GENETIC CODE

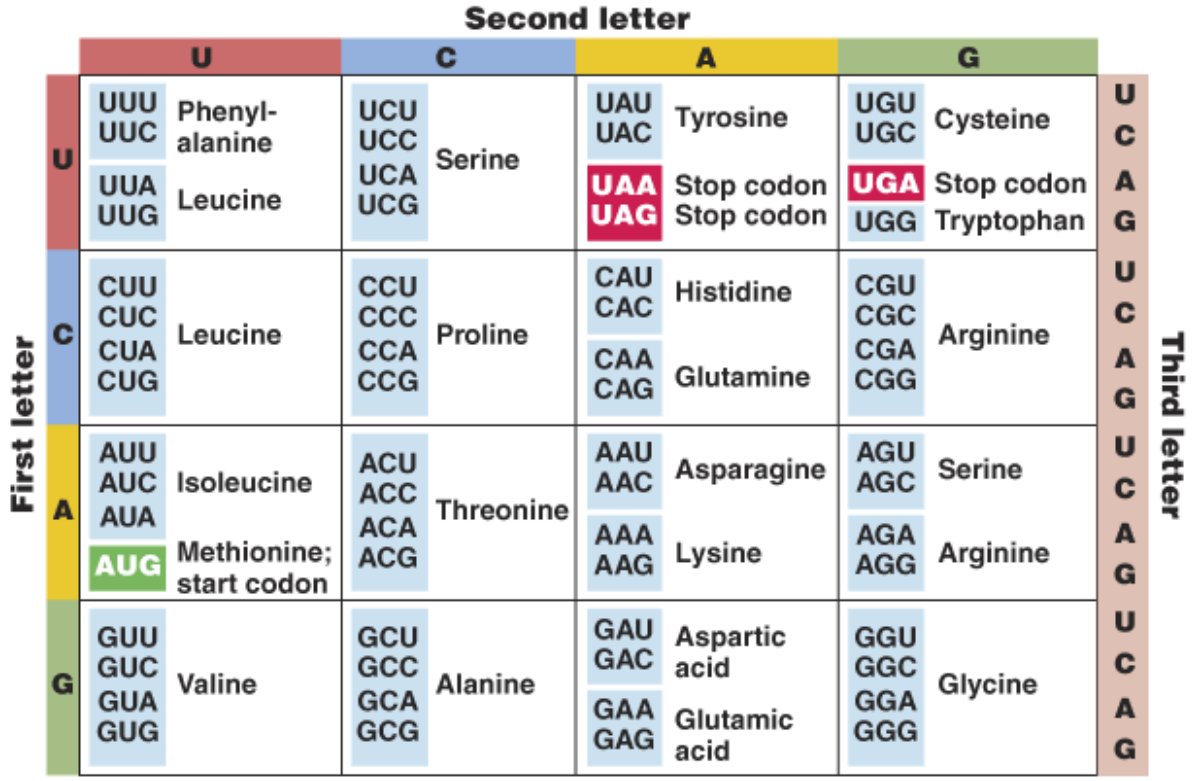
Gene is a DNA segment composed of a set of nucleotides. Nucleotides sequence within a gene controls amino acid sequence within a protein.
The genetic code is the set of rules by which the information encoded in genetic material (DNA or RNA) is translated into proteins (amino acid sequences) in living cells.
The genetic code is a set of three nucleotides. Each nucleotide is named on the basis of one of the four nitrogenous bases found in each of them. Nucleotide having nitrogenous bases Adenine is assigned the "A" letter, and nucleotide containing Guanine base is assigned letter "G". Similarly Cytosine and Thymine containing nucleotides are assigned letter "C" and "T" respectively. So A, G, C, and T are letters of Genetic code.

In 1961 Nirenberg and Matthaei "cracked" the first "word" of the genetic code (poly U)
In 1964 Nirenberg and Philip Leder discovered a way to determine the sequence of the letters in each triplet word for amino acids.
In 1968 Nirenberg won the **Nobel Prize** in Physiology or Medicine for his work on the genetic code. He shared the award with Har Gobind Khorana (University of Wisconsin), who mastered the synthesis of nucleic acids, and Robert Holley (Cornell University), who discovered the chemical structure of transfer-RNA. Collectively, the three were recognized "for their interpretation of the genetic code and its function in protein synthesis."

The codons are the hereditary units that contain the information coding for one amino acid. It consists of three nucleotides (a triplet). This information is first transcribed into messenger RNA, which has a sequence of bases complementary to the DNA from which it is copied. DNA and mRNA have only four different bases (A. G, C, T in DNA and A, G, C, U in RNA) whereas proteins contain 20 different amino acids. The code is read in groups of three bases, each triplet representing one amino acid. Three is the minimum number of bases needed to code for 20 amino acids.
The possible combination of four bases into a triplet is $4^3 = 64$
If the genetic code consisted of doublets, the number of codons would be $4^2 = 16$, which would be insufficient.
If groups of four bases were utilized, the possibilities would be many more than necessary ($4^4 = 256$).The organism uses its specific codon sequence in forming peculiar protein and hence these codes regulate and control organism at the cellular level.

20 amino acids and their short forms are as given under-

Ala=Alanine	Leu=Leucine
Asp=Aspartic acid	Phe=Phenylalanine
Glu=Glutamic acid	Thr=Threonine
Iso=Isoleucine	Glu=Glutamine
Meth=Methionine	Val=Valine
Ser=Serine	Asn=Asparagine
Tyr=Tyrosine	His=Histidine
Arg=Arginine	Lys=Lysine
Cys=Cysteine	Pro=Proline
Gly=Glycine	Trypto=Tryptophan



Source:http://biology.kenyon.adu/cources/biol114/Chap05/Chapter05.html

PROPERTIES OF GENETIC CODE

1. The code is triplet.
It is made of triplets of nucleotides which are called codons.

2. The genetic code is degenerate
Most amino acids have more than one codon, for example in the case of arginine, leucine, and serine amino acids each one of them has 6 different codons. In the case of leucine amino acid, these codons are CUA, CUC, CUG, CUU, UUA, and UUG. Similarly, GGG, GGA, GGC and GGU all code for glycine.

3. The genetic code is unambiguous
Each code has only one meaning, i.e. it codes only one amino acid. For example, AUG codes only one amino acid which is methionine. It can never code any other amino acid. GUG codes for Valine. But, in the absence of AUG, GUG can also code for valine.

4. The genetic code is universal
The same triplet in different species codes for the same amino acid, i.e. genetic code is universal for all organisms from simple bacteria to complex organisms.

5. The code is non-overlapping
Two adjoining codons cannot share a base, i.e. the three bases of one codon are distinct from the bases of adjacent codons.

6. The code is comma less

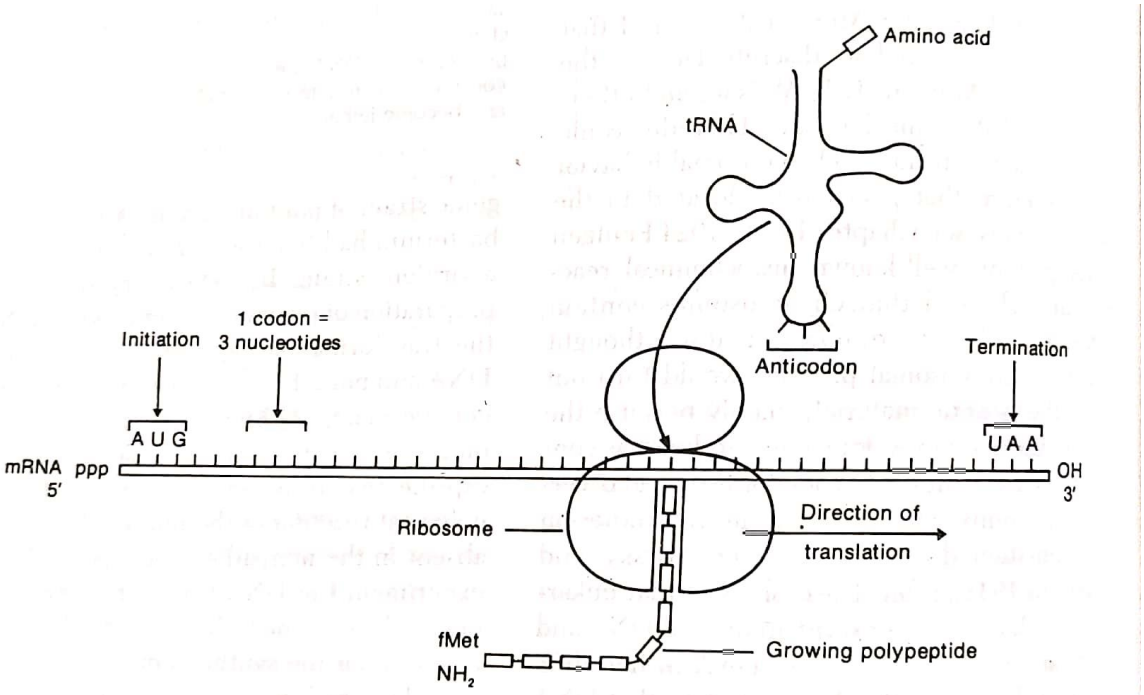
No punctuations are needed between two adjoining codons. After one amino acid is coded, the second amino acid will be automatically coded by the next three letters and no letters are wasted. There is no gap.

7. The genetic code has "start" and "stop" signals

There is only one start codon (AUG, initiation codon) which starts the translation process, but to stop this process three stop codons are present, i.e. UAA, UGA and UAG.

8. Polarity

Genetic code could be read only in 5' to 3' direction on m-RNA.



Group of three nucleotides coding for one amino acid, starting from Methionine
(Source: Cell and Molecular Biology, De Robertis)

Further Reading:

- Cell and Molecular Biology – E.D.P. De Robertis (B.I.Waverly Pvt. Ltd.)
- Molecular Biology and Biotechnology – P.K. Gupta
- Molecular Cell Biology – C.B. Powar
- Cell Biology – P.S Verma

जेनेटिक कोड

जीन एक डीएनए खंड है जो न्यूक्लियोटाइड के एक सेट से बना है। जीन के भीतर न्यूक्लियोटाइड्स अनुक्रम एक प्रोटीन के भीतर अमीनो एसिड अनुक्रम को नियंत्रित करता है।

आनुवांशिक कोड उन नियमों का समूह है जिनके द्वारा अनौपचारिक रूप से आनुवंशिक सामग्री (डीएनए या RNA) में एनकोडेड को जीवित कोशिकाओं में प्रोटीन (एमिनो एसिड अनुक्रम) में अनुवादित किया जाता है।

आनुवंशिक कोड तीन न्यूक्लियोटाइड का एक सेट है। प्रत्येक न्यूक्लियोटाइड का नाम उनमें से प्रत्येक में पाए जाने वाले चार नाइट्रोजनस आधारों में से एक के आधार पर किया जाता है। नाइट्रोजनीस आधार वाले न्यूक्लियोटाइड को एडेनिन को "ए" अक्षर सौंपा गया है, और गुआनाइन आधार वाले न्यूक्लियोटाइड को "जी" अक्षर सौंपा गया है। इसी तरह न्यूक्लियोटाइड युक्त साइटोसिन और थाइमिन को क्रमशः "सी" और "टी" अक्षर सौंपे जाते हैं। इसलिए A, G, C और T जेनेटिक कोड के अक्षर हैं।

1961 में निरेनबर्ग और मथाई ने जेनेटिक कोड (पॉली यू/poly U) का पहला "शब्द" खोजा

1964 में निरेनबर्ग और फिलिप लेडर ने अमीनो एसिड के लिए प्रत्येक ट्रिपल शब्द में अक्षरों के अनुक्रम को निर्धारित करने का एक तरीका खोजा।

1968 में निरेनबर्ग ने आनुवंशिक कोड पर अपने काम के लिए फिजियोलॉजी या मेडिसिन में नोबेल पुरस्कार जीता। उन्होंने हर गोबिंद खोराना (विस्कॉन्सिन विश्वविद्यालय) के साथ पुरस्कार साझा किया, जिन्होंने न्यूक्लिक एसिड के संश्लेषण में महारत हासिल की, और रॉबर्ट होली (कॉर्नेल विश्वविद्यालय), जिन्होंने स्थानांतरण-आरएनए (tRNA)की रासायनिक संरचना की खोज की। सामूहिक रूप से, तीनों को "आनुवांशिक कोड की व्याख्या और प्रोटीन संश्लेषण में इसके कार्य के लिए पहचाना गया।"

कोडन वंशानुगत इकाइयाँ हैं जिनमें एक अमीनो एसिड के लिए सूचना कोडिंग होती है। इसमें तीन न्यूक्लियोटाइड (एक ट्रिपलेट) होते हैं। यह जानकारी पहले मैसेंजर आरएनए में स्थानांतरित की जाती है, जिसमें डीएनए के पूरक अड्डों का अनुक्रम होता है, जहां से इसे कॉपी किया जाता है। डीएनए और एमआरएनए में केवल चार अलग-अलग आधार हैं (डीएनए में A, G, C, T और आरएनए में A, G, C, U) जबकि प्रोटीन में 20 अलग-अलग अमीनो एसिड होते हैं। कोड को तीन आधारों के समूहों में पढ़ा जाता है, प्रत्येक ट्रिपल एक अमीनो एसिड का प्रतिनिधित्व करता है। तीन अमीनो एसिड के लिए कोड करने के लिए आवश्यक आधारों की न्यूनतम संख्या है।

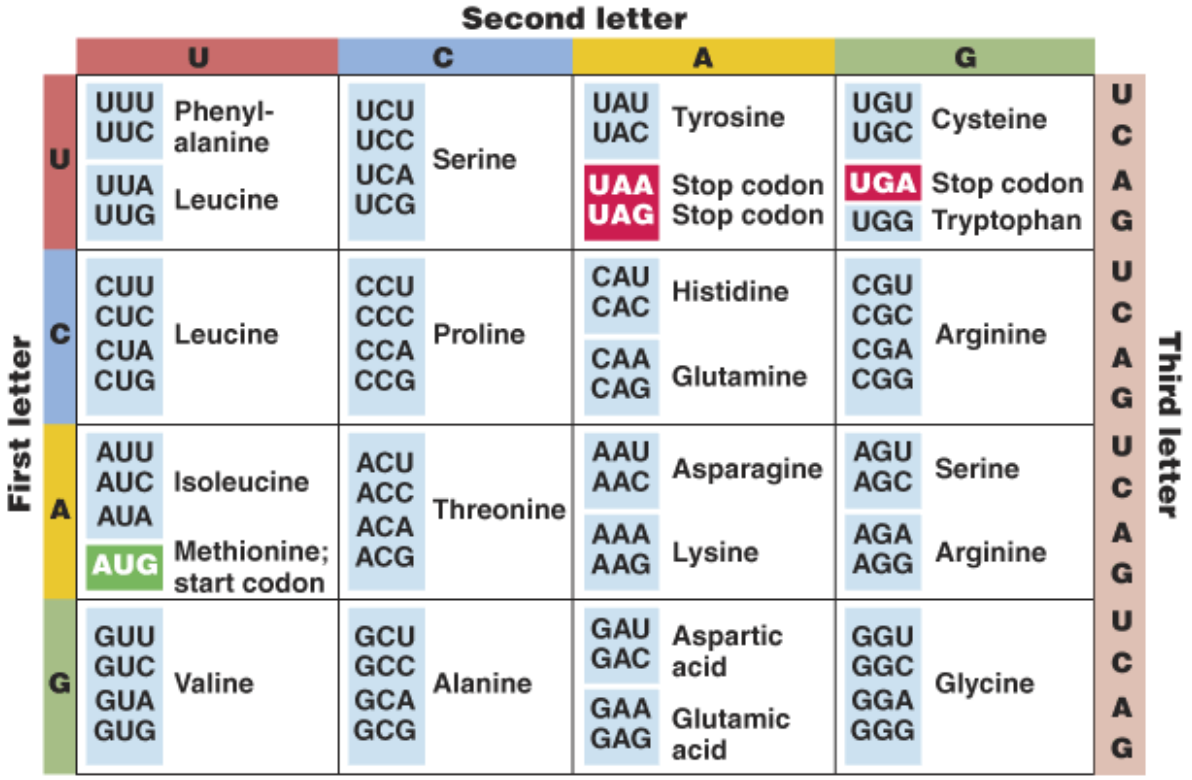
ट्रिपल में चार आधारों का संभावित संयोजन $4^3 = 64$ है

यदि आनुवंशिक कोड में दोहे होते हैं, तो कोडन की संख्या $4^2 = 16$ होगी, जो अपर्याप्त होगी।

यदि चार आधारों के समूहों का उपयोग किया गया था, तो संभावनाएं आवश्यकता से बहुत अधिक होंगी ($4^4 = 256$) जीव अजीब प्रोटीन बनाने में अपने विशिष्ट कोडन अनुक्रम का उपयोग करता है और इसलिए ये कोड सेलुलर स्तर पर जीव को नियंत्रित करते हैं।

20 अमीनो एसिड और उनके छोटे रूप निम्नानुसार हैं-

Ala=Alanine	Leu=Leucine
Asp=Aspartic acid	Phe=Phenylalanine
Glu=Glutamic acid	Thr=Threonine
Iso=Isoleucine	Glu=Glutamine
Meth=Methionine	Val=Valine
Ser=Serine	Asn=Asparagine
Tyr=Tyrosine	His=Histidine
Arg=Arginine	Lys=Lysine
Cys=Cysteine	Pro=Proline
Gly=Glycine	Trypto=Tryptophan



स्रोत: http: //biology.kenyon.adu/cources/biol114/Chap05/Chapter05.html

आनुवंशिक कोड के गुण

1. कोड त्रिक (ट्रिपलेट) है।

यह न्यूक्लियोटाइड्स के ट्रिपलेट्स से बना है जिन्हें कोडन कहा जाता है। (AUG, GCA, UGG, आदि)

2. आनुवंशिक कोड पतित है

अधिकांश अमीनो एसिड में एक से अधिक कोडन होते हैं, उदाहरण के लिए आर्गिनिन, ल्यूसीन और सेरीन अमीनो एसिड के मामले में, उनमें से प्रत्येक में 6 अलग-अलग कोडन होते हैं। ल्यूसीन अमीनो एसिड के मामले में, ये कोडन CUA, CUC, CUG, CUU, UUA और UUG हैं। इसी तरह, GGG, GGA, GGC और GGU ग्लाइसिन के लिए सभी कोड हैं।

3. आनुवंशिक कोड असंदिग्ध है

प्रत्येक कोड का केवल एक ही अर्थ होता है, अर्थात यह केवल एक एमिनो एसिड को कोड करता है। उदाहरण के लिए, AUG केवल एक एमिनो एसिड को कोड करता है जो मेथियोनीन है। यह किसी भी अन्य अमीनो एसिड को कभी भी कोड नहीं कर सकता है। Valine के लिए GUG कोड। लेकिन, AUG की अनुपस्थिति में, GUG वेलिन के लिए भी कोड कर सकते हैं।

4. आनुवंशिक कोड सार्वभौमिक है

एक ही अमीनो एसिड के लिए विभिन्न प्रजातियों के कोड में एक ही ट्रिपल होता है, यानी आनुवंशिक कोड सरल बैक्टीरिया से जटिल जीवों के लिए सभी जीवों के लिए सार्वभौमिक है।

5. कोड गैर-अतिव्यापी है

दो समीपवर्ती कोडन एक आधार साझा नहीं कर सकते हैं, अर्थात एक कोडन के तीन आधार आसन्न कोडन के आधार से अलग हैं

6. कोड अल्पविराम है

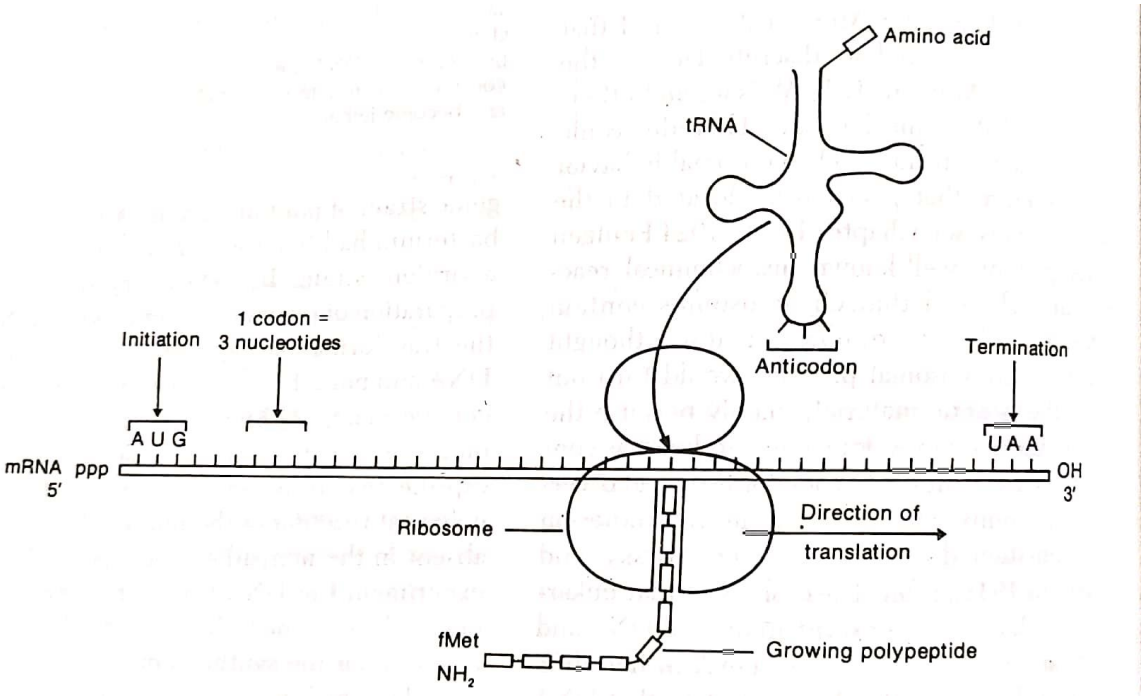
दो समीपवर्ती कोडन के बीच किसी विराम चिह्न की आवश्यकता नहीं होती है। एक अमीनो एसिड कोडित होने के बाद, दूसरा अमीनो एसिड स्वचालित रूप से अगले तीन अक्षरों द्वारा कोडित हो जाएगा और कोई भी अक्षर व्यर्थ नहीं जाता है। कोई फासला नहीं है।

7। आनुवंशिक कोड में "स्टार्ट" और "स्टॉप" सिग्नल होते हैं

केवल एक स्टार्ट कोडन (AUG, दीक्षा कोडन) है जो अनुवाद प्रक्रिया शुरू करता है, लेकिन इस प्रक्रिया को रोकने के लिए तीन स्टॉप कोडन मौजूद हैं , यानी UAA, UGA और UAG।

8. ध्रुवीयता/बहुरूपता

एम-आरएनए पर जेनेटिक कोड केवल 5' से 3' दिशा में पढ़ा जा सकता है।



मेथिओनिन से शुरू होने वाले एक अमीनो एसिड के लिए तीन न्यूक्लियोटाइड्स का समूह

(स्रोत: Cell and Molecular Biology, De Robertis)