Math 612: Single Cell Analysis **2019W Term 1**

Lecture 2: September 10

Lecturer: Geoffrey Schiebinger *Scribe: Isabela do Ó*

2.1 Cell Theory

The term "cell" was coined by Robert Hooke in 1665 [1], who named them due to the resemblance to cells in a christian monastery. The "Cell Theory" [2] states the following:

- Every organism is composed of one or more cells (The discovery of Virus brought this point into debate).
- Every cell comes from a preexisting cell (Not considering the origin of life).
- Cell is the basic structure of organizational unit of an organism.

2.2 Structure of a Cell

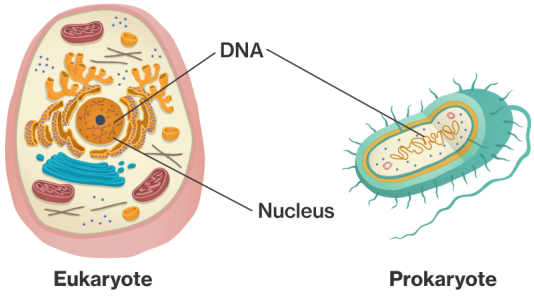

Figure 2.1: Comparison between eukaryotic cell and prokaryotic cell. Image from: www.coursehero.com/sg/microbiology/prokaryotes-vs-eukaryotes/

As we can see from the figure above, eukaryotic cells (like the ones we have in our body) have intracellular membranes. These are the structures that encapsulate the organelles, such as the Golgi apparatus, which in this image is illustrated in blue. While the eukaryotic in the cell the DNA is enclosed in a nucleus, in prokaryotes (for example, bacterias) the genetic information is floating in the cytoplasm.

For this class we will not be studying any organelle, and should only focus on understanding the DNA, RNA, Proteins and ribosomes. This brings us to our next section, "The Central Dogma".

2.3 Central Dogma

The Central Dogma of Molecular Biology explains how biological information flows from genes to proteins. DNA is **transcribed** into RNA which is **translated** to protein.

2.3.1 Transcription

During this process, a gene (which is made of DNA) is transcribed into the end result which is RNA. When this gene encodes for a protein, the RNA produced during the transcription is called "messenger RNA", or mRNA. There are other types of RNA, which we call non-coding RNA. Some examples are transfer RNAs and ribosomal RNAs. These three types are the most important for understanding the central dogma.

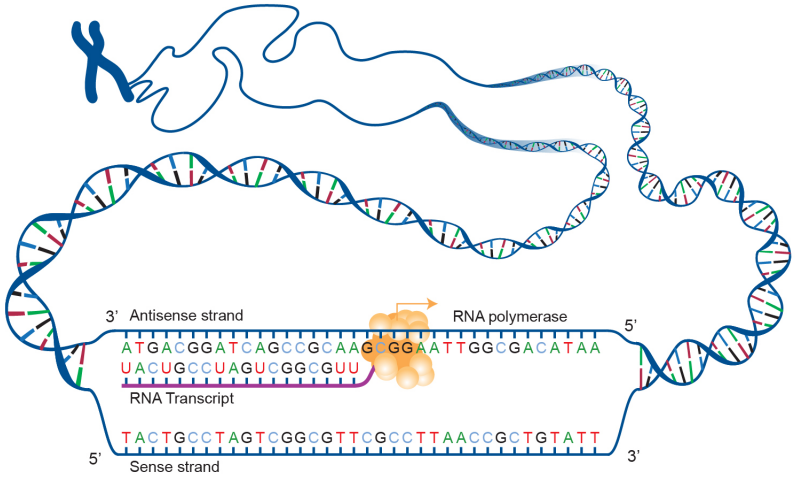

Figure 2.2: Illustration of the Transcription. Image from: The National Human Genome Research Institute (www.genome.gov)

.

The DNA is a double stranded helix, and each strand is formed by a backbone a polymer. The monomer is called nucleotides, which are composed of a deoxyribose, a phosphate group and a nitrogenous base. The two strands are connected by bonds between the nitrogenous bases. These bases can be one of four types: Adenine (A), Guanine (G), Cytosine (C) and Thymine (T). RNAs are quite similar. The two basic differences are that RNAs only have one strand, and instead of Thymine, there is Uracil (U).

To create a new RNA molecule, an enzyme (a kind of protein) called "RNA-polymerase", bind to a part of the DNA, the promoter (a physical region of the DNA close to the gene). However, for the RNA-polymerase to successfully bind, there must be Transcription Factord (another kind of protein) bound to the DNA. At this point the DNA is still a connected double strand.

After the RNA-polymerase is bound to the DNA, it separates the strands. One of the strands work as a template. The enzyme reads the DNA from 3' to 5', and produces the RNA, one base at a time from 5' to 3'.

Just like in the beginning of the gene there is a promoter that shows where the relevant information starts,

at the end there is a region called "Terminator". This region signals where translation should end, as the name suggests.

2.3.1.1 Control of Transcription

This refers to how the cell or organism controls how much a certain gene is being transcribed (gene activity). Since all cells in a multicellular organism have the same genetic information, this is a major part of the explanation to why these cells differ.

One important concept to know in order to understand different types of control of transcription is "open" and "closed" chromatin. Chromatin is the complex that consists of DNA and associated proteins. This Chromatin is a dynamic structure, varying in terms of compactness, which is given by how the DNA is wrapped on these proteins. We call "open" chromatin the one that is accessible for proteins related to the transcription. "Closed" chromatin does not allow for gene expression to happen. For this class we are separating this control in three types.

1. Transcription Factors: The protein produced by one gene can act as the transcription factor of another gene. Transcription Factors can act as inhibitor or as activators of genes.

2. DNA packing: How the DNA is presented in a cell is not static. It can be more or less tangled in histones (another protein), making it more or less compacted. Imagine the DNA as a rope that is a few hundreds of meters long, and the promoter is only a few centimeters long. It is much easier to associate anything to the promoter if the rope is straight and untangled as opposed to it being tangled. Closed or packed chromatin inhibits the association of proteins such as RNA-polymerase to specific regions of the DNA.

3. DNA methylation: The association of methyl with the DNA, in areas close to the promoters, usually inhibits transcription.

2.3.2 Translation

This is the process in which messenger RNA is translated into a protein. Proteins are polypeptides, which the monomeric unit is an amino acid.

The mRNA exists the nucleus through diffusion to the cytoplasm. There it encounters ribosomes that will be the main machinary to produce polypeptides. The RNA is read in triplets, called codons. It is a triplet because it is made of three bases that can be either Adenine (A), Guanine (G), Cytosine (C) and Uracil (U).

There are 61 different codons that mean amino acids, one of which that means "start", and three that means "stop". However there are only 21 amino acids. For this reason, we call this code "redundant", since more than one codon means the same amino acid.

Floating transfer RNAs encounters the ribosomes. These RNAs carry a specific amino acid, depending on what information they have in their anticodon region. The ribosome allows tRNAs to "try" to bind their anticodon region to the codon of the mRNA. If it works, the amino acid carried by the tRNA will be used to form the new polypeptide. As the polypeptide comes out of the ribosome it folds up depending on the biochemical reaction between the neighbouring amino acids and the environment.

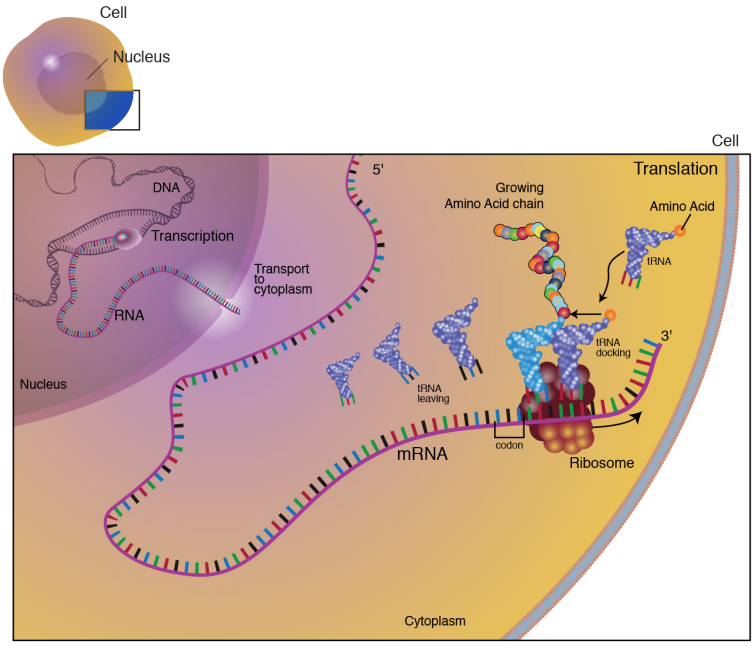

Figure 2.3: Illustration of the Translation. Image from: The National Human Genome Research Institute (www.genome.gov)

.

2.4 Cell Division

Here we are studying one type of cell division which is Mitosis. Mitosis is part of the Cell cylce. After a new cell is created through mitosis, it goes through a growth phase, then a DNA replication phase, another growth phase, and then mitosis.

During mitosis the DNA strands will duplicate, and become very condense. Then the nuclear envelope dissolves and "spindles" guide the chromosomes to the center of the cell. The two parts of each duplicated chromosome is pulled to one side of the cell. Now Each half of the cell should have the same genetic information and is ready to be a cell by itself.

An important part of this process is DNA duplication. This phenomenon is similar to the production of RNA through Transcription. However, in this case, both strands of the DNA will be used as templates, so at the end of this process there will be two DNA double-strands. Another difference are the enzymes involved. During this process, the enzyme helicase is responsible to unwind the double helix, while DNA polymarese attaches to the mother DNA strand and assembles a new complementary strand.

References

[1] R. Hooke, "1665.", *Micrographia*, 1961.

[2] M. Schleiden, "Beitrge zur Phytogenesis", *Archiv fr Anatomie, Physiologie und wissenschaftliche Medicin*, 1839, pp. 137176.

[3] T. Schwann, "Mikroskopische Untersuchungen ber die Uebereinstimmung in der Struktur und dem Wachsthum der Thiere und Pflanzen", *Berlin: Sander*, 1839.