1.4 GENETIC CODE

The **genetic code** is a set of rules by which the **nucleotide sequence of DNA or RNA** is translated into the **amino acid sequence of proteins**, the fundamental components of all living cells. It is a universal language of life, guiding the synthesis of proteins from genetic material and thereby directing the structure and function of every organism on Earth. The path to deciphering the genetic code was one of the most significant scientific achievements of the 20th century. Here's a chronological outline of key discoveries:

- **1953 – Watson and Crick** proposed the **double helix model of DNA**, establishing that genetic information is stored in the sequence of nitrogenous bases.
- **1957 – George Gamow**, a physicist, proposed that combinations of three nucleotides (a **triplet code**) could encode the 20 standard amino acids. Although his specific model was incorrect, it inspired important research.
- **1961 – Crick, Brenner, Barnett, and Watts-Tobin** demonstrated experimentally that the genetic code is read in **non-overlapping triplets**. This was achieved using **frameshift mutations** in bacteriophage experiments.

- **1961 – Marshall Nirenberg and Heinrich Matthaei** conducted the famous **poly-U experiment**, synthesizing artificial mRNA composed entirely of uracil (UUU) and showing that it coded for the amino acid **phenylalanine**. This was the first direct assignment of a codon to an amino acid.
- **1960s – Nirenberg, Har Gobind Khorana, and others** systematically determined the codon assignments for all 20 amino acids using synthetic mRNA sequences and **cell-free translation systems**.
- **1968 – Nirenberg, Khorana, and Holley** were awarded the **Nobel Prize** for their contributions to understanding the genetic code and protein synthesis.

Key Features of the Genetic Code

The genetic code has several fundamental characteristics that ensure its reliability and efficiency:

1. Triplet Nature: Each amino acid is encoded by a **codon** consisting of **three nucleotides**. For example, AUG codes for **methionine**.

2. Degeneracy (Redundancy): The code is **degenerate**, meaning that most amino acids are specified by **more than one codon**. For instance, leucine is encoded by six different codons (e.g., CUU, CUA, CUG).

3. Non-Overlapping: The code is read in **non-overlapping triplets**, with each nucleotide being part of only one codon in a given reading frame.

Second letter

First letter	U	C	A	G	Third letter
U	UUU, UUC: Phe UUA, UUG: Leu	UCU, UCC, UCA, UCG: Ser	UAU, UAC: Tyr **UAA Stop** **UAG Stop**	UGU, UGC: Cys **UGA Stop** UGG Trp	U C A G
C	CUU, CUC, CUA, CUG: Leu	CCU, CCC, CCA, CCG: Pro	CAU, CAC: His CAA, CAG: Gln	CGU, CGC, CGA, CGG: Arg	U C A G
A	AUU, AUC, AUA: Ile **AUG** Met	ACU, ACC, ACA, ACG: Thr	AAU, AAC: Asn AAA, AAG: Lys	AGU, AGC: Ser AGA, AGG: Arg	U C A G
G	GUU, GUC, GUA, GUG: Val	GCU, GCC, GCA, GCG: Ala	GAU, GAC: Asp GAA, GAG: Glu	GGU, GGC, GGA, GGG: Gly	U C A G

Fig. 1.7. *Genetic Code Table.*

4. Comma-less: Codons are read **continuously**, with no punctuation or spacing between them once the start codon is recognized.

5. Universal: With few rare exceptions (e.g., in mitochondrial DNA or some protozoa), the genetic code is **virtually universal** across all organisms, from bacteria to humans—suggesting a **common evolutionary origin**.

6. Specificity: Each codon always specifies the **same amino acid**, ensuring the **fidelity** of protein synthesis.

7. Start and Stop Codons:

- The **start codon** (AUG) initiates translation and codes for methionine.
- There are **three stop codons**—UAA, UAG, and UGA—which signal the **termination** of translation and do not encode any amino acid.

The genetic code holds immense scientific and biological significance as it lies at the core of the central dogma of molecular biology—DNA is transcribed into RNA, which is then translated into protein. Understanding this code has paved the way for transformative advances in genetic engineering, synthetic biology, gene therapy, and molecular diagnostics. One of its key features, degeneracy, offers mutational robustness, allowing certain mutations—particularly in the third base of a codon—to be silent or synonymous, thus minimizing harmful effects. The near universality of the genetic code across species enables the creation of transgenic organisms, where genes from one species can function in another, exemplified by the production of human insulin in bacteria. Scientists have even expanded the genetic code, introducing non-natural amino acids through engineered codons, thereby broadening the functional diversity of proteins. Additionally, natural variations in mitochondrial genomes and some ciliates reveal evolutionary adaptations to the code. The phenomenon of codon usage bias, where certain codons are preferred over others for the same amino acid, also plays a crucial role in regulating gene expression and translation efficiency across different organisms.