

Central Dogma of Molecular Biology

M. Kurka

github.com/miroslavkurka/Alignment-lecture

June 9, 2021

Overview

- DNA
- Protein
- Sequence
- The Dogma
- Conclusion

What is DNA?

DNA

DNA is long linear polymer that carry information in a form that can be passed from one generation to the next. DNA consist of a large number of linked nucleotides, each composed of a sugar, a phosphate, and a base. Sugars linked by phosphates form a common backbone that plays a structural role, whereas the sequence of bases along a nucleic acid strand carries genetic information.

nucleotides \approx bytes

What is a Protein?

Protein

Proteins are linear polymers built of monomer units called amino acids. The sequence of linked amino acids is called the primary structure. Proteins spontaneously fold up into three-dimensional structures that are determined by the sequence of amino acids in the protein polymer. Three-dimensional structure formed by hydrogen bonds between amino acids near one another is called secondary structure, whereas tertiary structure is formed by long-range interactions between amino acids. Protein function depends directly on this three-dimensional structure

protein \approx function

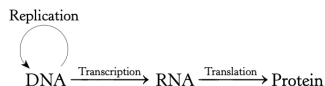
What are Sequences?

DNA sequence

$$X = x_1 x_2 \dots x_n, x_i \in \{C, A, G, T\}$$

100

© 2006 The Authors
Journal compilation © 2006 Blackwell Publishing Ltd



[C](#) [G](#) [U](#) [T](#) [I](#) [N](#) [G](#)

The Actual Dogma

“The central dogma of molecular biology deals with the detailed residue-by-residue transfer of sequential information. It states that such information can not be transferred from protein to either protein or nucleic acid. ”

– Francis Crick, Central Dogma of Molecular Biology¹

- Postulated in 1958 paper
- Restated by Crick in 1970 as a response to Temin's paper²
- The common misconception comes from Watson's 1965 text

¹CRICK, F. Central Dogma of Molecular Biology. Nature 227, 561–563 (1970).

²TEMIN, H. Central Dogma Reversed. Nature 226, 1198–1199 (1970) ▶

The Essence of the Problem

A systematic discussion of our present knowledge of protein synthesis could usefully be set out under three headings, each dealing with a flux: the flow of energy, the flow of matter, and the flow of information. I shall not discuss the first of these here. I shall have something to say about the second, but I shall particularly emphasize the third - the flow of information.

By information I mean the specification of the amino acid sequence of the protein. It is conventional at the moment to

Figure: Archived copy of the 1958 paper from Cold Spring Harbor

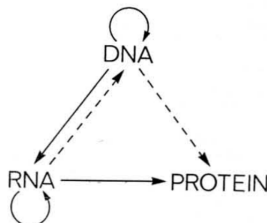
[illegible]

Fig. 2. The arrows show the situation as it seemed in 1958. Solid arrows represent probable transfers, dotted arrows possible transfers. The absent arrows (compare Fig. 1) represent the impossible transfers postulated by the central dogma. They are the three possible arrows starting from protein.

Three Categories Now¹

We have three categories defined as followed

- General Transfer
- Special Transfer
- Unknown Transfer

DNA→DNA	RNA→RNA	Protein→Protein
DNA→RNA	RNA→DNA	Protein→DNA
RNA→Protein	DNA→Protein	Protein→RNA

Figure: General, Special and Unkown Transfer respectively

So far, however, there is no evidence for the first two of these except in a cell infected with an RNA virus. In such a cells the central dogma demands that at least one of the first two special transfers should occur this statement, incidentally, shows the power of the central dogma in making theoretical predictions.