Basics of Molecular Biology

Your DNA is like the assembly instructions that came with the toys you bought for your children last Christmas and hastily put together after they went to bed on Christmas Eve. However, the DNA is vastly more complex (hard to believe, huh?) Inside almost every cell of your body is a central structure called a nucleus. In that nucleus are 23 pairs of chromosomes. One of each pair came from each of your biologic parents. A chromosome consists of a very long, double stranded molecule of DNA (deoxyribonucleic acid). Every link of the chain on both strands is connected to one of four possible bases: adenine (A), thymine (T), guanine (G) and cytosine (C). Each of these bases is connected to a corresponding base on the opposite chain by a non-covalent bond, hence the double-stranded nature of DNA. Each of the two linked bases is known as a base pair. This double strand is then twisted into a spiral, yielding the double-stranded helix structure that Dr. James Watson and Dr. Francis Crick discovered in 1953 and for which they received the Nobel Prize in Medicine in 1962. The length of the DNA in each of the human chromosomes ranges from 50 million base pairs in the shortest to 250 million base pairs in the longest. The human genome (all the chromosomes together) contains as many as 3 billion base pairs.

Along the DNA chain, we can identify various runs of base pairs that work together. We call this unit a gene and assign a name to it such as ACTN3, MMP3 or ALDH2. A set of all 46 human chromosomes contains approximately 25,000 genes. While the DNA itself does no work, many of these genes contain the instructions for the cell to produce an enzyme or other protein to function in the body. These proteins may be part of a muscle cell or ligaments, a blood component, brain tissue, sensors in the wall of blood vessels, etc. That protein may also be an enzyme that facilitates various reactions in the body such as digesting food, eliminating toxins, processing and storing fats, etc.

The DNA is able to code for a protein because each individual triplet of base pairs codes for one specific amino acid out of the twenty naturally occurring amino acids. Amino acids are the building blocks of all proteins. There is a complex process in the cell that builds proteins, amino acid by amino acid. The sequence of amino acids and the length of the protein molecule are dictated by the blueprint of base pairs found in the DNA.

Sometimes an “error” occurs and one of the bases on the DNA chain is substituted by another base. This substitution, if not repaired by natural processes found in the cell, can result in the wrong amino acid being incorporated into the protein at that particular position. Sometimes the substitution causes no noticeable difference. Sometimes it causes the protein to function differently, either more effectively, less effectively or sometimes the protein will not function at all. Depending on when and where this substitution occurs, the change might be passed down to the offspring and continue on down the line. These “substitutions” are actually not uncommon. The good news is that the majority of them have little to no effect on the individual.

The purpose of molecular genetic testing is to identify specific base pairs at certain positions on specific genes. It then becomes the goal of clinical researchers to determine which substitutions are significant and how so.