THE GENETIC CODE -2

Dr Anita Kumari

Assistant Professor (Guest faculty)

Dept of Zoology

L S College

Genetic code, is the sequence of nucleotides in deoxyribonucleic acid (DNA) and ribonucleic acid (RNA) that determines the amino acid sequence of proteins. Though the linear sequence of nucleotides in DNA contains the information for protein sequences, proteins are not made directly from DNA. Instead, a messenger RNA (mRNA) molecule is synthesized from the DNA and directs the formation of the protein. RNA is composed of four nucleotides: adenine (A), guanine (G), cytosine (C), and uracil (U).

Three adjacent nucleotides constitute a unit known as the codon, which codes for an amino acid. For example, the sequence AUG is a codon that specifies the amino acid methionine. There are 64 possible codons, three of which do not code for amino acids but indicate the end of a protein. The remaining 61 codons specify the 20 amino acids that make up proteins. The AUG codon, in addition to coding for methionine, is found at the beginning of every mRNA and indicates the start of a protein. Methionine and tryptophan are the only two amino acids that are coded for by just a single codon (AUG and UGG, respectively). The other 18 amino acids are coded for by two to six codons. Because most of the 20 amino acids are coded for by more than one codon, the code is called degenerate.

The genetic code, once thought to be identical in all forms of life, has been found to diverge slightly in certain organisms and in the mitochondria of some eukaryotes. The deciphering of the genetic code was accomplished by American biochemists Marshall W. Nirenberg, Robert W. Holley, and Har Gobind Khorana in the early 1960s.

Nucleotide triplets (codons) specifying different amino acids are shown in the table.

TABLE 4.3. The 64 triplet codons (in mRNA) and their meanings in terms of amino acids

First position (5′ end)	Second position (middle or central base)				Third position (3′ end)
	U	C	A	G	
U	UUU phe	UCU ser	UAU tyr	UGU cys	U
	UUC phe	UCC ser	UAC tyr	UGC cys	C
	UUA ile	UCA ser	UAA stop	UGA stop	A
	UUG ile	UCG ser	UAG stop	UGG trp	G
C	CUU leu	CCU pro	CAU his	CGU arg	U
	CUC leu	CCC pro	CAC his	CGC arg	C
	CUA leu	CCA pro	CAA gln	CGA arg	A
	CUG leu	CCG pro	CAG gln	CGG arg	G
A	AUU ile	ACU thr	AAU asn	AGU ser	U
	AUC ile	ACC thr	AAC asn	AGC ser	C
	AUA ile	ACA thr	AAA lys	AGA arg	A
	AUG met	ACG thr	AAG lys	ACG arg	G
G	GUU val	GCU ala	GAU asp	GGU gly	U
	GUC val	GCC ala	GAC asp	GGC gly	C
	GUA val	GCA ala	GAA glu	GGA gly	A
	GUG val	GCG ala	GAG glu	GGG gly	G

Degeneracy of genetic code:

Crick's work is suggested that the genetic code is **degenerate.** It simply means that each of the 64 triplets must have some meaning within the code; so at least some amino acids must be specified by two or more different triplets. If only 20 triplets are used (with the other 44 being nonsense, in that they do not code for any amino acid), then most frameshift mutations can be expected to produce nonsense words, which presumably stops the protein-building process. If this were the case, then the suppression of frameshift mutations would rarely, if ever, work. However, if all triplets specified some amino acid, then the changed words would simply result in the insertion of incorrect amino acids into the protein. Thus, Crick reasoned that many or all amino acids must have several different names in the base-pair code; this hypothesis was later confirmed biochemically.

The genetic code is non-overlapping.Three bases encode an amino acid. termed codons.The code is read from a fixed starting point and continues to the end of the coding sequence. We know this because a single frame shift mutation anywhere in the coding sequence alters the codon alignment for the rest of the sequence.The code is degenerate in that some amino acids are specified by more than one codon.

Codon degeneracy.

There are only 20 natural amino acids used as building blocks for proteins, many amino acids are specified by more than one codon, a phenomenon known as codon degeneracy.

There are two methods by which the same amino acid is specified by two or more codons**:**

1. The tRNAs accepting the same amino acid are different for different synonymous codons. Such tRNAs are called "isoacceptor tRNAs" and they differ in anticodons. For example, one of the tRNAs carrying leucine is $tRNA_1^{leu}$ with

anticodon 3′ GAC5′, while the other is $tRNA_2^{leu}$ with anticodon 3′ GAG5′.
2. A single type of tRNA pairs with two or more synonymous codons. For example, tRNA. Accepting the amino acid alanine in yeast ($tRNA^{aIa}$) bears the anticodon 3′ CGI5′ that can pair with the codons 5′ GCU3′, 5′ GCC3 and 5′ GCA3′ on mRNA Crick in 1966 proposed the "wobble hypothesis" to explain the pairing of a single type anticodon with synonymous codons.
According to the Wobble hypothesis, the base position at the 5′-end of anticodon is the **"wobble position".** Two bases of anticodon from 3′-end are complementary to the two bases of the codon (in mRNA). The base at the wobble position can pair with different bases. For example, a single type of $tRNA^{gly}$ with the anticodon 3′ CCI5′ can pair with the codons 5'GGU3′, 5'GGC3′ and 5'GGA3′ specifying the amino acid glycine.
Thus inosine (I) at the wobble position can pair, with U, C and A in the codon. Similarly, U can pair with A and G, while G at the wobble position can pair with C and/U.

The genetic code is **universal** since similar codons are assigned to identical amino acids along with similar START and STOP signals in the majority of genes in microorganisms and plants. However, a few exceptions,as stop codon.

The codons GUG and AUG may code for methionine as a starting codons, although GUG is meant for valine. This breaks the property of non-ambiguousness. Thus, it can be said that few codes often differs from the universal code or non-ambiguous code.

Thankyou