Replication transcription and translation crash course

lecture1_transcript.html-CentralDogma DNA → RNA → protein The central dogma of molecular biology describes how genetic information is stored and interpreted in the cell: An organism's genetic code is stored in DNA, which is transcribed into RNA, which is eventually translated into protein. Proteins perform the majority of cellular functions such as motility, DNA regulation and replication. Although central dogma is true in most situations, there are a number of notable exceptions to the model. For example, retroviruses are able to generate DNA from RNA through reverse transcription. Also, some viruses are so primitive that they don't even have DNA, instead only using RNA to protein. 1.4.2 DID you know DNA? Central dogma is sometimes misinterpreted too strongly to mean that DNA stores only immutable information from one generation to the next that remains the same in a generation, RNA is used only as a temporary means of information transfer, and proteins are the only molecule that can perform complex actions. Again, there are many exceptions to this interpretation, for example: somatic mutations can alter DNA in a generation, and different cells may have different content than DNA. Some cells undergo programmed DNA alterations during maturation, resulting in a different DNA content, more famous B and T immunity while blood cells Epigenetic changes in DNA can be inherited from one generation to the next RNA can play many different roles in gene regulation, metabolic detection, and enzymatic reac-tions, functions that were previously thought to be reserved for proteins. The proteins themselves may undergo conformational changes that are epigenetically inherited non-tably prion states that have been famously responsible for mad cow disease DNA → RNA → function of protein DNA The DNA molecule stores the genetic information of an organism. DNA contains regions called genes, which encode for the proteins to be produced. Other regions of DNA contain regulatory elements, which partially influence the level of expression of each gene. In the genetic code of DNA are both the data on the proteins that need to be coded, and the control circuits, in the form of regulatory patterns. DNA structure DNA is composed of four nucleotides: A(adenin), C (cytosine),T (thymine), and G (guanine). A and G are purines, which have two rings, while C and T are pyrimidines, with a ring. A and T are connected by two hydrogen bonds, while C and G connected by three links. Therefore, the A-T pairing is weaker than the C-G pairing. (For this reason, the genetic composition of bacteria that live in hot springs is 80% G-C). lecture1_transcript.html-Complementarity The two strands of DNA of the double helix are complementary, which means that if there is an A on one strand, it will be linked to one T on the other, and if there is a C C one strand, it will be tied to one G on the other. DNA strands also have directionality, which refers to the positions of the pentosis ring where the phosphate backbone connects. This convention of directionality comes from the fact that DNA and RNA polymerese synthesize in the direction of 5' to 3'. With this in mind, we can say that the DNA strands are anti-parallel, as the 5' end of one strand is adjacent to the 3' end of the other. As a result, DNA can be read in both the 3' to 5' direction and the direction of 5' to 3', and genes and other functional elements can be found in each. By convention, DNA is written from 5' to 3'. The 5' and 3' directions refer to positions on the pentosis ring where the phosphate backbone connects. The basic pairing between DNA nucleotides is its primary and secondary structure. In addition to the secondary structure of DNA, there are several additional levels of structure that closely compact DNA and influence gene expression (Figure 3). The tertiary structure describes the twist in the DNA scale that forms a helical shape. In the quaternary structure, DNA is tightly wrapped around small proteins called histones. These histone DNA complexes are more coiled into tighter structures observed in chromatin. Before DNA can be reproduced or transcribed into RNA, the chromatin structure must be locally unwrapped. Thus, gene expression can be regulated by changes in the structure of chromatin, making it easier or more difficult to unpack DNA. This regulation of gene expression by the modification of chromatin is an example of epigenetics. © Zephyris on Wikipedia. Some rights reserved. License: CC BY-SA. This content is excluded from our Creative Commons license. For more information, see . Figure 1.2: DNA double helix structure. The nucleotides are in the center, and the sugar-phosphate backbone is outside. DNA Replication The structure of DNA, with its weak hydrogen bonds between the bases of the center, makes it easy to separate the strands for DNA replication (the ability to separate DNA strands also allows transcription, translation, recombination and DNA repair, among others). Watson and Crick noted this as it has not escaped our opinion that the specific matching we postulated immediately suggests a possible copying mechanism for genetic material. In DNA replication, the two complementary strands are separated, and each strand is used as models for the construction of a new strand. DNA polymers to each of the strands causing the replication, reading each existing strand of the direction from 3' to 5' and placing down complementary bases so that the new strand develops in the direction of 5' to 3'. Because the new strand must from 5' to 3', one strand (the head strand) can be continuously copied, while the other (the trolling strand) pushes into pieces that are then glued together by DNA ligase. The end result is 2 pieces of double-stranded DNA, each of which is composed of an old strand, and 1 new strand; for this reason, DNA replication is semiconservative. Many organisms have their DNA divided into several chromosomes. Each chromosome contains two strands of DNA, which are complementary to each other but are read in opposite directions. Genes can occur on either strand of DNA. DNA before a gene (in the 5' region is considered upstream while DNA after a gene (in the 3' region is considered downstream . 1.4.3 Transcript lecture1_transcript.html #Transcription DNA → RNA → Protein Courtesy of Macmillan Publishers Limited. Used with permission. Source: Qiu, Jane. Epigenetics: Unfinished Symphony. Nature 441, 7090 (2006): 143-45. Figure 1.3: DNA is packaged over multiple layers of organization in a compact mRNA generation chromosome Transcription is the process by which RNA is produced using a DNA model. DNA is partially untied to form a bubble, and RNA polymerrase is recruited at the transcription start site (TSS) by regulatory protein complexes. Polymer RNA reads the DNA of the direction from 3' to 5' and places complementary bases to form messenger RNA (mRNA). RNA uses the same nucleotides as DNA, except that Uracil is used in place of Thymine. Post-transcriptional changes in mRS in eukaryotes undergo post-translational changes or processes that further alter the mRN strand. In particular, a process called splicing removes introns, intervening regions that do not encode for proteins, so that only the coding regions, the exons, remain. Different regions of the primary transcription can be episeled to lead to different protein products (alternative splicing). In this way, a huge number of different molecules can be generated on the basis of different splicing permutations. In addition to splicing, both ends of the mRA molecule are treated. The 5' end is crowned with a modified guanine nucleotide. At the 3' end, approximately 250 adenine residues are added to form a poly(A) tail. lecture1_transcript.html RNA → RNA → Protein RNA is produced when DNA is transcribed. It is structurally similar to DNA, with the following major differences: 1. Nucleotide uracil (U) is used instead of THE DNA thymine (T). 2. RNA contains ribose instead of deoxyribosis (deoxyribosis does not have the oxygen molecule on the 2' position ribose). 3. RNA is single-stranded, while DNA is double-stranded. RNA molecules are the intermediate step to encode a protein. RNA molecules also have catalytic and regulatory functions. An example of catalytic function is protein synthesis, where RNA is part of the ribosome. It many different types of RNA, including: Courtesy of Forluvoft on Wikipedia. Images in the public domain. Figure 1.4: RNA is produced from a DNA model during transcription. A bubble is opened in the DNA, allowing polymer RNA to enter and place bases complementary to DNA. a) Introduction to transcription b) Transcript elongation c) Transcription ending mRNA contains information to make a protein and is translated into protein sequence. tRNA (RNA transfer) specifies the translation of amino-codon acid. It contains a pair of 3 pairs of base anti-codon complementary to a codon on the mRN, and carries the amino acid corresponding to its anticodon attached to its 3' end. rRNA (RIBosomal RBA) forms the nucleus of the ribosome, the organeitus responsible for translating mRNA into proteins. snRNA (small nuclear RNA) is involved in splicing (removing introns from) pre-RNA, as well as other functions. Other functional types of RNA exist and are still being discovered. Although proteins are generally thought to perform essential cellular functions, RNA molecules can have complex three-dimensional structures and perform various functions in the cell. According to the RNA world hypothesis, early life was entirely based on RNA. RNA served as both a repository of information (like DNA today) and a functional workhorse (like today's protein) in early organisms. It is thought that the protein subsequently appeared via ribosomes, and DNA is thought to have appeared last, by reverse transcription. lecture1_transcript.html DNA translation → RNA → protein translation Unlike transcription, in which nucleotides remained the means of encoding information in DNA and RNA, when RNA is translated into protein, the primary structure of the protein is determined by the sequence of amino acids it is composed of. Since there are 20 amino acids and only 4 nucleotides, 3 sequences of nucleotides in mRNE, known as codons, encode for each of the 20 amino acids. Each of the 64 3 possible sequences of nucleotides (codon) uniquely specifies either a particular amino acid or is a stop codon that stops the translation of proteins (the starting codon also encodes methionine). Since there are 64 possible codon sequences, the code is degenerated, and some amino acids are specified by multiple codings. Most of the degeneration occurs in the 3rd codon position. Post-translational changes Like mRN, proteins also undergo other changes that affect its structure and function. A guy post-translational modification (PTM) involves introducing new functional groups into amino acids. In particular, phosphorylation is the process by which a group of phosphate is added to an amino acid that can fully activate or disable the protein. Another type of PTM is the cleavage of peptide bonds. For example, the hormone insulin is split twice the formation of disulfides in the original protein. Figure 1.5: This codon table shows which of the 20 amino acids each of the 3 nucleotide codons into mNR is translated. In red are the stop codons, which complete the translation. DNA → RNA → protein is the molecule responsible for performing most of the cell's tasks, and can have many functions, such as enzymatic, contractile, transport, immune system, signal and receptor to name a few. Like RNA and DNA, proteins are polymers made from repetitive sub-units. Instead of nucleotides, however, proteins are made up of amino acids. Each amino acid has particular properties of size, load, shape and acidity. As such, the additional structure emerges beyond the simple sequence of amino acids (the primary structure), as a result of interactions between amino acids. As such, the three-dimensional shape, and therefore the function, of a protein is determined by its sequence. However, determining the shape of a protein from its sequence is an unresolved problem in computational biology. lecture1_transcript.html-Regulation All genes are not expressed at the same time in a cell. For example, cells would waste energy if they produced a lactose carrier in the absence of lactose. It is important for a cell to know which genes it should express and when. A regulatory network is involved to control the level of gene expression in a specific circumstance. Transcription is one of the steps at which protein levels can be regulated. The promoter region, a segment of DNA found upstream (after the end of 5' genes, works in transcriptional regulation. The promoter region contains patterns that are recognized by proteins called transcription factors. When linked, transcription factors can recruit RNA polymerrase, leading to gene transcription. However, transcription factors can also be involved in complex regulatory interactions. There may be several binding sites in a promoter, which can act as a logical gateway for gene activation. The regulation of eucaryokes can be extremely complex, as the expression of genes has been affected not only by the neighbouring region of the promoter, but also by extenders and distant repressors. We can use probabilistic models to identify genes that are regulated by a given transcription factor. For example, given all the known reasons for linking a given transcription factor, we can calculate the probability that a candidate motive will also link the (see notes for precept #1). Comparative sequence analysis can also be used to identify regulatory grounds, as regulatory grounds show characteristic patterns of evolutionary conservation. Lake operon in E. coli and other bacteria is an example of a simple regulatory circuit. In bacteria, genes with related functions are often located side by side, controlled by the same regulatory region, and transcribed together; this group of is called an operon. The lake operon works in the metabolism of sugar lactose, which can be used as an energy source. However, bacteria prefer to use glucose as an energy source, so if there is glucose present in the environment the bacteria do not want to make proteins that are encoded by the lake operon. Therefore, the transcription of the lake operon is regulated by an elegant circuit in which transcription occurs only if there is lactose but not glucose present in the environment. © unknown source. All rights reserved. This content is excluded from our Creative Commons license. For more information, see . Figure 1.6: Operon Lac illustrates a simple biological regulatory system. In the presence of glucose, lactose metabolism genes are proven because glucose inactivates an activating protein. In the absence of lactose, a repressive protein is also the operon. The genes of lactose metabolism are expressed only in the presence of lactose and in the absence of glucose. © unknown source. All rights reserved. This content is excluded from our Creative Commons license. For more information, see . Figure 1.7: Metabolic pathways and regulation can be studied by computational biology. The models are made from genome-wide information and used to predict metabolic function and metabolic engineering. An example of biological engineering is modifying the genome of bacteria to overproduce artemesenin, an antibiotic used to treat malaria. lecture1_transcript.html living organisms are made from self-organized building blocks. The energy source is needed to organize the blocks. The basic mechanism involved in building blocks is the degradation of small molecules to obtain energy to build large molecules. The process of degradation of molecules to release energy is called catabolism and the process of using energy to assemble more complex molecules is called anabolism. Anabolism and catabolism are both metabolic processes. Metabolism regulates the flow of mass and energy in order to maintain an organism in a state of low entropy. Enzymes are a critical component of metabolic reactions. The vast majority of enzymes (but not all!) are proteins. Many biologically critical reactions have high activation energies, so that the non-catalyzed reaction would occur extremely slowly or not at all. Enzymes accelerate these reactions, so that they can occur at a rate that is sustainable for the cell. In living cells, are organized into metabolic pathways. A reaction can have many steps, with the products of one stage serving as substrate for the next. In addition, metabolic reactions often require an energy investment (especially as a molecule called ATP), and the energy released by a reaction can be captured by a subsequent reaction in the pathway. Metabolic pathways are also important for regulating metabolic reactions inhibited, later steps may lack the substrate or energy they need to proceed. Often, regulatory checkpoints appear early in the metabolic pathways, because if the reaction is to be stopped, it is obviously better to stop it before a lot of energy has been invested. lecture1_transcript.html-SystemsBiology Systems biology strives to explore and explain the behavior that emerges from complex interactions between the components of a biological system. An interesting recent paper in systems biology is Metabolic gene regulation in a dynamically changing environment (Bennett et al., 2008). This work hypothesizes that yeast is a linear, time-invariant system, and executes a signal (glucose) through the system to observe the response. A periodic response to low-frequency fluctuations in glucose levels is observed, but there is little response to high-frequency fluctuations in glucose levels. Thus, this study finds that yeast acts as a low pass filter for fluctuations in glucose levels. lecture1_transcript.html-SyntheticBiology Not only can we use computational approaches to model and analyze biological data collected from cells, but we can also design cells that implement specific logic circuits to perform new functions. The task of designing new biological systems is known as synthetic biology. A particularly notable success of synthetic biology is the improvement of artemesenin production. Arteme-senin is a drug used to treat malaria. However, artemisinin was quite expensive to produce. Recently, a yeast strain was designed to synthesize a precursor of artemisininic acid at half the previous cost. There are various model organisms for all aspects of human biology. The importance of using model organisms at an appropriate level of complexity. Note: In this particular book, we will focus on human biology, and we will use examples of yeast baker Saccharomyces cerevisiae, fruit fly Drosophila melanogaster, nematode worm Coenorhabditis elegans, and mouse from the house Mus musculus. We will only deal with bacterial evolution in the context of the metagenomics of the human microbiome. microbiome.