The Central Dogma and Basic Transcription*

Andrew Hughes

1 The Central Dogma

The Central Dogma

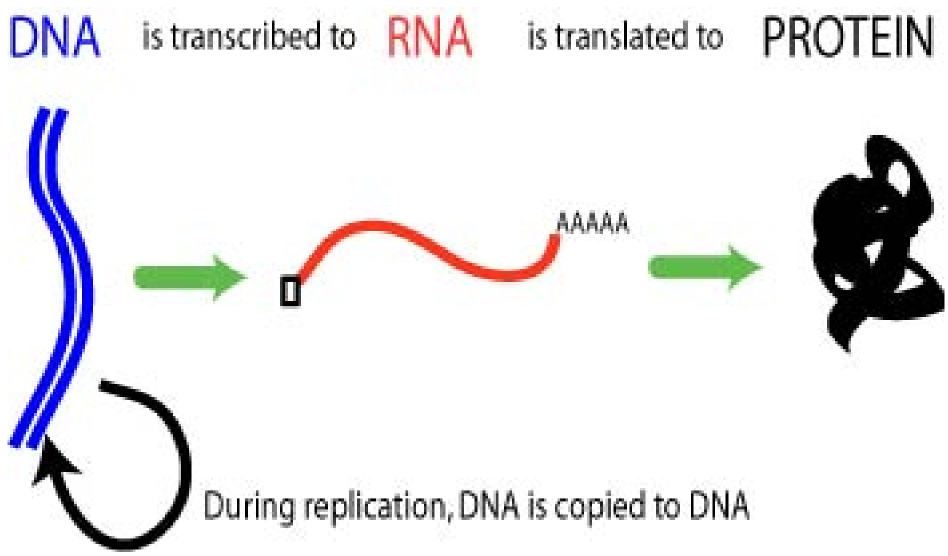

Figure 1

The Dogma is: DNA -> RNA -> Protein

The Central Dogma of genetics is: DNA is transcribed to RNA which is translated to protein. Protein is never back-translated to RNA or DNA; and except for retroviruses, DNA is never created from RNA. Furthermore, DNA is never directly translated to protein. DNA to RNA to protein.

*Version 1.5: Jul 27, 2003 5:37 pm -0500

†http://creativecommons.org/licenses/by/1.0

DNA is the long term hard-copy of the genetic material; by way of analogy it is similar to information stored on a computer's hard-disk drive. DNA is very stable and inert (that's kinda the point of DNA).

RNA is a temporary intermediary between DNA and the protein making factories, the ribosomes. RNA could be compared to information stored in a cache in that the lifetime of RNA is much shorter than that of either DNA or the average protein, and also RNA serves to carry information from the genome, located in the nucleus of the cell, to the ribosomes, which are located outside of the nucleus either in the cytosol or on the endoplasmic reticulum (which is a large set of folded membranes proximal to the nucleus that help manufacture proteins for extra-cellular export).

Proteins could be viewed as the programs of the cell (to complete our analogy). They are the physical representation of the abstract information contained within the genome. Proteins vary greatly in their activity and halflife. Trying to classify proteins is like classifying programs, they come in all shapes and sizes (once again, that's kinda the point). For example, collagen is a ubiquitous protein that makes up almost all of your bodies main structural tissue. It is not particularly active, per se, but is capable of self-polymerization, complete with tertiatry structure, into different forms depending on the particular type of collagen being expressed. Collagen makes up the bulk of you hair, tendons, skin, and ligaments (and much much more). In contrast to collagen, take lactase, the enzyme that allows you to digest lactose. Lactose is the sugar found in milk and dairy products. Lactose intolerant people have a faulty lactase gene and cannot digest lactose. Lactose has no structural function; it's function is entirely enzymatic. Hormones are proteins as well. Insulin, for example, is a protein that your body uses to signal various organs to take up sugar from the blood stream.

DNA, RNA, and Proteins

- DNA is long-term storage. It is stable, packaged, and inert.
- RNA is short-term storage. It is unstable and lacks secondary structure. Some RNA has enzymatic activity.
- Proteins are the 'programs' of the cells. They are the physical manifestations of the abstract information recorded in the genome.

DNA, RNA, and Proteins

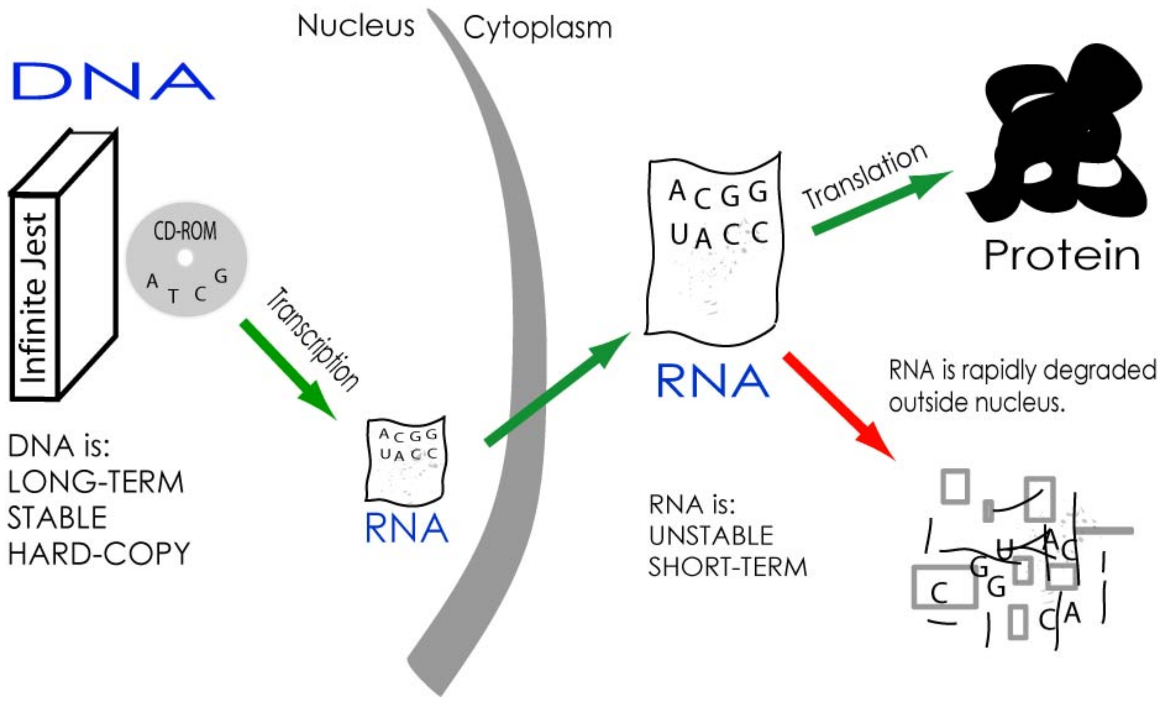

Figure 2

An Analogy

Imagine you want to build the worlds best paper airplane. While doing research on the topic in your local library you discover a book called "How to Build the Worlds Best Paper Airplane!". You are shocked and enthralled by your fortuitous discovery and you hurry to the circulation counter of the library to check the book out. However, to your dismay the book is too valuable to be checked out. "This book must not leave the library," says the very stern librarian who eyes you suspiciously, then continues, "None of our books can leave the library. They are too valuable and the outside world is too dangerous!" Meekly, you return to your desk in the library and flip through the book to see if you could possibly build the airplane inside the library. While doing so you realize that the paper airplane is much too complicated for you to build yourself. You think in dismay, "What am I going to do? I can't remove the book from the library and I can't even begin to build the plane by myself." Then you remember that your friendly neighbor, Mr. R, mentioned to you the other day that he is a professional orgami expert. You call him and beg his help, explaining the situation, and he replies, "No problem, just bring me a copy of the plans and we'll get it build in a jiffy." "Great", you shout aloud through the phone breaking library and cell-phone etiquite and earning yourself nasty looks from the poor med-students and law-students busy engraving large books into their respective memories. After you regain your composure you quietly copy the relevant sections of the book, leave the library, and hurry over to your neighbors house with the plans. You and Mr. R work patiently through the night and by morning you have created the worlds best paper airplane!

This analogy illustrates the process of gene expression. The library is the nulceus of the cell and the books are the DNA. The books (DNA) do not leave the library (nucleus). Information from the books must be copied and transported out of the library for use. Transcription is the process of copying the books and

RNA would be analagous to the copies themselves. Mr. R is, of course, our wonderful ribosomes which are the protein factories that translate the genetic code into physical products (more proteins). The paper airplane would be analagous to the translated protein product.

2 Introduction to Transcription

Transcription

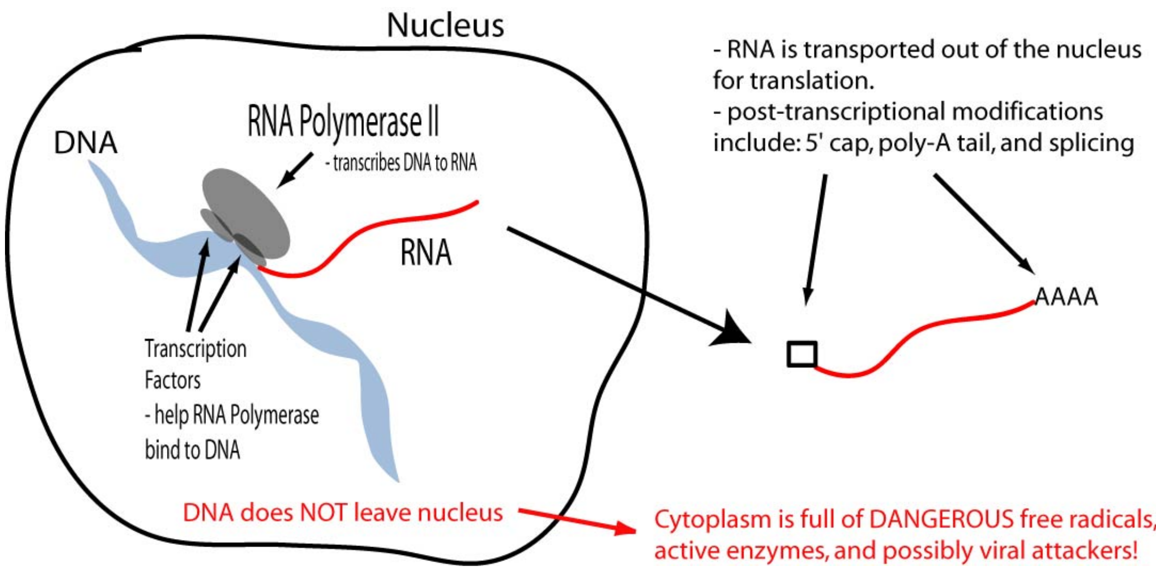

Figure 3

Transcription is the process of creating RNA from DNA.

Transcription is the first major step in gene expression (translation, the process of making a polypeptide from an mRNA transcript, is the second major step). DNA is sequestered inside the nucleus for organizational and security reasons. The nucleus provides a safe, controlled environment in which to store and access the cell's genetic material. To a computer scientist, this organization should make intuitive sense. The cell's valuable data is sequestered in specific compartment and is accessed in a controlled and isolated manner. This level of abstraction protects the genetic code from dangerous enzymes and free radicals floating around in the cytoplasm. It also helps to insulate the genome from viral attack. However, because the creation of proteins occurs outside the nucleus in the cytoplasm, the cell needs some way of transporting the genetic information from the nucleus to the cytoplasm. RNA is the cell's intermediary between the DNA in the nucleus and the ribosomes in the cytoplasm, where translation occurs. Geneticists are just recently realizing that RNA is actually much more than a simple messenger, it appears to have many interesting active properties as well, however there is no doubt that transcription is one of it's main functions. Transcription is particularly important in terms of gene expression because it is the point where most of the regulation of gene expresion occurs. Generally, the amount of a gene product produced depends directly on how much RNA transcript is created. If a cell wants to make more of a gene, it typically makes more RNA from the gene which is then exported to the cytoplasm and translated into more protein. If the cell want's to stop the expression of a gene, it stops making RNA transcripts and the gene expression turn off. The expression turns off because RNA's are rapidly degraded by nucleases inside the cell (and elsewhere for that matter, nucleases are ubiquitous), thus once the cell stops making RNA, any leftover RNA in the cytoplams is degraded and expression stops.

- Transcription occurs in the cell's nucleus.
- RNA polymerase is the protein molecule that reads the DNA and creates the RNA intermediary.
- Transcription requires: DNA, RNA polymerase, ribonucleotides, and some ATP for energy.
- Uracil (U) is substituted for thymine (T) in RNA.
- Transcription initiation is the main point of regulation of gene expression.