SOS in Biochemistry, Jiwaji University, Gwalior
M.Sc. II Semester (2019-20)
Paper BCH 201: Fundamentals of Molecular Biology (Unit IV)

The Genetic Code

The genetic code is a nonoverlapping code, with each amino acid plus polypeptide initiation and termination specified by RNA codons composed of three nucleotides.

Nobel Prize in Physiology or Medicine for 1968

Robert W. Holley
Born: 28 January 1922, Urbana, IL, USA
Died: 11 February 1993, Los Gatos, CA, USA
Field: Genetics, Molecular Biology
Prize share: 1/3

Har Gobind Khorana
Born: 9 January 1922, Raipur, India
Died: 9 November 2011, Concord, MA, USA
Field: Genetics, Molecular Biology
Prize share: 1/3

Marshall W. Nirenberg
Born: 10 April 1927, New York, NY, USA
Died: 15 January 2010, New York, NY, USA
Field: Genetics, Molecular Biology
Prize share: 1/3

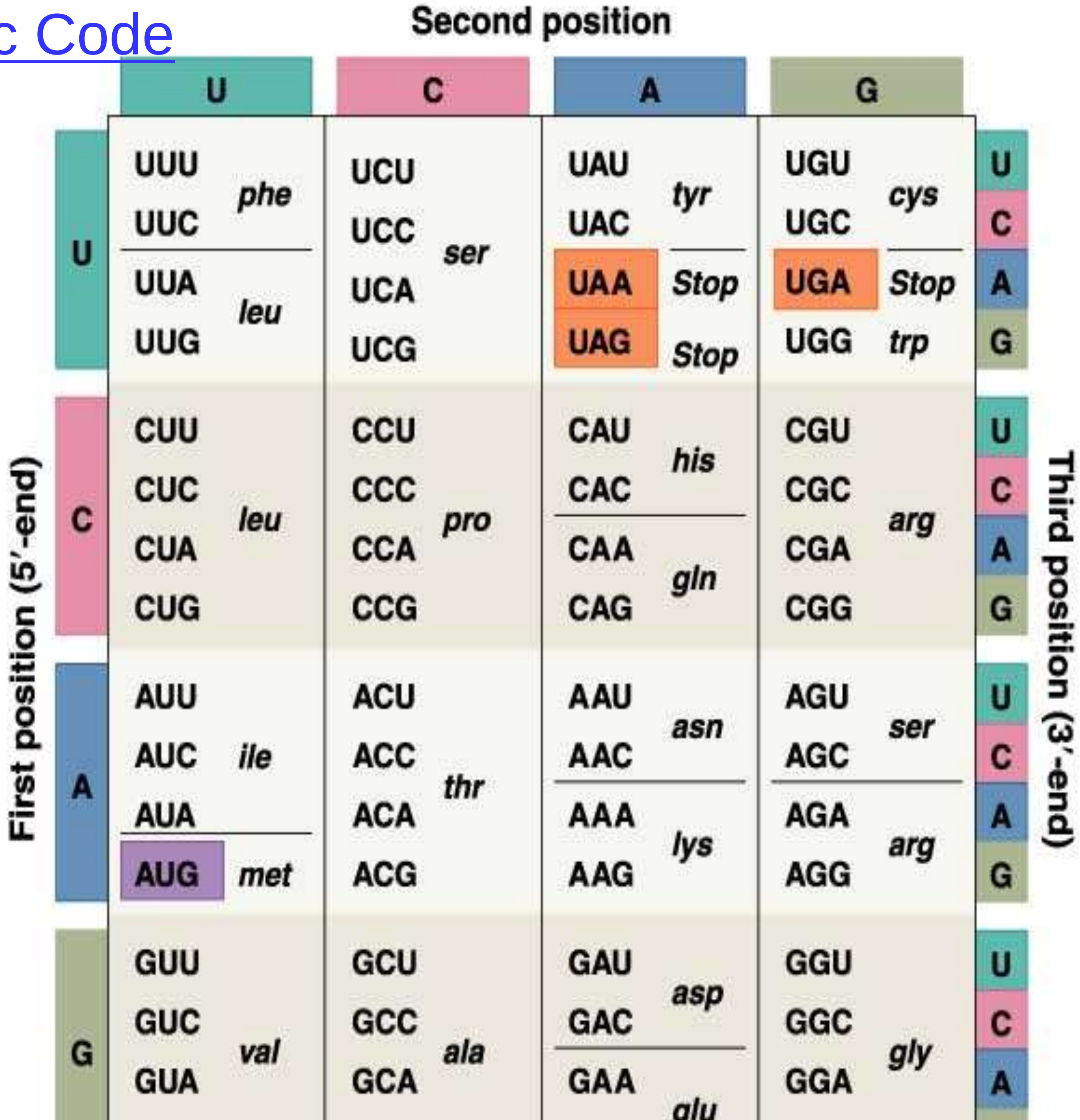
Work: “For their interpretation of the Genetic Code and its function in Protein Synthesis"

THE GENETIC CODE

- Most genetic code tables designate the codons for amino acids as mRNA sequences. Important features of the genetic code include:
- Each codon consists of three bases (triplet). There are 64 codons. They are all written in the 5' to 3' direction.
- 61 codons code for amino acids. The other three (UAA, UGA, UAG) are stop codons (or nonsense codons) that terminate translation.
- There is one start codon (initiation codon), AUG, coding for methionine. Protein synthesis begins with methionine (Met) in eukaryotes, and formylmethionine (fmet) in prokaryotes.
- The code is unambiguous. Each codon specifies no more than one amino acid.

- The code is degenerate. More than one codon can specify a single amino acid.
- All amino acids, except Met and tryptophan (Trp), have more than one codon.
- For those amino acids having more than one codon, the first two bases in the codon are usually the same. The base in the third position often varies.
- The code is almost universal (the same in all organisms). Some minor exceptions to this occur in mitochondria and some organisms.
- The code is commaless (contiguous). There are no spacers or "commas" between codons on an mRNA.
- Neighboring codons on a message are non-overlapping.

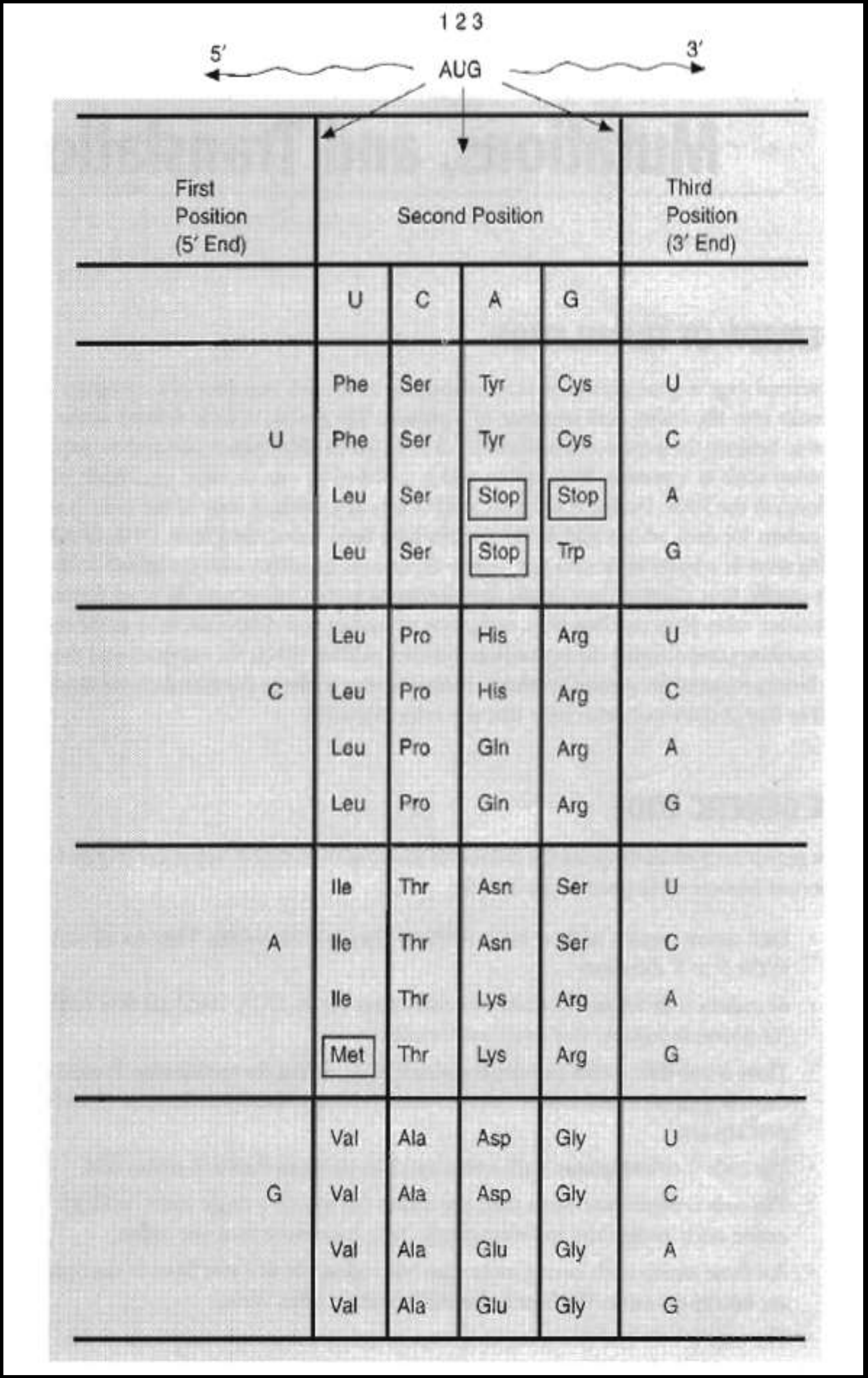
Genetic Code



Second position

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

The genetic code



1 2 3

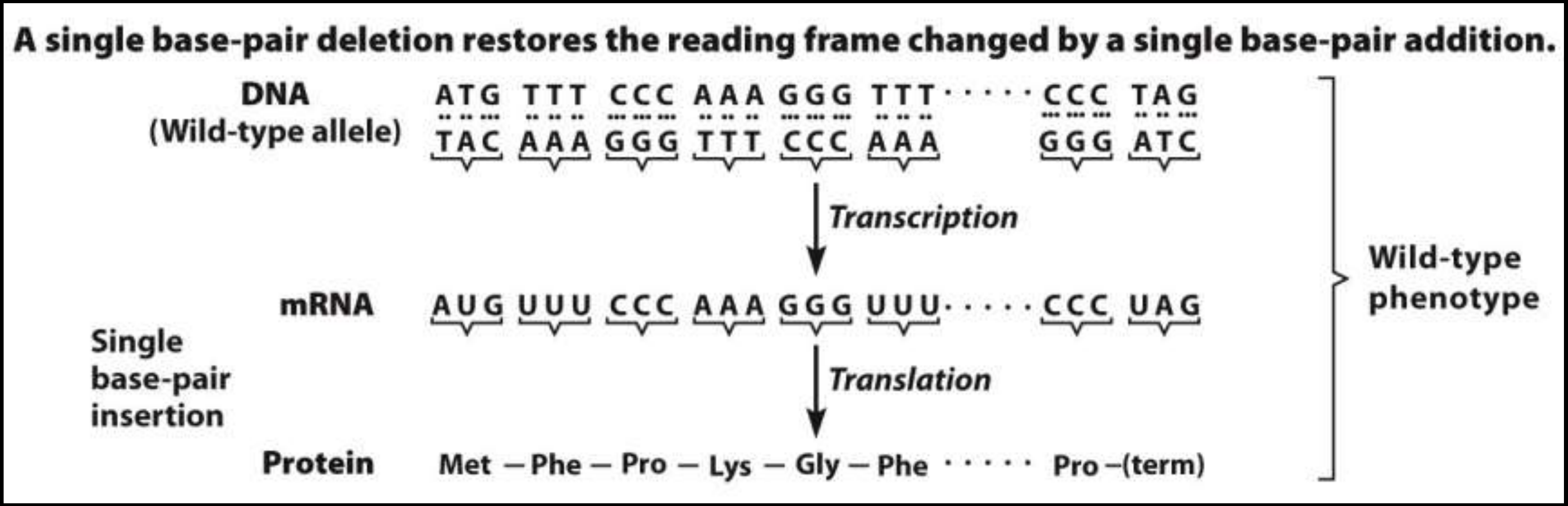
5' AUG 3'

First Position (5′ End)	Second Position				Third Position (3′ End)
	U	C	A	G	
U	Phe	Ser	Tyr	Cys	U
	Phe	Ser	Tyr	Cys	C
	Leu	Ser	Stop	Stop	A
	Leu	Ser	Stop	Trp	G
C	Leu	Pro	His	Arg	U
	Leu	Pro	His	Arg	C
	Leu	Pro	Gln	Arg	A
	Leu	Pro	Gln	Arg	G
A	Ile	Thr	Asn	Ser	U
	Ile	Thr	Asn	Ser	C
	Ile	Thr	Lys	Arg	A
	Met	Thr	Lys	Arg	G
G	Val	Ala	Asp	Gly	U
	Val	Ala	Asp	Gly	C
	Val	Ala	Glu	Gly	A
	Val	Ala	Glu	Gly	G

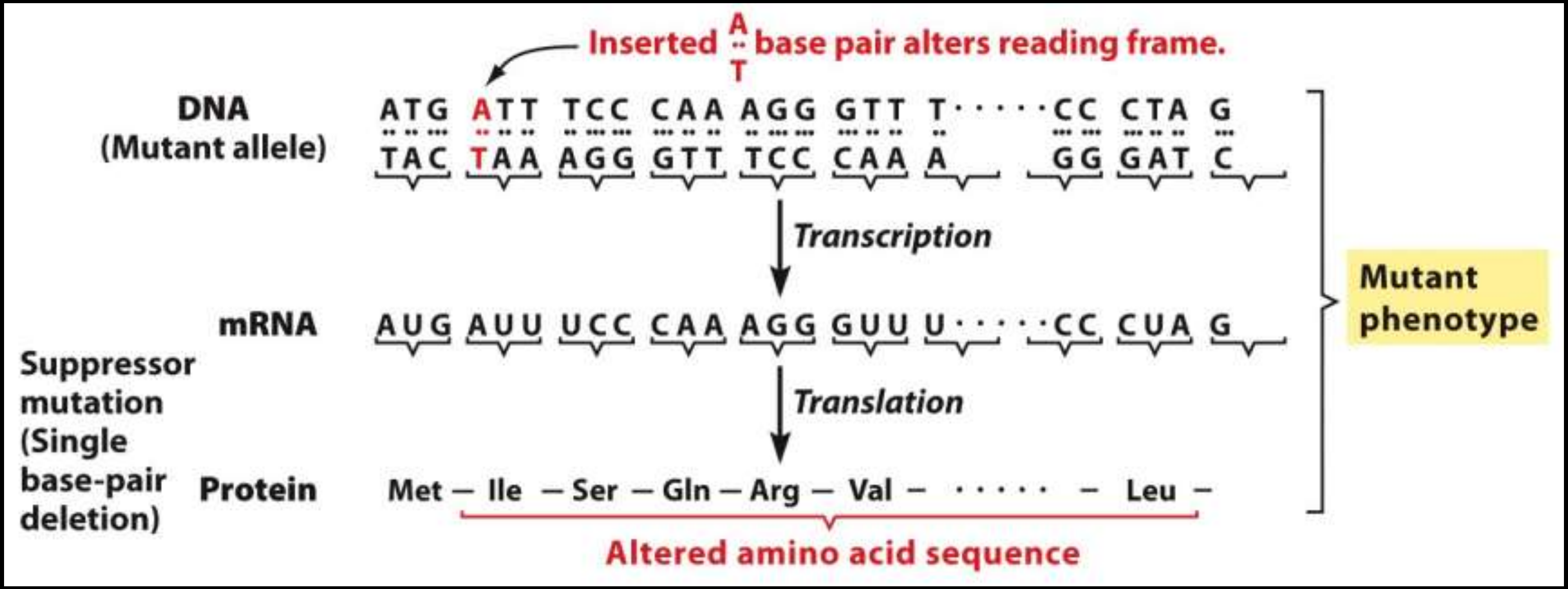
Properties of the Genetic Code

- The genetic code is composed of nucleotide triplets.
- The genetic code is nonoverlapping.
- The genetic code is comma-free.
- The genetic code is degenerate.
- The genetic code is ordered. (5’ to 3’)
- The genetic code contains start and stop codons.
- The genetic code is nearly universal.

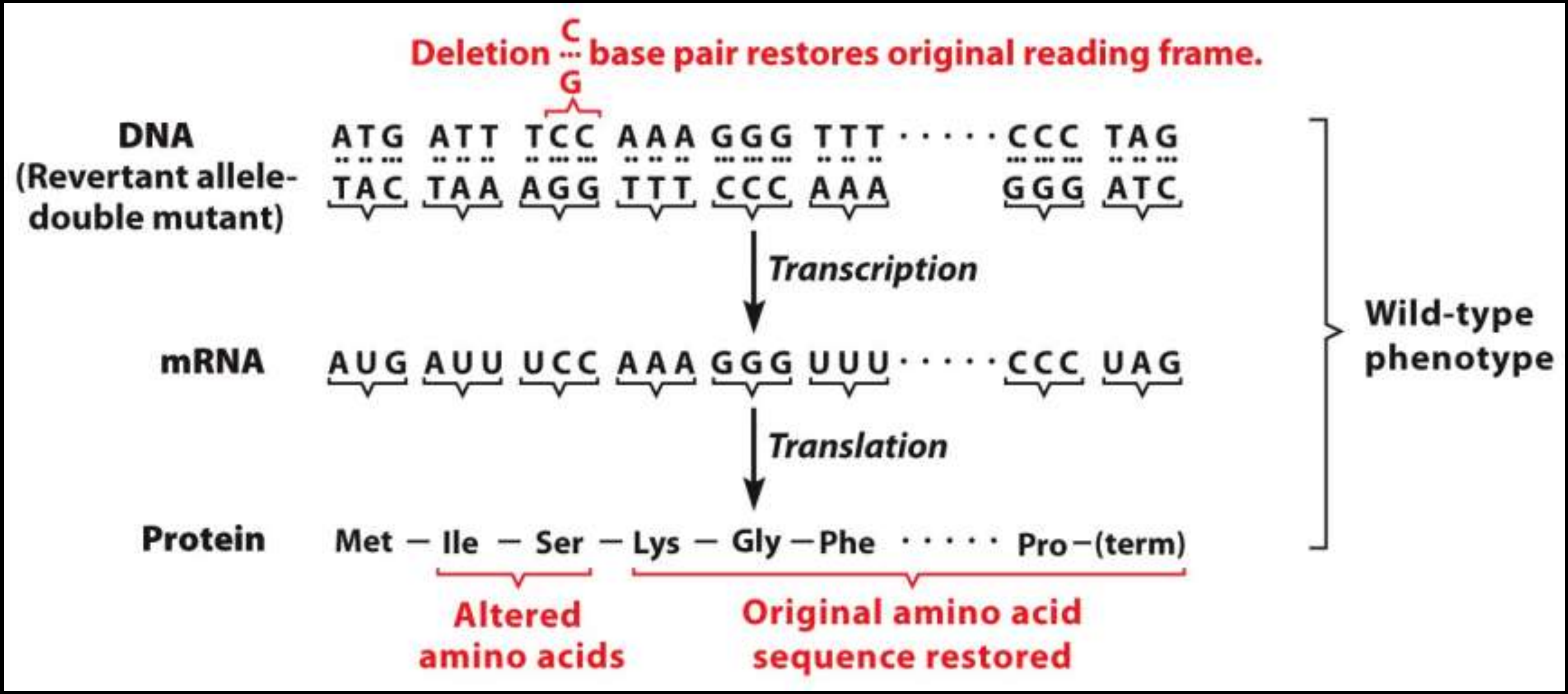
A Triplet Code*


A single base-pair deletion restores the reading frame changed by a single base-pair addition.

A Single-Base Pair Insertion Alters the Reading Frame*

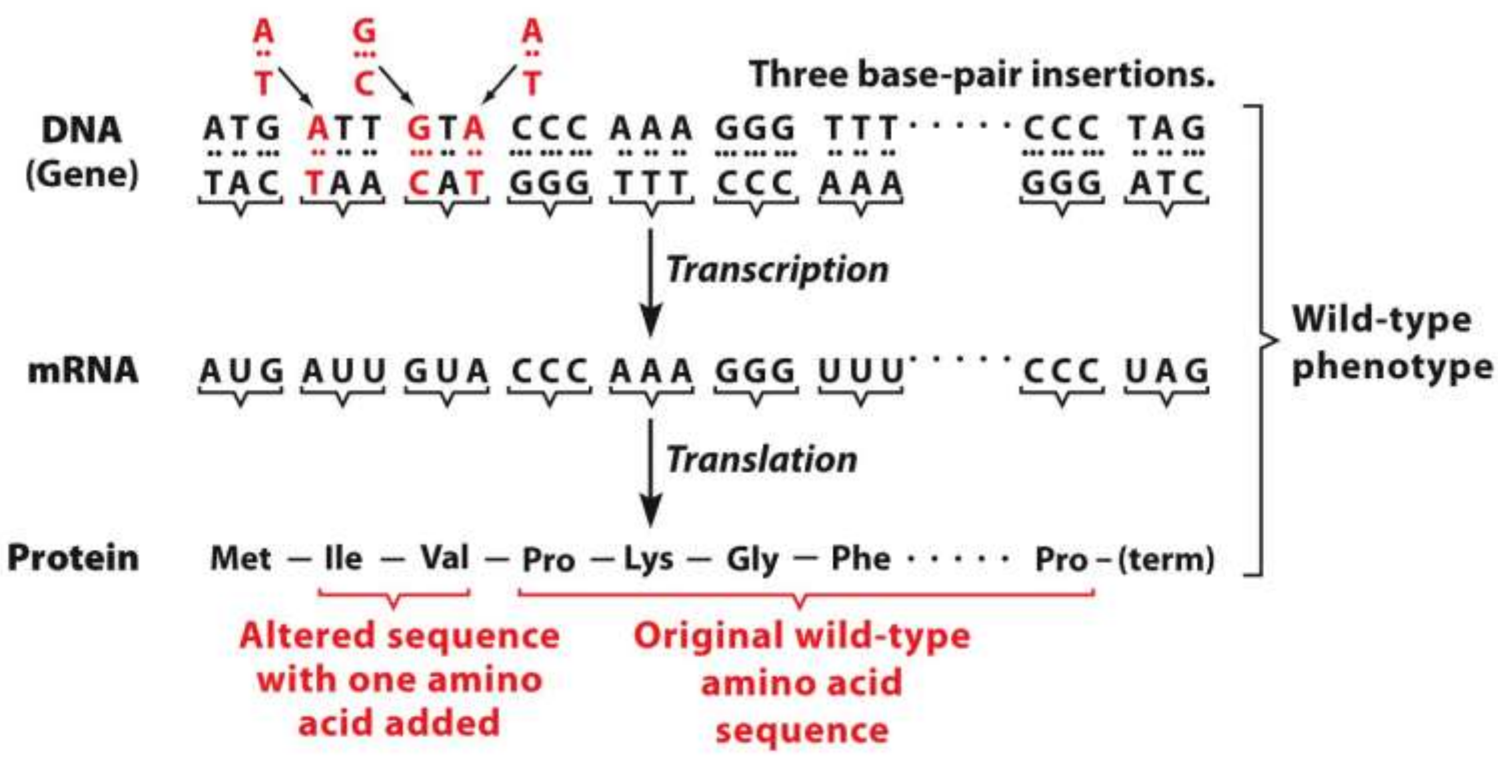


A suppressor mutation restores the original reading frame.*



Insertion of 3 base pairs does not change the reading frame.*

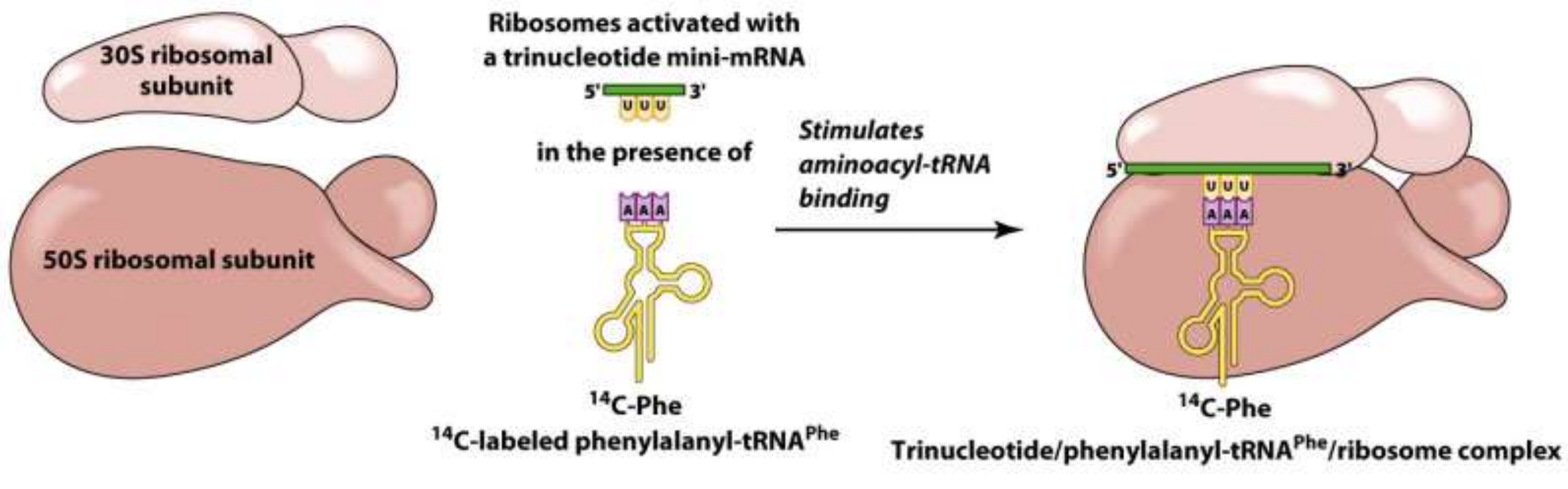
Recombinant containing three single base-pair additions has the wild-type reading frame.



Evidence of a Triplet Code: *In Vitro* Translation Studies

- Trinucleotides were sufficient to stimulate specific binding of aminoacyl-tRNAs to ribosomes.
- Chemically synthesized mRNAs containing repeated dinucleotide sequences directed the synthesis of copolymers with alternating amino acid sequences.
- mRNAs with repeating trinucleotide sequences directed the synthesis of a mixture of three homopolymers.

Deciphering the Genetic Code



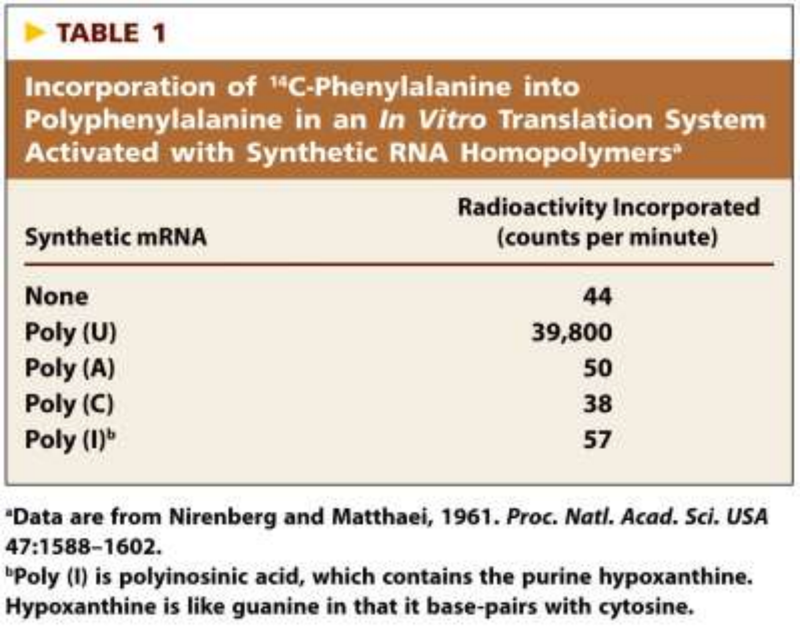


TABLE 1

Incorporation of ^{14}C-Phenylalanine into Polyphenylalanine in an *In Vitro* Translation System Activated with Synthetic RNA Homopolymers[a]

Synthetic mRNA	Radioactivity Incorporated (counts per minute)
None	44
Poly (U)	39,800
Poly (A)	50
Poly (C)	38
Poly (I)[b]	57

[a] Data are from Nirenberg and Matthaei, 1961. *Proc. Natl. Acad. Sci. USA* 47:1588–1602.

[b] Poly (I) is polyinosinic acid, which contains the purine hypoxanthine. Hypoxanthine is like guanine in that it base-pairs with cytosine.

You must know single letter codes and some triplets!

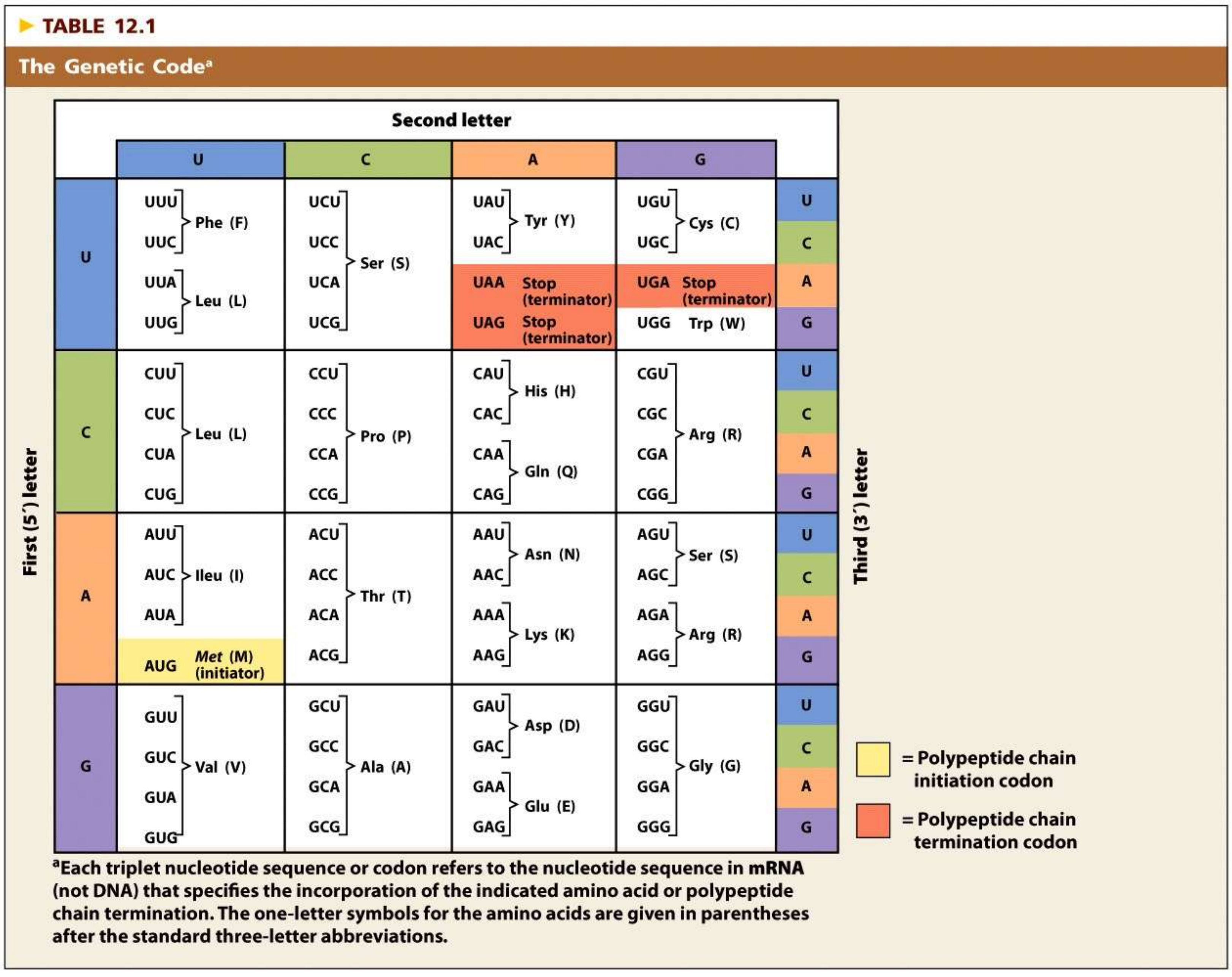


TABLE 12.1

The Genetic Code[a]

	Second letter				
First (5′) letter	**U**	**C**	**A**	**G**	**Third (3′) letter**
U	UUU Phe (F)	UCU Ser (S)	UAU Tyr (Y)	UGU Cys (C)	**U**
	UUC	UCC	UAC	UGC	**C**
	UUA Leu (L)	UCA	UAA Stop (terminator)	UGA Stop (terminator)	**A**
	UUG	UCG	UAG Stop (terminator)	UGG Trp (W)	**G**
C	CUU Leu (L)	CCU Pro (P)	CAU His (H)	CGU Arg (R)	**U**
	CUC	CCC	CAC	CGC	**C**
	CUA	CCA	CAA Gln (Q)	CGA	**A**
	CUG	CCG	CAG	CGG	**G**
A	AUU Ileu (I)	ACU Thr (T)	AAU Asn (N)	AGU Ser (S)	**U**
	AUC	ACC	AAC	AGC	**C**
	AUA	ACA	AAA Lys (K)	AGA Arg (R)	**A**
	AUG *Met* (M) (initiator)	ACG	AAG	AGG	**G**
G	GUU Val (V)	GCU Ala (A)	GAU Asp (D)	GGU Gly (G)	**U**
	GUC	GCC	GAC	GGC	**C**
	GUA	GCA	GAA Glu (E)	GGA	**A**
	GUG	GCG	GAG	GGG	**G**

🟨 = Polypeptide chain initiation codon

🟥 = Polypeptide chain termination codon

[a] Each triplet nucleotide sequence or codon refers to the nucleotide sequence in mRNA (not DNA) that specifies the incorporation of the indicated amino acid or polypeptide chain termination. The one-letter symbols for the amino acids are given in parentheses after the standard three-letter abbreviations.

The Genetic Code

- Initiation and termination Codons
  - Initiation codon: AUG
  - Termination codons: UAA, UAG, UGA
- Degeneracy: partial and complete
- Ordered
- Nearly Universal (exceptions: mitochondria and some protozoa)

Key Points

- Each of the 20 amino acids in proteins is specified by one or more nucleotide triplets in mRNA. (20 amino acids refers to what is attached to the tRNAs!)

- Of the 64 possible triplets, given the four bases in mRNA, 61 specify amino acids and 3 signal chain termination. (have no tRNAs!)

Key Points

- The code is nonoverlapping, with each nucleotide part of a single codon, degenerate, with most amino acids specified by two to four codons, and ordered, with similar amino acids specified by related codons.

- The genetic code is nearly universal; with minor exceptions, the 64 triplets have the same meaning in all organisms. (this is funny)

Do all cells/animals make the same Repertoire of tRNAs?

The Wobble Hypothesis: Base-Pairing Involving the Third Base of the Codon is Less Stringent.

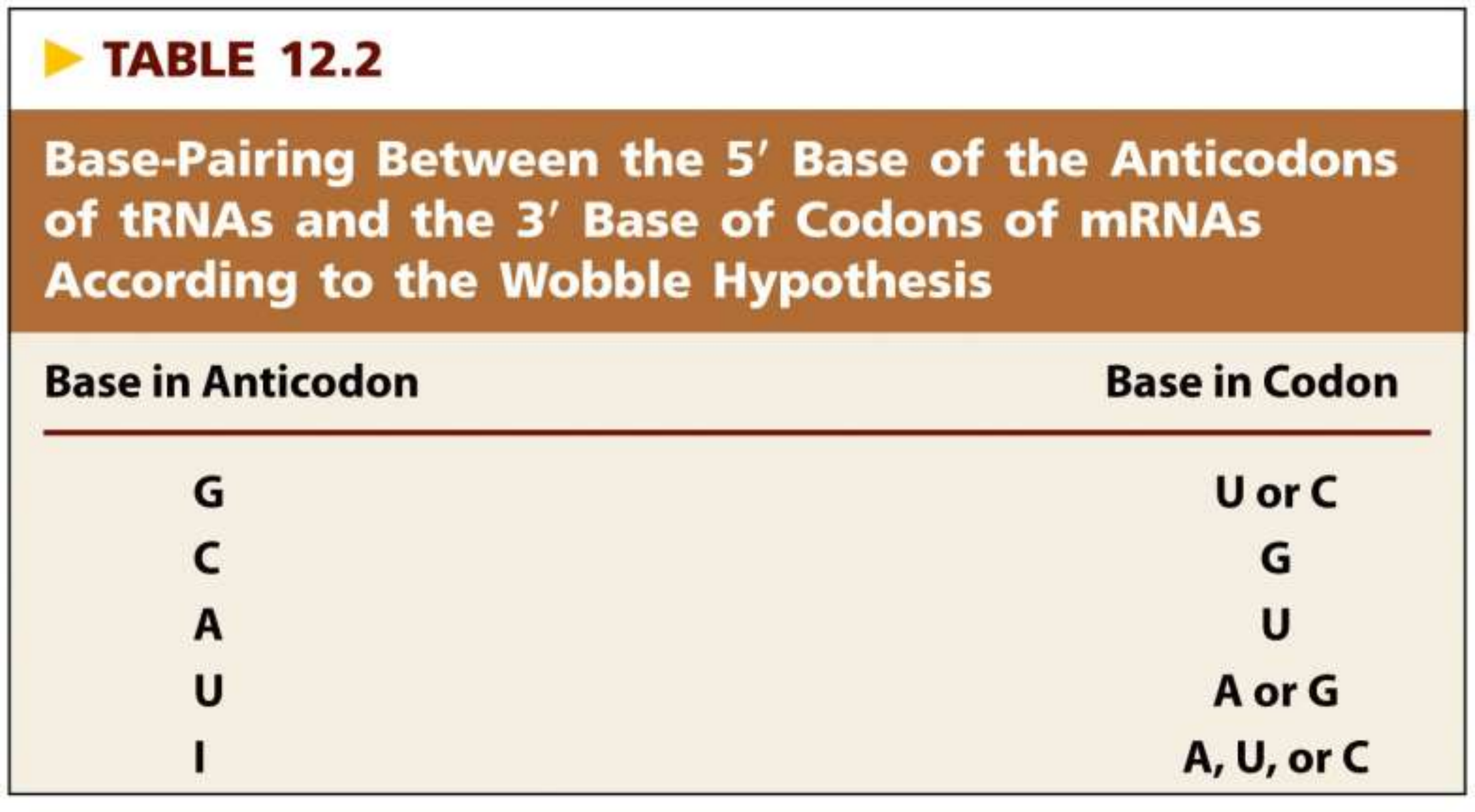


TABLE 12.2

Base-Pairing Between the 5′ Base of the Anticodons of tRNAs and the 3′ Base of Codons of mRNAs According to the Wobble Hypothesis

Base in Anticodon	Base in Codon
G	U or C
C	G
A	U
U	A or G
I	A, U, or C

Base-Pairing with Inosine at the Wobble Position