Genetic code and its characteristics
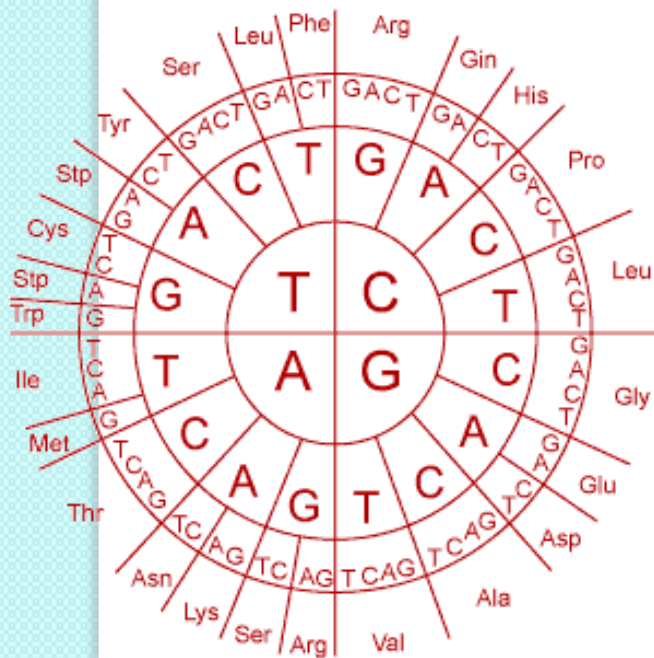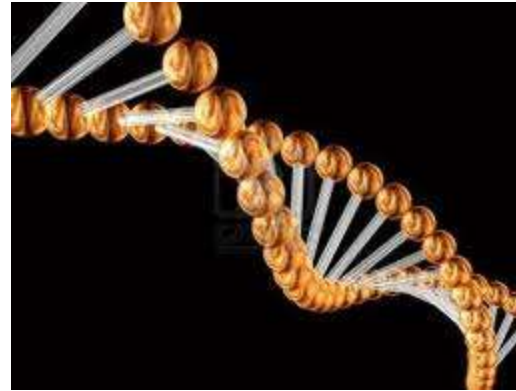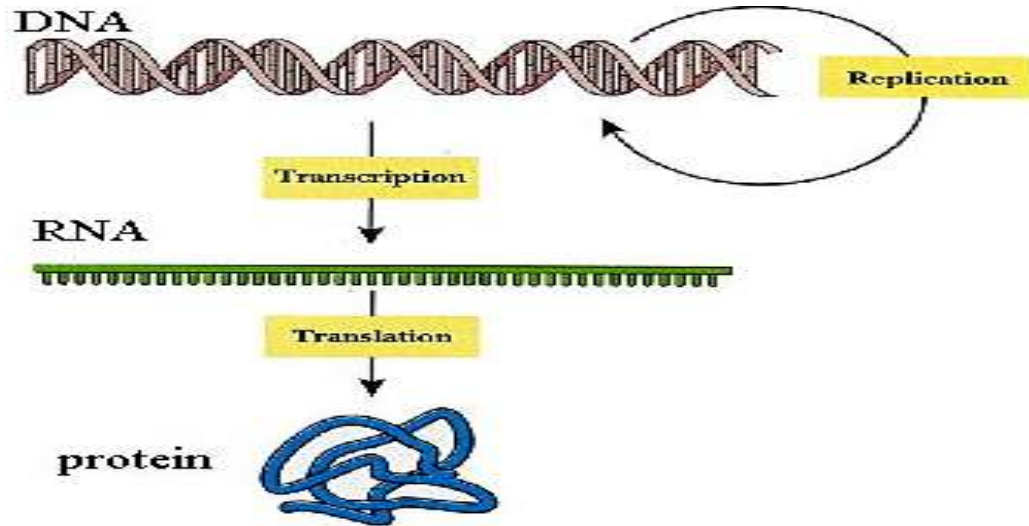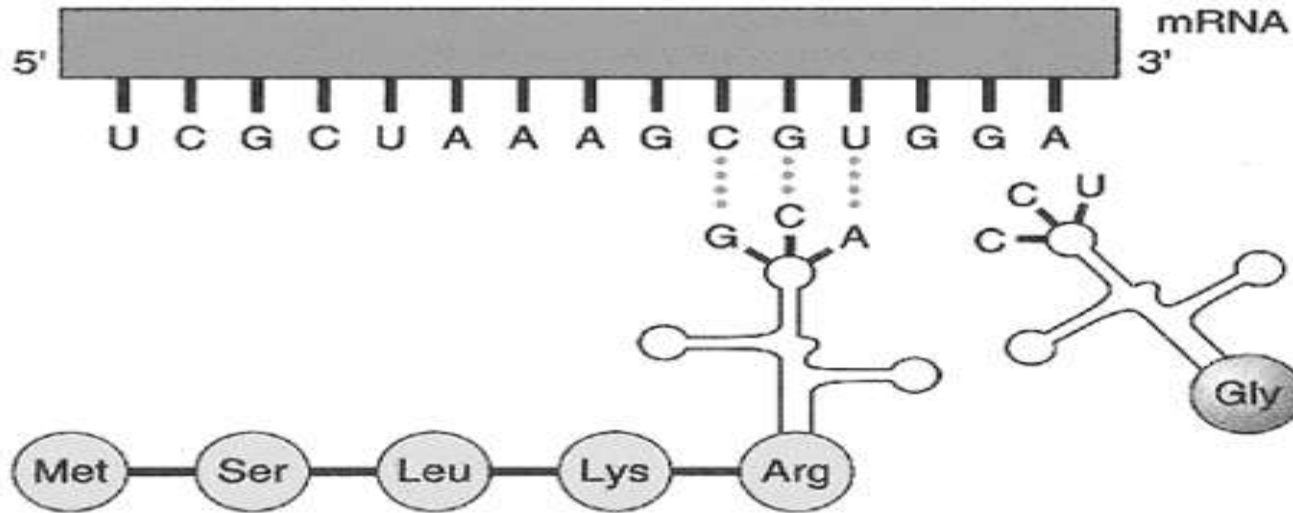

Biochemistry For Medics
www.namrata.co

Flow of genetic information

The central dogma of molecular biology describes the two-step process, transcription and translation, by which the information in genes flows into proteins: DNA → RNA → protein.

Introduction

The pathway of protein synthesis is called Translation because the language of nucleotide sequence on mRNA is translated in to the language of an amino acid sequence. The process of Translation requires a Genetic code, through which the information contained in nucleic acid sequence is expressed to produce a specific sequence of amino acids.

Introduction – Genetic Code

The letters A, G, T and C correspond to the nucleotides found in DNA. They are organized into codons.
The collection of codons is called Genetic code.

For 20 amino acids there should be 20 codons.
Each codon should have 3 nucleotides to impart specificity to each of the amino acid for a specific codon
1 Nucleotide- 4 combinations
2 Nucleotides 16 combinations
3 Nucleotides- 64 combinations (Most suited for 20 amino acids)
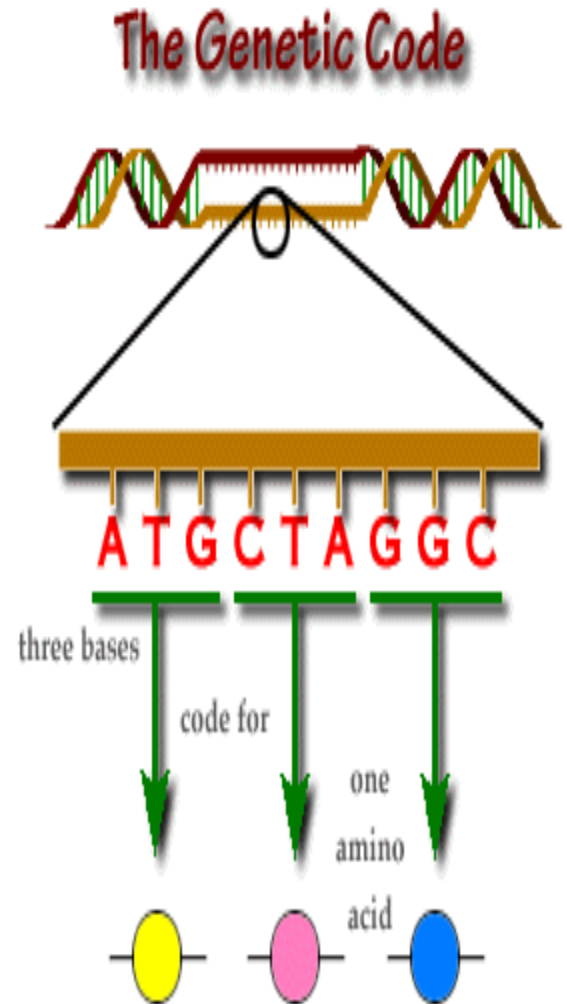
Genetic Code

- Genetic code is a dictionary that corresponds with sequence of nucleotides and sequence of Amino Acids.

- Words in dictionary are in the form of codons

- Each codon is a triplet of nucleotides

- 64 codons in total and three out of these are Non Sense codons.

- 61 codons for 20 amino acids

Genetic Code- Table

Second Letter

1st letter		U		C		A		G		3rd letter
U	UUU UUC	**Phe**	UCU UCC	**Ser**	UAU UAC	**Tyr**	UGU UGC	**Cys**	U C	
	UUA UUG	**Leu**	UCA UCG		UAA UAG	**Stop** **Stop**	UGA UGG	**Stop** Trp	A G	
C	CUU CUC CUA CUG	**Leu**	CCU CCC CCA CCG	**Pro**	CAU CAC	**His**	CGU CGC CGA CGG	**Arg**	U C A G	
					CAA CAG	**Gln**				
A	AUU AUC AUA	**Ile**	ACU ACC ACA ACG	**Thr**	AAU AAC	**Asn**	AGU AGC	**Ser**	U C A	
	AUG	**Met**			AAA AAG	**Lys**	AGA AGG	**Arg**	G	
G	GUU GUC GUA GUG	**Val**	GCU GCC GCA GCG	**Ala**	GAU GAC	**Asp**	GGU GGC GGA GGG	**Gly**	U C A G	
					GAA GAG	**Glu**				

Genetic Code-Characteristics

- **Specificity-** Genetic code is specific (Unambiguous)

- A specific codon always codes for the same amino acid.

e.g. UUU codes for Phenyl Alanine, it can not code for any other amino acid.

Genetic Code-Universal

Universal- In all living organism Genetic code is the same.

- The exception to universality is found in mitochondrial codons where AUA codes for methionine and UGA for tryptophan, instead of isoleucine and termination codon respectively of cytoplasmic protein synthesizing machinery.

- AGA and AGG code for Arginine in cytoplasm but in mitochondria they are termination codons.

Genetic Code-Redundant

Redundant- Genetic code is Redundant, also called Degenerate.
Although each codon corresponds to a single amino acid but a single amino acid can have multiple codons. Except Tryptophan and Methionine each amino acid has multiple codons.

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

Genetic Code- Non Overlapping and Non Punctuated

- All codons are independent sets of 3 bases.

- There is no overlapping ,

- Codon is read from a fixed starting point as a continuous sequence of bases, taken three at a time.

- The starting point is extremely important and this is called Reading frame.

Non Sense Codons

- There are 3 codons out of 64 in genetic code which do not encode for any Amino Acid.

- These are called **termination codons or stop codons** or nonsense codons. The stop codons are UAA, UAG, and UGA. They encode no amino acid. The ribosome pauses and falls off the mRNA.

Second letter

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

Initiator codon

- AUG is the initiator codon in majority of proteins-

- In a few cases GUG may be the initiator codon

- Methionine is the only amino acid specified by just one codon, AUG.

Wobbling phenomenon

- The rules of base pairing are relaxed at the third position, so that a base can pair with more than one complementary base.

- Some tRNA anticodons have **Inosine a**t the third position.

- Inosine can pair with U, C, or A. This means that we don't need 61 different tRNA molecules, only half as many are required.

Wobbling phenomenon

- First two bases in Codon in m RNA(5'-3') base pair traditionally with the 2 nd and 3rd base of the Anticodon in t RNA(5'-3')

- Non traditional base pairing is observed between the third base of the codon and 1st base of anticodon.

- The reduced specificity between the third base of the codon and the complementary nucleotide in anticodon is responsible for wobbling.

Wobbling phenomenon

| t RNA (first base)	m RNA (Third base)	Base pairing
C	G	Traditional
A	U	Traditional
U	A	Traditional
U	G	Nontraditional
G	C	Traditional
G	U	Nontraditional
I	U	Nontraditional
I	C	Nontraditional
I	A	Nontraditional

Wobbling phenomenon

- Proline has 4 codons(5'-3')
- CCU
- CCC
- CCA
- CCG
- The first three codons can be recognized by a single t RNA having Inosine at the first place.(IGG- 5'-3')

Clinical Significance

- Mutations can be well explained using the genetic code.

  A) Point Mutations

  1) Silent

  2) Misense

  3)Nonsense

  B) Frame shift mutations
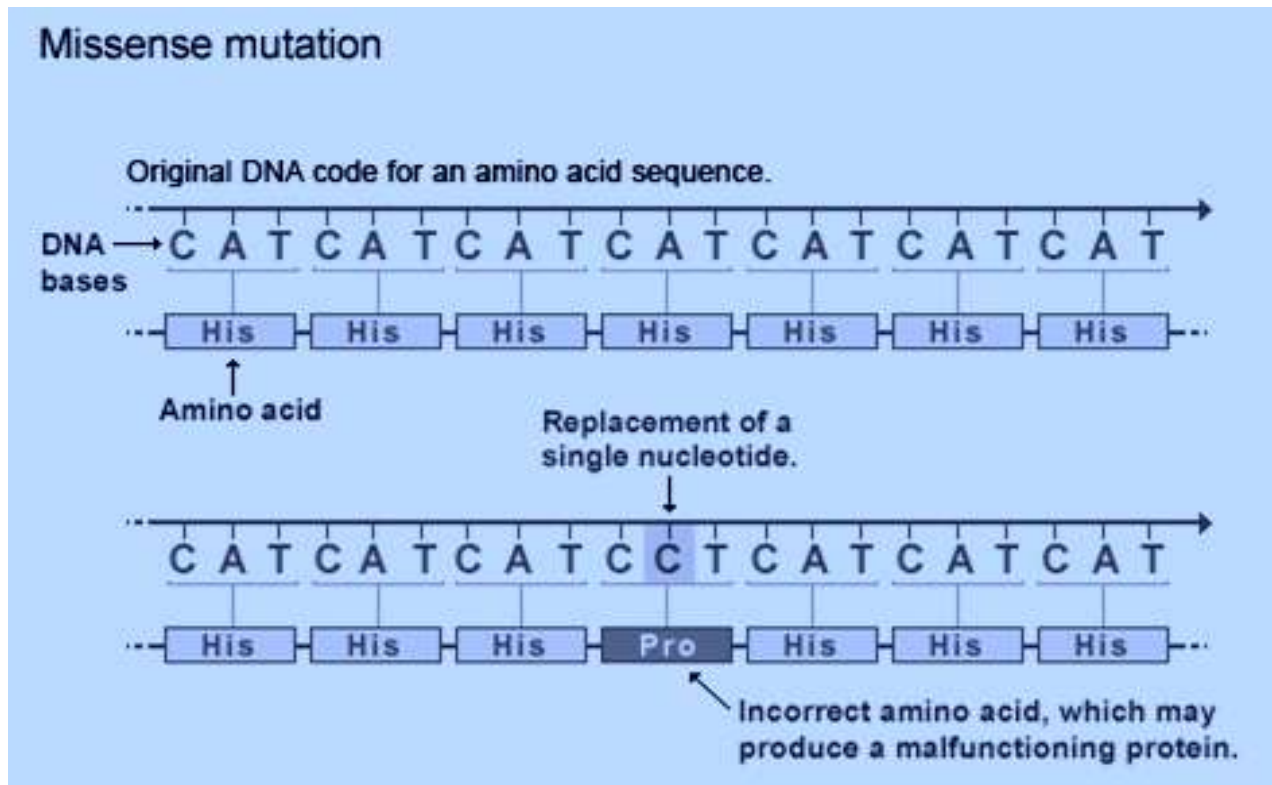
Silent Mutations

Single nucleotide change-A to G, same amino acid is incorporated. Mutation goes unnoticed.
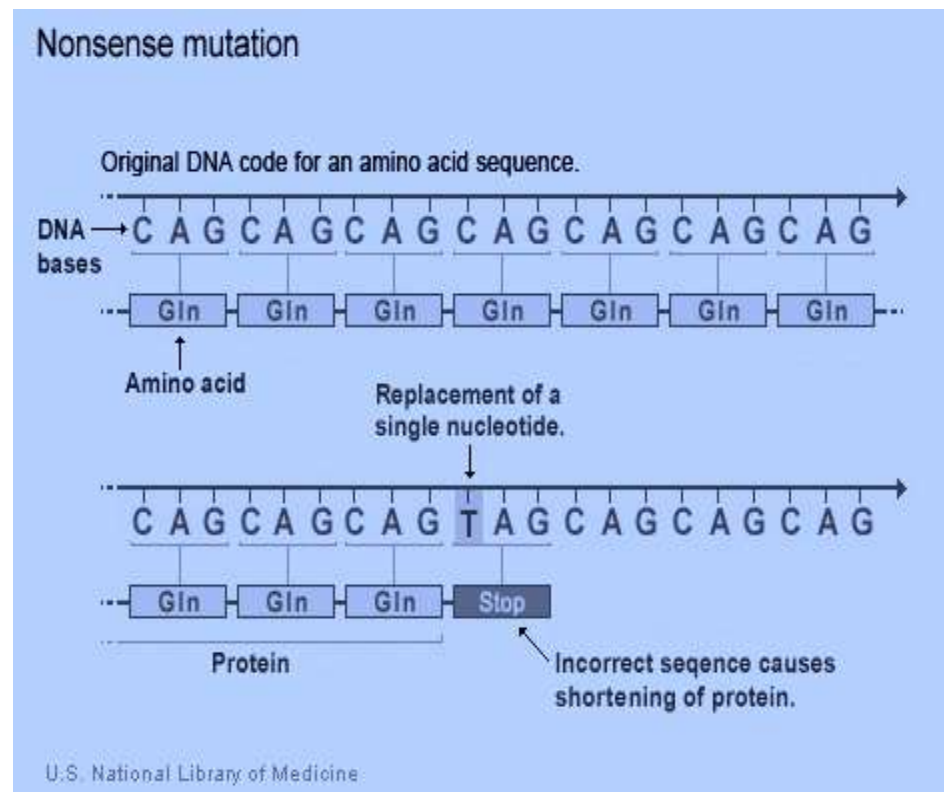
Silent Mutations

ATG	GAA	GCA	CGT
Met	Glu	Ala	Gly

↓

ATG	GAG	GCA	CGT
Met	Glu	Ala	Gly
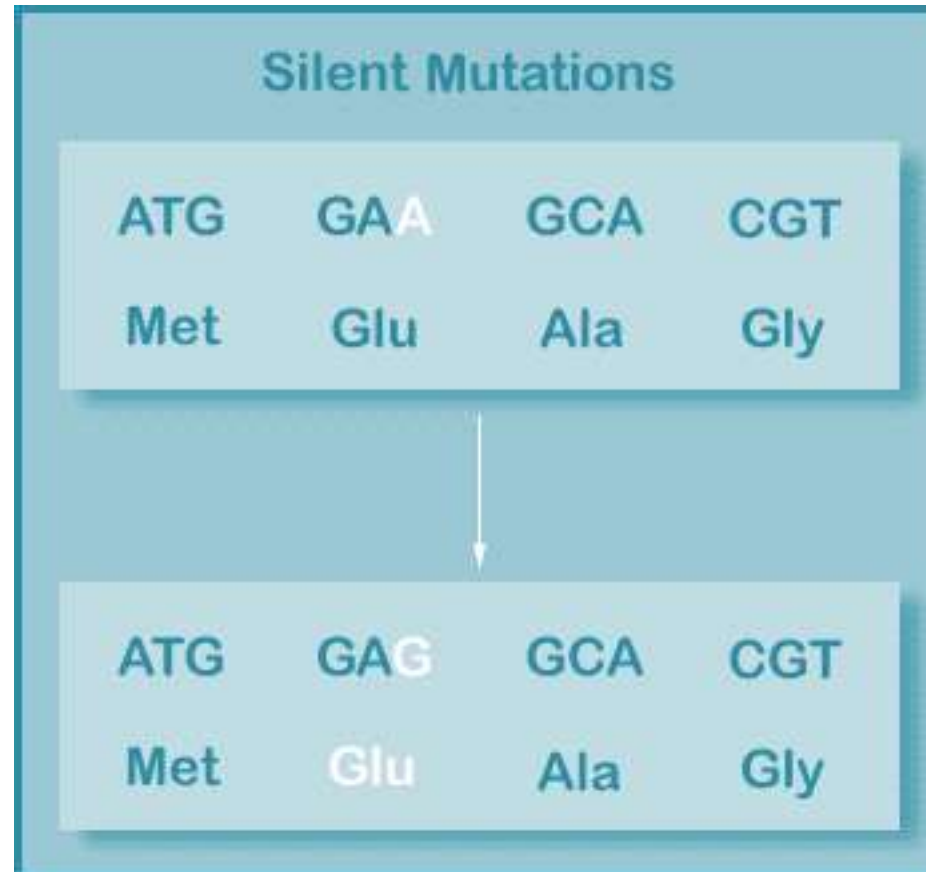
Missense mutations

Single nucleotide change A to C – different amino acid incorporated. Loss of functional capacity of protein.

Non sense Mutation

Single nucleotide change from C to T, stop codon is generated (In m RNA represented by UAG), premature termination of chain, may be incompatible with life.

Nonsense mutation

Original DNA code for an amino acid sequence.

DNA bases → C A G C A G C A G C A G C A G C A G C A G

Gln Gln Gln Gln Gln Gln Gln

↑ Amino acid

Replacement of a single nucleotide.
↓

C A G C A G C A G T A G C A G C A G C A G

Gln Gln Gln Stop

Protein

Incorrect seqence causes shortening of protein.

U.S. National Library of Medicine

Frame sift Mutations

- Insertion or removal of a base/bases can alter the reading frame with the resultant incorporation of different amino acids
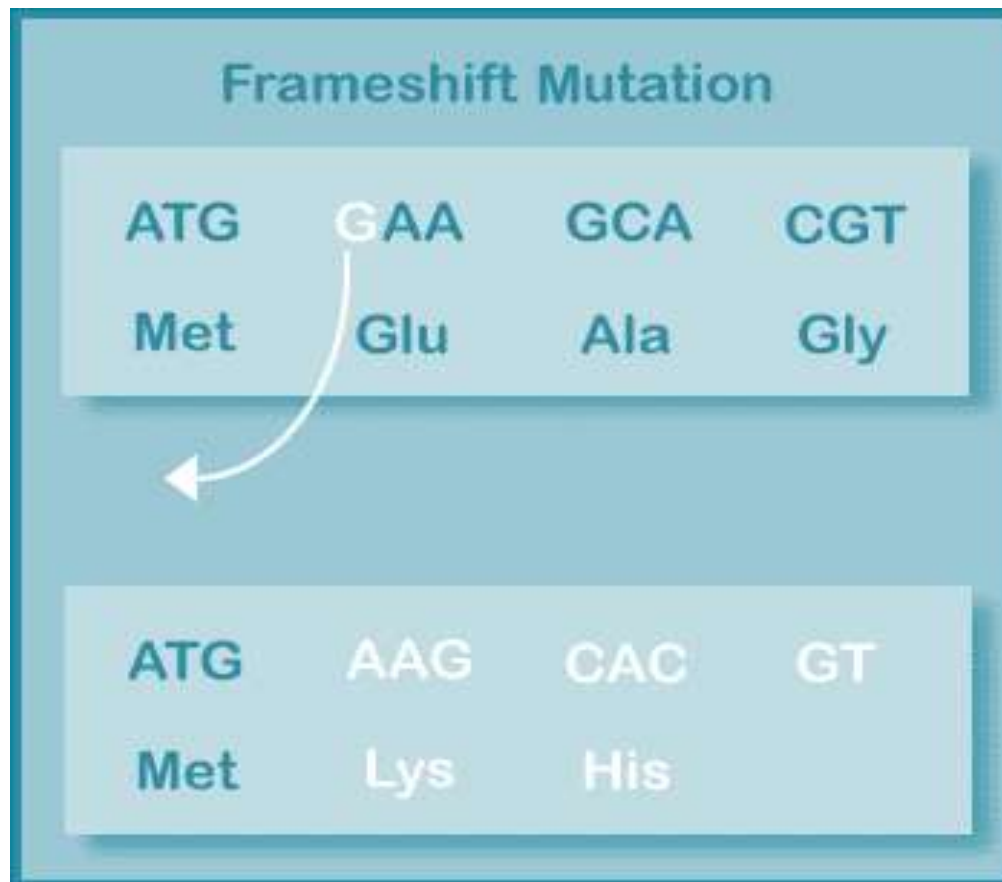
Genetic code- Summary

- Genetic code is unambiguous, universal, degenerate, commaless and non overlapping.

- UAG, UGA and UAA are termination codons .

- AUG is the start codon

- Mutations can be well described using genetic code