This article was published in "«Advances in Bioinformatics and its Applications" (editors – M.He, G.Narasimhan, S.Petoukhov), Proceedings of the International Conference (Florida, USA, 16-19 December 2004), Series in Mathematical Biology and Medicine, 2005, v.8, pp.512-532, New Jersey-London-Singapore-Beijing, World Scientific.

THE RULES OF DEGENERACY AND SEGREGATIONS IN GENETIC CODES. THE CHRONOCYCLIC CONCEPTION AND PARALLELS WITH MENDEL'S LAWS

SERGEY V. PETOUKHOV

Department of Biomechanics, Mechanical Engineering Research Institute, Russian Academy of Sciences, M. Kharitonyevsky pereulok 4, Moscow, Russia 101990.

E-mail: petoukhov@hotmail.com

Systems of elements of genetic code are studied by their cognitive presentation in a form of mathematical matrices of symbolic and numerical kinds. This cognitive form of data presentation permits to discover new phenomenological rules of evolution of genetic codes, to reveal two branches of evolution within genetic code, to present hidden interrelations between the golden section and parameters of genetic polyplets, to disclose matrices of a hyperbolic turn in genetic matrices, *etc*. Mysterious sets of structures, realized by the nature in a hierarchical system of genetic codes, can be confronted by a heuristic manner with families of mathematical matrices, which contain elements of these structures. A few rules of degeneracy and segregations of genetic codes are revealed in this direction. A new answer on the fundamental question - "why 20 amino acids?" - is proposed as well.

.

1. Introduction.

This article is devoted to author's results of investigation of systems of genetic codes from the viewpoint of a cognitive form of a matrix presentation of these systems. The form is based at symbolical and numerical matrices of genetic polyplets. These matrices, which are conditionally named as "genomatrices", constitute special families, which are generated and investigated by classical matrix calculus. Numerical genomatrices are originated from symbolical genomatrices, when symbols of genetic elements are replaced by their real quantitative parameters. This cognitive form of data presentation has already led to a discovery of new phenomenological rules of evolution of genetic codes, to revelation of two branches of evolution within genetic code, to presentation of hidden interrelations between the golden section and parameters of genetic polyplets, to disclose matrices of a hyperbolic turn in genetic matrices, to substantiation of new approaches in the field of genetic codes, *etc*. Mysterious sets of structures, realized by the nature in a hierarchic

system of genetic codes, can be confronted by a heuristic manner with families of mathematical matrices, which contain elements of these structures. A few rules of degeneracy and segregations of genetic codes are revealed in this direction. As well, a new answer on the fundamental question - "why 20 amino acids?" - is proposed.

2. A Cognitive Matrix Presentation of a Set of Genetic N-plets

2.1. *Symbolic and Numerical Genomatrices*

Polyplets are one of the main peculiarities of genetic code. Really, the alphabet of genetic code is a set of four monoplets: A (adenine), C (cytosine), G (Guanine), U/T (uracil in RNA or thymine in DNA); 64 triplets encode amino acids; each protein is encoded by more or less long polyplets (each protein with n amino acids is encoded by a 3n-plet).

Information is stored in computers in a form of matrices. The author considers a set of four "letters" of genetic alphabets in a form of the (2x2) matrix P (Fig. 1). Besides that, he considers a family of matrices $P^{(n)}$, where n = 1, 2, 3,.. [1]. The symbol $P^{(n)}$ means, that the matrix P is raised to the power "n" in a sense of tensor (Kronecker) multiplication. For example, if P is raised to the tensor power 3, a result is the (8x8) matrix of 64 triplets (Fig. 1). Tensor multiplication is applied in mathematics and physics widely. Its symbol is "⊗". The brief article by B.G. Konopelichenko and Y.B. Rumer [2] was the first attempt to apply ideas of tensor multiplication of genetic matrices in the field of genetic code. This work had no significant continuation during a quarter of a century in a connection with a death of Y.B. Rumer in 1976.

P =

	1	0
1	C	A
0	U	G

,

$P^{(2)} = P \otimes P =$

	11	10	01	00
11	CC	CA	AC	AA
10	CU	CG	AU	AG
01	UC	UA	GC	GA
00	UU	UG	GU	GG

$P^{(3)} =$

	111	110	101	100	011	010	001	000
111	CCC	CCA	CAC	CAA	ACC	ACA	AAC	AAA
110	CCU	CCG	CAU	CAG	ACU	ACG	AAU	AAG
101	CUC	CUA	CGC	CGA	AUC	AUA	AGC	AGA
100	CUU	CUG	CGU	CGG	AUU	AUG	AGU	AGG
011	UCC	UCA	UAC	UAA	GCC	GCA	GAC	GAA
010	UCU	UCG	UAU	UAG	GCU	GCG	GAU	GAG
001	UUC	UUA	UGC	UGA	GUC	GUA	GGC	GGA
000	UUU	UUG	UGU	UGG	GUU	GUG	GGU	GGG

Fig. 1. The beginning of the family of genomatrices $P^{(n)}$ for n = 1, 2, 3 [1,3].

Formally constructed genomatrix $P^{(3)}$ contains the complete set of 64 triplets (Fig. 2). Each matrix column represents one of eight classical octets by Wittmann [3], which reflect real biochemical properties of triplets. It is a first confirmation of accuracy of such matrix approach, which reveals an order within the genetic system.

A matrix $P^{(n)}$ consists of the complete set of n-plets as its matrix elements. $P^{(n)}$ has an order (2^n x 2^n). All matrices $P^{(n)}$ will be named conditionally as genetic matrices (or genomatrices) because of their connections with elements of the genetic code. Each matrix $P^{(n)}$ has different binary numerations of its column and rows. These different numerations are connected with different "binary sub-alphabets" of genetic alphabet (see details in [1,5]). The family of genomatrices $P^{(n)}$, when "n" is sufficiently great, represents by unified manner the complete system of genetic polyplets beginning from alphabetic letters (monoplets) and triplets, which encode amino acids, till longest polyplets, which encode the longest proteins in biological bodies.

Each quadrant of a genomatrix $P^{(n)}$ contains a complete set of n-plets with the same initial letter. If one doesn't pay attention to this first letter, then each quadrant of $P^{(n)}$ reproduces a matrix $P^{(n-1)}$ of previous "generation" or of previous value "n-1" (it is provided by the properties of tensor multiplication of matrices). In other words, a genomatrix of each new generation contains information about all previous generations in a hidden form (genomatrices with "a memory of generations"). The longest genomatrix $P^{(n)}$, n-plets of which encode the longest protein, is named archmatrix. Since such archmatrix contains information about all genomatrices with shorter coding polyplets, the task of investigation of the complete system of genetic polyplets is formally bring to investigation of this archmatrix. Many results, described in this paper, were obtained due to analysis of this archmatrix or the family of $P^{(n)}$.

If one replaces each symbol of nitrogenous bases in symbol genomatrices $P^{(n)}$ by there real quantitative parameters, respective numerical genomatrices are appeared. Let us consider a concrete example with quantities of hydrogen bonds of complementary nitrogenous bases: C and G have three hydrogen bonds, A and U/T have two ones. Many authors suspect, that these hydrogen bonds have an information meaning. Let us replace each polyplet in $P^{(n)}$ by the product of these numbers of hydrogen bonds: C=G=3, A=U=2. For instance, due to such operation, the triplet CGA is replaced by number 3x3x2=18 in the genomatrix $P_{MULT}^{(3)}$. As a result, multiplicative nonsingular numerical matrices

$P_{MULT}^{(n)}$ are appeared. Figure 2 demonstrates the multiplicative matrix $P_{MULT}^{(3)}$ constructed in this way.

All matrices $P_{MULT}^{(n)}$ are symmetrical relative to both diagonals and they can be named "bi-symmetrical matrices". The sums of all numbers in the cells of each row and of each column in a matrix $P_{MULT}^{(n)}$ are equal to 5^n ; a total sum of numbers is equal 10^n. According to [6-8], such matrices after their special normalization belong to a class of stochastic matrices, which are used in new approaches. For instance, in the case of the matrix $P_{MULT}^{(3)}$, these sums are equal to $125 = 5^3$ and to $1000=10^3$ respectively. A rank of this matrix is equal to 8. The matrix $P_{MULT}^{(3)}$ has four kinds of numbers only: 8, 12, 18 and 27. The certain regularities are observed in their disposition.

27	**18**	**18**	**12**	**18**	**12**	**12**	**8**	*125*
18	**27**	**12**	**18**	**12**	**18**	**8**	**12**	*125*
18	**12**	**27**	**18**	**12**	**8**	**18**	**12**	*125*
12	**18**	**18**	**27**	**8**	**12**	**12**	**18**	*125*
18	**12**	**12**	**8**	**27**	**18**	**18**	**12**	*125*
12	**18**	**8**	**12**	**18**	**27**	**12**	**18**	*125*
12	**8**	**18**	**12**	**18**	**12**	**27**	**18**	*125*
8	**12**	**12**	**18**	**12**	**18**	**18**	**27**	*125*
125	*125*	*125*	*125*	*125*	*125*	*125*	*125*	***1000***

Fig. 2. The multiplicative matrix $P_{MULT}^{(3)}$ with cells, which contain products of numbers of complementary hydrogen bonds for triplets: C=G=3, A=U=2. The right column shows sums of numbers of each row. The lower row shows sums of numbers of each column. Bold frames mark diagonal cells.

2.2. *A Golden Section and Golden Genomatrices*

The author discovered, that these genomatrices $P_{MULT}^{(n)}$ are connected with a famous golden section. This golden section (or "divine proportion") is a mathematical symbol of self-reproduction for many centuries (see website www.goldenmuseum.com). A golden section is a value $\varphi = (1+5^{0.5})/2 = 1{,}618..$ (sometimes the inverse of this value φ^{-1} is called a golden section in literature). Many authors around the world publish articles about manifestation of golden section in different physiological systems: cardio-vascular systems, respiratory systems, electric activities of brain, locomotion activity, *etc*. As a result, the golden section appears as a possible important element in a mysterious phenomenon of heritable integration of physiological subsystems in biological body.

The discovered hidden connection between the golden section φ and basic parameters of genetic codes lies in the mathematical fact, that each genomatrix $P_{MULT}^{(n)}$ is identical to the square of a respective "golden" matrix $\Phi_{MULT}^{(n)}$: $P_{MULT}^{(n)} = (\Phi_{MULT}^{(n)})^2$ [1,9]. For instance, Fig. 3 demonstrates the matrix $\Phi_{MULT}^{(3)} = (P_{MULT}^{(3)})^{1/2}$. This matrix has only two pairs of inverse numbers, generated from a single value of the golden section: φ^1 and φ^{-1}, φ^3 and φ^{-3}. Matrices with matrix elements, all of which are equal to golden section φ in different powers only, are conditionally named as "golden matrices".

φ^3	φ^1	φ^1	φ^{-1}	φ^1	φ^{-1}	φ^{-1}	φ^{-3}
φ^1	φ^3	φ^{-1}	φ^1	φ^{-1}	φ^1	φ^{-3}	φ^{-1}
φ^1	φ^{-1}	φ^3	φ^1	φ^{-1}	φ^{-3}	φ^1	φ^{-1}
φ^{-1}	φ^1	φ^1	φ^3	φ^{-3}	φ^{-1}	φ^{-1}	φ^1
φ^1	φ^{-1}	φ^{-1}	φ^{-3}	φ^3	φ^1	φ^1	φ^{-1}
φ^{-1}	φ^1	φ^{-3}	φ^{-1}	φ^1	φ^3	φ^{-1}	φ^1
φ^{-1}	φ^{-3}	φ^1	φ^{-1}	φ^1	φ^{-1}	φ^3	φ^1
φ^{-3}	φ^{-1}	φ^{-1}	φ^1	φ^{-1}	φ^1	φ^1	φ^3

Fig. 3. The golden octet genomatrix $\Phi_{MULT}^{(3)} = (P_{MULT}^{(3)})^{1/2}$. Here φ is a golden section.

If the simplest genomatrix $P_{MULT}^{(1)}$ is raised to the power ½ in ordinary sense (that is, if we take the square root), the result is a bi-symmetrical golden matrix $\Phi = (P_{MULT}^{(1)})^{1/2}$, the matrix elements of which are equal to the golden section and to its inverse value. Fig. 4 demonstrates the matrices $P_{MULT}^{(1)} = \Phi_{MULT}^2$, $(P_{MULT}^{(1)})^{1/2} = \Phi$, $(P_{MULT}^{(2)})^{1/2} = \Phi_{MULT} \otimes \Phi_{MULT} = \Phi^{(2)}$.

$$P_{MULT} = \begin{vmatrix} 3 & 2 \\ 2 & 3 \end{vmatrix} ; \quad (P_{MULT})^{1/2} = \Phi_{MULT} = \begin{vmatrix} \varphi & \varphi^{-1} \\ \varphi^{-1} & \varphi \end{vmatrix} ; \quad (P_{MULT}^{(2)})^{1/2} = \Phi_{MULT}^{(2)} = \begin{vmatrix} \varphi^2 & \varphi^0 & \varphi^0 & \varphi^{-2} \\ \varphi^0 & \varphi^2 & \varphi^{-2} & \varphi^0 \\ \varphi^0 & \varphi^{-2} & \varphi^2 & \varphi^0 \\ \varphi^{-2} & \varphi^0 & \varphi^0 & \varphi^2 \end{vmatrix}$$

Fig. 4. The beginning of a family of golden genetic matrices $(P_{MULT}^{(n)})^{1/2} = \Phi_{MULT}^{(n)}$, where "φ" is the golden section.

The author has proposed in principle a new definition to a golden section on the basis of the matrix specifics of genetic code systems: a golden section φ and its inverse value φ^{-1} are single matrix elements of a bi-symmetrical matrix Φ_{MULT}, which is the square root from such bi-symmetrical matrix P_{MULT} (2x2), elements of which are genetic numbers of hydrogen bonds (C=G=3, A=U=2) and which has a positive determinant [1]. This definition does not use conceptions of line segments, their proportions, *etc.*, which are traditional for the definition of golden section. Besides, it has a binary character: it defines a

system pair of φ and φ^{-1} simultaneously. In our opinion, many realizations of the golden section in the nature are connected with its matrix essence and with its matrix representation. Probably, golden section phenomena in many physiological systems are connected with a realization of golden section in matrices of genetic code. The idea of golden matrices can also be useful for non-biological phenomena.

2.3. *Other Mathematical Properties and Forms of Genomatrices*

Golden and other bi-symmetrical genomatrices have a group-invariant property, which is connected with invariance of matrix numerical mosaics under multiplications of matrices (Petoukhov 2001b, 2003c, 2004). We will explain this property through the example of the golden matrix $\Phi_{MULT}^{(3)}$ from Fig. 3. The matrix consists of four numbers: φ^1 and φ^{-1}, φ^3 and φ^{-3} only with their special disposition. The numbers φ^3 and φ^{-3} are disposed at matrix diagonals separately in the form of a diagonal cross. Fig. 5 demonstrates a disposition of number φ^{-1} in matrix cells, a set of which produces a special mosaic in a form of a symbol "69" (these cells are marked by dark in the left matrix). The number φ are disposed in matrix cells, a set of which produces a mirror-symmetrical mosaic in comparison with a 69-mosaic of previous case (these cells are marked by dark in the right matrix).

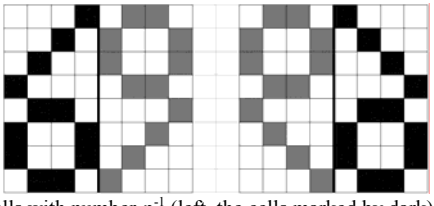

Fig. 5. A mosaic of cells with number φ^{-1} (left, the cells marked by dark) and a mosaic of cells with number φ (right) from the matrix $\Phi_{MULT}^{(3)}$ (Fig. 3).

If the octet matrix $\Phi_{MULT}^{(3)}$ is raised to the power of n (n = 2, 3, …), the resulting octet matrix $(\Phi_{MULT}^{(3)})^n$ has a new set of four numbers in each case again with the same their disposition within of octet matrices, that is with the same numerical mosaics. In such genomatrices, the invariance of numerical mosaics is independent of value of numbers. This property is realized for such matrices with the arbitrary set of four numbers a, b, c, d, if they are disposed by the same manner within a matrix. Moreover, if we have one matrix *X* with a set of four numbers *a, b, c, d* and another matrix *Y* with another set of four numbers *k, m, p, q*, then the product of these matrices will be the matrix *Z=X*Y* with a set of new four numbers *r, g, v, z* and with the same mosaic of

their disposition [1,9]. It is obvious that the four symbols (for example, *a, b, c, d*) in such matrices can be not only ordinary numbers, but also arbitrary mathematical objects: complex numbers, matrices, functions of time (for example, it can be that $a=R*\cos wt$, $b=T*\sin wt$, …), *etc*. Each genomatrix $P_{MULT}^{(n)}$ can be interpreted as a metrical tensor of Euclidean 2^n-dimansional space, where respective matrix $\Phi_{MULT}^{(n)}$ is an appropriate matrix of coordinates of vectors of an affine basis. The author has proposed an idea of genetic inheritance of a system of such metrical genotensors, which are related with biophysical genotensor fields and with Riemannian geometry.

A genomatrix $\Phi_{MULT}^{(n)}$ has "n+1" kinds of numbers and respectively "n+1" numerical mosaics for them. An appearance of such mosaics becomes more complicated with enlarged "n". Mosaics of all matrices $\Phi_{MULT}^{(n)}$ have a mutual interrelation (mosaics with "a memory of generations") [1,9]. Any bi-symmetrical (2x2)-genomatrix, which is divided by a value of a square root from its determinant (for example, the genomatrix $(5^{-1/4})*\Phi_{MULT}$), belongs to a famous class of bi-symmetrical matrices of a hyperbolic turn, which have an important meaning in the special theory of relativity and in the relativistic solitonic sine-Gordon equation (the author supposed earlier that this equation plays a role of a fundamental solitonic equation of living matter in a connection with common phenomena of supramolecular solitons in different biological objects [3]). A hyperbolic turn lies in the basis of a theory of logarithms, important physiological meaning of which is demonstrated, for instance, by the psychophysical Weber-Fechner's law of logarithmic dependence between feelings and irritations. Therefore a formal connection exists between genomatrices and logarithmic physiological phenomena. The development of these analogies is conducted now.

Four letters A, C, G, U can be disposed in a (2x2) genomatrix by 16 variants. The genomatrix P (Fig. 1), proposed by the author, differs from other 15 variants by a set of two peculiarities: 1) its numerical presentation on the basis of quantity of hydrogen bonds reveals a connection of the genomatrix with the golden section, golden matrices (Fig. 3) and a hyperbolic turn; 2) columns of the matrix $P^{(3)}$ reproduce classical Wittmann's octets (Fig. 1).

A selection of a form of basic (2x2) genomatrix is an arbitrary (creative) act. Along with the matrix form P, considered above (Fig. 1), the author has proposed to use other possible forms in heuristic aims [1,10]. Figure 6 demonstrates examples of such genomatrices M_K, each of which generates a respective family of genomatrices $M_K^{(n)}$. These matrices are quasi-bisymmetrical ones because of small asymmetries in relative to second diagonal. These quasi-bisymmetrical matrices have many mathematical

properties of bisymmetrical ones. For instance, they have described mosaic-invariance property and have the same connection with a golden section.

These new forms differ from the initial matrix P by arithmetic signs or imaginary values in matrix presentation of separate genetic letters, *etc*. They permit to reveal interesting parallelisms among genomatrices and famous matrices from different fields of sciences. For instance, the matrix M_{1MULT} belong to a class of matrices of spiral transformation and it is used now for modeling of widespread spirals in biological morphology in connection with genetic structures. The matrix M_{2MULT} has analogous traits with the matrix by Pauli from theoretical physics. The matrices $M_{2MULT}^{(n)}$ are essential for a new branch of genetic code investigations from the viewpoint of statistic quantum mechanics [1,10]. When $M_{2MULT}^{(n)}$ is normalized into a respective matrix with a trace, which is equal to 1, it can be interpreted as a density matrix. It is useful, when we try to find connections between physical structures of genetic code and quantum physics. It is known, that the concept of a wave function from quantum mechanics can be used for systems with a few particles only. If we have a genetic system with many particles, we need to operate with density matrices as a basic formalism of statistical quantum mechanics.

$$M_1 = \begin{vmatrix} C & A \\ -U & G \end{vmatrix} ; \quad M_{1MULT} = \begin{vmatrix} 3 & 2 \\ -2 & 3 \end{vmatrix} ; \quad M_2 = \begin{vmatrix} C & -iA \\ iU & G \end{vmatrix} ; \quad M_{2MULT} = \begin{vmatrix} 3 & -2i \\ 2i & 3 \end{vmatrix}$$

Fig. 6. Examples of alternative forms of basic genomatrices M_1 and M_2 and their numerical presentation M_{1MULT} and M_{2MULT} for the case of hydrogen bonds C=G=3, A=U=2.

2.4. *A Disposition of Amino Acids in the Genomatrix $P^{(3)}$*

It seems improbable, that 20 amino acids could be arranged automatically by symmetrical order into 64 cells of bi-periodic matrix $P^{(3)}$. One has at least two important reasons to think so. Firstly, the matrix $P^{(3)}$ for 64 triplets was constructed by the author for triplets only without any initial information about amino acids. This matrix was constructed in this paper by absolutely formal manner in a form of a third tensor power of the matrix P. The second reason is that a great set of intermediate biochemical agents (ferments, nucleic acids, etc.) and of processes exist between DNA's sequence of triplets and a final action of assembling of different amino acids into a protein chain. But such improbable fact of symmetrical disposition of 20 amino acids with their different coded degeneracy in $P^{(3)}$ comes true really (see details in [1,3]). This fact is a serious argument for the benefit of such matrix approach. Figure 7

demonstrates such disposition for the vertebrate mitochondria genetic code, which is considered by many authors as the most ancient and “perfect” variant of genetic code [12, pp. 65-68].

Each sub-quadrant (2x2) of the matrix $P^{(3)}$ contains a subfamily of those four triplets, which are equivalent to each other by two first letters. Such quadruple of triplets will be named “a subfamily of NN-triplets”. A complete set of 16 subfamilies of NN-triplets is segregated by the nature into two subsets with 8 subfamilies in each. The first subset, marked by dark on Fig. 7, contains those subfamilies of NN-triplets, coded values of which are independent of their third letter. In this reason, all four triplets of such subfamily encode the same amino acids. On the contrary, the second subset, marked by white color, contains those subfamilies of NN-triplets, coded values of which are dependent of their third letter. In this reason, each such subfamily has triplets, which encode different amino acids or stop-signals.

	111	**110**	**101**	**100**	**011**	**010**	**001**	**000**
111	**CCC Pro 63**	CCA **Pro 62**	**CAC His 61**	**CAA Gln 60**	**ACC Thr 59**	**ACA Thr 58**	**AAC Asn 57**	**AAA Lys 56**
110	**CCU Pro 55**	**CCG Pro 54**	**CAU His 53**	**CAG Gln 52**	**ACU Thr 51**	**ACG Thr 50**	**AAU Asn 49**	**AAG Lys 48**
101	**CUC Leu 47**	**CUA Leu 46**	**CGC Arg 45**	**CGA Arg 44**	**AUC Ile 43**	**AUA Met 42**	**AGC Ser 41**	**AGA Stop 40**
100	**CUU Leu 39**	**CUG Leu 38**	**CGU Arg 37**	**CGG Arg 36**	**AUU Ile 35**	**AUG Met 34**	**AGU Ser 33**	**AGG Stop 32**
011	**UCC Ser 31**	**UCA Ser 30**	**UAC Tyr 29**	**UAA Stop 28**	**GCC Ala 27**	**GCA Ala 26**	**GAC Asp 25**	**GAA Glu 24**
010	**UCU Ser 23**	**UCG Ser 22**	**UAU Tyr 21**	**UAG Stop 20**	**GCU Ala 19**	**GCG Ala 18**	**GAU Asp 17**	**GAG Glu 16**
001	**UUC Phe 15**	**UUA Leu 14**	**UGC Cys 13**	**UGA Trp 12**	**GUC Val 11**	**GUA Val 10**	**GGC Gly 9**	**GGA Gly 8**
000	**UUU Phe 7**	**UUG Leu 6**	**UGU Cys 5**	**UGG Trp 4**	**GUU Val 3**	**GUG Val 2**	**GGU Gly 1**	**GGG Gly 0**

Fig. 7. An enriched representation of the bi-periodic octet matrix $P^{(3)}$ for a case of the vertebrate mitochondria genetic code. The matrix consists of binary numeration numbers of its columns and

rows, 64 triplets with their coordination numbers from 0 up to 63 in decimal system, and 20 amino acids with their traditional abbreviations. Stop-codons are marked as “stop”.

A disposition of these “black” and “white” subfamilies of NN-triplets in the matrix $P^{(3)}$ is very symmetric. For instance, left and right halves of this matrix are mirror-anti-symmetric to each other in its colors: any pair of cells, disposed by mirror-symmetrical manner in these halves, has opposite colors. Diagonal quadrants of the matrix are equivalent to each other from the viewpoint of their mosaic. The rows 1-2, 3-4, 5-6, 7-8 are equivalent to each other from the viewpoint of a disposition of the same amino acids in their appropriate cells, *etc*. From the set of 20 amino acids, 8 amino acids (Ala, Arg, Gly, Leu, Pro, Ser, Thr, Val) belong to the “black” sub-families of NN-triplets and other 12 amino acids (Asn, Asp, Cys, Gln, Glu, His, Ile, Lys, Met, Phe, Trp, Tyr) are presented in the “white” sub-families. The two subsets of these amino acids will be named “canonical” ones.

2.5. *The Rules of Degeneracy; Two Branches of Evolution within Genetic Codes*

At the biological level, where 64 triplets code 20 amino acids, modern science discovered the existence of many evolutionary variants of genetic codes. These genetic codes differ through their specifics of degeneracy (or by concrete relations between 20 amino acids and 64 triplets). Each amino acid is coded in a concrete genetic code by a certain quantity of triplets. This quantity of its triplets will be named as number of degeneracy of this amino acid in this genetic code. For example, the amino acid Thr is encoded by four triplets ACC, ACA, ACU, ACG in the case of the vertebrate mitochondria code. In other words, its number of degeneracy is equal to 4 in this genetic code. In another genetic code, this amino acid can have another number of degeneracy, for example, this number can be equal to 8. Different genetic codes have different sets of numbers of degeneracy.

Modern science knows 17 genetic codes, which are realized in different kinds of organisms or of their subsystems. Figure 8 demonstrates a table, constructed by the author, with numbers of degeneracy of amino acids for all 17 genetic codes. The initial data about these codes were taken from the website of the National Center for Biotechnological Information (USA): http://www.ncbi.nlm.nih.gov/Taxonomy/Utils/wprintgc.cgi. Numbers of degeneracy of amino acids are equal to numbers from 1 to 8. For example, the first genetic code in Fig. 8 has 12 amino acids, which number of degeneracy is equal to 2; 6 amino acids, which number of degeneracy is equal to 4; and 2 amino acids, which number of degeneracy is equal 6. At first it seems, that the

distribution of numbers of degeneracy of amino acids in a set of 17 genetic codes is chaotic or non-regular on the whole. But a series of all numbers of degeneracy can be divided into two regular categories: a category of low-degeneracy amino acids (with their numbers of degeneracy from 1 to 3) and a category of high-degeneracy amino acids (with their numbers of degeneracy

Genetic code	Distribution of numbers of degeneracy from 1 to 8 among 20 AA								ΣAA with ND from 1 to 3	ΣAA with ND from 4 to 8
	1	2	3	4	5	6	7	8		
1) The Vertebrate Mitochondrial Code		12		6		2			12	8
2) The Standard Code	2	9	1	5		3			12	8
3) The Mold, Protozoan, and Coelenterate Mitochondrial Code and the Mycoplasma /Spiroplasma Code	1	10	1	5		3			12	8
4) The Invertebrate Mitochondrial Code		12		6		1		1	12	8
5) The Echinoderm Mitochondrial Code	2	8	2	6		1		1	12	8
6) The Euplotid Nuclear Code	2	8	2	5		3			12	8
7) The Bacterial and Plant Plastid Code	2	9	1	5		3			12	8
8) The Ascidian Mitochondrial Code		12		5		3			12	8
9) The Flatworm Mitochondrial Code	2	7	3	6		1		1	12	8
10) Blepharisma Nuclear Code	2	8	2	5		3			12	8
11) Chlorophycean Mitochondrial Code	2	9	1	5		2	1		12	8
12) Trematode Mitochondrial Code	1	10	1	6		1		1	12	8
13) Scenedesmus obliquus mitochondrial Code	2	9	1	5	1	1	1		12	8
14) Thraustochytrium Mitochondrial Code	2	9	1	5	1	2			12	8
15) The Alternative Yeast Nuclear Code	2	9	1	5	1	1	1		12	8
16) The Yeast Mitochondrial Code		13		5		1		1	13	7
17) The Ciliate, Dasycladacean and Hexamita Nuclear Code	2	8	1	6		3			11	9

Fig. 8. 17 genetic codes and distributions of their numbers of degeneracy (ND) among 20 amino acids (AA). „Stop" means stop-codons. A set of observed numbers of degeneracy of amino acids in genetic codes consists of numbers from 1 to 8. Bold frames mark two categories of

numbers of degenerations. The two right columns show quantities of amino acids (ΣAA), which have numbers of degeneracy from 1 to 3 (low-degenerate amino acids) and which have numbers of degeneracy from 4 to 8 (high-degenerate amino acids).

from 4 to 8). The author revealed that each from these categories has a permanent quantity - 12 and 8 - of corresponding amino acids for all genetic codes practically [13]. The phenomenological rule № 1 can be formulated on this basis: in genetic codes, the set of 20 amino acids contains two opposite subsets: the first subset consists of 12 low-degeneracy amino acids (with their numbers of degeneracy from 1 to 3), and the second subset consists of 8 high-degeneracy amino acids (with their numbers of degeneracy from 4 to 8).

Rule №1 has small exceptions: two codes from a set of 17 ones have ratios 11:9 and 13:7 between low-degeneracy and high-degeneracy amino acids (see Fig. 8). These non-typical ratios are the nearest integers ones, which encircle the canonical ratio “12:8” from the contrary sides of numerical axis. They demonstrate additionally the main role of the ratio 12:8 as that centre, around which minimal numerical fluctuations are realized.

The following rule №2 was revealed as well: if a triplet encodes different amino acids in different genetic codes, then these amino acids belong to the same canonical subset of amino acids (see the previous paragraph). In other words, it is practically forbidden for those triplets, which encode amino acids from one canonical subset of degeneracy, to pass during biological evolution into the group of triplets, which encode amino acids from another canonical subset.

A single exception from this rule is the triplet UAG, which can encode amino acids Leu or Gln from different canonical subsets. The rule says nothing about stop-codons, and so it does not consider those evolutionary cases, when triplets, which encode stop-codons (or amino acids) in one genetic code, begin to encode amino acids (or stop-codons respectively) in another code.

Described phenomenological rules testify that two independent branches of evolution of genetic code at billions biological species exist: one branch – for canonical set of high-degeneracy amino acids, and another branch - for canonical set of low-degeneracy amino acids. These evolutionary branches within the consolidated code system can be compared with a parallel evolution of male and female organisms within a frame of one biological species, or with a parallel evolution of consonants and vowels in a language. It reveals simultaneously that, instead of unified set of 20 amino acids, the nature realizes association of two very different subsets of 8 and 12 amino acids (the way of Yin-Yang categories).

Remark 1: number 24 is the least divisible integer for numbers 8 and 12.

Remark 2: main numbers of degeneracy of amino acids in all genetic codes are 1, 2, 3, 4, 6, 8; all of them are divisors of number 24 (four genetic codes have a single amino acids with its number of degeneracy 5 or 7; a rate of each of these non-typical numbers of degeneracy is equal to 0,88 %).

Prediction: unknown genetic code exists, where one amino acid has a number of degeneracy 12 (it is a divisor of 24 as well) [13].

In connection with these remarks 1 and 2, number 24 can be considered as a hidden constant of coordination among numbers of degeneracy in genetic codes. Number 24 is well known in biology for a long time due to phenomena of chronobiology. Chronomedicine, which has successful ancient history, declares that, during a day-night period, physiological systems of an organism have regular changes of physiological activity and passivity, which are correlated with a division of this daily cycle into 24 equal parts. These chronobiological phenomena are connected with a cyclic arrival of solar energy on earth and with those mechanisms of photosynthesis, which are provided a cyclic production of living substance itself in autotrophic organisms. In chronobiology, number 24 is not an arbitrary division of daily period, but a phenomenological constant of coordination between physiological cycles and a daily period of solar energy arrival. This knowledge is one of the main parts of Oriental medicine: acupuncture, pulse diagnostics of Tibetan medicine, etc.

On the basis of rules 1 and 2 and a revealing of hidden constant 24 in a genetic codes phenomenology, the author proposed a chronobiological conception of genetic codes structure [13]. According to this conception, genetic code structures relate with 24-hours period of solar energy arrival and with the most basic mechanisms of photosynthesis. Physiological cycles in organisms are provided by not only self-organizing cyclic behavior of protein systems (as usually considered), but they are originated from structures of genetic code systems. Genetic code systems encode not only sequence of amino acids in proteins, but they determine important aspects of chronobiological phenomena also.

Additional phenomenological rules about numbers of degeneracy were published in [13].

3. Genetic Tetra-sets

3.1. *Segregations in Genetic Codes and Parallelisms with Mendel's Laws*

What types of laws underlie genetic codes? Attempts of answering on this question have led the author to put forward a working hypothesis about a

realization in genetic system special biological principles of segregations. These principles have a formal relationship with Mendel's laws, which are sometimes named in literature as the main laws of biology.

Let us define a concept of biological tetra-set. Tetra-set is a such set of elements with biologically significant different attributes (traits), which is segregated in accordance with these different attributes into two or three subsets, quantitative compositions of which relate to each other in "canonical" ratios 1:3 or 2:2 or 4:0 or 1:2:1. The author paid attention to wide spread of tetra-sets in genetic code system [1,10]. A few examples are demonstrated below.

A transfer of genetic information to a next generation occurs through gametal cells. Gametes are produced as a result of meiosis, which is a special division of cells. E. Schrodinger in his famous book [14] noted specially a mysterious fact of reproduction of a typical tetra-set of four cells as a result of meiosis for very different biological species. In the case of reproduction of female gamete (ovogenesis), four non-equal cells are produced; one of them differ essentially from the other three cells by its enlarged size and mass. Other three puny cells, which are united in biology by their mutual name "reducing corpuscles", are degenerated in the course of time. So, a biological tetra-set with the ratio of segregation 1:3 is generated in the result of ovogenesis. In the case of meiosis with a formation of male gametes (spermatogenesis), all four gametal cells are equal to each other. One can say that a tetra-set with the ratio 4:0 is formed in this case.

On the absolutely another end of biological phenomena, a creator of analytic psychology C. Jung, studying archetypes of human consciousness, discovered a similar universal archetype: "Quaternary set. A universal archetype, which is a logical prerequisite of every entire judgment. It frequently has a structure 3+1, one of the elements of which takes a special place or possesses a different nature. Just the fourth element, adding to the others, makes them a single whole, that symbolizes universal set" [15, p. 140].

In the field of genetic code itself, the four-letter alphabet of genetic code gives an example of a tetra-set with a segregation ratio 3:1. Really, three letters A, C, G form one subset according to an attribute of their equivalence for molecules DNA and RNA. The fourth letter U forms an opposite subset because it has an opposite attribute: letter U (uracil) works in DNA, but the letter T (thymine) replaces it in RNA.

Black-and-white mosaics on Fig. 7 gives another example. Each quadrant of the matrix $P^{(3)}$ on Fig. 7 has three subfamilies of NN-triplets, marked by one color, and one subset of NN-triplets, marked by an opposite color. In other

words, each family of N-triplets, disposed in such quadrant, is a tetra-set of four subfamilies of NN-triplets with a segregation ratio 3:1.

Each subfamily of NN-triplets is a tetra-set as well, because it can be characterized in different genetic codes by one of the canonical segregation rations – 1:3 or 2:2 or 4:0 or 1:2:1. But a segregation of a subfamily of NN-triplets into subsets with a different code meaning for each triplet is forbidden in biological evolution: 17 genetic codes have no one subfamily of NN-triplets with a non-canonical segregation ratio 1:1:1:1. A list of such examples can be continued [10].

Mendelian genetics considers a crossing of biological forms, which differ by alternative traits. Such crossing reproduces groups of organisms with one or another alternative trait or with their mixture. According to Mendel's laws, in first and second generations of such hybrids, these groups form tetra-sets with canonical segregation ratios 3:1 or 2:2 or 4:0 or 1:2:1. The author paid attention to these structural parallelisms in tetra-sets phenomena between genetic codes and Mendelian genetics to develop a theory of tetra-sets in genetic code systems [10]. This theory explains certain cases of tetra-sets in genetic code system from a viewpoint of a concept of "pro-alleles", which have dominant or recessive sense by analogy with a concept of alleles in Mendelian genetics. Note that modern understanding of Mendel's laws is related with a splitting of pairs of homological chromosomes. Mendel's laws are not destined for levels of microstructures, which are less than chromosomes. Therefore segregations regularities of elements sets of genetic codes, described in this paper, are independent of Mendel's laws and cannot be deduced from them.

3.2. *Genetic Sequences as Tetra-Sets*

Can analogous principles of tetra-segregations influence peculiarities of composition of genetic compositions (texts), which define structures of proteins? For instance, can genetic sequences from a certain wide class be segregated by a pair of binary-opposite attributes into subsets of triplets with a ratio 3:1 between their quantitative compositions? The author made first investigations of it on separate examples with positive results.

The first tested example is a sequence of triplets in DNA [16], which encodes the simplest protein insulin. This protein has α-chain and β-chain and is encoded by the sequence of 51 triplets in sum (in view of a start-codon ATG, a total sum of triplets is equal to 52=4x13):

- α-chain (21 triplets with an indication of a current number before each triplet): 1GGC→2ATC→3GTT→4GAA→5CAG→6TGT→7TGC→8ACT→9TCT→10ATC→ →11TGC→12TCT→13CTT→14TAC→15CAG→16CTT→17GAG →18AAC→19TAC→ →20TGT→21AAC;
- β-chain (30 triplets with an indication of a current number before each triplet): 1TTC→2GTC→3AAT→4CAG→5CAC→6CTT→ 7TGT→8GGT→9TCT→ →10CAC→ 11CTC→12GTT→13GAA→14GCT→15TTG→16TAC→→17CTT→ 18GTT→ →19TGC→20GGT→21GAA→22CGT→23GGT→24TTC→ 25TTC→26TAC→27ACT→ →28CCT→29AAG→30ACT .

It is easy to calculate that this set of 51 triplets is divided into two subsets. One of these subsets contains 13 G-triplets and the second subset contains 38 A-, C- and U-triplets (in view of a start-codon ATG, a total sum of triplets in the second subset is equal to 39). The quantitative ratio between these two subsets is equal 1:3. Therefore one can say about the insulin sequence as about a tetra-set with the ratio of segregation 1:3.

For second example, the author took a genetic sequence EG13077 at Escherichia from the "EcoGene WEB Site" in a random way:

ATG-GTT-CAG-**AAG-**CCC-CTC-**ATT-AAG-**CAG-GGA-TAT-TCA-CTG-GCA-GAG-GAA-**ATA-**GCC-**AAC-AGC-**GTC-**AGT-**CAC-GGC-**ATT-**GGG-TTG-GTG-TTT-GGT-**ATC-**GTT-GGG-CTG-GTG-TTG-CTA-CTG-GTT-CAG-GCG-GTG-GAT-CTT-**AAT-**GCC-**AGC-**GCC-**ACA-**GCG-**ATA-ACC-AGC-**TAC-**AGC-**CTC-TAT-GGC-GGC-**AGT-ATG-ATC-**CTG-CTG-TTC-CTC-GCT-TCG-**ACG-**CTC-TAT-CAC-GCC-**ATT-**CCT-CAT-CAA-CGG-GCA-**AAA-ATG-**TGG-CTG-**AAG-AAA-**TTT-GAC-CAT-TGC-GCT-**ATT-**TAC-CTG-TTG-**ATT-**GCC-GGA-**ACC-**TAC-**ACG-**CCG-TTT-TTG-CTG-GTG-GGG-CTG-GAT-TCT-CCG-TTA-GCG-CGC-GGG-TTG-**ATG-ATT-**GTT-**ATC-**TGG-**AGC-**CTG-GCA-TTG-CTG-GGT-**ATT-**CTG-TTT-**AAA-**CTG-**ACC-ATC-**GCG-CAC-CGA-TTC-**AAA-ATT-**TTA-TCT-CTG-GTG-**ACC-**TAT-CTG-GCG-**ATG-**GGC-TGG-CTG-TCG-CTG-GTG-GTA-**ATT-**TAT-GAA-**ATG-**GCA-GTT-**AAG-**CTC-GCG-GCG-GGC-**AGC-**GTT-**ACC-**TTA-CTG-GCG-GTA-GGC-GGC-GTG-GTT-TAT-TCG-CTC-GGG-GTG-**ATT-**TTC-TAC-GTC-TGC-**AAA-**CGC-**ATT-**CCA-TAC-**AAC-**CAT-GCC-**ATC-**TGG-CAC-GGC-TTC-GTG-CTC-GGC-GGT-**AGT-**GTG-TGC-CAC-TTT-CTG-GCG-**ATC-**TAT-TTG-TAT-**ATT-**GGG-CAG-GCG-TAA

This sequence contains a set of 220 triplets with a subset of 55 A-triplets (they are marked out by us specially to demonstrate their non-regular or random disposition in the sequence) and with a second subset of other 165 triplets in sum (50 C-triplets, 69 T-triplets, 46 G-triplets). The ratio between these two subsets is equal 55:165=1:3. In other words, one can say that this genetic sequence is a tetra-set (on basis of the four families of N-triplets) with a segregation ratio 1:3.

Of course, individual examples don't permit to determine a degree of generality of such principles of organization for all set of genetic sequences. This analysis of genetic sequences from the viewpoint of tetra-sets should be continued, including an application of an idea of dominant and recessive pro-alleles [10]. In the case of a wide distribution of similar principles of tetra-sets in genetic sequences, a quantity of admissible variants of sequences is limited by these principles additionally. Perhaps, an existence of these principles is one of the reasons of the famous fact, that not all thinkable variants of sequences encode proteins.

3.3. *Why 20 Amino Acids?*

Many attempts to answer on this fundamental question are known. The author proposes a new answer: the set 20 amino acids are presented in genetic code in that reason, that it is formed by two alternative subsets of 8 and 12 amino acids [10,13]. In this way, the initial question is turned into another deeper question: why the groups of 8 and 12 amino acids are existed in the united genetic set?

A possible answer on this new fundamental question is related with the revealed fact, that these two groups constitute two branches of independent evolution within genetic code. Just a biological mechanism of tetra-segregations can be responsible for the realization of such two groups of amino acids in a correlation with an evident tetra-presentation of these numbers: 8 = 4x2 and 12 = 4x3. In this tetra-presentation, number 8 contains number 2 in a role of modular block, and number 12 contains number 3 in an analogous role. In our opinion, a principle of tetra-sets is one of the most fundamental ones in biology.

Our investigations revealed that the considered pair of alternative attributes ("high-degenerative and low-degenerative") is not a single pair for a division of the set of 20 amino acids into subsets of 8 and 12 amino acids. Genetic code is constructed so, that such division is a typical one for many other pairs of real binary-opposite attributes, in relation to which such division is considered. Similar divisions, but with different groups of 8 and 12 amino acids, are given by such binary-opposite attributes as "complementary-uncomplementary" amino acids, "high-carbon or low-carbon" amino acids, "hydrophobic or hydrophilic" amino acids, "eightfold or non-eightfold quantity of protons" in amino acids [1,3,10]. In our opinion, this phenomenon of a multiversion realization of groups of 8 and 12 amino acids is related with a phenomenon of parallel information canals in organisms.

Figure 9 demonstrates a confirmative example with groups of complementary and uncomplimentary amino acids: a realization of groups of 8 and 12 amino acids is provided by a principle of their tetra-construction from typical modular blocks with 2 units and with 3 units (8=4x2 and 12=4x3). Complementary amino acids are those, which are encoded by groups of codons and their anti-codons. One can see from Fig. 7, that 8 amino acids form the four pairs of complementary amino acids (Pro-Gly, Arg-Ala, Lys-Phe, Met-Tyr), but other 12 amino acids are uncomplementary ones. The subset of 8 complementary amino acids is divided into those four pairs (or four modular blocks with two amino acids), each of which is encoded by triplets from a separate family of N-triplets. And the subset of 12 non-complementary amino acids is divided into those four pairs (or four modular blocks with three amino acids), each of which is encoded by triplets from a separate family of N-triplets (in the case of each of the amino acids Ser and Leu, we take here into account those family of N-triplets, all four triplets of which encode it in Fig. 7). So, each of four families of N-triplets encodes 2 complementary amino acids and 3 uncomplementary ones.

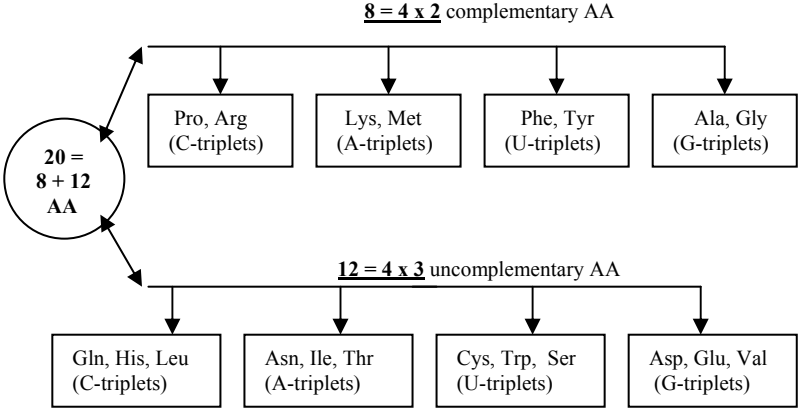

Fig. 9. An example of a presentation of the set of 20 amino acids (AA) as a sum of two subsets with 8 = 4x2 and 12 = 4x3 ones of complementary and uncomplementary types according to Fig. 7. These tetra-subsets have four pairs of amino acids of the complementary type and four triples of the uncomplementary amino acids correspondingly. An appropriate family of N-triplets, which encode the amino acids, is shown in each case.

4. The Stent's Hypothesis.

G. Stent published in his book [17, p. 64] an interesting consideration about a possible connection between genetic code structures and a symbolical system

of the Ancient Chinese "Book of Changes"". This consideration can be named as the Stent's hypothesis. A few authors supported it later. For example, the Nobel prize winner in molecular genetics F. Jacob wrote as well: "Perhaps, for seizing of relations between genetics and language it would be necessary to study namely through the Ancient Chinese "Book of Changes" [18]. So, in the field of molecular genetics, a special problem of investigation of parallelisms between genetic structures and the ancient symbolical system exists. The matrix form of a presentation of genetic polyplets, described above, gave new additional materials to this problem.

It is known that this ancient symbolical system is based on four basic objects – Old Yang, Young Yang, Old Yin and Young Yin (for example, see [19]).

Old Yang (*1*1)	Young Yin (*1*0)
Young Yang (*0*1)	Old Yin (*0*0)

; $P_C = \begin{vmatrix} \underline{1}1 & \underline{1}0 \\ \underline{0}1 & \underline{0}0 \end{vmatrix}$

$$P^{(3)}_C = \begin{vmatrix} \underline{111}111 & \underline{111}110 & \underline{111}101 & \underline{111}100 & \underline{111}011 & \underline{111}010 & \underline{111}001 & \underline{111}000 \\ \underline{110}111 & \underline{110}110 & \underline{110}101 & \underline{110}100 & \underline{110}011 & \underline{110}010 & \underline{110}001 & \underline{110}000 \\ \underline{101}111 & \underline{101}110 & \underline{101}101 & \underline{101}100 & \underline{101}011 & \underline{101}010 & \underline{101}001 & \underline{101}000 \\ \underline{100}111 & \underline{100}110 & \underline{100}101 & \underline{100}100 & \underline{100}011 & \underline{100}010 & \underline{100}001 & \underline{100}000 \\ \underline{011}111 & \underline{011}110 & \underline{011}101 & \underline{011}100 & \underline{011}011 & \underline{011}010 & \underline{011}001 & \underline{011}000 \\ \underline{010}111 & \underline{010}110 & \underline{010}101 & \underline{010}100 & \underline{010}011 & \underline{010}010 & \underline{010}001 & \underline{010}000 \\ \underline{001}111 & \underline{001}110 & \underline{001}101 & \underline{001}100 & \underline{001}011 & \underline{001}010 & \underline{001}001 & \underline{001}000 \\ \underline{000}111 & \underline{000}110 & \underline{000}101 & \underline{000}100 & \underline{000}011 & \underline{000}010 & \underline{000}001 & \underline{000}000 \end{vmatrix}$$

Fig. 10. Basic tables of "Book of Changes" for four digrams and for 64 hexagrams in Fu-Xi's order. The genomatrices P_C and $P^{(3)}_C$ coincide with them.

Each of these objects is symbolized by a digram of respective binary symbols in a form of a solid line or a broken line (Fig. 10). These four objects are presented in a form of a matrix (2x2). Let us return to Fig. 1 and consider matrices P and $P^{(3)}$. Each element of them has a binary coordinate number, which is an association of binary numbers of its row and column. If one replaces each matrix element by its binary coordinate number, "binary-coordinate" genomatrices P_C and $P^{(3)}_C$ are appeared (Fig. 10). It is astonishing that both of these binary genomatrices P_C and $P^{(3)}_C$ are formally well known already for a few thousand years from the "Book of Changes". Really, the genomatrix P_C coincides with the Chinese matrix of four binary digrams (Fig. 10, top). The genomatrix $P^{(3)}_C$ coincides with a famous matrix of 64 binary hexagrams, disposed according to Fu-Xi's order. Each hexagram in the

Chinese matrix is arranged from two independent trigrams, which symbolize its row and its column (these pair of independent trigrams within one hexagram were named historically as “trigram of earth and trigram of heaven” or “trigram of space and trigram of time”, *etc*. [20, p.101]). But each element of the similar genomatrix $P^{(3)}_C$ is also a binary hexagram with two independent trigrams, originated from two different binary sub-alphabets of the genetic alphabet [1,3,5,9]. Ancient Chineses declared that both of these matrices P_C and $P^{(3)}_C$ are universal natural archetypes. It should be emphasized, that the author didn’t use the tables of “Book of Changes” in his academic researches of genetic code, but as a result he collided with these ancient tables unexpectedly.

Young and Old Yin and Yang are symbolized by numbers 6, 7, 8 and 9 in the system of “Book of Changes” [Sczhutskii, 1997, p. 22, 522]. Exactly these numbers characterize a quantity of protons in chemical elements, form which nitrogenous bases of genetic alphabet are constructed: carbon C has 6 protons, nitrogen N – 7 protons, oxygen O – 8 protons, amids NH_2 – 9 protons [3,5].

Numbers 2 and 3, which are realized in quantities of hydrogen bonds of complementary genetic letters, are basic numbers in ancient Chinese numerical systems [21, p.15]. The pair of numbers 8 and 12, related with described two subsets of amino acids, has an important meaning in this Chinese system: numbers 8 and 12 “are a typical measure of alternative divisions of space-time in Chinese chronotopogrames… In their stereometric representations, numbers 8 and 12 are interrelated as well in the role of basic parameters of a cube and an octahedron. A cube has 8 vertices and 12 edges, an octahedron has 8 facets and 12 edges. Both of these regular polyhedrons were basic cosmography models for ancient Chinese wise men” [21, p. 39-40].

One else example is related with a published numerical presentation of the genomatrix $P^{(3)}$ from the viewpoint of quantities of atoms in rings of nitrogenous bases: A=G=9, C=U=T=6 [3, p. 61]. In such numerical genomatrix, sums of numbers in each of eight columns, symbolized by eight trigrams, form a numerical series 144, 168, 168, 192, 168, 192, 192, 216. The author was informed recently, that this series for a set of the same eight trigrams was known for a long time in considered Chinese system and it had a name “series of countable wands” [22, p. 269-292]. A limited volume of this article does not permit to continue this list of new impressive parallelisms.

Acknowledgements

The author is grateful to Professors K. Frolov, M. He, G. Darvas for their help. This work is presented in the frame of cooperation between Russian and Hungarian Academies of Sciences.

References

1. S.V. Petoukhov, *Symmetry in Genetic Information* (guest ed. S. Petoukhov), a special issue of *"Symmetry: Culture and Science",* **12 (3-4)**, 255 (2001).
2. B.G. Konopelichenko, Y.B. Rumer, *DAN*, **223 (2)**, 471 (1975) (in Russian).
3. S.V. Petoukhov, *The Biperiodic Table of Genetic Code and Number of Protons,* (MKC, 2001) (in Russian).
4. H.G. Wittmann, *Die Naturwissenschaften,* **48 (24)**, 55 (1961).
5. S.V. Petoukhov, *Symmetry: Culture and Science,* **10 (3-4)**, 211 (1999).
6. M. He, *Symmetry in Genetic Information* (guest ed. S. Petoukhov), a special issue of *"Symmetry: Culture and Science",* **12 (3-4)**, 307 (2001).
7. M. He, *Bull. Mathem. Biology*, **66 (5),** 965 (2004).
8. M. He M., Petoukhov S., Ricci P., *Bulletin for Mathematical biology*, (2004) (in press).
9. S.V. Petoukhov, *Symmetry: Culture and Science,* **14 (1-4)**, 90 (2003).
10. S.V. Petoukhov, *Supplement to the book:* A.V. Shubnikov, V.A. Koptsik, *Symmetry in Science and Art,* 3d edition (ICI, 2004) (in Russian).
11. S.V. Petoukhov, *Questions of Metaphysics*,(ed. Y.S.Vladimirov), (Binom, 2005) (in print in Russia)
12. M. D. Frank-Kamenetskey, *The most principal molecule,* (Nauka, 1988).
13. S.V. Petoukhov, *Symmetry in Genetic Information* (guest ed. S.Petoukhov), a special issue of *"Symmetry: Culture and Science",* **12 (3-4)**, 275 (2001).
14. E. Schrodinger, *What is Life?,*(Atomizdat, 1972) (in Russian).
15. G. Wehr, *C.Jung in Selbstzeugnissen und Bilddokumenten,* (Rowohlt, 1969).
16. S.G. Inge-Vechtomov, *Lectures on molecular genetics,* (Nauka, 1983).
17. G.S. Stent, *The Coming of the Golden Age,* (Natural History Press, 1969).
18. F. Jacob, *Critique,* **XXX (322),** 197, (1974).
19. Jou Tsung Hwa, *The Tao of I Ching*, (Tai Chi Foundation,1984).
20. Y.K. Tshutskey, *The Chinese classical "Book of Changes",*(Vostochnaia Literatura, 1997) (in Russian).
21. A.I. Kobzev, *Studies on symbols and numbers in Chinese classical philosophy*, (Vostochnaia Literatura, 1997) (in Russian).
22. B.B. Vinogrodskey, *Taoism's practices of immortality,* (Sophia, 2002) (in Russian).