What Is This DNA Stuff and Why Should I Care?
Part 1: Cell Structure and DNA

David R. Neithamer

DNA is an acronym for deoxyribonucleic acid. It is a very large bio-polymer found in the cells of all living things on earth including plants, animals, fungus, bacteria, and viruses. It is the blueprint for life as it holds the instructions for your body to grow, heal, and function. The longest strand of human DNA is over 247 million base pairs long and the shortest strand is over 49 million base pairs. I will discuss base pairs in just a moment, but think of them as links in a chain, or better yet, rungs on a ladder. That's a very long chain or an extremely tall ladder! The combined human DNA in each cell is over 3 billion base pairs and if you were to remove all the DNA from one of your cells and stretch it end to end it would be about 6 feet in length. Human DNA is about 99.9% similar to all other human beings - that's what makes us human. That 0.1% difference in our DNA, which is about 3 million base pairs, is what makes us all individuals with a wide variety of physical, physiological, and behavioral differences. In essence, it is what makes each of us unique.

DNA has a double helical structure (***A***) as shown on the left side of **Figure 1**. If you unwind the double helix you will see the corresponding ladder structure of DNA shown in the middle (***B***). The unique double helical structure of DNA was discovered in 1953 by James Watson and Francis Crick, based on X-ray crystallography experiments by Rosalind Franklin, and marked a milestone in the history of science. Watson and Crick received the Noble Prize in Physiology or Medicine in 1962 for their discovery.

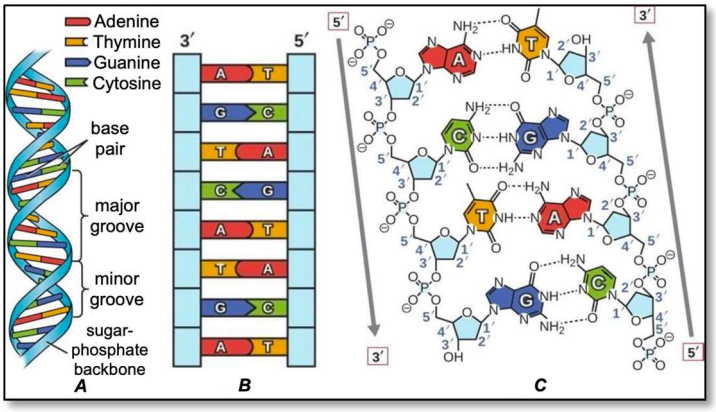

Figure 1. Double Helix Structure, Ladder View, and Detailed Chemical Structure
Source: https://openstax.org/books/microbiology/pages/10-2-structure-and-function-of-dna

The actual chemical structure of the DNA molecule (***C***) is shown on the right side of Figure 1. Two backbone chains of alternating deoxyribose sugar units (S) and phosphate groups (PO_4) form the uprights of the ladder and the rungs of the ladder are formed by paired linkages of 4 nitrogen-containing bases called nucleotides with the names adenine (A), thymine (T), cytosine (C), and guanine (G). The name nucleotide "base pair" is given to each paired combination. Nucleotide A always pairs with T and nucleotide C always pairs with G because of the matched hydrogen bonding between pairs: A&T can only form 2 hydrogen-bonds (dashed lines) whereas C&G form 3 hydrogen-bonds to each other. Understanding the basic chemical structure of DNA helps you to better understand some of the jargon used later in this series of articles.

Earlier I said that DNA is found in the cells of all living things, but where is it actually found in the cell and is all DNA the same? **Figure 2** shows a picture of a typical animal cell, and the DNA is only found in two locations within the cell - the nucleus (purple sphere) which is the control center for the cell and the mitochondria (orange-red structures) which are the powerhouses for the cell. The DNA found in the nucleus is called nuclear DNA and it is sometimes referred to as nuDNA, whereas, the DNA found in the mitochondria is called mitochondrial DNA and is referred to as mtDNA.

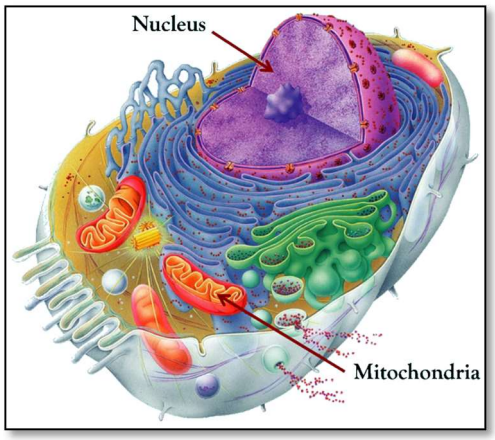

Figure 2. Structure of a Typical Animal Cell
Source: Science - The Folly of Human Conceits Copyright @ 2009 Pearson Education, Inc., publishing as Pearson Benjamin Cummings

Nuclear DNA does not just float around in the nucleus as long, twisted ladders, it ultimately self-assembles into structures called chromosomes. **Figure 3** shows how the DNA strands wrap around proteins called histones to form small, cyclindrical structures which then self-assemble. Think of histones as sewing spools with the DNA wrapped around the spool like threads. The combined histone/DNA structures are called a nucleosomes or a chromatins. These histone "spools" condense the DNA and ultimately pack it into the larger chromosomal structures. Packing the DNA into these chromatin structures not only allows the DNA to fit neatly into the nucleus of the cell, just like storing spools of thread in a sewing box, but it also protects the DNA strands from unintended cleavage and break-down.

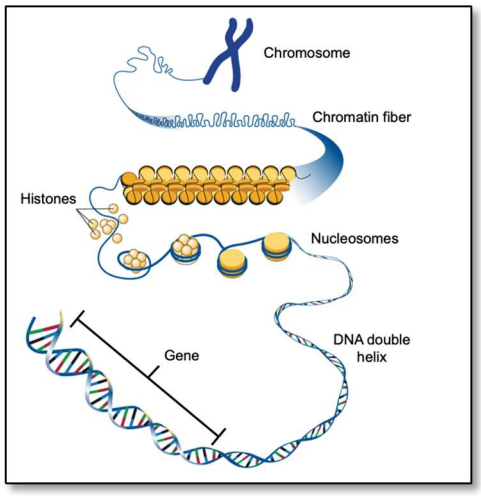

Figure 3. DNA Structure and Self-Assembly into a Chromosome
Source: Phys.org https://phys.org/news/2013-01-aging-cells-dna-rogues.html

Chromosomes normally have an X-shape with a single connection point, called the centromere, half-way to two-thirds of the way along the chromosome. The two short ends are called the P-arms and the long ends are called the Q-arms. This X-shaped chromosome is actually a paired combination of two chromosomes - one chromosome comes from the father and the other chromosome comes from the mother. Humans have a genome with 46 chromosomes, 23 each from the father and mother. As shown in the karotype in **Figure 4**, our genome contains 22 pairs of autosomes that have the same paired morphology (size and shape). Autosomal pairs are labeled with numbers (1–22) roughly in order of their sizes based on their number of base pairs. Nuclear DNA found on anyone of these 22 pairs of chromosomes is called autosomal DNA and is referred to as atDNA. Our genome also contains 1 allosome pair of sex chromosomes. Allosomes are labeled with the letters X or Y and the DNA that comes from these chromosomes is called X-DNA or Y-DNA respectively. I will talk more about the chromosomes in Parts 3-6 when I discuss the different types of DNA tests (atDNA, Y-DNA, and X-DNA).

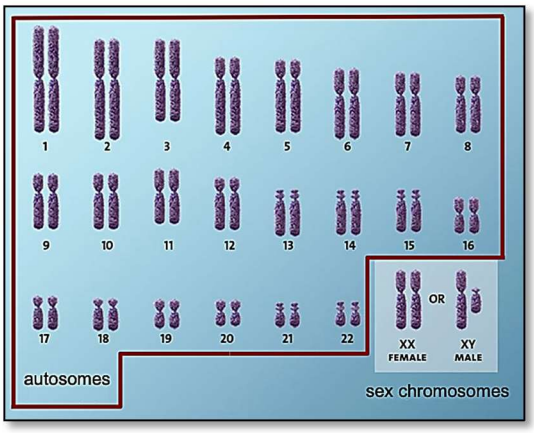

Figure 4. Karotype of 46 Human Nuclear DNA Chromosome Source
***Source:* *https://medium.com/analytics-vidhya/the-last-explanation-you-will-ever-need-to-answer-this-conditional-probability-question-f39f8e93086c*.**
Analytics Vidhya Article by Gal Gilor. Fair Use.

The second type of DNA is found in the mitochondria of the cell and is referred to as mtDNA. Human mtDNA forms unique circular molecules, like a twisted rubber ladder with its ends connected to itself, that contain 16,569 base pairs **(see Figure 5**). Mitochondria produce energy for the cell using oxygen and glucose to create adenosine triphosphate (ATP), the cell's main energy source. In addition, they help regulate the self-destruction of cells and are necessary for the production of cholesterol and heme (a component of hemoglobin). Human cells average about 100 mitochondria per cell with a range of 10-2000 per cell. Each human mitochondrion within a cell contains about 5 mtDNA molecules with a range of 1-15. Therefore, on average, each human cell contains about 500 mtDNA molecules.

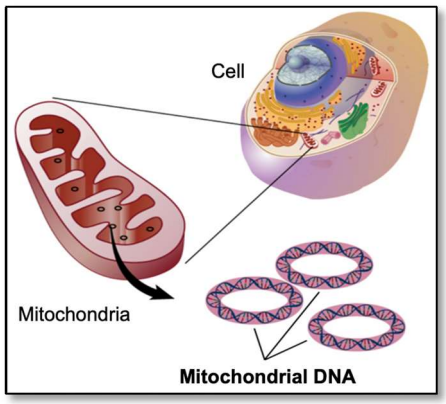

Figure 5. Mitochondrial DNA
Source: Wikimedia Commons, the free media repository.
https://en.wikipedia.org/wiki/Mitochondrial_DNA

We have all heard the saying, "It's all in our genes" when discussing the personality traits or physical characterists of a person, but what does that mean? What are genes? Very simply, genes are specific regions of DNA found along the DNA strands, as shown previously in Figure 3, that code for the production of proteins. In other words, genes are regions along the DNA strands that control the production of all the proteins in your body. Proteins are large, complex molecules that are responsible for nearly every task of cellular life, including cell formation, shape, movement, environmental response, DNA maintenance, and regulation of cellular growth. Basically, they are responsible for everything a cell can do. Examples of proteins are hormones such as insulin and estrogen and enzymes such as lipases and oxidases.

Although we do not know exactly how many genes humans have in their DNA, there are about 20,000-25,000 human genes that have been identified or are suspected. Most of the genes are found in the nuclear DNA; only 37 protein-coding genes are found in the mitochondrial DNA. Genes are typically between 1,000-38,000 base pairs long and account for approximately 1.5% of our total DNA. The remaining regions of our DNA (about 98.5%) that do not code for protein production are called "Junk DNA". The term "Junk DNA" is misleading, however, as recent evidence suggests that it

contains "functional elements" that are necessary for regulating the genes. For us, the "Junk DNA" is the most interesting part of the DNA because this is where most of the genealogical and forensics testing occurs.

Throughout our lifetime the cells in our bodies are constantly replenishing themselves through a process called mitosis. As shown at the top of **Figure 6**, mitosis is the normal division of cells to create two identical diploid "daughter" cells, which have the same DNA and same 46 chromosomes as the parent cell. (Note: for clarity in the figure, only one pair of chromosomes is shown in the parent cell for mitosis cell division). Any errors that occur during mitosis can result in mutations or cell death, however, these are not important from a genealogical standpoint because these cells are not passed down to any offspring.

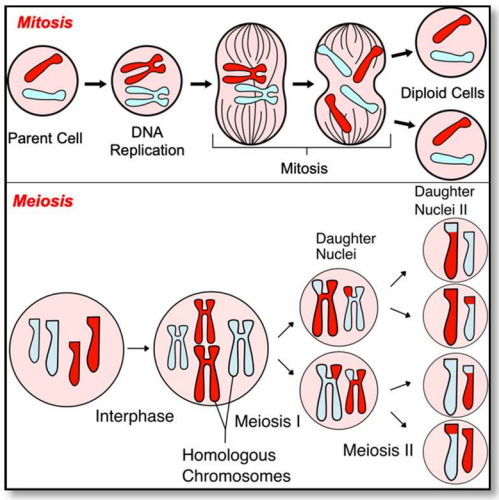

Figure 6. Normal Diploid Cell Division (Mitosis) and Haploid Sex Cell Formation (Meiosis).
Source: Wikimedia Commons, the free media repository.
https://commons.wikimedia.org/wiki/File:Major_events_in_mitosis.svg
https://commons.wikimedia.org/wiki/File:Meiosis_Overview_NL.svg

When the body creates the sex cells, sperm in men and ova (egg cells) in women, it uses a cell division process called meiosis that reduces the chromosome number from 46 to 23, creating four daughter cells, each genetically distinct from the parent cell that gave rise to them (see the bottom of Figure 6). (Note: for clarity in the figure, only two pairs of chromosomes are shown in the parent cell for meiosis cell division). During meiosis, the chromosomes pair with each other and undergo genetic swapping or reshuffling of their DNA called genetic recombination. This genetic recombination occurs only in atDNA and X-DNA; it does not occur in Y-DNA or mtDNA, which are passed to their offspring unchanged. Recombination is what leads to the great variety observed between offspring from the same parents. Any mutations that occur during meiosis are important because they are passed down from parent to child.

When we talk about mutations, they are not the cause of mutant half human-half fly creatures - that is only found in Hollywood. Mutations are naturally occurring changes to the DNA strand (or mtDNA circle) over time. All DNA undergoes mutation, this is the basis for evolution, but the rates of mutation are variable; some nucleotides mutate after 10 generations while others mutate after 200 generations. The three most common types of DNA mutations are: Insertion, where an extra base pair is added to the DNA strand, Deletion, where a base pair is removed from the DNA strand, and Substitution, where one of the nucleotides in a base pair is replaced by another nucleotide (i.e. if a C is replaced by a T or an A is replaced by a G). Most mutations are benign because they occur in the "Junk DNA" regions, but this can actually be very useful in identifying family connections if two individuals contain the same mutation. However, some mutations can lead to disease or abnormalities if they originate in the genes, or protein coding regions. Mutations of this type lead to diseases such as cystic fibrosis, sickle cell anemia, Tay-Sachs, and breast cancer (brca1/brca2).

We have discussed the basic structure of our cells, where DNA is found in the cells, and the structure of DNA itself. The next AFAOA newsletter article will focus on how the DNA testing companies actually test your DNA. How they use your saliva sample to identify your genetic cousin matches is rather fascinating. Until then, happy genealogical and genetic hunting!

