

Course- B.Sc. (Botany Honours), Part -3

Paper-VI (Group-B), Molecular Biology

Topic- Genetic Code.

PDF

Binod Kumar Pandey

Assistant Professor & HOD, Botany Department,

Ram Ratan Singh College, Mokama, Patna.

[E.mail- binodkp14@gmail.com](mailto:binodkp14@gmail.com)

Genetic code

The genetic code may be defined as the exact sequence of DNA nucleotides read as three letter words or codons, that determines the sequence of amino acids in protein synthesis. In other words, the genetic code is the set of rules by which information encoded in genetic material (DNA or RNA sequences) is translated into proteins (amino acid sequences) by living cells.

The main points related to genetic code are given below:

1. The genetic code is 'read' in triplets of bases called codons. In other words, a set of three nucleotide bases constitutes a codon.
2. In a triplet code, three RNA bases code for one amino acid.
3. There are 64 codons which correspond to 20 amino acids and to signals for the initiation and termination of transcription.
4. The code uses codons to make the amino acids that, in turn, constitute proteins.
5. Each triplet [codon] specifies one amino acid in a protein structure or a start signal or stop signal in protein synthesis.
6. The code establishes the relationship between the sequence of bases in nucleic acids (DNA and the complementary RNA) and the sequence of amino acids in proteins.
7. The code explains the mechanism by which genetic information is stored in living organisms.

Types of Genetic Code:

The genetic code is of two types. The genetic code can be expressed as either RNA codons or DNA codons. RNA codons occur in messenger RNA (mRNA) and are the codons that are actually "read" during the synthesis of polypeptides (the process called translation).

But each mRNA molecule acquires its sequence of nucleotides by transcription from the corresponding gene [DNA]. Because DNA sequencing has become so rapid and because most genes are now being discovered at the level of DNA before they are discovered as mRNA or as a protein product, it is extremely useful to have a table of codons expressed as DNA. Both tables are given here.

DNA Codons:

These are the codons as they are read on the sense (5' to 3') strand of DNA. Except that the nucleotide thymine (T) is found in place of uracil (U), they read the same as RNA codons. However, mRNA is actually synthesized using the antisense strand of DNA (3' to 5') as the template.

Types of Codon:

The genetic code consists of 64 triplets of nucleotides. These triplets are called codons. With three exceptions, each codon encodes for one of the 20 amino acids used in the synthesis of proteins. This produces some redundancy in the code.

Most of the amino acids are encoded by more than one codon. One codon that is AUG serves two related functions. It signals the start of translation and codes for the incorporation of the amino acid methionine (Met) into the growing polypeptide chain.

The codons are of two types

1. Sense Codon:

Those codons that code for amino acids are called sense codons. There are 61 sense codons in the genetic code which code for 20 amino acids.

2. Signal Codons:

Those codons that code for signals during protein synthesis are known as signal codons. There are four codons which code for signal. These are AUG, UAA, UAG and UGA.

Signal codons are of two types

(i) Start Codons:

The codon which starts the translation process is known as start codon. It is also known as initiation codon because it initiates the synthesis of polypeptide chain. Example of this codon is AUG. This codon also codes for the amino acid methionine. In some cases, valine (GUG) codes for start signal. In eukaryotes, the starting amino acid is methionine, while in prokaryotes it is N-formyl methionine.

(ii) Stop Codons:

Those codons that provide signal for termination of polypeptide chain are known as stop codons. These codons are also known as termination codons because they provide signal for the termination and release of polypeptide chain. Examples of stop codons are UAA, UAG and UGA. Since stop signal codons do not code for any amino acid they were earlier called as non-sense codons.

Signals of stop or termination codons are read by proteins called release factors. Stop signals are not read by tRNA molecules. In prokaryotes, release factors are RF1, RF2 and RF3. The factor RF1 recognizes stop codons UAA and UAG, while RF2 recognizes UAA and UGA. The function

of RF3 is to stimulate RFI and RF2. In eukaryotes, a single release factor (RF) recognizes all three stop codons.

Properties of Genetic Code:

Genetic code has some important properties.

1. The Code is Triplet:

The genetic code is triplet. The triplet code has 64 codons which are sufficient to code for 20 amino acids and also for start and stop signals in the synthesis of polypeptide chain. In a triplet code three RNA bases code for one amino acid.

2. The Code is Universal:

The genetic code is almost universal. The same codons are assigned to the same amino acids and to the same START and STOP signals in the vast majority of genes in animals, plants, and microorganisms. However, some exceptions have been found.

Most of these involve assigning one or two of the three STOP codons to an amino acid instead. Some exceptions have been reported for mitochondrial genome and in unicellular eukaryotes for synthesis of nonstandard proteins such as selenocysteine and pyrrolysine.

3. The Code is Commaless:

It is believed that the genetic code is commaless. In other words, the codons are continuous and there are no demarcation lines between codons. Deletion of a single base in a commaless code alters the entire sequence of amino acids after the point of deletion as given below.

RNA Bases	UUU	[C]UC	GUA	UCC	ACC
Amino Acids	Phe	Leu	Val	Ser	Thr

The deletion of base C from leucine will change the genetic message in the following manner:

RNA Bases	UUU	UCG	UAU	CCA	CC
Amino Acids	Phe	Ser	Tyr	Pro	-

Experimental evidences also reveal that the genetic code is commaless. Khorana and coworkers have also demonstrated that the genetic code is commaless.

4. The Code is Non-Overlapping:

Three nucleotides or bases code for one amino acid. In a non-overlapping code, six bases will code for two amino acids. In a non-overlapping code, one letter is read only once. In overlapping code, six nucleotides or bases will code for 4 amino acids, because each base is read three times

Example: There are Bases : CATGAT

Non-overlapping Code : 2 that is CAT and GAT;

Overlapping Code : 4 that is CAT, GAT, ATG and TAT

If mutation of one base into another leads in alteration of one amino acid only, it indicates that the code is non-overlapping. Mutation experiments with TMV gave similar results which indicated that the code is non-overlapping.

5. The Code is Non-ambiguous:

The genetic code has 64 codons. Out of these, 61 codons code for 20 different amino acids. However, none of the codons codes for more than one amino acid. In other words, each codon codes only for one amino acid. This clearly indicates that the genetic code is non-ambiguous. In case of ambiguous code, one codon should code for more than one amino acid. In the genetic code there is no ambiguity.

6. The Code is Redundant:

In most of the cases several codons code for the same amino acid. Only two amino acids, viz. tryptophan and methionine are coded by one codon each. Nine amino acids are coded by two codons each, one amino acid [Isoleucine] by three codons, five amino acids by 4 codons each, and three amino acids by 6 codons each (Table 28.4).

This multiple system of coding is known as degenerate or redundant code system. Such system provides a protection to the organism against many harmful mutations, because if one base of a codon is mutated, there are other codons which will code for the same amino acid and there will be no alteration in the polypeptide chain.

The redundancy or degeneracy of the code is not random except for serine, leucine and arginine. All codons coding for same amino acid are in the same box (except above three). Thus the first two letters are GC in all four codons of alanine and GC and GU in all four codons of valine (Table 28.4).

TABLE 28.4. Differences between codons and anticodons

<i>S.No.</i>	<i>Particulars</i>	<i>Codons [mRNA] AUG</i>	<i>Anticodons [tRNA] UAC</i>
1.	Number in the code	64	Lesser than 64
2.	Amino acids coded	One by one codon	Some pair with more than 1 codon
3.	Polarity	5 to 3	3 to 5
4.	Base number	1 2 3	3 2 1

7. The Code Has Polarity:

The code has a definite direction for reading of message, which is referred to as polarity. Reading of codon in opposite direction will specify for another amino acid due to alteration in the base sequences in the code.

In the following codons, reading of message from left to right and right to left will specify for different amino acids. Because the codon in the following case will be read as UUG from left to right and as GUU from right to left which codes for another amino acid.

Codons	UUG	GUC	UCG	CCA	AGG
Reading left to right	Leu	Val	Ser	Pro	Arg
Reading Right to left	Val	Leu	Ala	Thr	Gly

This is well known that the message in mRNA is read in the 5' - 3' direction. Thus the polarity of genetic code is from 5' end to 3' end.