Microbial Genetics
Introduction

Microbes are very small organisms can not see without a microscope, so they are abundant on Earth. They live everywhere in water ,air, rock, and soil . Some live in freezing cold while others thrive in searing heat. Some microbes need oxygen to live, but others do not need it. These microscopic organisms are found in plants and animals as well as in the human body.

Some microbes cause disease for humans, plants, and animals. Microbial diseases undoubtedly played a major role in historical events, it was in the year 1347 when **plagu** or '**black death**' struck Europe and within 4 years killed 25 million people, that is, one third of the population.

Others microbes are essential for a healthy life, and we could not exist without them. Indeed, the relationship between microbes and humans is delicate and complex. Most microbes belong to one of four major groups: bacteria, viruses, fungi, or protozoa. A common word for microbes that cause disease is "germs." Some people refer to disease-causing microbes as "bugs.

If we want to understand the growth of these organisms, their reproduction and development, so, it is necessary to understand its genetics.

Microbial genetics means genetics of microbes (bacteria, Archaea, viruses, including bacterial viruses i.e., bacteriophages and unicellular or mycelial eukaryotes including yeasts, fungi, algae and protozoa).

Geneticsis: the study of how genes carry information, how they are replicated and passed to other generations, and how they affect the characteristics of an organism.

Applications of Microbial genetics include in medicine, veterinary, agriculture, food and pharmaceutical industries.

The first study on microbial genetics started by George W. Beadle (1903–1989) and Edward L. Tatum (1909–1975) while investigating genetics of tryptophan metabolism and nicotinic acid synthesis in *Neurospora*, a fungus postulating the "one gene one enzyme" hypothesis. However, studies on bacterial genetics started in 1947 (Joshua Lederberg) with demonstration of exchange of genetic factors through conjugation in *Escherichia coli*, mediated through plasmids, "fertility factors". Later, by process such as transformation, transduction and chromosomal gene mobilization lead to genome (chromosome) mapping in bacteria. These techniques need to restriction enzyme analysis for sequencing, cloning and expression of several genes (in prokaryotic and eukaryotic).

Bacterial Genetics
Nucleic Acid Structure

DNA (deoxyribonucleic acid) and **RNA** (ribonucleic acid) are polymers of nucleotides linked in a chain through phosphodiester bonds. In biological systems, they serve as information-carrying molecules or, in the case of some RNA molecules, catalysts.

Nucleotides (building blocks of nucleic acids) have a distinctive structure composed of three components covalently bound together:

a 5-carbon sugar - ribose or deoxyribose
a nitrogen base : pyrimidine (one ring) or purine (two rings)
a phosphate group

The combination of a base and sugar is called a nucleoside.

Nucleotides also exist in activated forms containing two or three phosphates, called nucleotide diphosphates or triphosphates. If the sugar in a nucleotide is deoxyribose, the nucleotide is called a deoxynucleotide; if the sugar is ribose, the term ribonucleotide is used.

DNA is composed of repeating *nucleotides* containing the bases adenine =A, thymine =T, cytosine =C, and guanine =G; a phosphate group and a deoxyribose sugar. Bases found in specific *complementary base pairs*, the hydrogen bonds from which connect strands of DNA: adenine with thymine, and cytosine with guanine. G-C base pairs have Three hydrogen bonds, whereas A-T base pairs have Two hydrogen bonds.

RNAs are usually single stranded and the base pairs that form are A-U and G-C.

Bacterial Genome

Chromosomes are cellular structures made up of genes that carry hereditary information. A ***gene*** is not just a segment of DNA, a gene is a specific sequence of nucleotides that codes for a functional product, usually a protein. All the genetic information in a cell is the ***genome****.*

The **genome** of an organism can be defined as the total DNA content of the cell , and as such it contains all the genetic information required to direct the growth and development of the organism.

In prokaryotes, most of the genome (**85-90%**) is non-repetitive DNA, which means coding DNA mainly forms it, while non-coding regions only take a small part. In bacteria there is little repetitive DNA as seen in higher eukaryotes.

The relatively small genome size of bacterial genomes, together with the fact that they contain very little non-coding or repeat DNA and that they do not contain introns, has made bacterial genomes ideal candidates for whole genome sequencing projects.

Identification of potential genes within bacterial genomes is also much more reliable, because firstly bacterial genes do not contain introns and secondly the genes are much more closely packed.
Usually bacteria have only one chromosome (while yeast have 7 , humans 46), but some bacteria have more than one chromosome. DNA in chromosomes is in the form of one, long double helix associated with many of proteins that regulate the activity of genes. In prokaryotes, DNA of bacteria is circular and it is not found within a nuclear membrane. The chromosome of *E. coli*, *for example*, contains about 4 million base pairs and is approximately 1000 times longer that the cell . Because the DNA is supercoiled by an enzyme called ***topoisomerase II*** or ***DNA gyrase*** ,The chromosome takes up only about **10%** of the cell's volume (Figure 1).

The nucleoid (meaning nucleus-like) is an irregularly-shaped region within the cell of a prokaryote that contains all or most of the genetic material. In contrast to the nucleus of a eukaryotic cell, it is not surrounded by a nuclear membrane.

The genome of prokaryotic organisms generally is a circular, double-stranded DNA. The nucleoid is largely composed of about **60% DNA**, plus a small amount of RNA and protein.

A genophore is the DNA of a prokaryote. It is commonly referred to as a prokaryotic chromosome. The term "chromosome" is misleading, because the genophore lacks chromatin. The genophore is compacted through a mechanism known as supercoiling, but a chromosome is additionally compacted through the use of chromatin.

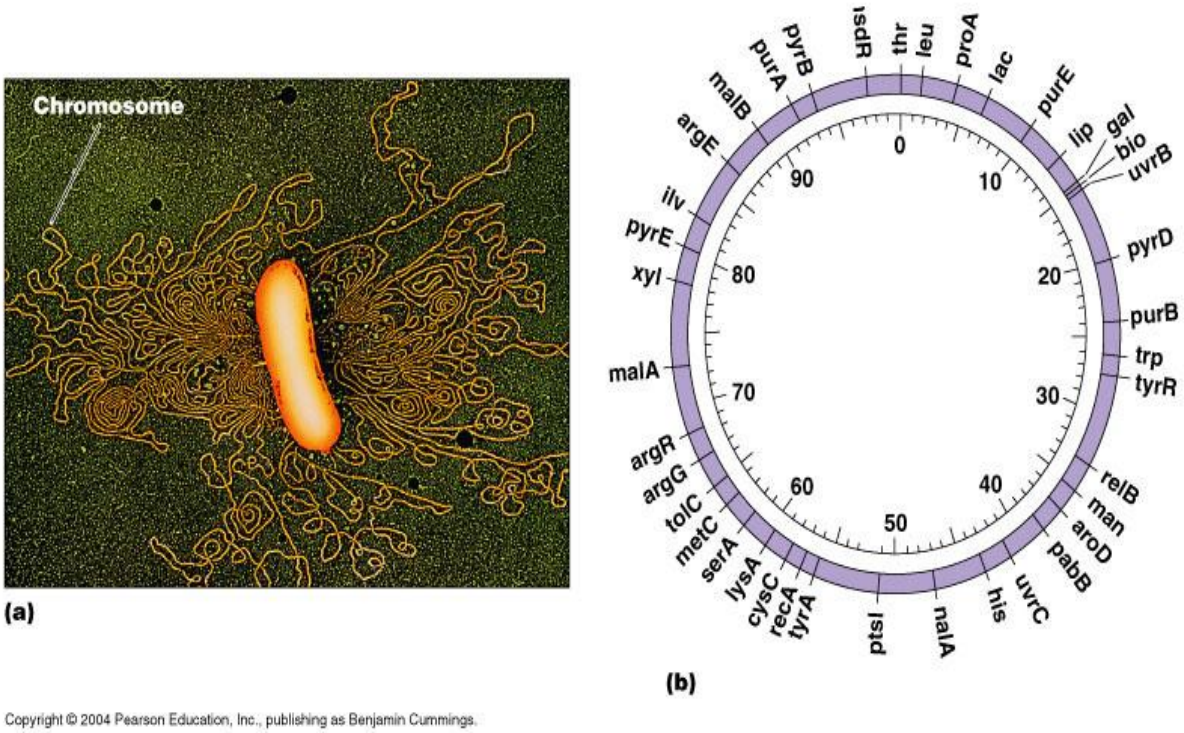


Figure 1:a.Chromosome of *E. coli* , b. The chromosomal map.

Gene names

Gene named with distinguished name from three letters followed by a hyphen and a number. These letters and number are always italicized. The chosen letters are usually either abbreviations of a longer description (such as ***lin*** refer to lineage defective or ***unc*** which refer to uncoordinated) or it may be acronym (such as ***sur*** for suppressor of **ras**). A number then follows the letters (such as ***lin-31***) to indicate the approximate order in which the mutations were discovered. Originally, most or all gene names were derived from genetic screens in which mutant alleles were isolated.

DNA Replication

- DNA replication employs a large number of proteins and enzymes, each of which plays a critical role during the process. One of the key players is the enzyme **DNA polymerase**, which adds nucleotides one by one to the growing DNA chain that are complementary to the template strand. In prokaryotes, three main types of polymerases are known: ***DNA pol I***, ***DNA pol II***, and ***DNA pol III***.
- DNA pol III is the enzyme required for DNA synthesis; DNA pol I and DNA pol II are primarily required for repair.
- There are specific nucleotide sequences called **origins of replication** where replication begins. The origin of replication is recognized by certain proteins that bind to this site.
- An enzyme called **helicase** unwinds the DNA by breaking the hydrogen bonds between the nitrogenous base pairs.

ATP hydrolysis is required for this process. As the DNA opens up, Y-shaped structures called replication forks are formed. Two replication forks at the origin of replication are extended bi-directionally as replication proceeds. Single-strand binding proteins coat the strands of DNA near the replication fork to prevent the single-stranded DNA from winding back into a double helix.

DNA polymerase is able to add nucleotides only in the 5' to 3' direction (a new DNA strand can be extended only in this direction). It also requires a free 3'-OH group to which it can add nucleotides by forming a phosphodiester bond between the 3'-OH end and the 5' phosphate of the next nucleotide. This means that it cannot add nucleotides if a free 3'-OH group is not available.

Another enzyme, **RNA primase**, synthesizes an RNA primer that is about five to ten nucleotides long and complementary to the DNA, priming DNA synthesis.

A **primer** provides the free 3'-OH end to start replication. DNA polymerase then extends this RNA primer, adding nucleotides one by one that are complementary to the template strand.

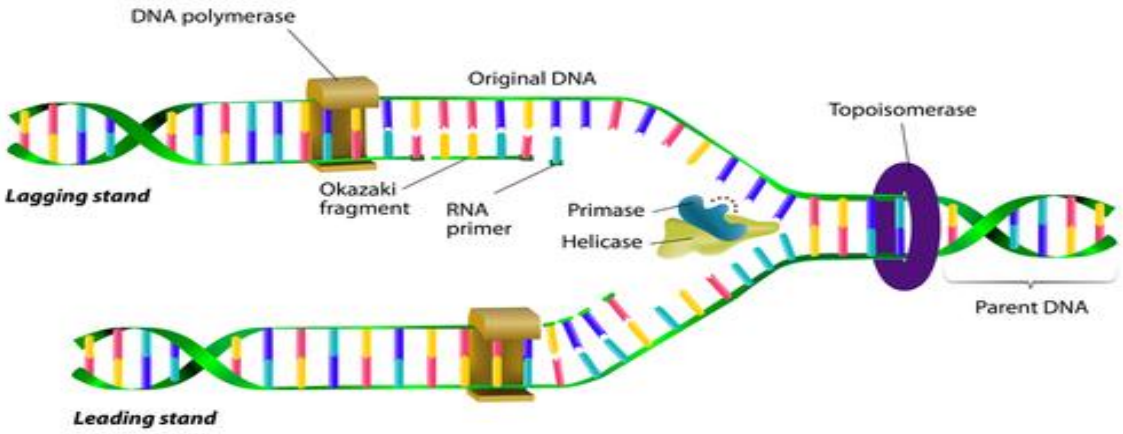
DNA polymerase can only extend in the 5' to 3' direction, which poses a slight problem at the replication fork. As we know, the DNA double helix is anti-parallel;

that is, one strand is in the 5' to 3' direction and the other is oriented in the 3' to 5' direction.

One strand (**the leading strand**), complementary to the 3' to 5' parental DNA strand, is synthesized continuously towards the replication fork because the polymerase can add nucleotides in this direction.

The other strand (**the lagging strand**), complementary to the 5' to 3' parental DNA, is extended away from the replication fork in small fragments known as ***Okazaki fragments***, each requiring a primer to start the synthesis. Okazaki fragments are named after the Japanese scientist who first discovered them.

- The leading strand can be extended by one primer alone, whereas the lagging strand needs a new primer for each of the short Okazaki fragments. The overall direction of the lagging strand will be 3' to 5', while that of the leading strand will be 5' to 3'.
- **Topoisomerase** prevents the over-winding of the DNA double helix ahead of the replication fork as the DNA is opening up; it does so by causing temporary nicks in the DNA helix and then resealing it.
    - The primers are removed by the ***exonuclease*** activity of **DNA pol I**, while the gaps are filled in by deoxyribonucleotides.



Before the cell divides , The DNA in a cell is duplicated (DNA replication) so each daughter cell receives the same number of genetic information.

This information is passed on to the next generations when the DNA is replicated and the cell divides as well as it may be passed between cells of the same generation by recombination (Figure 2)

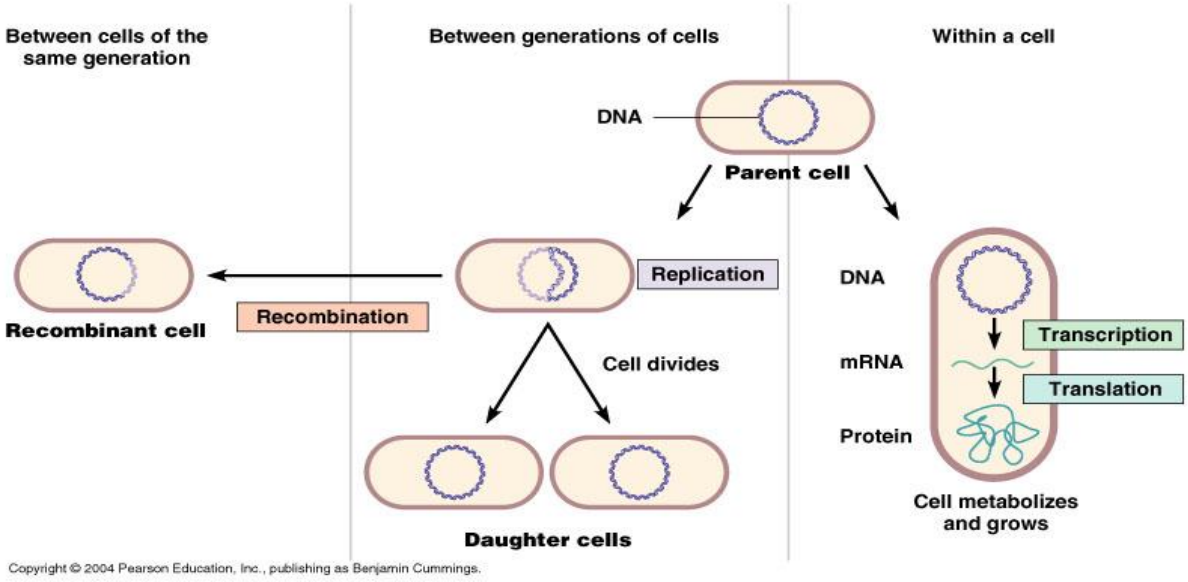


Figure 2:Transferring of genetic material either vertically from mother cells to daughter or horizontally between cell to cell by recombination.

Gene Expression

Gene expression is the process by which information from a gene is used in the synthesis of a functional gene product. These products are often proteins, via Two main steps Transcription and translation, but in non-protein coding genes such as transfer RNA (tRNA) or small nuclear RNA (snRNA) genes, the product is a functional RNA.

In genetics, gene expression is the most fundamental level at which the genotype gives rise to the phenotype, i.e. observable trait.

RNA Transcription

In process called (transcription). The information in DNA can be transcribed into RNA and this information, then, translated into protein (translation). When the gene expression occur the DNA is transcribed to produce a single-stranded molecule of mRNA. Then, the mRNA is translated into proteins.

Only the coding strand or the sense strand of the gene is transcribed. The opposite strand is non-coding strand or the antisense.

- Prokaryotic transcription is the process in which messenger RNA transcripts of genetic material in prokaryotes are produced, to be translated for the production of proteins.

- Transcription like replication need free 3' end to add the complementary nucleotide, also the direction of movement and polymerization of RNA polymerase from 5' →3'.
- The transcript sequence (complete gene) start with promoter region where the DNA will opened and the RNA polymerase bind to start transcription.
- **Promoter** : a regulatory nucleotide sequences of DNA (**40-60** nts) at the beginning of every gene located upstream (towards the 5' region) of a gene. In prokaryotes, the promoter is recognized by RNA polymerase (**δ** sigma sub unite). The promoter consists of two short sequences know as **-10 box** and **-35 box** positions upstream from the transcription start site.
- **Coding region**: Nucleotide sequences which will detriment the genetic code then will translated to Amino acid. It start with **ATG** triplet initiation codon (**AUG** in m RNA). The length of coding region depend on type of produced protein.
- **Terminator** : Nucleotide sequences exist after the coding region rich with poly G followed by poly C then poly A .

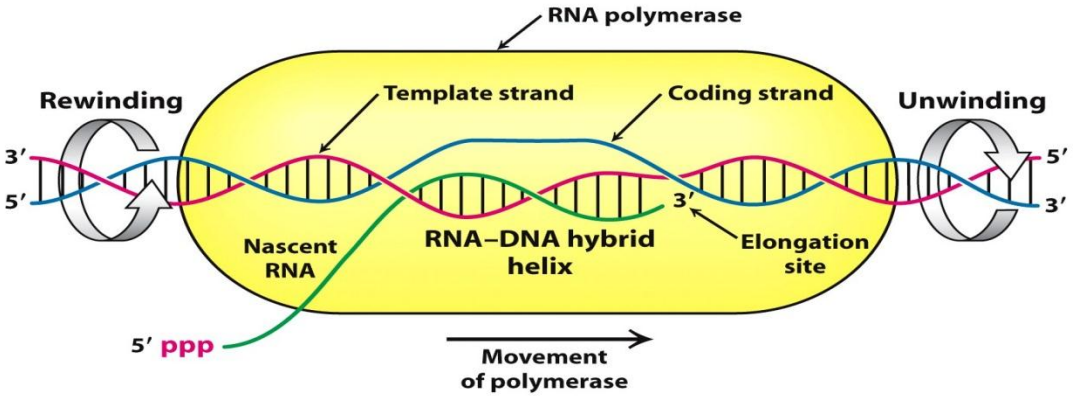


Figure 29.13
***Biochemistry*, Seventh Edition**

- The RNA polymerase enzyme consist from core (**5 units**) and another **6th** unit called **sigma δ**, after core binding with sigma it will convert to **holoenzyme**.
- sigma subunit play significant role in recognition promoter region then it will released leaving core continue transcription RNA from template.
- The **Promoter** and **Terminator** are directions for RNA polymerase to indicate the location of the gene to be transcribed.
- The start and stop codons are directions for the ribosome to indicate where the amino acid information for translation begins and ends.

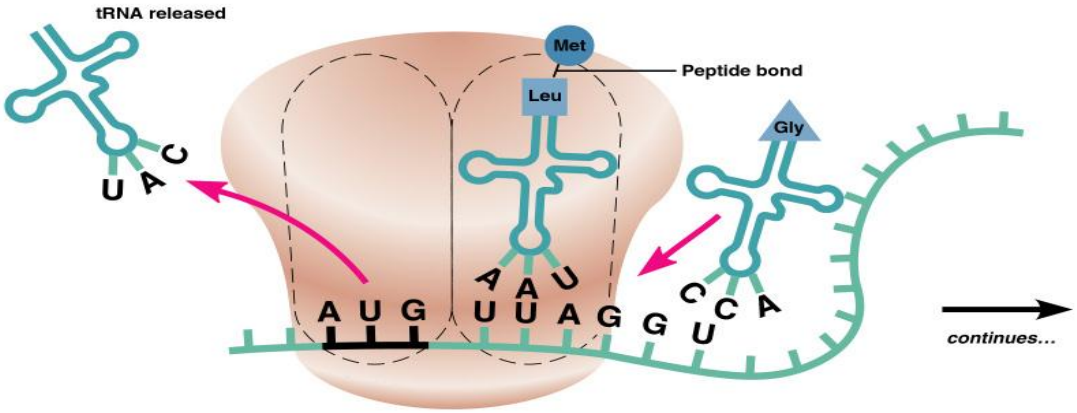
Stages of Transcription

- **Initiation stage** involves recognizing core promoter region by sigma factor of RNA polymerase .
- **Elongation stage** involves the further addition of ribonucleotides
- **Termination stage**: When Transcription process reached a region called **Terminator**, termination stage start by: a mechanism called **Rho-independent termination**: is a mechanism in prokaryotes that causes RNA transcription to be stopped without the aid of rho protein.

Rho-dependent termination: Rho factor is a prokaryotic ATP-dependent unwinding enzyme involved in termination transcription.

Translation

The translation of genetic information from the 4-letter (A ,C ,G and T) of polynucleotides into the 20-amino acid of proteins is a complex process. The information in the sequence of a messenger RNA molecule is read out in groups of three nucleotides at a time: each triplet of nucleotides, or codon, specifies (codes for) a single amino acid in a corresponding protein.



4 The first amino acid joins to the second by a peptide bond, and the first tRNA is released.

- Initiation codon (AUG): signals the start of translation. Lies just downstream of the Shine Dalgarno sequence .
- Termination codon (UAG,UGA,UAA): signals the end of translation.
- mRNA stands for messenger RNA. It's the product of transcription and the template for translation.
- Prokaryotic mRNAs don't have the 5' cap or polyA tail.
- Both tRNA (transfer RNA) and rRNA (ribosomal RNA) are products of transcription. However, they do not serve as the template of translation.

- tRNA is responsible for bringing in the correct amino acid during translation. rRNA makes up the ribosome, which is the enzyme responsible for translation.

  Ribosomes are the sites of protein synthesis. Bacterial ribosomes differ from those of eukaryotic cells in both size and chemical composition. They are organized in units of **70S**.