but they do occur. Sometimes the changes are induced by outside factors such as ionizing radiation or chemicals that interfere with the copying process. Changes to the chromosomes are called ***mutations***. Because each chromosome contains two duplicate strands, if one of them is damaged, the other can take over. The mutated chromosome can still be copied to daughter cells and if the damage occurred in the ovary of a woman or the testes of a man, the mutation can be passed on to children.

Each cell in our bodies has a complete set of chromosomes. When cells reproduce, they make duplicate copies of the chromosomes. When a woman's body prepares an egg or a man's body prepares a sperm—meiosis—the pairs of chromosomes are separated and the egg or sperm only get one of each pair (23 chromosomes). When the egg is fertilized by the sperm, the 23 chromosomes from the egg and the corresponding 23 chromosomes from the sperm pair up to for 23 pairs of chromosomes. The fertilized egg uses the instructions stored in the chromosomes to make a new person. If a chromosome from one parent is damaged or contains a flaw but the chromosome from the other parent is functional, the person is not affected. We say the problem is ***recessive***. It can be present but not apparent, as shown in Figure 8.7.



8.7 Inheriting a recessive trait

This is why cultures discourage marriage of close relatives to reduce the chance of children receiving recessive flaws from both parents. However if the chromosomes from both parents have the same flaw the child might

suffer from a disease. One of the most common types of inherited diseases in North America is cystic fibrosis.

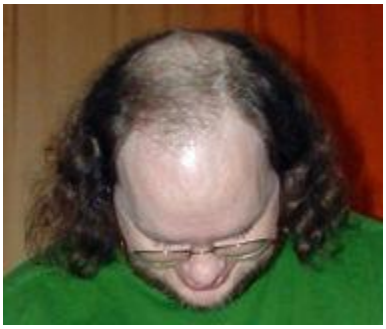
Approximately one out of every 30 people in North America has the cystic fibrosis flaw in one of their chromosomes. Because it is a recessive trait, it does not affect people if they have a healthy chromosome from their other parent. It only becomes a problem if both parents have a flawed chromosome.

Chromosomes and Sex-Linked Traits

Most of the chromosomes are pairs of duplicates except the 23rd pair which determine the child's sex. The chromosome labeled Y in Figure 8.8 is the male chromosome which comes from the father and is shorter than the corresponding X chromosome that comes from the mother.

8.8 Pair #23

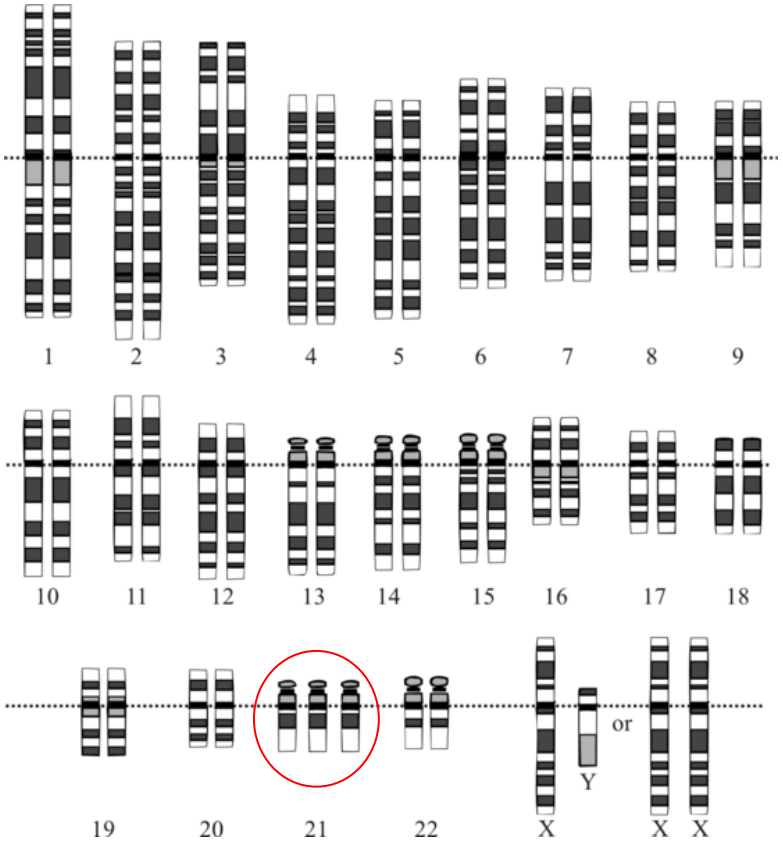
If there is an instruction in the female X chromosome that isn't correct in the section for which there is no backup copy on the male Y chromosome, the child will be affected. The instructions that have to do with color vision are found on the part of the X chromosome that is not present on the shorter Y chromosome. If an error in this part of the chromosome causes color blindness, male children can suffer from it but female children will not unless they inherit it from both parents. For example, approximately 1 in 20 males have a problem distinguishing red from green. Only the females can pass on this trait. If a man's father has typical male-patterned baldness, as shown in Figure 8.9, he doesn't have to worry about inheriting it from his father. However, he should check out his mother's father and her uncles to see if he might inherit it from her.



8.9 Male pattern baldness

Mutations Caused by Error during Meiosis

Chromosomes may be dyed to reveal more detail about their structure. If the pairs of chromosomes are died and arranged side-by-side in pairs, it is easier to spot differences in the pairs. This type of arrangement is called a **keryogram**. Some problems are caused by errors that occur during meiosis. For example, if the chromosomes to not separate correctly, an egg or sperm could have both of one of the chromosomes and then when the egg is fertilized, the embryo has three chromosomes instead of two. For example, children with Downs Syndrome have three strands of chromosome #21 as shown in Figure 8.10.



8.10 Keryogram of person with Downs Syndrome

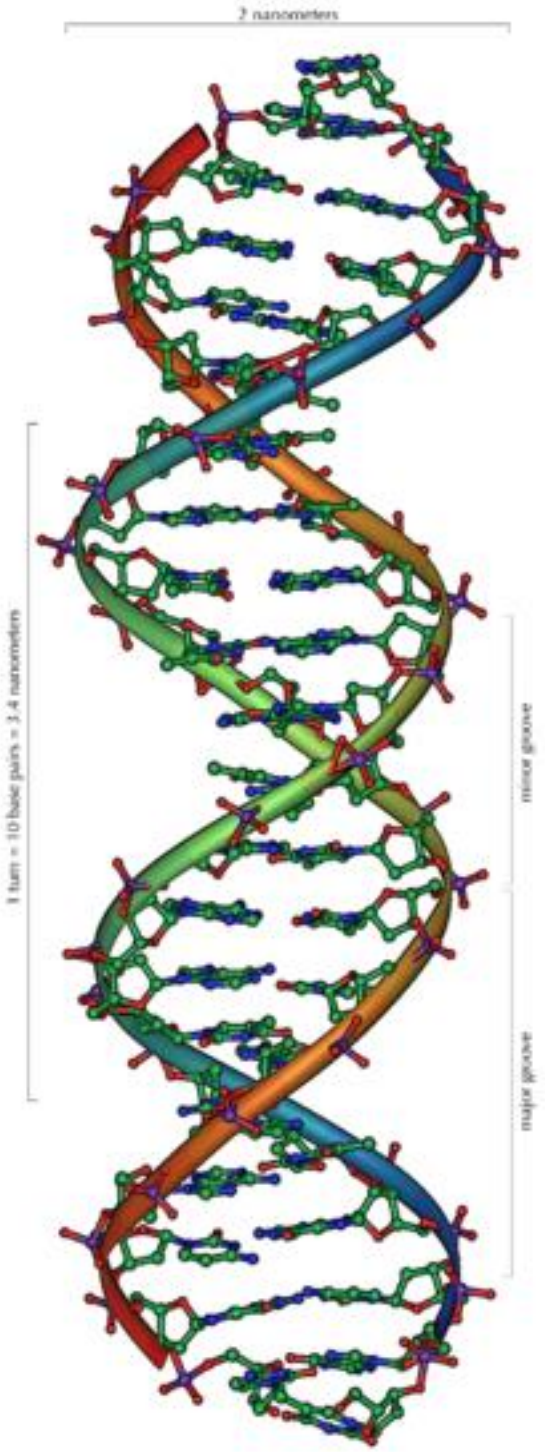
DNA

The structure of the strands of material that make up the chromosomes was not well understood until the mid-1950s when it was described by James Watson and Francis Crick as a double-helix. This was a milestone in

understanding how chromosomes work. Now we know that chromosomes are made of DNA. The DNA molecules are like a computer program that contains the instructions for assembling all the molecules that make up all the cells in the human body. Since this discovery, scientists have been trying to determine how this programming works which begins with the structure of the DNA molecule.

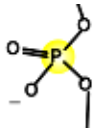
The Structure of DNA

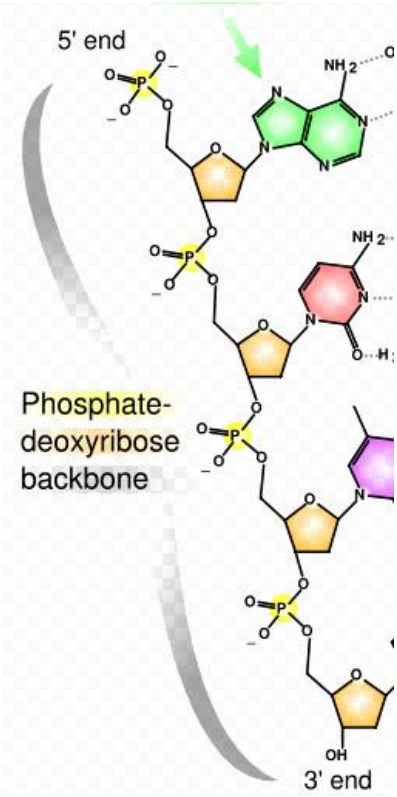
The strand of material that makes up a chromosome turned out to be a very long molecule with a complex shape. It is like a ladder that is twisted. The vertical side supports—the "rails" of the ladder—form spirals that wind around each other and the 'steps' of the ladder are made of two pieces that join in the middle. The name DNA is an abbreviation of its chemical name—**D**eoxyribo**n**ucleic **A**cid. See Figure 8.11.



8.11 DNA

The "Rails" of the DNA "Ladder"

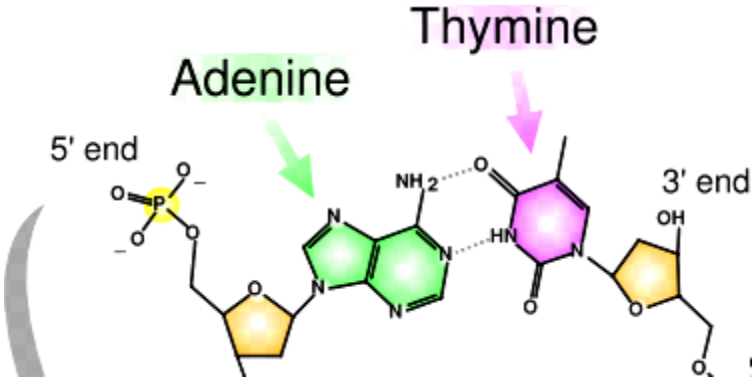
The sides of the DNA molecule that make up the "rails" of the ladder are made of two molecules that alternate. One of them is made of Phosphorous and Oxygen atoms.  The other is made of four carbon atoms and an oxygen atom in a five sided ring.  (the carbon atoms are at the corners and extra hydrogen atoms aren't shown). The "rail" is made by alternating these molecules as shown in Figure 8.12. These two molecules alternate to make up the "rail" on one side of the ladder. The "rail" on one side begins with the phosphate atom—called the ***5 prime end***— and ends with the ring—called the ***3 prime end***. The "rail" on the other side begins with the ring and ends with the phosphate molecule. See Figure 8.12.



8.12 The "rail" of the ladder

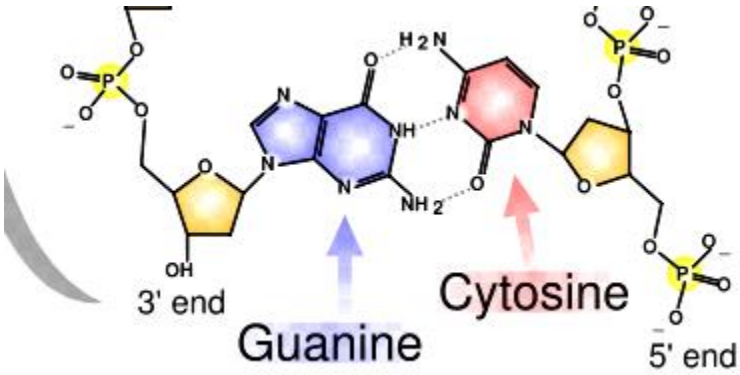
The “Steps” of the DNA “Ladder”

The “steps” of the ladder are made of only four different molecules called ***bases*** and they always come in pairs called ***base pairs***. One type of “step” is the combination of ***Adenine*** and ***Thymine*** as shown in Figure 8.13. The dotted lines between the adenine and thymine indicate weak bonds that come apart easily. For convenience, these two molecules are just referred to as A and T. The step can be AT or TA depending on which side the molecules are on.



8.13 The AT "Step"

The other type of “step” in the DNA “ladder” is the combination of ***Guanine*** and ***Cytosine*** as shown in Figure 8.14. The dotted lines between the Guanine and cytosine indicate weak bonds that come apart easily. For convenience, these two molecules are just referred to as G and C. The step can be GC or CG depending on which side the molecules are on.



8.14 The GC "step"

Role of Amino Acids and Proteins

To understand the role of DNA we need to know a few more things about cells and how they function. ***Proteins*** are complex molecules that can control cell functions or make up the structures of the cells. For example, hemoglobin is a protein that transfers oxygen from the lungs to the cells and insulin is a protein that regulates the use of sugar. There are many different types of proteins. If we continue to use the metaphor of the DNA being like a programmed manufacturing machine, the proteins are the finished products. Proteins are assemblies of 20 basic building block molecules called ***amino acids***. Humans get amino acids from the food they eat which supplies the raw materials for making proteins.

Amino Acid	*Abbreviation*
Alanine	Ala
Arginine	Arg
Asparagine	Asn
Aspartic acid	Asp
Cysteine	Cys
Glutamic acid	Glu
Glutamine	Gln
Glycine	Gly
Histidine	His
Isoleucine	Ile
Leucine	Leu
Lysine	Lys
Methionine	Met
Phenylalanine	Phe
Proline	Pro
Serine	Ser
Threonine	Thr
Tryptophan	Trp
Tyrosine	Tyr
Valine	Val

8.15 Amino acid building blocks

To create proteins, the DNA molecules need to contain directions for specifying any one of those 20 amino acids plus directions for how many of them are needed and in what order to assemble a protein like insulin or hemoglobin.

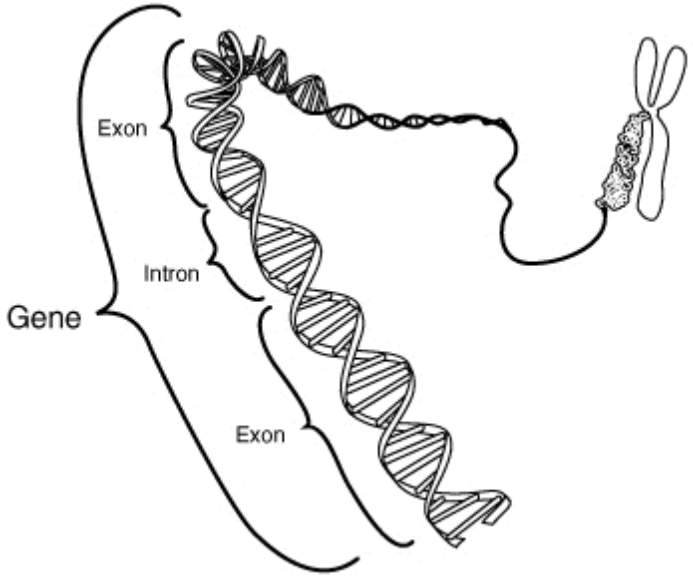
Specifying an Amino Acid

Subsequent research has found that it takes three of the *steps* of the DNA *ladder* to specify one of the amino acids. Each step of the ladder can be one

of four arrangements. If we represent each of the molecules by its first letter, the possibilities are: AT, TA, CG, and GC. If three of these steps are involved, there are 64 possible combinations (4x4x4) which is more than enough to assign each amino acid a code using three *steps* of the ladder. A three-step group that identifies an amino acid is called a ***codon***.

Genes

Some steps on the ladder are used to write instructions for assembling the molecule, some turn the function on or off, and others mark the beginning and end of the instruction. A combination of steps on the DNA ladder that contain all of the instructions necessary for assembling a protein is called a ***gene***. (Note: the definition of a gene is changing as we know more about it. This is a simplified version of the general definition.) The diagram in Figure 8.16 is oversimplified. The set of instructions necessary to assemble a typical protein consist of thousands of steps.



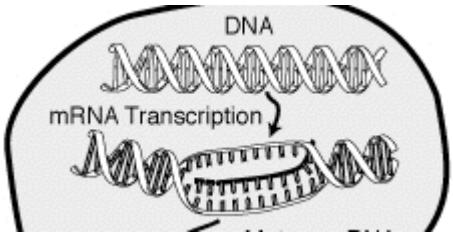
8.16 Genes in DNA in a chromosome

Each cell in our body contains a complete set of genes with all the necessary instructions for creating all of the different types of cells in our bodies. At some point in the development of an embryo, the cells take on specific functions and the genes that contain the code for other functions are turned off. Similarly, when we reach our full size, the genes that control growth reduce their activity to simply replace the cells that are lost. We don't know how this works but it is an area of intense research.

Messenger and Transfer RNA

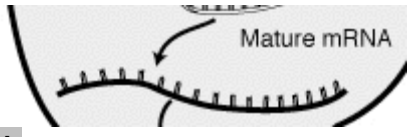
The DNA is the set of master instructions for assembling proteins. The actual assembly is done using another molecule called a ***messenger RNA*** or ***mRNA***.

- Messenger RNA begins as a relatively short piece of the *rail* of a DNA molecule.
- The section of the DNA molecule that contains the code for a particular protein splits at the middle of each step, separating the AT and CG connections and spreads apart to make room for the RNA.
- The RNA, which is like a single rail of DNA, fits between the steps of the split DNA.



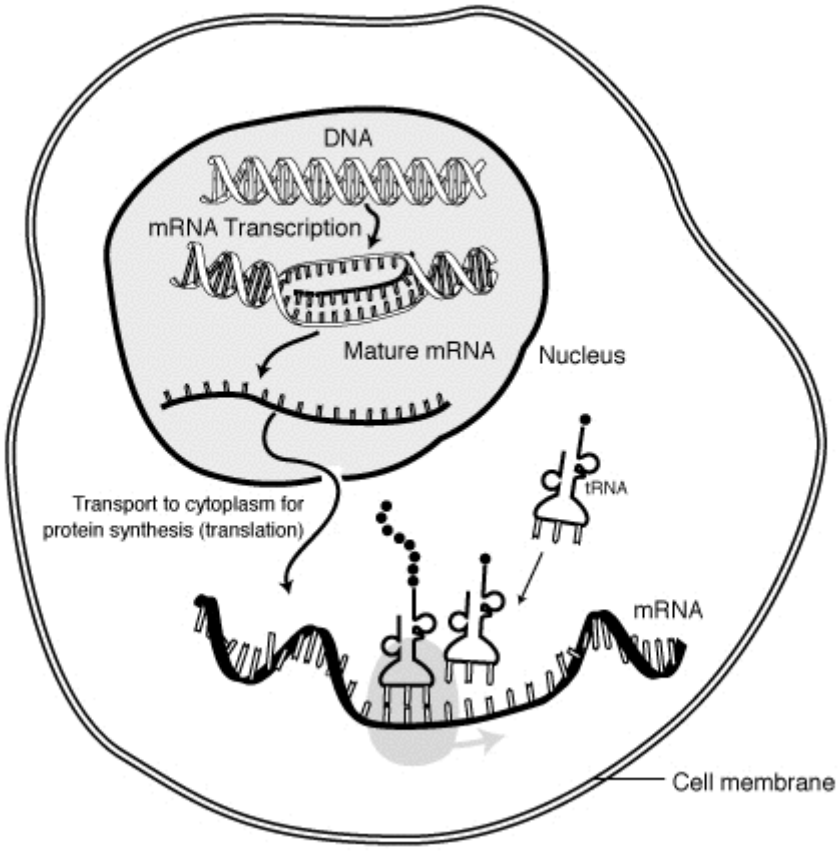
8.17 RNA copies part of the DNA

- The DNA and RNA are surrounded by a soup of available A,T,C, and G molecules. The RNA acts like the other side of the DNA ladder and attracts its own set of A,T,C, or G molecules that match up with one side of the DNA. This process of making a copy of part of the DNA molecule is called ***transcription***.
- The RNA molecule moves out of the DNA molecule. The two sides of the DNA molecule reattach. The RNA molecule is like half a ladder.



8.18 RNA leaves the DNA

- The messenger RNA moves to another part of the cell where the amino acids are waiting to be assembled.
- In the next stage, another molecule named ***transfer RNA*** or ***tRNA*** copies the code from the mRNA, uses it to select an amino acid, and then attaches the amino acid to the protein that is being assembled.



8.19 Transcription process

- When the process is complete, the mRNA breaks down so its parts can be reused

Sequencing Human DNA

Once it was discovered that DNA contained genes that controlled all of the functions of the cell, the obvious question was what part of each chromosome controlled what function? To answer this question, scientists started examining portions of the DNA molecule. They discovered that the task would be enormous.

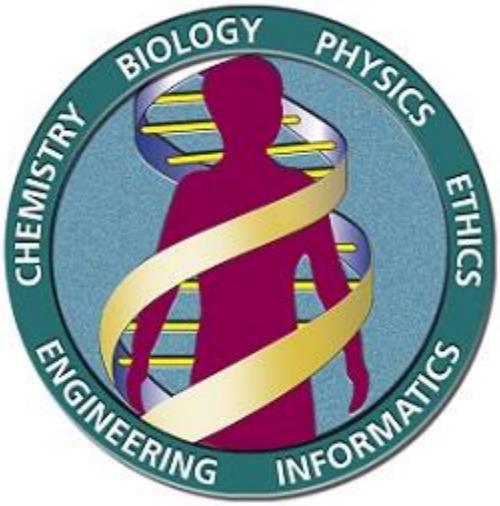
The width of a human hair is about 100,000 nanometers but the width of a DNA molecule is about 2 nanometers and base pairs are spaced about .34 nanometers apart. This is much too small to manipulate directly with the smallest tools. In the 1980s, the methods available for determining the base pairs in the DNA molecule used toxic chemicals that could break the molecule into segments, radioactive tracer elements, and ultra-thin gels.

(Collins, 2006) It would take a university research lab 18 months to determine the sequence of few hundred base pairs (steps on the twisted ladder).

Many diseases are caused by dysfunctional genes. One of the most common in North America is cystic fibrosis (CF). After months of research, scientists were able to say that the dysfunctional gene that was responsible for CF was on chromosome 7 and they had narrowed it down to a segment that was 2 million base-pairs long. Scientists invented new computerized search techniques but it still took four more years to finally identify the one base-pair that was different in people who had the disease. To make this discovery, two dozen teams around the world worked for ten years at a cost of $50 million. A cure hasn't been found yet, but they can test people to see who has the defective gene as a recessive.

Like many research projects, they discovered useful information along the way in addition to what they originally sought. They found that significant amounts of the DNA molecule don't have a discernable function. Mutations in those sections don't appear to cause any problems. They also found that the definition of a gene was not as simple as they thought. One gene could create several different but similar proteins. (Collins, 2006)

To decode all of the human DNA would be a much bigger challenge. There are approximately 3 billion base pairs in human chromosomes that make up the complete instruction manual for how to make a human—known as the ***human genome***. The ***human genome project (HGP)*** was founded in 1990 U.S. National Institutes of Health in cooperation with the Wellcome Trust in London and labs in France, Germany, China, and Japan. It was expected to take 15 years and cost $3-billion. The head of the project was James D. Watson who was one of the discoverers of the DNA structure.



8.20 Human Genome Project logo

One of the issues that arose early in the project was ownership of the information. Watson and most scientists thought the information should be public knowledge but the U.S. was allowing private companies who were doing similar research to patent the information. Recall that the court ruling on ownership of oil beneath the ground—the rule of capture—had a serious negative impact on the development of oil resources. Many scientists were concerned that allowing private companies to claim ownership of the knowledge about the human genome would restrict development. Proponents of patents argued that private companies could discover this information faster and cheaper. To illustrate this point, one of the private companies—Celera Genomics owned by Craig Venter—declared that it would pursue the same goal and would do it faster and for a tenth of the cost. (Human Genome Project, 2008) <Link>

Advances in computer technology were greater than the teams expected and the competition between the public and private efforts spurred both of them to greater efforts. Instead of taking 15 years, they were able to meet their target of mapping more than 80% of the genome in only ten years. The achievement was announced jointly on June 26, 2000 by President Clinton and Prime Minister Blair. Since that time, they have worked on some of the more difficult sections and have about 92% done. Two areas of the DNA are not done. The section at the intersection of the two strands of DNA—the ***centromere***—is still unmapped and the ends of the DNA strands—***telomers***—appear to be very repetitive and they are not well mapped. Scientists found that there are about 22,000 to 23,000 genes in the human genome.

The U.S. patent office has granted patents on more than 20 percent of the human genes. (Lovgren, 2008) <Link> Critics of this practice claim that it is limiting the use of these discoveries and discouraging research on cures for disease. (Crichton, 2007) <Link> Bills have been introduced in the U.S. House of Representatives in 2002 and in 2007 to ban this practice but they did not get out of committee and were not even scheduled for debate. (H.R. 977: Genomic Research and Accessibility Act, 2007)

8.21 U.S. Patent office logo

Advances in computer and laboratory technologies have reduced the price of performing a DNA sequence dramatically. Instead of costing $3billion dollars and taking fifteen years, it can now be done for less than $10,000 in days. Because of the lower price, new studies can be done such as the one for which they mapped the DNA of one thousand people who were more than 100 years old to see if they had any special genes that would account for their long lives. (Smith, 2010) <Link>

Glossary

Term	Description
3 prime end (3')	The end of the DNA backbone that ends with a ring of four carbons and an oxygen molecule
5 prime end (5')	The end of the DNA backbone that ends with a phosphate molecule
adenine	One of the molecules that comprise a “step” of the DNA

	ladder that always combines with thymine
bioinformatics	field of study that combines biology and computer science
chromosomes	literally means colored bodies—they contain DNA
codon	a group of three base-pairs that can identify an amino acid
cytosine	One of the molecules that comprise a "step" of the DNA ladder that always combines with guanine
daughter cells	duplicate cells created by mitosis
guanine	One of the molecules that comprise a "step" of the DNA ladder that always combines with cytosine
karyogram	arrangement of chromosomes
mitosis	process of one cell dividing into two duplicate cells
mRNA	messenger RNA is a molecule that transcribes a copy of part of the DNA instruction
recessive	inherited trait that does not affect the child unless it is inherited from both parents
thymine	One of the molecules that comprise a "step" of the DNA ladder that always combines with adenine
transcription	copying part of the DNA molecule to make a protein
tRNA	transfer RNA is a molecule that uses the code on the mRNA to assemble amino acids into proteins

References

Cell nucleus. (2008, March 6). Retrieved March 8, 2008, from Wikipedia: http://en.wikipedia.org/wiki/Cell_nucleus

Chimpanzee genome project. (2008, February 14). Retrieved March 8, 2008, from Wikipedia: http://en.wikipedia.org/wiki/Chimpanzee_genome_project

Chromosome. (2008, March 7). Retrieved March 8, 2008, from Wikipedia: http://en.wikipedia.org/wiki/Chromosomes

Collins, F. S. (2006). *The Language of God*. New York: Free Press.

Crichton, M. (2007, February). *Patenting Life*. Retrieved March 9, 2008, from Michael Crichton: The Official Site: http://www.crichton-official.com/essay-nytimes-patentinglife.html

H.R. 977: Genomic Research and Accessibility Act. (2007, March 1). Retrieved March 11, 2008, from GovTrack.us: http://www.govtrack.us/congress/bill.xpd?bill=h110-977

Human Genome Project. (2008, March 9). Retrieved March 9, 2008, from Wikipedia: http://en.wikipedia.org/wiki/Human_Genome_Project

Lovgren, S. (2008, October 13). *One-Fifth of Human Genes Have Been Patented, Study Reveals*. Retrieved March 9, 2008, from National Geographic News: http://news.nationalgeographic.com/news/2005/10/1013_051013_gene_patent.html

Smith, M. (2010, July 1). *Longevity? It's in the genes for centenarians*. Retrieved July 24, 2010, from ABC News / Health: http://abcnews.go.com/Health/Wellness/longevity-genes-centenarians-live-100/story?id=11065380

Last edited on June 14, 2016