Genetics

Biology Dep.3rd Stage

Asst. Professor. Dr. Kazhal M. Sulaiman

Lec-10-

What is DNA sequencing?

Sequencing DNA means determining the order of the four chemical building blocks - called "bases" - that make up the DNA molecule. The sequence tells scientists the kind of genetic information that is carried in a particular DNA segment. For example, scientists can use sequence information to determine which stretches of DNA contain genes and which stretches carry regulatory instructions, turning genes on or off. In addition, and importantly, sequence data can highlight changes in a gene that may cause disease.

DNA sequencing is the process of determining the nucleic acid sequence – the order of nucleotides in DNA. It includes any method or technology that is used to determine the order of the four bases: adenine, guanine, cytosine, and thymine. The advent of rapid DNA sequencing methods has greatly accelerated biological and medical research and discovery.

Knowledge of **DNA sequences** has become indispensable for basic biological research, and in numerous applied fields such as medical diagnosis, biotechnology, forensic biology, virology and biological systematics. Comparing healthy and mutated DNA sequences can diagnose different diseases including various cancers, characterize antibody repertoire, and can be used to guide patient treatment.[5] Having a quick way to sequence DNA allows for faster and more individualized medical care to be administered, and for more organisms to be identified and cataloged.

DNA sequencing may be used to determine the sequence of individual genes, larger genetic regions (i.e. clusters of genes or operons), full chromosomes, or entire genomes of any organism. DNA sequencing is also the most efficient way to indirectly sequence RNA or proteins (via their open reading frames). In fact, DNA sequencing has become a key technology in many areas of biology and other sciences such as medicine, forensics, and anthropology.

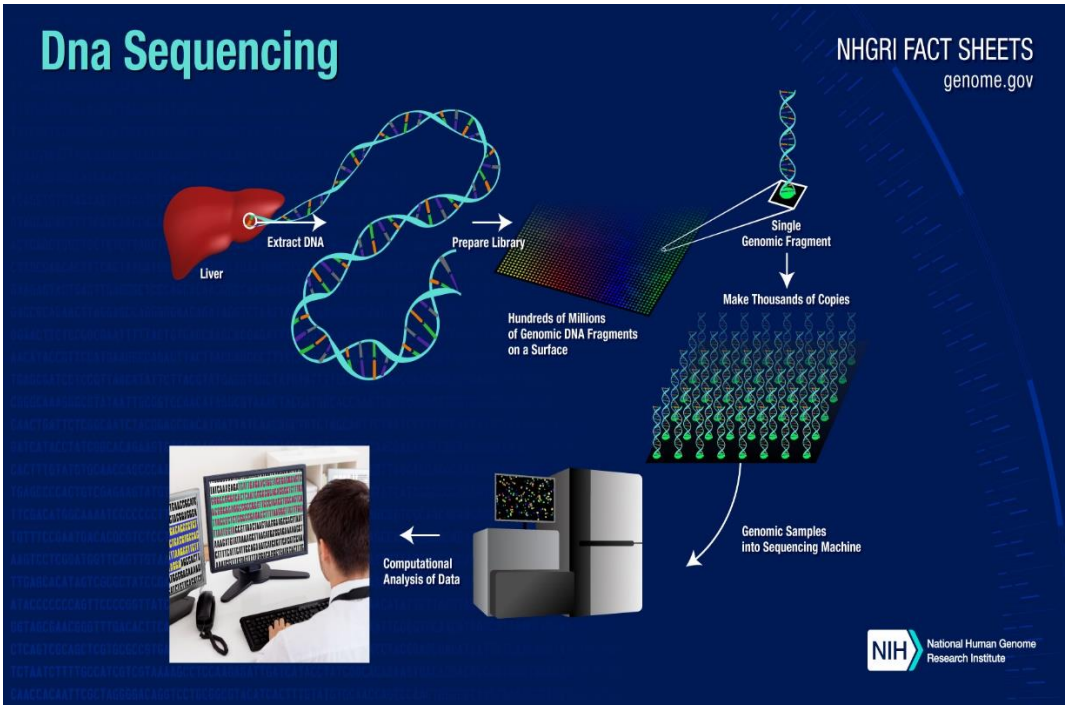
In the DNA double helix, the four chemical bases always bond with the same partner to form "base pairs." Adenine (A) always pairs with thymine (T); cytosine (C) always pairs with guanine (G). This pairing is the basis for the mechanism by which DNA molecules are copied when cells divide, and the pairing also underlies the methods by which most DNA sequencing experiments are done. The human genome contains about 3 billion base pairs that spell out the instructions for making and maintaining a human being.

How new is DNA sequencing?

Since the completion of the Human Genome Project, technological improvements and automation have increased speed and lowered costs to the point where individual genes can be sequenced routinely, and some labs can sequence well over 100,000 billion bases per year, and an entire genome can be sequenced for just a few thousand dollars.

Many of these new technologies were developed with support from the National Human Genome Research Institute (NHGRI) Genome Technology Program and its Advanced DNA Sequencing Technology awards. One of NHGRI's goals is to promote new technologies that could eventually reduce the cost of sequencing a

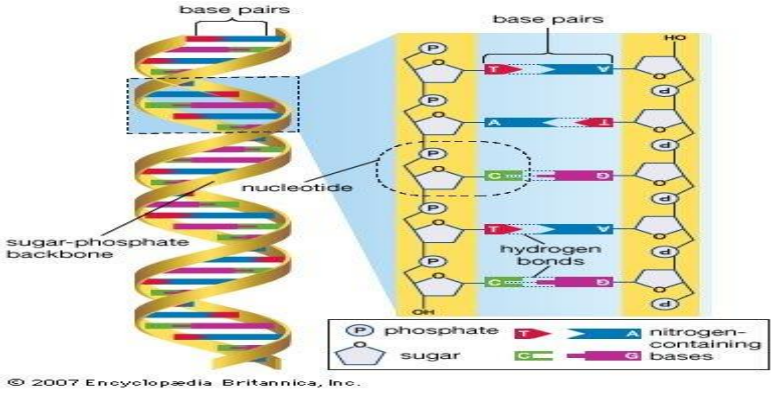
human genome of even higher quality than is possible today and for less than $1,000.



What new sequencing methods have been developed?

Since the completion of the Human Genome Project, technological improvements and automation have increased speed and lowered costs to the point where individual genes can be sequenced routinely, and some labs can sequence well over 100,000 billion bases per year, and an entire genome can be sequenced for just a few thousand dollars.Many of these new technologies were developed with support from the National Human Genome Research Institute (NHGRI) Genome Technology Program and its Advanced DNA Sequencing Technology awards. One of NHGRI's goals is to promote new technologies that could eventually reduce the cost of sequencing a human genome of even higher quality than is possible today and for less than $1,000.

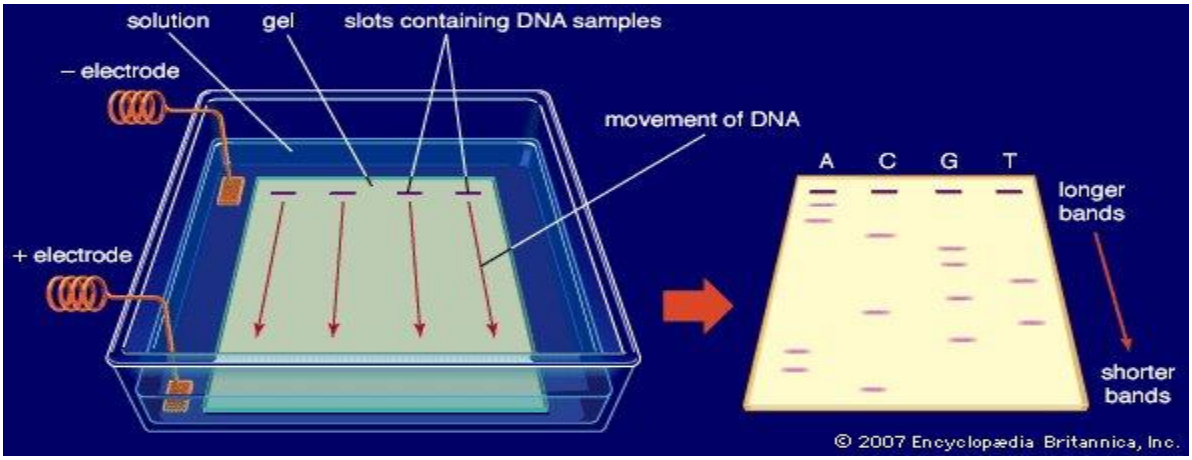
DNA sequencing, technique used to determine the nucleotide sequence of DNA (deoxyribonucleic acid). The nucleotide sequence is the most fundamental level of knowledge of a gene or genome. It is the blueprint that contains the instructions for building an organism, and no understanding of genetic function or evolution could be complete without obtaining this information.



First-Generation Sequencing Technology

So-called first-generation sequencing technologies, which emerged in the 1970s, included the **Maxam-Gilbert method,** discovered by and named for American molecular biologists **Allan M. Maxam and Walter Gilbert**, and the Sanger method (or dideoxy method), discovered by English biochemist Frederick Sanger. **In the Sanger method**, which became the more commonly employed of the two approaches, DNA chains were synthesized on a template strand, but chain growth was stopped when one of four possible dideoxy nucleotides, which lack a 3’ hydroxyl group, became incorporated, thereby preventing the addition of another nucleotide. A population of nested, truncated DNA molecules was produced that represented each of the sites of that particular nucleotide in the template DNA. The

molecules were separated according to size in a procedure called electrophoresis, and the inferred nucleotide sequence was deduced by a computer. Later, the method was performed by using automated sequencing machines, in which the truncated DNA molecules, labeled with fluorescent tags, were separated by size within thin glass capillaries and detected by laser excitation.



n gel electrophoresis an electric field is applied to a buffer solution covering an agarose gel, which has slots at one end containing DNA samples. The negatively charged DNA molecules travel through the gel toward a positive electrode and are separated based on size as they advance.

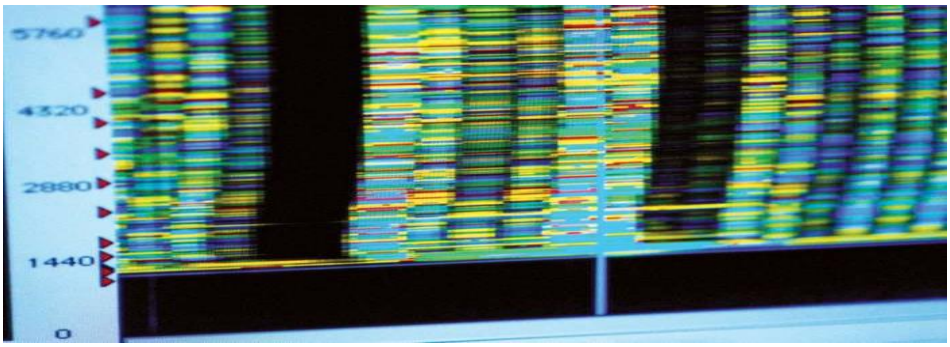
Next-Generation Sequencing Technology

Next-generation (massively parallel, or second-generation) sequencing technologies have largely supplanted first-generation technologies. These newer approaches enable many DNA fragments (sometimes on the order of millions of fragments) to be sequenced at one time and are more cost-efficient and much faster than first-generation technologies. The utility of next-generation technologies was improved significantly by advances in bioinformatics that allowed for increased

data storage and facilitated the analysis and manipulation of very large data sets, often in the gigabase range (1 gigabase = 1,000,000,000 base pairs of DNA).

Applications of DNA Sequencing Technologies

Knowledge of the sequence of a DNA segment has many uses. First, it can be used to find genes, segments of DNA that code for a specific protein or phenotype. If a region of DNA has been sequenced, it can be screened for characteristic features of genes. For example, open reading frames (ORFs)—long sequences that begin with a start codon (three adjacent nucleotides; the sequence of a codon dictates amino acid production) and are uninterrupted by stop codons (except for one at their termination)—suggest a protein-coding region. Also, human genes are generally adjacent to so-called CpG islands—clusters of cytosine and guanine, two of the nucleotides that make up DNA.. In order to determine the function of a gene, various domains can be identified that are common to proteins of similar function. For example, certain amino acid sequences within a gene are always found in proteins that span a cell membrane; such amino acid stretches are called transmembrane domains. If a transmembrane domain is found in a gene of unknown function, it suggests that the encoded protein is located in the cellular membrane. Other domains characterize DNA-binding proteins. Several public databases of DNA sequences are available for analysis by any interested individual.



DNA sequencing

A nucleotide sequence determined using DNA sequencing technologies.

The applications of next-generation sequencing technologies are vast, Using these technologies, scientists have been able to rapidly sequence entire genomes (whole genome sequencing) of organisms, to discover genes involved in disease, and to better understand genomic structure and diversity among species generally.

Applications of DNA Sequencing Technologies in Molecular biology

Sequencing is used in molecular biology to study genomes and the proteins they encode. Information obtained using sequencing allows researchers to identify changes in genes, associations with diseases and phenotypes, and identify potential drug targets.

Applications of DNA Sequencing Technologies in Evolutionary biology

Since DNA is an informative macromolecule in terms of transmission from one generation to another, DNA sequencing is used in evolutionary biology to study how different organisms are related and how they evolved.

Applications of DNA Sequencing Technologies in Medicine

Medical technicians may sequence genes (or, theoretically, full genomes) from patients to determine if there is risk of genetic diseases. This is a form of genetic testing, though some genetic tests may not involve DNA sequencing

Applications of DNA Sequencing Technologies in Forensics

DNA sequencing may be used along with DNA profiling methods for forensic identification and paternity testing. DNA testing has evolved tremendously in the last few decades to ultimately link a DNA print to what is under investigation. The DNA patterns in fingerprint, saliva, hair follicles, etc. uniquely separate each living organism from another.