Symbiosis

www.symbiosisonline.org
www.symbiosisonlinepublishing.com
ISSN Online: 2377-4274

Research Article | International Journal of Genetic Science | Open Access

The 24-Quadruplet Genetic Code: Merits Of The Projected Quadruplet Codons, From Combinatorial Perspective

W.B. Bozegha*

**Numeration science literature, Development research project, No. 9 Otiotio road, Yenagoa, Bayelsa state, Nigeria.*

Received: March 05, 2018; Accepted: September 26, 2018; Published: September 29, 2018

***Corresponding author:** W.B. Bozegha, Numeration science literature, Development research project, No. 9 Otiotio road, Yenagoa, Bayelsa state, Nigeria. Tel: +2348068029860; E-mail: wbbozegha@yahoo.com*

Abstract

Statement of the Problem: Triplet codons have held sway since 1954, when the 64-triplet genetic code emerged as the mathematical solution to the molecular biologists' quest for producing at least 20 code words from 4 nucleotide bases to account for their observation in 1953 that the sequence of the four bases in the nucleus of a cell influenced the sequence of the twenty amino acids of a protein in the surrounding cellular cytoplasm. The 64 triplet codons then formed the genetic code, derived from a sequence of the four bases. The genetic code is ridden with several irregularities which are only widely discussed in genetics literature without remedies, apart from bogus[a1] theories[a2], like the wobble phenomenon, frozen accident etc for their explanation.

Methodology and theoretical orientation: Square kinematics scheme, a new technique for computing permutations of 4 from 4 by view mixing is presented in consideration of the fact that the problem of generating at least 20 code words from 4 bases is combinatorial, bordering on permutations of 4 from 4 i.e ${}_4P_4$ = 4!=4x3x2x1=24 quadruplets.

Findings: 24-quadruplet codons, representing a 24-quadruplet genetic code sequence (structure) produced[a3]..

Conclusion and significance: A 24-quadruplet genetic code free from irregularities produced from a sequence of the four nucleotide bases in response to the molecular biologists' quest of 1953 to raise a code to typify the observed relationship between the four nucleotide bases and the twenty amino acids of protein. The 24 quadruplets of this new genetic code structure represent the 'workforce' in protein synthesis, where 20 codons take charge of the placement of 20 amino acids in a sequence corresponding to theirs at one codon per amino acid, with four spare codons for four place/time based start/stop control signals.

Recommendation: Experimental experts to take the challenge of spelling the new 24-quadruplet genetic code to render it fit for adoption [a4].

Keywords: Quadruplet Codons; Permutations; Four; Twenty; Twenty-Four.

Introduction

The scientific observation by molecular biologists in 1953[a5] that the sequence of the DNA four bases A, T, G, C, (Adenine, Thymine, Guanine, Cytosine) in the nucleus of the cell influenced the sequence of the twenty amino acids of protein in the surrounding cytoplasm of the cell is the crux of protein synthesis, bordering on **protein type proliferation and diversification.**

Following this discovery, molecular biologists sought for ways by which the ATGC four-base combination regarded as four-letter alphabet could be made to generate enough code words to attend to the 20 amino acids of protein individually in protein synthesis and studies. This is a functional relationship also called the genetic code. Thereupon, they set about a quantitative reasoning that gave them 64 triplets which they also produced by the indirect base-four neo-digibreed[a6] method using Punnett Square[a7]. Unfortunately, they went astray in their interpretation of the following combinatorial terms of selections for permutation synthesis involved therein.

(i) The selection 1 from 4 as 4^1 = 4 singlets, instead of

$$ {}_4P_1 = \frac{4!}{(4-1)!} = \frac{4!}{3!} = \frac{4x3x2x1}{3x2x1} = 4 \quad \text{singlets.} $$

(ii)The selection 2 from 4 as 4^2 = 4x4=16 duplexes, instead of

$$ {}_4P_2 = \frac{4!}{(4-2)!} = \frac{4!}{2!} = \frac{4x3x2x1}{2x1} = 12 duplexes $$

(iii)The selection 3 from 4 as 4^3 = 4x4x4 = 64 triplets (adopted), instead of $ {}_4P_3 = \frac{4!}{(4-3)!} = \frac{4!}{(4-3)!} = \frac{4!}{1!} = \frac{4x3x2x1}{1} = 24$ triplets.

(iv)The selection 4 from 4 as 4^4 = 4x4x4x4 = 256 quadruplets (ignored (a8)), instead of ${}_4P_4 = 4! = 4x3x2x1 = 24$ quadruplets.

This mistake led to their acceptance of the 64triplets as the code words. The 64-triplet code structure (a mixture of 24 permutations and 40 non-permutations) accepted and adopted after 'spelling' in 1968(a9) has been a thorn in the flesh of molecular biology (a10)studies of protein synthesis to date,

and scientists are now researching to see, if the bases could be increased to more than 3 per codon[a11]. (Internet on Francis Crick).http/en.wikipedia.org/wiki/Francis-Crick.

Materials and Methods

Materials

The materials consist of the RNA four bases in the sequence of A,U,G,C (Adenine, Uracil, Guanine, Cytosine) and the DNA four bases in the order of A, T, G, C (Adenine, Thymine, Guanine, Cytosine) as carried in a particular rung of the double helix and are used as input set of 4 bases in the multiplicative replication input/output system in computational combinatorics developed by this author in the 1990's[a12].

Method

Direct Method, Designated as Square Kinematics View Mixing Technique.

The input set of RNA four bases A,U,G,C .[a13]are loaded at the corners of the square in clockwise direction as depicted in Fig.1. The loaded square is deployed in three ways as depicted in Fig. 1 (a), (b), and (c) to generate 8 combinatorially [a14]valid quadruplets per deployment per section using kinematics and view mixing as shown in Chart 1 carrying a genetic code structure of 24 quadruplet codons from lines (1) to (24). [a15].

Square Kinematics Technique for generating permutations of 4 from 4: input[a16]set AUGC (Chart 1).

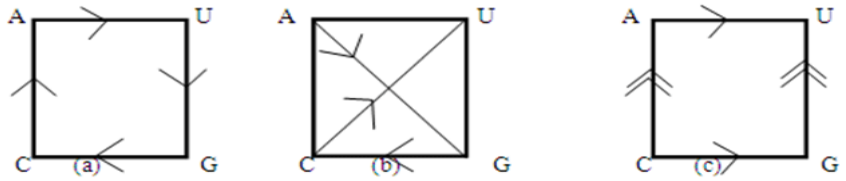
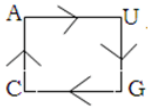
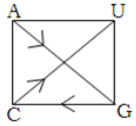
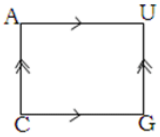


Figure 1: (a) Sides deployment
(b) Diagonals deployment
(c) Parallels deployment

Chart 1: Output of $_4P_4$Permutation

 Step 1 (a) Sides Deployment	Viewing along sides From A	Clockwise	AUGC	Line1
		Fro	CGUA	Line 2
	Viewing along sides From U	Clockwise	UGCA	Line 3
		Fro	ACGU	Line 4
	Viewing along sides From G	Clockwise	GCAU	Line 5
		Fro	UACG	Line6
	Viewing along sides From C	Clockwise	CAUG	Line 7
		Fro	GUAC	Line 8
 Step 2, (b) Diagonals Deployment	Viewing along diagonals From A	Clockwise	AGCU	Line 9
		Fro	UCGA	Line 10
	Viewing along diagonals From U	Clockwise	UCAG	Line 11
		Fro	GACU	Line 12
	Viewing along diagonals From G	Clockwise	GAUC	Line 13
		Fro	CUAG	Line 14
	Viewing along diagonals From C	Clockwise	CUGA	Line 15
		Fro	AGUC	Line 16
 Step 3, (c) Parallels Deployment	Viewing the parallels AU//CG	Horizontals left to right	AUCG	Line 17
		Fro	GCUA	Line 18
	Viewing the parallels UA//GC	Horizontals right to left	UAGC	Line 19
		Fro	CGAU	Line 20
	Viewing the parallels CA//GU	Verticals upward	CAGU	Line 21
		Fro	UGAC	Line 22
	Viewing the parallels AC//UG	Verticals downward	ACUG	Line 23
		Fro	GUCA	Line 24

Summary of valid products: lines 1-24 = 24 quadruplets
Factorial $_4P_4$= 4! = 4x3x2x1 = 24 quadruplets
Production = prediction = 24 quadruplets

Result

(a) A genetic code structure of 24 permutation quadruplets is presented in Table I as a computational reality for the result, being a combinatorial derivation. The list of 20 amino acids of protein[1] is adapted from the book, *The World of the Cell* by Becker, Wayne M. (1986), Fig. 17.4, p. 529.

(b) In the parlance of computer science both ***hardware*** (square kinematics scheme) ***and software (***view mixing using square kinematics scheme) for the computation of permutations of 4 from 4 associated with the production of the ***quadruplet codons*** are made available.[a17]

Table 1: New genetic code of 24-quadruplet codon structure

Input set of RNA four bases	Output permutations of quadruplet codons					Remarks
	Serial no. of 24 quadruplet codons	20 amino acids[a]/ 4 signals	Amino acids 20 /signals 4 to be specified by codons upon spelling by the experimental experts	Direct Square Kinematics, Ref. Chart 1 lines 1-24		Some salient points
		[1]		Genetic Code Sequence		
AUGC	1	Alanine	AYTBD	AUGC		
	2	Arginine	"	CGUA		(1) Output sequence of permutations per method is unique .
	3	Asparagine	"	UGCA		
	4	Aspartic acid	"	ACGU		
	5	Cysteine	"	UAGC		
	6	Glutamic acid	"	UACG		
	7	Glutamine	"	CAUG		(2) All 24 codons per sequence possess integrity, potency and uniqueness.
	8	Glycine	"	GUAC		
	9	Histidine	"	AGUC		
	10	Isoleucine	"	UCGA		
	11	Leucine	"	UCAG		(3) Collinearity between 24 codons and 20 amino acids/ 4 signals evident.
	12	Lysine	"	GACU		
	13	Methionine	"	GAUC		
	14	Phenylalanine	"	CAUG		
	15	Proline	"	CUGA		
	16	Serine	"	AGUC		
	17	Threonine	"	AUCG		(4) All codons are convertible to equivalents of DNA rungs in base content by the replacement of U by T showing that the genetic code is actually the RNA transcribed from the DNA as intimated.[2] (b)
	18	Tryptophan	"	GCUA		
	19	Tyrosine	"	UAGC		
	20	Valine	"	CGAU		
	21	Signal 1	"	ACUG		
	22	Signal 2	"	GUCA		
	23	Signal 3	"	CAGU		
	24	Signal 4	"	UGAC		
Total	24	24	24	24	24	

[a] List of 20 amino acids of protein.(a18) adapted from Fig. 17.4, The World of the Cell, p.529 by Becker, Wayne M. (1986).[a19]

[b] By Jill Wright et al (1988) in their book, Prentice Hall Life Science at page 63 with regard to protein synthesis, where it is stated that the RNA in the ribosomes, along with the RNA sent out from the nucleus directs the production of proteins.

Key

AYTBD = Allocation yet to be determined

Signal 1 = Place start signal

Signal 2 = Time start signal

Signal 3 = Place stop signal

Signal 4 = Time stop signal

Discussion

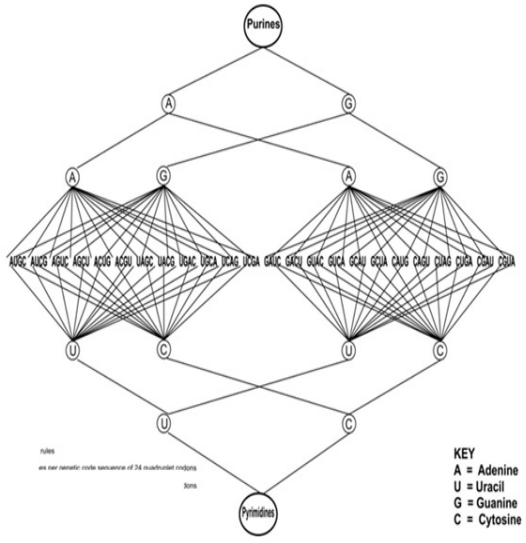
The discussion is geared to highlight the merits of the ***quadruplet codons*** of the new genetic code produced by combinatorics. In addition to the presentation of an irregularity-.[a20]free 24-quadruplet genetic code, a new technique for permutation of 4 from 4 using a square kinematics scheme based on view mixing is made available for an extra to the merits [a21] of the ***quadruplet codons.*** That is to say, in computer parlance both ***hardware*** (square kinematics scheme) ***and software*** (view mixing using square kinematics scheme) for the computation of permutations of 4 from 4 are associated with the production of the quadruplet codons whose merits are being discussed now.

This newly produced 24-quadruplet genetic code represents the natural (true) [a22] genetic code and serves its purpose in protein synthesis. It has 24-quadruplet codons as a workforce of 24 workers comprising 20 'labourers' and 4 "supervisors'. The 20 labourers are meant to be responsible for the placement of the 20 amino acids in a sequence at one labourer per amino acid turn by turn and the 4 supervisors to serve as four signals for start/stop control in respect of place and time during the building of the sequence of amino acids for a protein type at one supervisor per specific signal.

Whence two basic functions per quadruplet codon emerge to the effect that a quadruplet codon (i) can serve as incumbent prototype codon in the place of a seed for the reproduction of the whole genetic code plant of 24 quadruplet permutation codons [a23]; (ii) be responsible for the placement of a specific amino acid or be a specific signal in protein synthesis and studies. This portrayal is suited to the design and purpose of the natural genetic code that is efficiently engaged in protein synthesis throughout nature since creation.

Merits of the genetic code quadruplet codon intrinsic qualities thereof

(a) When a particular sequence of the RNA four bases e.g. AUGC is used as input set in the *input/output format of the combinatorial multiplicative replication system* for permutation of 4 from 4, an output of factorial complement of 24 quadruplets is obtained, given by ${}_4P_4$ =4! =4x3x2x1=24 quadruplets. These are permutations which are traditionally called **codons individually** and **genetic code collectively.** Each quadruplet codon is a *true copy* of the RNA input set because of bearing the complete set of base types in a unique sequence of its own, except the *incarnation* codon which has the same base sequence and therefore identical to the input set. This presence of all four base types per quadruplet codon confers a quality known as codon integrity upon all the quadruplets belonging to the genetic code. There is **no single case of absent base type to cause under utilization** of base types or repeated base type to bring about redundancy of base type in the quadruplet structure and texture of the new codon, unlike the old triplet codon where these flaws[a24] abound.



Illustrating
(1)Chargaff's Rules
2(A=U=12) lines per genetic code sequence of 24 quadruplet codons
2(G=C=12) lines per genetic code sequence of 24 quadruplet codons
(2) Watson-Crick's base pairing rules
2(A/U x 12) lines and 2(G/C x 12) lines per genetic code sequence of 24 quadruplet codons

(b) The workings of both Chargaff's rules and Watson-Crick's base pairing[a25] rules can be illustrated in every genetic code sequence of 24-quadruplet codons as per Diagram 1, New Genetic Code Structure in Dendritic Dichotomization.

Functional perspective

(a) Each quadruplet codon can be deployed as incumbent input prototype for the production of the genetic code of 24 quadruplets, reflecting potency.

(b) In addition, each quadruplet codon in the setting of the genetic code is functionally responsible for either the placement of a specific amino acid in the building of a protein type or the actuation of one of four specific signals in protein synthesis, showing uniqueness.

(c) The much desired *collinearity* between the 24 quadruplet code words of the genetic code and 20 amino acids of protein and 4 codon-size empty compartments left by 4 signals is evident in one to one correspondence.

(d) The four quadruplet codons in the genetic code which serve as four signals for start/ stop controls for place and time and convey no amino acids in the formation of protein type occasion four corresponding empty compartments amidst the sequence of 20 amino acids of a protein type. Each unit compartment is equivalent to the length of a quadruplet codon. These four empty compartments in their rightful places or positions in any protein type sequence are beneficial in two ways as follows:-

(i) They bring to perfection the collinearity between genetic code and the protein type it codes by ensuring that the quadruplet codons on the one hand and the amino acids/unit empty compartments on the other hand maintain s**erial positional parity,** as illustrated in the twin rows per chamber of Chart 2, depicting protein type proliferation and diversification as being diametrically opposite across the two parallels.

(ii) These empty compartments exist as flexible portions of protein types for protein folding, necessary for protein packaging for eventual disposal from factory. This brings us to the threshold of understanding protein folding and packaging. The appearance of the folded or packaged protein is seen to be in block form, of which the content per block can be surmised as being made up of 24 sequences of diverse protein types bearing 480 amino acids and 96 empty unit compartments corresponding to a batch of 24 consecutive input quadruplet codons in permutation synthesis in terms of proliferation and diversification.

(e) The four base types per quadruplet codon are usually motile to the effect of causing variation of sequence of a codon, which is responsible for the uniqueness of codons in the new genetic code.

CORRIDOR (A)	TRUNK (B)																								MARGIN (C)
ATGC	INITIAL INPUT SET USING SQUARE KINEMATICS TECHNIQUE (SEE APPENDIX)																								PRO-DUCTS
ATGC 1	AUGC	CGUA	UGCA	ACGU	GCAU	UACG	CAUG	GUAC	AGCU	UCGA	UCAG	GACU	GAUC	CUAG	CUGA	AGUC	ACUG	GCUA	UAGC	CGAC	CAUG	UGAC	ACUG	GUCA	GS 1
	S 1	S 2	AC 1	AC 2	AC 3	AC 4	AC 5	AC 6	AC 7	AC 8	AC 9	AC 10	AC 11	AC 12	AC 13	AC 14	AC 15	AC 16	AC 17	AC 18	AC 19	AC 20	S 3	S 4	PT 1
CGTA 2	CGUA	AUGC	GUAC	CAUG	UACG	GCGA	ACGU	UGCA	CUAG	GAUC	GACU	UCAG	UCGA	AGCU	AGUC	CUGA	CGAU	UAGC	GCUA	AUCG	ACUG	GUCA	CAGU	UGAC	GS 2
	S 2	S 1	AC 6	AC 5	AC 4	AC 3	AC 2	AC 1	AC 12	AC 11	AC 10	AC 9	AC 8	AC 7	AC 14	AC 13	AC 18	AC 17	AC 16	AC 15	S 3	S 4	AC 19	AC 20	PT 2
TGCA 3	CGCA	ACGU	GCAU	UACG	CAUG	GUAC	AUGC	CGUA	UCAG	GACU	GAUC	CUAG	CUGA	AGUC	AGCU	UCGA	UGAC	CAGU	GUCA	ACUG	AUCG	GCUA	UAGC	CGAU	GS 3
	AC 1	AC 2	AC 3	AC 4	AC 5	AC 6	S 1	S 2	AC 9	AC 10	AC 11	AC 12	AC 13	AC 14	AC 7	AC 8	AC 20	AC 19	S 4	S 3	AC 15	AC 16	AC 17	AC 18	PT 3
ACGT 4	ACGU	UGCA	CGUA	AUGC	GUAC	CAUG	UACG	GCAU	AGUC	CUAG	CUAG	GAUC	GACU	UCAG	UCGA	AGCU	ACUG	GUCA	CAGU	UGAC	UAGC	CGAU	AUCG	GCUA	GS 4
	AC 2	AC 1	S 2	S 1	AC 6	AC 5	AC 4	AC 3	AC 14	AC 13	AC 12	AC 11	AC 10	AC 9	AC 8	AC 7	S 3	S 4	AC 19	AC 20	AC 17	AC 18	AC 15	AC 16	PT 4
GCAT 5	GCAU	UACG	CAUG	GUAC	AUGC	CGUA	UGCA	ACGU	GAUC	CUAG	CUGA	AGUC	AGCU	UCGA	UCAG	GACU	GCUA	AUCG	CGAU	UAGC	UGAC	CAGU	GUCA	ACUG	GS 5
	AC 3	AC 4	AC 5	AC 6	S 1	S 2	AC 1	AC 2	AC 11	AC 12	AC 13	AC 14	AC 7	AC 8	AC 9	AC 10	AC 16	AC 15	AC 18	AC 17	AC 20	AC 19	S 4	S 3	PT 5

TACG 6	UACG	GCAU	AUGC	UGCA	CGUA	AUGC	GUAC	CAUG	UCGA	AGCU	AGUC	CUGA	CUAG	GAUC	GACU	UCAG	UAGC	CGAU	AUCG	GCUA	GUCA	ACUG	UGAC	CAGU	GS 6
	AC 4	AC 3	AC 2	AC 1	S 2	S 1	AC 6	AC 5	AC 8	AC 7	AC 14	AC 13	AC 12	AC 11	AC 10	AC 9	AC 17	AC 18	AC 15	AC 16	S 4	S 3	AC 20	AC 19	PT 6
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	
CATG 7	CAUG	GUAC	AUGC	CGUA	UGCA	ACGU	GCAU	UACG	CUGA	AGUC	AGCU	UCGA	UCAG	GACU	GAUC	CUAG	CAGU	UGAC	ACUG	GUCA	GCUA	AUCG	CGAU	UAGC	GS 7
	AC 5	AC 6	S 1	S 2	AC 1	AC 2	AC 3	AC 4	AC 13	AC 14	AC 7	AC 8	AC 9	AC 10	AC 11	AC 12	AC 19	AC 20	S 3	S 4	AC 16	AC 15	AC 18	AC 17	PT 7
GTAC 8	CAUG	CAUG	UACG	GCAU	ACGU	UGCA	CGUA	AUGC	GACU	UCAG	UCGA	AGCU	AGUC	CUGA	CUAG	GACU	GUCA	ACUG	UGAC	CAGU	CGAU	UAGC	GCUA	AUCG	GS 8
	AC 6	AC 5	AC 4	AC 3	AC 2	AC 1	S 2	S 1	AC 10	AC 9	AC 8	AC 7	AC 14	AC 13	AC 12	AC 11	S 4	S 3	AC 20	AC 19	AC 18	AC 17	AC 16	AC 15	PT 8
AGCT 9	AGCU	UCGA	GCUA	AUCG	CUAG	GAUC	UAGC	CGAU	ACUG	GUCA	GUAC	CAUG	CAGU	UGCA	UGCA	ACGU	AGUC	CUGA	GACU	UCAG	UACG	GCAU	AUGC	CGUA	GS 9
	AC 7	AC 8	AC 16	AC 15	AC 12	AC 11	AC 17	AC 18	S 3	S 4	AC 6	AC 5	AC 19	AC 20	AC 1	AC 2	AC 14	AC 13	AC 10	AC 9	AC 4	AC 3	S 1	S 2	PT 9
TCGA 10	UCGA	AGCU	CGAU	UAGC	GAUC	CUAG	AUCG	GCUA	UGAC	CAGU	CAUG	GUAC	GUCA	ACUG	ACGU	UGCA	UCAG	GACU	CUGA	AGUC	AUGC	CGUA	UACG	GCAU	GS 10
	AC 8	AC 7	AC 18	AC 17	AC 11	AC 12	AC 15	AC 16	AC 20	AC 19	AC 5	AC 6	S 4	S 3	AC 2	AC 1	AC 9	AC 10	AC 13	AC 14	S 1	S 2	AC 4	AC 3	PT 10
TCAG 11	UCAG	GACU	CAGU	UGAC	AGUC	CUGA	GUCA	ACUG	UAGC	CGAU	CGUA	AUGC	AUCG	GCUA	GCAU	UACG	UCGA	AGCU	CUAG	GBUC	GUAC	CAUG	UGCA	ACGU	GS 11
	AC 9	AC 10	AC 19	AC 20	AC 14	AC 13	S 4	S 3	AC 17	AC 18	S 2	S 1	AC 15	AC 16	AC 3	AC 4	AC 8	AC 7	AC 12	AC 11	AC 6	AC 5	AC 1	AC 2	PT 11
GACT 12	GACU	UCAG	ACUG	GUCA	AGUC	UGAC	CAGU	GCUA	AUCG	AUGC	CGUA	CGAU	UAGC	UACG	UACG	GCAU	GAUC	CUAG	AGCU	ACGA	UGCA	ACGU	GUAC	CAUG	GS 12
	AC 10	AC 9	S 3	S 4	AC 13	AC 14	AC 20	AC 19	AC 16	AC 15	S 1	S 2	AC 18	AC 17	AC 4	AC 3	AC 11	AC 12	AC 7	AC 8	AC 1	AC 2	AC 6	AC 5	PT 12
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	
GATC 13	GAUC	CUAG	AUCG	GCUA	UCGA	AGCU	CGAU	UAGC	GUCA	ACUG	ACGU	UGCA	UGAC	CAGU	CAUG	GUAC	GACU	UCAG	AGUC	CUGA	CGUA	AUGC	GCAU	UACG	GS 13
	AC 11	AC 12	AC 15	AC 16	AC 8	AC 7	AC 18	AC 17	S 4	S 3	AC 2	AC 1	AC 20	AC 19	AC 5	AC 6	AC 10	AC 9	AC 14	AC 13	S 2	S 1	AC 3	AC 4	PT 13
CTAG 14	CUAG	GAUC	UAGC	CGAU	AGCU	UCGA	GCUA	AUCG	CAGU	UGAC	UGCA	ACGU	ACUG	GUCA	AUAC	CAUG	CUAG	AGUC	UCAG	GACU	GCAU	UACG	CGUA	AUGC	GS 14
	AC 12	AC 11	AC 17	AC 18	AC 7	AC 8	AC 16	AC 15	AC 19	AC 20	AC 1	AC 2	S 3	S 4	AC 6	AC 5	AC 13	AC 14	AC 9	AC 10	AC 3	AC 4	S 2	S 1	PT 14

CTGA 15	CUGA	AGUC	UGAC	CAGU	GACU	UCAG	ACUG	GUCA	CGAU	UAGC	UACG	GCAU	GCUA	AUGC	AUGC	CGUA	CUAG	GAUC	UCGA	AGCU	ACGU	UGCA	CAUG	GUAC	GS 15
	AC 13	AC 14	AC 20	AC 19	AC 10	AC 9	S 3	S 4	AC 18	AC 17	AC 4	AC 3	AC 16	AC 15	S 1	S 2	AC 12	AC 11	AC 8	AC 7	AC 2	AC 1	AC 5	AC 6	PT 15
AGTC 16	AGUC	CUGA	GUCA	ACUG	UCAG	GACU	CAGU	UGAC	AUCG	GCUA	GCAU	UACG	UAGC	CGAU	CGUA	AUGC	AGCU	UCGA	GAUC	CUAG	CAUG	GUAC	ACGU	UGCA	GS 16
	AC 14	AC 13	S 4	S 3	AC 9	AC 10	AC 19	AC 20	AC 15	AC 16	AC 3	AC 4	AC 17	AC 18	S 2	S 1	AC 7	AC 8	AC 11	AC 12	AC 5	AC 6	AC 2	AC 1	PT 16
ATCG 17	AUCG	GCUA	UCGA	AGCU	CGAU	UAGC	GAUC	CUAG	ACGU	UGCA	UGAC	CAGU	CAUG	GUAC	GUCA	ACUG	AUGC	CGUA	UACG	GCAU	GACU	UCAG	AGUC	CUGA	GS 17
	AC 15	AC 16	AC 8	AC 7	AC 18	AC 17	AC 11	AC 12	AC 2	AC 1	AC 20	AC 19	AC 5	AC 6	S 4	S 3	S 1	S 2	AC 4	AC 3	AC 10	AC 9	AC 14	AC 13	PT 17
GCTA 18	GCUA	AUCG	CUAG	GAUC	UAGC	CGAU	AGCU	UCGA	GUAC	CAUG	CAGU	UGAC	UGCA	ACGU	ACUG	GUCA	GCAU	UACG	CGUA	AUGC	AGUC	CUGA	GACU	UCAG	GS 18
	AC 16	AC 15	AC 12	AC 11	AC 17	AC 18	AC 7	AC 8	AC 6	AC 5	AC 19	AC 20	AC 1	AC 2	S 3	S 4	AC 3	AC 4	S 2	S 1	AC 14	AC 13	AC 10	AC 9	PT 18
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	

TAGC 19	UAGC	CGAU	AGCU	UCGA	GCUA	AUCG	CUAG	GAUC	UGCA	ACGU	ACUG	GUCA	GUAC	CAUG	CAGU	UGAC	UACG	GCAU	AUGC	CGUA	CUGA	AGUC	UCAG	GACU	GS 19
	AC 17	AC 18	AC 7	AC 8	AC 16	AC 15	AC 12	AC 11	AC 1	AC 2	S 3	S 4	AC 6	AC 5	AC 19	AC 18	AC 4	AC 3	S 1	S 2	AC 13	AC 14	AC 9	AC 10	PT 19
CGAT 20	CGAU	UAGC	GAUC	CUAG	AUCG	GCUA	UCGA	AGCU	CAUG	GUAC	GUCA	ACUG	ACGU	UGCA	UGAC	CAGU	CGUA	AUGC	GCAU	UACG	UCAG	GACU	CUGA	AGUC	GS 20
	AC 18	AC 17	AC 11	AC 12	AC 15	AC 16	AC 8	AC 7	AC 5	AC 6	S 4	S 3	AC 2	AC 1	AC 20	AC 19	S 2	S 1	AC 3	AC 4	AC 9	AC 10	AC 13	AC 14	PT 20
CAGT 21	CAGU	UGAC	AGUC	CUGA	GUCA	ACUG	UCAG	GACU	CGUA	AUGC	AUCG	GCUA	GCAU	UACG	UAGC	CGAU	CAUG	GUAC	ACGU	UGCA	UCGA	AGCU	CUAG	GAUC	GS 21
	AC 19	AC 18	AC 14	AC 13	S 4	S 3	AC 9	AC 10	S 2	S 1	AC 15	AC 16	AC 3	AC 4	AC 17	AC 18	AC 5	AC 6	AC 2	AC 1	AC 8	AC 7	AC 12	AC 11	PT 21
TGAC 22	UGAC	CAGU	GACU	UCAG	ACUG	GUCA	CUGA	AGUC	UACG	GCAU	GCUA	AUCG	AUGC	CGUA	CGAU	UAGC	UGCA	ACGU	GUAC	CAUG	CUAG	GAUC	UCGA	AGCU	GS 22
	AC 20	AC 19	AC 10	AC 9	S 3	S 4	AC 13	AC 14	AC 4	AC 3	AC 16	AC 15	S 1	S 2	AC 18	AC 17	AC 1	AC 2	AC 6	AC 5	AC 12	AC 11	AC 8	AC 7	PT 22
ACTG 23	ACUG	GUCA	CUGA	AGUC	UGAC	CAGU	GACU	UCAG	AUGC	CGUA	CGAU	UAGC	UACG	GCAU	GCUA	AUCG	ACGU	UGCA	CAUG	GUAC	GAUC	CUAG	AGCU	UCGA	GS 23
	S 3	S 4	AC 13	AC 14	AC 18	AC 19	AC 10	AC 9	S 1	S 2	AC 18	AC 17	AC 4	AC 3	AC 16	AC 15	AC 2	AC 1	AC 5	AC 6	AC 11	AC 12	AC 7	AC 8	PT 23

GTCA 24	GUCA	ACUG	UCAG	GACU	CAGU	UGAC	AGUC	CUGA	GCAU	UACG	UAGC	CGAU	CGUA	AUGC	AUCG	GCUA	GUAC	CAUG	CAUG	UGCA	AGCU	UCGA	GAUC	CUAG	GS 24	
	S 4	S 3	AC 9	AC 10	AC 19	AC 20	AC 14	AC 13	AC 3	AC 4	AC 17	AC 18	S 2	S 1	AC 15	AC 16	AC 6	AC 5	AC 1	AC 10	AC 7	AC 8	AC 11	AC 12	PT 24	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	PT	GS

Key to Chart 2:

S1	=	Signal 1	=	Place Start Signal	GS	=	24-quadruplet genetic code sequence
S2	=	Signal 2	=	Time Start Signal	PT	=	Protein type sequence of 20 amino acids/4 codon
S3	=	Signal 3	=	Place Stop Signal			empty compartments left by 4 control signals
S4	=	Signal 4	=	Time Stop Signal			

The climax of the merits of the quadruplet codon is **the protein type proliferation and diversification** engineered by the genetic code which is the offspring of the quadruplet codon. The attributes of the genetic code in nature are well represented by this newly derived genetic code of 24 quadruplet codons as illustrated in Chart 2 captioned "**Protein type proliferation and diversification…**"

Chart 2 configuration and content.

Chart 2 titled "Protein Type Proliferation and Diversification: Climax of Merits of Quadruplet

Codons" is basically divided into three sections: Corridor, Trunk and Margin from left to right

and labelled A, B, C, respectively. The Corridor A, is divided into 24 segments or chambers and numbered 1-24. These 24 chambers are extended into the adjacent two sections, Trunk B and Margin C. But in these two sections, each chamber is subdivided into two subrows, upper and lower. Corridor A and Trunk B of Chart 2 represent the nucleus and cytoplasm of a cell respectively, in terms of content. The nucleus of a cell we are told houses the DNA double helix with rungs, each bearing the nucleotide set of four bases: Adenine, Thymine, Guanine, Cytosine (A, T, G, C) in diverse sequences. The basic number of rungs along the double helix due to one quadruplet input set of the DNA bases is given by the combinatorial input/output multiplicative replication system for 4 from 4 permutation is ${}_4P_4$ = 4! = 4x3x2x1 = 24 rungs. In effect the DNA double helix could be divided into basic lengths, each comprising 24 consecutive rungs for carriage of the DNA base quadruplet per rung. The RNA equivalent of the DNA base quadruplet per rung is Adenine, Uracil, Guanine, Cytosine (A, U, G, C) we are told! So Corridor A in the nucleus of a cell carries the DNA 24 rungs per basic length of the DNA double helix. The Trunk B, divided into 24 segments or chambers along its length and divided also into 24 columns across its breadth is the main centre of display of the outcome of the 24 quadruplet genetic code engagement in protein synthesis, wherein each of the 24 columns of the upper subrow is occupied by a quadruplet codon of unique sequence, while the lower subrow holds one of 20 amino acids of protein or one of four empty codon-size compartments left by the four signals for place and time based start/stop control during protein synthesis. The Margin C at the right-hand side of Chart 2 is a display for the count of the number of protein types synthesized as well as the number of 24 quadruplet genetic sequences responsible for the synthesis done. So Chart 2 and its content of 24 protein types synthesized by 24 diverse sequences of the 24 quadruplet genetic code is a portrayal of the outcome of the carriage of DNA base quadruplets on 24 consecutive rungs at one quadruplet per rung along the DNA double helix and linked by 24 RNA base quadruplets thereto across its (Chart 2) breadth.

The much needed collinearity between a genetic code sequence (in upper subrow) and a protein type (in lower subrow) is evident in the paired upper and lower subrows of all 24 segments (chambers) of Trunk B in support of protein synthesis geared to protein type proliferation and diversification.

The quadruplet codon in a collective sense is represented by the genetic code which is concise, composite and precise. The newly derived genetic code is concise in having a factorial complement formulary of ${}_4P_4$ = 4!=4x3x2x1=24 quadruplets. The genetic code is composite in its workforce of 24 consisting of "20 labourers" and "4 supervisors" under one management. It is precise in application affording collinearity of one to one correspondence with the 20 amino acids of protein and 4 codon-size empty compartments left by four signals in protein synthesis.

Finally, this rendition conveys a fourfold breakthrough as follows:-

a. Square kinematics (one of twelve systematic techniques for generating permutations and combinations in fulfillment of predicted factorial complements in combinatorics for the first time) for producing the permutations of 4 from 4 dissimilar objects.

b. Application of permutation synthesis to the successful derivation of the true genetic code structure of 24 quadruplets from an input set of the four RNA bases A.U.G.C (Adenine, Uracil, Guanine, Cytosine) that confronted molecular biologists from the 1950s without solution until now.

c. Offering a genetic code that exhibits collinearity with all protein types of one to one correspondence between its 24 quadruplet codons and the 20 amino acids/4 codon-size empty compartments left by 4 signals in protein types during protein synthesis. This genetic code can best be described as

a replica of the natural genetic code operating smoothly in protein synthesis in plants and animals [a26] since creation till date.

d. Presenting the novel theoretical finding that the sequence of 20 amino acids that makes a protein type is interspersed with 4 empty compartments corresponding to the 4 operational signals in the genetic code responsible for its formation. By unit compartment, it is meant, the equivalence of the length of the quadruplet codon. In effect the sequence of any protein type is discontinuous, thinking of the contiguity of the 20 constituent amino acids; unlike the genetic code sequence which codes it, that has 24 contiguous codons and therefore continuous.

Conclusion

The new 24-quadruplet genetic code as produced, being combinatorially fit is a worthy proposal for consideration for adoption. That the quadruplet codon is the seed source of the genetic code of 24 quadruplets servicing protein synthesis in plants and animals since creation is a remarkable credit to the honour of the new quadruplet codon. More highlights are offered on the theoretical aspects of protein synthesis up to the threshold of protein folding and packaging, from where we see that the folded or packaged protein consists of a number of blocks: each made up of 24 sequences of diverse protein types bearing a total of 480 amino acids and 96 empty unit compartments corresponding to the output of a batch of 24 consecutive input quadruplet codons in permutation synthesis. In other words Chart 2 reflects the content of a block.

Recommendation

This precise genetic code structure of 24 quadruplets is therefore recommended for spelling by experimental experts to render it fit for adoption in coding application in protein synthesis and studies.

References (Part One)

1. Jason W. Chin, Kaihang Wang, Mag. Wolfgang, Reprogramming the Genetic Code: From Triplet to Quadruplet Codes. Angewante Minireviews Angwe. Chem. Int. Ed. 2012; 51(10):2288-2297
2. J. Christopher Anderson, Ning Wu, Stephen W. Santoro, Vishva Lakshman, David S. King et al, An Expanded Genetic Code With A Functional Quadruplet Codon PNAS, 2004; 101(20):7566-7571.
3. Dieter Soll, Uttam L RajBhandary, The genetic code – Thawing the "Frozen accident" J. Biosci,2006;31(4):459-463
4. Agris, P.F, Decoding the genome: A modified view Nucleic Acids Res. 2004;32(1):223-238
5. Nirenberg M and Leder P, RNA codewords and proteins synthesis: The effect of trinucleotides upon the binding of sRNA to ribosomes, 1964; 145(3639):1399-1407
6. Becker, W, M, The World of the Cell. The Benjamin and Cummings Publishing Company, Inc.1986.
7. Crick, F.R.C, The Genetic Code II, Scientific American, 1962;
8. Crick, F.R.C, The Genetic Code II, Scientific American, 1966;
9. Roberts, M. B. V. Biology, A Functional Approach, The English Language Book Society and Nelson, 1971;
10. Yanofsky, Charles. The Gene Structure and Protein Structure, Scientific American, 1967;216(5):80-95

References (Part Two)

a1 - Mine, original term

a2 - Mine, original term

a3 - Mine, original term

a4 - Truth in Science recommended for use by all stake holders.

a5 - George Gamow: https://en.wikipedia.org/wiki/George Gamow

a6 - Mine from Numeration Science Literature development

a7 - The World of the Cell by Wayne M. Becker, (1986). The Benjamin Publishing Company, Inc.

a8 - Biology A Functional Approach (1971) Page 492, by M. B. V. Roberts. The English Language Book Society and Nelson.

a9 - Genetics A Molecular Approach 2nd Edition (1992) Page 124 By T. A. Brown, Chapman and Hall London U.K.

a10 - Opinion, mine because of degeneracy and other irregularities associated with it.

a11 - Internet on Francis Crick, http/en.wikipedia.org/wiki/Francis.Crick.

a12 - Maiden idea illustrated in Chart 2 captioned "Protein Type Proliferation and Diversification…" in view of input quadruplet in Corridor A and 24 output quadruplets in Trunk B.

a13 - Maiden idea on the application of the four RNA bases in the new technique of computing 4 from 4 permutations by Square Kinematics View Mixing Scheme.

a14 - Maiden idea validated by the output of eight unique permutation quadruplets (non-isodigitals) from each of three pathways in the working of the technique.

a15 - Maiden demonstration of performance of the new technique of generating 4 from 4 permutations numbering 24 quadruplets as displayed in lines 1-24 of Chart 1 in fulfilment of 4P4 = 4! = 4x3x2x1 = 24 quadruplets.

a16 - Maiden idea on the framing of components of the input/output multiplicative replication system using the Square Kinematics View Mixing technique invented by this author.

a17 - Maiden presentation of the result of 4 from 4 permutation i.e. 4P4 = 4! = 4x3x2x1 = 24 quadruplets as displayed in Chart 1, lines 1-24 as produced by the Square Kinematics technique.

a18 - List of 20 amino acids in Table 1 col. 1 adapted from Fig. 17.4 page 529 of The World of the Cell (1986) by Becker, Wayne M.

a19 - Ibid.

a20 - Irregularity-free 24-quadruplet genetic code produced in Chart 1 lines 1-24, and presented in Table 1 under Results.

a21 - Maiden opinion identifying the production of the 24-quadruplet genetic code from a quadruplet input codon as a merit of the quadruplet codons.

a22 - Maiden opinion based on collinearity between genetic code and protein type evident in Chart 2 in support of the Primordial choice of RNA four bases A, U, G, C as substitute for 20 amino acids of protein for input set in the input/output multiplicative replication system aimed at protein type proliferation and diversification being required of the working of the genetic code in protein building in Nature.

a23 - Maiden observation of the performance of the Square Kinematics View Mixing technique illustrated in Fig. 1 and Chart 1, involving one quadruplet input set yielding output sequence of 24 quadruplets representing the new 24-quadruplet genetic code.

a24 - Maiden categorization of certain base types missing in some triplet codons amounting to underutilization of them as flaws of the 64 triplet genetic code by this author based on combinatorial examination of the reigning 64 triplet genetic code.

a25 - "The World of the Cell" page 409, by Wayne M. Becker (1986), The Benjamin Publishing Company, Inc.

a26 - Maiden opinion in favor of the 24-quadruplet genetic code in efficacy and efficiency.