



Numbers in Life: A Statistical Genetic Approach

Obubu Maxwell*

Department of Statistics, Nnamdi Azikiwe University, Awka, Nigeria

Ikediwa Udoka Chinedu

Department of Statistics, Nnamdi Azikiwe University, Awka, Nigeria

Anabike Charles Ifeanyi

Department of Statistics, Nnamdi Azikiwe University, Awka, Nigeria

Abstract

In this paper, we briefly reviewed the numbers in life from a statistical genetic approach. The human genome comprises of 6 billion chemical bases of DNA. The DNA encodes 30,000 genes. It consists of two parts; the nuclear genome; which consists of 3,200,000,000 nucleotides of DNA, divided into 24 linear molecules, the shortest 50,000,000 nucleotides in length and the longest 260,000,000 nucleotides, each contained in a different chromosome and the mitochondrial genome; which contains approximately 16,600 base pairs encoding 37 genes. Most human cells have 46 chromosomes. However, the number of chromosomes in the nuclei of a person with Down syndrome is 47. The DNA of any two people on Earth is 99.6 percent identical, the 0.4 percent variation represents about 20 million base pairs. Almost all 98 percent of the human DNA is noncoding, while in bacteria, only 2% of the genetic material does not code for anything.

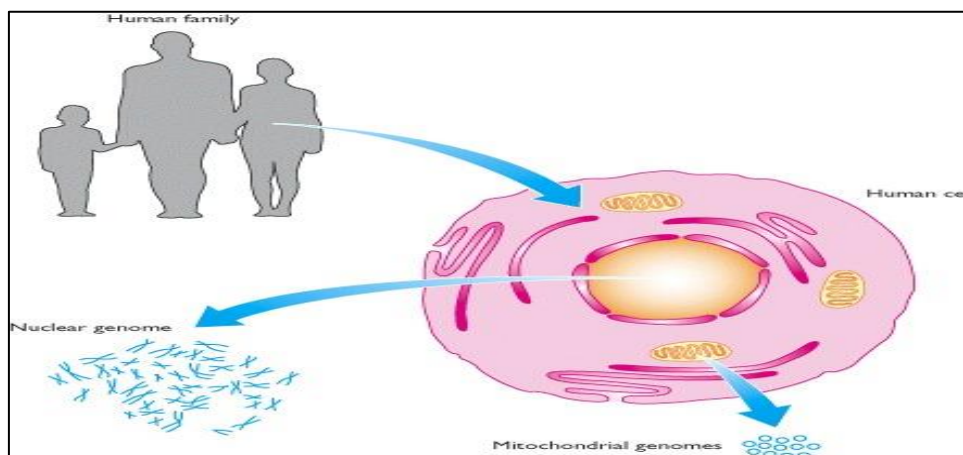
Keywords: Numbers in life; Statistical genetics; DNA; Chromosomes; Base pair; Double helix; Down syndrome; Genes.



CC BY: [Creative Commons Attribution License 4.0](https://creativecommons.org/licenses/by/4.0/)

1. Introduction

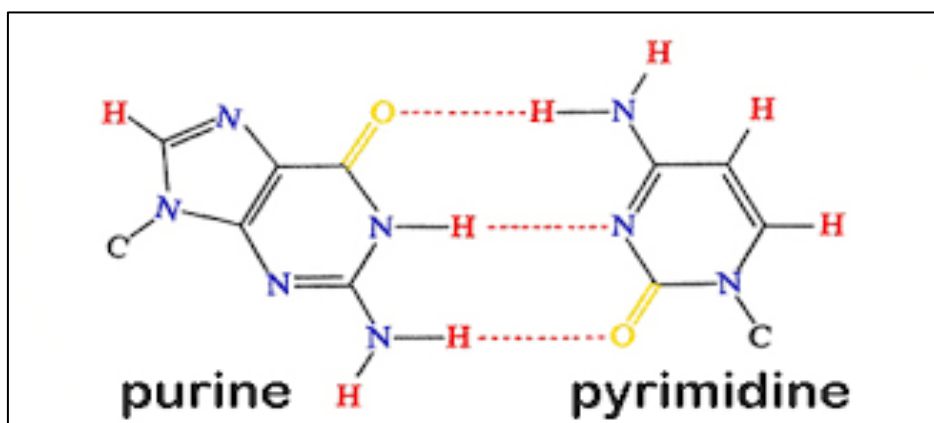
Life as we know it is specified by the genomes of the myriad organisms with which we share the planet. Every organism possesses a genome that contains the biological information needed to construct and maintain a living example of that organism. Most genomes, including the human genome and those of all other cellular life forms, are made of DNA (deoxyribonucleic acid) but a few viruses have RNA (ribonucleic acid) genomes. DNA and RNA are polymeric molecules made up of chains of monomeric subunits called nucleotides. The human genome, which is typical of the genomes of all multicellular animals, consists of two distinct parts (Figure 1): The nuclear genome, and the mitochondrial genome. It is comprised of 6 billion chemical bases (or nucleotides) of DNA packaged into two sets of 23 chromosomes, one set inherited from each parent. The DNA encodes 30,000 genes. The 30,000 genes are usually present in two copies. It is made up of 23 chromosome pairs with a total of about 3 billion DNA base pairs. There are 24 distinct human chromosomes: 22 autosomal chromosomes, plus the sex-determining X and Y chromosomes. Chromosomes 1-22 are numbered roughly in order of decreasing size. Somatic cells usually have one copy of chromosomes 1-22 from each parent, plus an X chromosome from the mother and either an X or Y chromosome from the father, for a total of 46. There are estimated 20,000-25,000 human protein-coding genes. The nuclear genome; The DNA molecules present in the nucleus of a eukaryotic cell, comprises approximately 3,200,000,000 nucleotides of DNA, divided into 24 linear molecules, the shortest 50,000,000 nucleotides in length and the longest 260,000,000 nucleotides, each contained in a different chromosome. The mitochondrial genome is a circular DNA molecule of 16,569 nucleotides, multiple copies of which are located in the energy-generating organelles called mitochondria. Each of the approximately 10^{13} cells in the adult human body has its own copy or copies of the genome, the only exceptions being those few cell types, such as red blood cells, that lack a nucleus in their fully differentiated state. In this research, we qualitatively examine the Numbers in life from a statistical genetic approach.



2. The Numbers

2.1. Base Pair

A base pair is a unit consisting of two nucleobases bound to each other by hydrogen bonds. They form the building blocks of the DNA double helix and contribute to the folded structure of both DNA and RNA. The human genome is made up of *3.2 billion* base pair. The Marble lungfish (*Protopterus aethiopicus*) also known as the leopard lung fish, found in Africa has about *133 billion* base pairs in their genome, thus making it the species with the most base pair.



The Marble lungfish



2.2. Genes

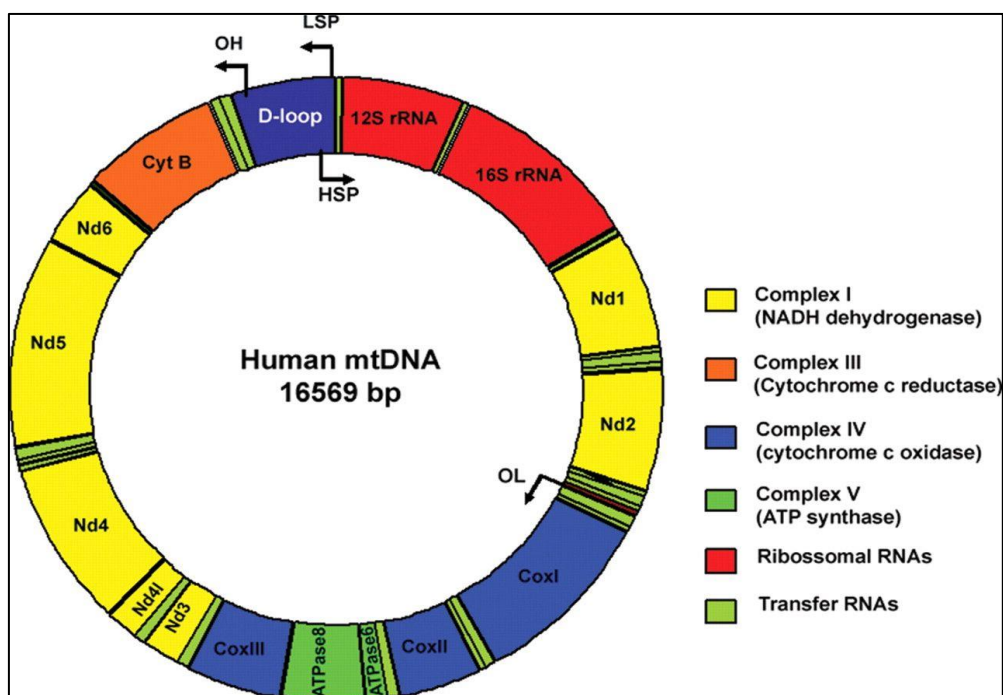
A gene is the basic physical and functional unit of heredity. Genes are made up of DNA. Some genes act as instructions to make molecules called proteins. However, many genes do not code for proteins. The Human Genome Project estimated that humans have between *20,000* and *25,000* genes. Every person has two copies of each gene, one inherited from each parent. Most genes are the same in all people, but a small number of genes (less than 1 percent of the total) are slightly different between people. Alleles are forms of the same gene with small differences in their sequence of DNA bases. These small differences contribute to each person's unique physical features. Scientists keep track of genes by giving them unique names. Because gene names can be long, genes are also assigned symbols, which are short combinations of letters (and sometimes numbers) that represent an abbreviated version of the gene name. For example, a gene on chromosome 7 that has been associated with cystic fibrosis is called the cystic fibrosis transmembrane conductance regulator; its symbol is CFTR. Our genes provide cells with

information on how to make proteins. Scientists have estimated that humans may produce up to *100,000* proteins, so they thought there were about as many human genes. Today, they know that some genes contain the code for making multiple proteins.



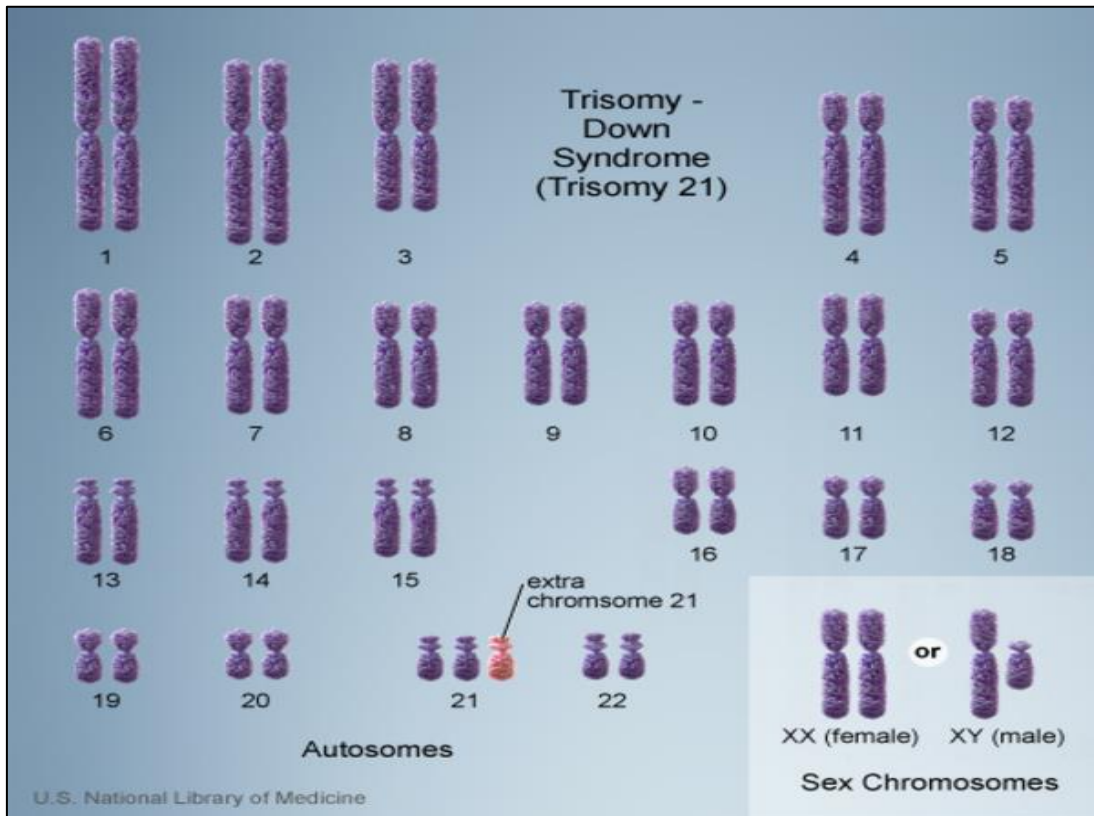
2.3. The Mitochondrial Genome

The mitochondrial genome is the full genetic complement of a mitochondrion. Mitochondrial DNA is only a small portion of the total DNA of a eukaryotic cell and in most species is solely inherited from the mother. In humans mitochondrial DNA contains approximately 16,600 base pairs encoding 37 genes. Mitochondria are the cell's power plants, and many of their genes are involved in production of cellular energy. They have their own set of genes because they are thought to have evolved from bacteria that were engulfed by eukaryotic cells (cells containing a nucleus) some 1.5 billion years ago, during the Precambrian period.



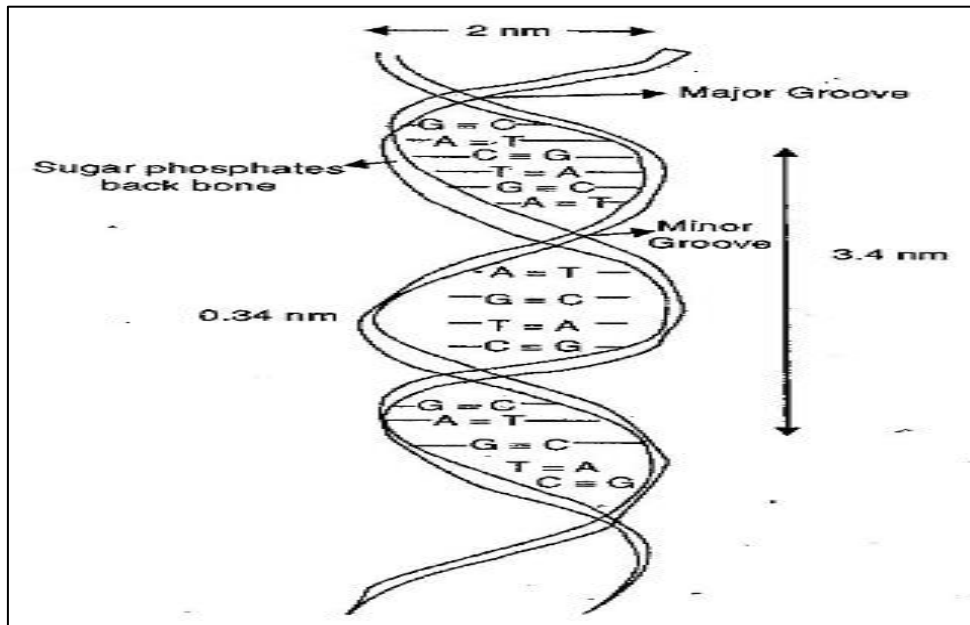
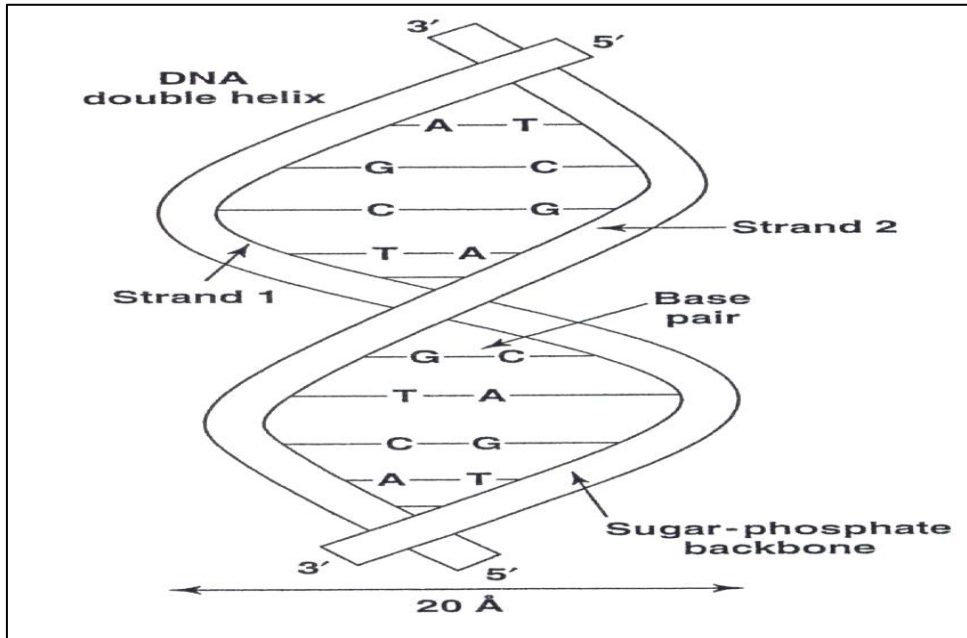
2.4. Down's Syndrome

Down's syndrome is a genetic disorder caused when abnormal cell division results in extra genetic material from chromosome 21. Down's syndrome causes a distinct facial appearance, intellectual disability and developmental delays. It may be associated with thyroid or heart disease. The number of chromosomes in the nuclei of a person with Down syndrome is therefore **47**. Most human cells have **46** chromosomes, but occasionally, a glitch in cell division results in a cell with too few or too many chromosomes. When this happens in egg or sperm cells, the child can have an abnormal number of chromosomes. People with Down syndrome have an extra copy of chromosome 21, one of the smallest chromosomes in the genome.



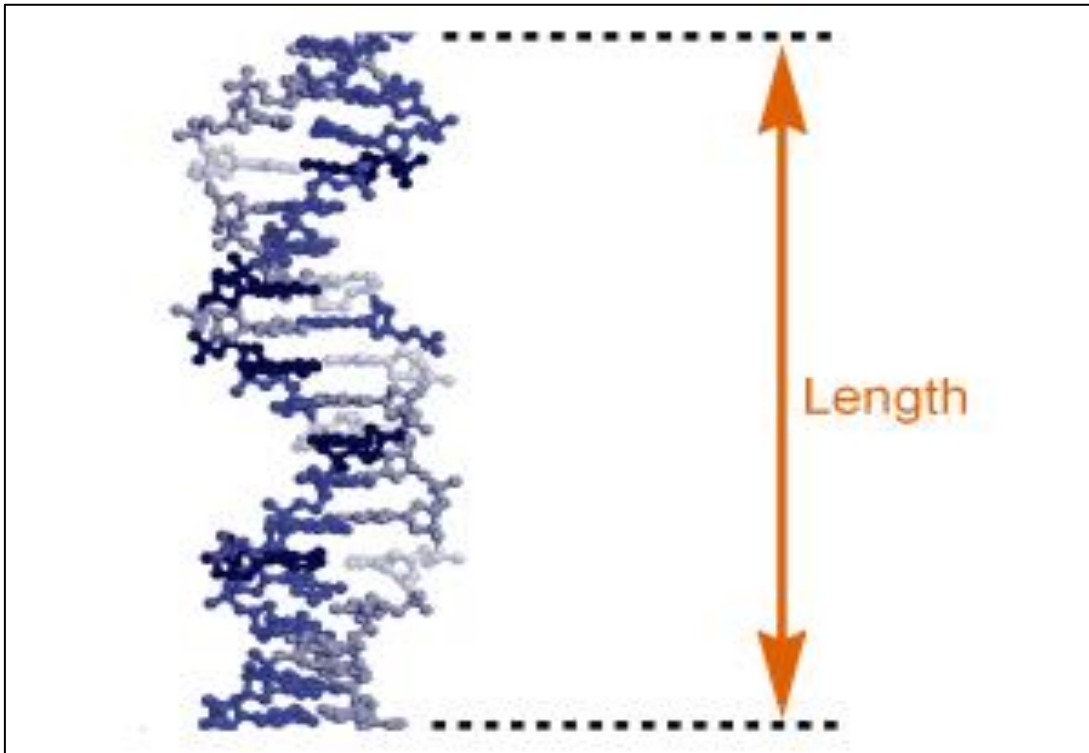
2.5. Double Helix; 1953

Double helix is a term used to describe the structure of DNA. A DNA double helix consists of two spiral chains of deoxyribonucleic acid. The shape is similar to that of a spiral staircase. The *DNA double helix* biopolymer of nucleic acid, held together by nucleotides which base pair together. In *B-DNA*, the most common *double helical* structure found in nature, the *double helix* is right-handed with about 10–10.5 base pairs per turn. The *double helix* structure of DNA contains a major groove and minor groove. Scientists uncovered the double-helical structure of DNA in the year 1953. Until then, scientists knew that traits were passed down to offspring in predictable ways, but they didn't understand how. All that changed when James Watson and Francis Crick showed that DNA is shaped like a spiral staircase that can be split, copied and passed on to future generations. Watson and Crick received a Nobel.

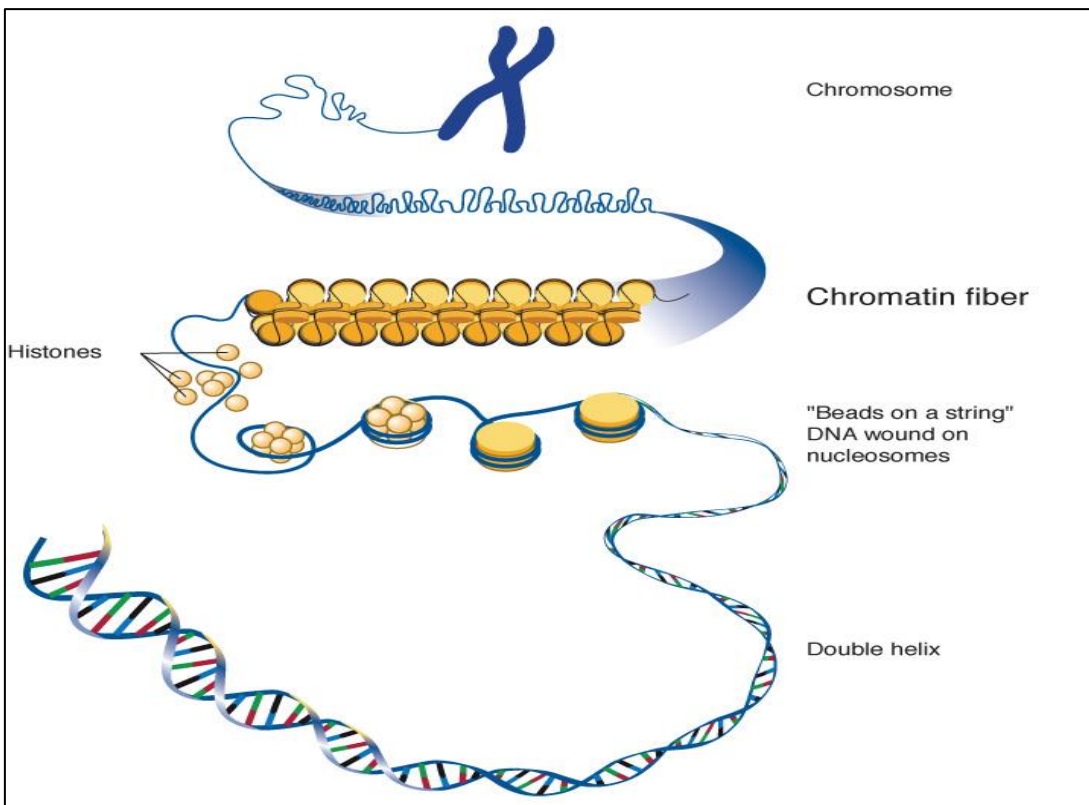


That's how many feet long the DNA from one of your cells would be if you uncoiled each strand and placed them end to end. Do this for all your DNA, and the resulting strand would be 67 billion miles long—the same as about 150,000 round trips to the Moon.





That's the length in inches across a cell's nucleus, which holds your DNA. If you sliced human hair into tenths lengthwise, each slice would be about that big around. To keep the space tidy, DNA spools around a group of proteins called histones. The resulting taut package of wound-up DNA is called chromatin, which winds up even tighter to form your chromosomes.



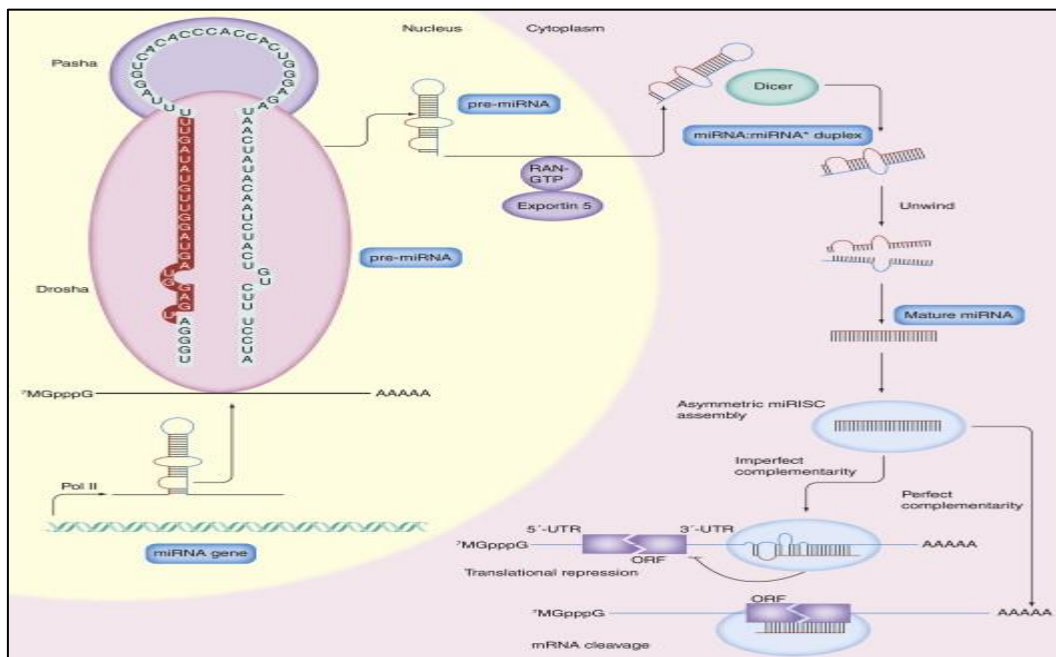
3. Human Genetic Variation

Human genetic variation is the genetic differences in and among populations. There may be multiple variants of any given gene in the human population (alleles), a situation called polymorphism. No two humans are genetically identical. The DNA of any two people on Earth is 99.6 percent identical. But 0.4 percent variation represents about 20 million base pairs (or 0.6% of the total of 3.2 billion base pairs), which can explain many of the differences between individuals, especially if the changes lie in key genes. Our environment also contributes to our individuality.



3.1. Micro RNAs

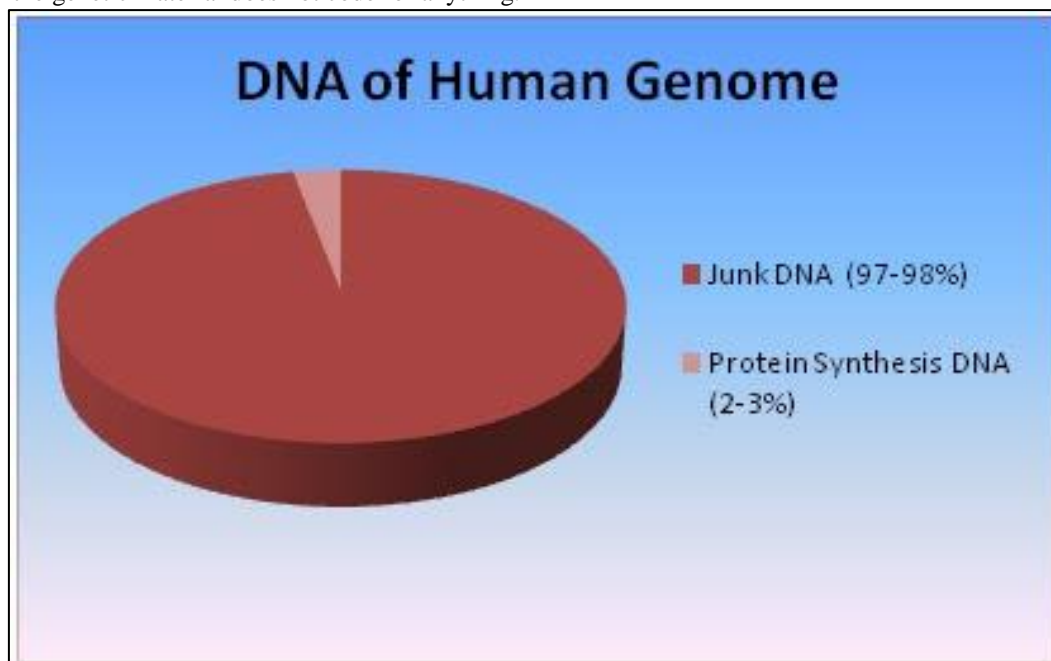
One-third (1/3) of human genes are estimated to be regulated by microRNAs. MicroRNAs are classes of small noncoding RNAs, 18- to 28-nucleotide-long, until now, 940 members of the family were identified in humans. These genetic "micromanagers" consist of only about 22 RNA units called nucleotides, their major role is in the posttranscriptional regulation of protein expression, and their involvement was demonstrated in normal and in pathological cellular processes. I.e. they can stop a gene from producing the protein it encodes.



3.2. Junk DNA

In genetics, the term junk DNA refers to regions of DNA that are noncoding. DNA contains instructions (coding) that are used to create proteins in the cell. However, the amount of DNA contained inside each cell is vast and not all of the genetic sequences present within a DNA molecule actually code for a protein. Some of this noncoding DNA is used to produce non-coding RNA components such as transfer RNA, regulatory RNA and ribosomal RNA. However, other DNA regions are not transcribed into proteins, nor are they used to produce RNA molecules and their function is unknown. The proportion of coding versus noncoding DNA varies significantly

between species. In the human genome for example, almost all (98%) of the DNA is noncoding, while in bacteria, only 2% of the genetic material does not code for anything.



Bibliography

- [1] Anderson, S., Bankier, A. T., and Barriell, A. G., 1981. "Sequence and organization of the human mitochondrial genome." *Nature*, vol. 290, pp. 457–474.
- [2] Avery, O. T., MacLeod, C. M., and McCarty, M., 1944. "Studies on the chemical nature of the substance inducing transformation of pneumococcal types." *J. Exp. Med.*, vol. 79, pp. 137–158.
- [3] Baltimore, D., 2001. "Our genome unveiled." *Nature*, vol. 409, pp. 814–816.
- [4] Bork, P. and Copley, R., 2001. "Filling in the gaps." *Nature*, vol. 409, pp. 818–820.
- [5] Griffith, F., 1928. "The significance of pneumococcal types." *J. Hygiene*, vol. 27, pp. 113-159.
- [6] Hershey, A. D. and Chase, M., 1952. "Independent functions of viral protein and nucleic acid in growth of bacteriophage." *J. Gen. Physiol.*, vol. 36, pp. 39–56.
- [7] Landers, E. S., Linton, L. M., and Birren, B., 2001. "International human genome sequencing consortium." *Initial Sequencing and Analysis of the Human Genome Nature*, pp. 409860–921.
- [8] Li, W. H., Gu, Z., Wang, H., and Nekrutenko, A., 2001. "Evolutionary analyses of the human genome." *Nature*, vol. 409, pp. 847–849.
- [9] Pennisi, E., 2000. "And the gene number is ...?" *Science*, vol. 288, pp. 1146–1147.
- [10] Rowen, L., Koop, B. F., and Hood, L., 1996. "The complete 685-kilobase DNA sequence of the human β T cell receptor locus." *Science*, vol. 272, pp. 1755–1762.
- [11] Sachidanandam, R., Weissman, D., and Schmidt, S., 2001. "The International SNP Map Working Group. A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms." *Nature*, vol. 409, pp. 928–933.
- [12] Strachan, T. and Read, A. P., 1999. *Human molecular genetics*. 2nd edition ed. Oxford: BIOS Scientific Publishers.
- [13] Turner, P. C., McLennan, A. G., Bates, A. D., and White, M. R. H., 1997. *Instant notes in molecular biology*. Oxford: BIOS Scientific Publishers.
- [14] Venter, J. C., Adams, M. D., and Myers, E. W., 2001. "The sequence of the human genome." *Science*, vol. 291, pp. 1304–1351.
- [15] Watson, J. D. and Crick, F. H. C., 1953. "Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid." *Nature*, vol. 171, pp. 737–738.