Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 May 2019 doi:10.20944/preprints201904.0011.v2

Nucleotide epi-chains and new nucleotide probability rules in long DNA sequences

Sergey V. Petoukhov[1]

[1]-Mechanical Engineering Research Institute of Russian Academy of Sciences.
Russia, 101990, Moscow, M. Kharitonievskiy pereulok, 4,
http://eng.imash.ru/, info@imash.ru

Comment: Some elements of this article were presented by the author in his keynote speeches at the following conferences: the International Belgrade Bioinformatics Conference 2018 (Belgrade, Serbia, 18-22 June, 2018, http://belbi.bg.ac.rs/); the Second International Conference "Artificial Intelligence, Medical Engineering, Education" (Moscow, Russia, 6-8 October 2018, http://www.ruscnconf.org/aimee2018/index.html); the First Congress of the International Society of Natural Medicine (Samorin, Slovak Republic, 12-14 October 2018, https://www.simn.info/congress-of-natural-medicine-2018/); the Second International Conference on Computer Science, Enginnering and Education Applications (Kiev, Ukraine, 27-28 January 2019, http://www.uacnconf.org/iccseea2019/index.html).

Abstract. One of creators of quantum mechanics P. Jordan in his work on quantum biology claimed that life's missing laws were the rules of chance and probability of the quantum world. The article presents author's results of studying frequencies (or probabilities) of nucleotides on so-called epi-chains of long DNA sequences of various eukaryotic and prokaryotic genomes. DNA epi-chains are algorithmically constructed subsequencies of DNA nucleotide sequences. According to the algorithm of construction of any epi-chain of the order n, the epi-chain is such nucleotide subsequence, in which the numerations of adjacent nucleotides differ by natural number n (n = 1, 2, 3, 4,…). Correspondingly each epi-chain of order $n \geq 2$ contains n times less nucleotides than the original DNA sequence. The presented results unexpectedly discover that in long single-stranded and double-stranded DNA of any tested genome its DNA epi-chains of different orders n (values n are not too large) have practically identical frequencies (or probabilities) of each kind of nucleotides. These data allow considering DNA as a regular rich set of epi-chains, which can play a certain role in genetic and epigenetic phenomena as the author belives. Appropriate rules of nucleotide frequencies on epi-chains of long DNA sequences are formulated for further their tests on a wider set of genomes. These results testify on existence of long-range coherence in long DNA and remind the Fröhlich's theory of long-range coherence in biological systems. The phenomenological data are discussed from different standpoints: the DNA double helices and helical antennas with circular polarizations of electromagnetic waves; relations with the Fröhlich's theory; numerical analysis of DNA epi-chains under binary representations of nucleotides. Results are useful for developing quantum and algebraic biology.

Keywords: DNA sequence, helix, nucleotide frequencies, DNA epi-chains, helical antennas, Fröhlich's theory, long-range coherence, epigenetics, quantum biology, binary representation.

Content

 doi:10.20944/preprints201904.0011.v2

1. Introduction

One of creators of quantum mechanics P. Jordan, credited with the authoring the first work on quantum biology, postulated the following. The mechanisms of living organisms are associated with what he referred to as his 'amplifier theory', based on Bohr's notion of the 'irreversible act of amplification', required to bring the fuzzy quantum reality into sharp focus by 'observing' it [Jordan, 1932]. Jordan claimed that «*life's missing laws were the rules of chance and probability (the indeterminism) of the quantum world that were somehow scaled up inside living organisms*» [McFadden, Al-Khalili, 2018]. It is these laws of chance and probability, postulated by Jordan, that the author is looking for in his analysis of percentage (or probabilistic, or frequencies) characteristics of long DNA sequences of hydrogen bonds and nitrogenous bases [Petoukhov, 2018a,b, 2019a,b; Petoukhov, Petukhova, Svirin, 2018; Petoukhov, Svirin, 2012, 2018; Darvas, 2018]. Long DNA sequences are a very important and convenient for such computerized research of possible biological rules of frequenices (or percentages, or probabilities), which in Jordan's day was impossible. The main idea of all quantum mechanics is that everything in the world of appropriate objects is described only as probabilistic. Correspondingly the author has undertaken study of nucleotide frequencies (probabilities) in long DNA of many species on the basis of a new author's approach.

Regarding frequencies of nucleotides in DNA strands, the so-called Chargaff's rules have been known since the middle of the last century [Chargaff, Lipshitz, Green,1952; Chargaff, 1971]. In DNA molecules, genetic information is written in very long sequential texts using only 4 letters: adenine A, cytosine C, guanine G, thymine T. For example, the human genome consists of several billion such letters. The first Chargaff's rule holds that in double-stranded DNA the amount of guanine (G) is equal to the amount of cytosine (C) and the amount of adenine (A) is equal to the amount of thymine (T). This rule was theoretically confirmed by the double-helix model of the DNA [Watson, Crick, 1953].

The second Chargaff's rule states that the same parity is approximately valid (%C ≈ %G and %A ≈ %T) for each of the two long DNA strands alone (the term "long" usually refers to DNA sequences containing ≥ 100000 nucleotides). According to [Albrecht-Buehler, 2006], this rule applies to the eukaryotic chromosomes, the bacterial chromosomes, the double stranded DNA viral genomes, and the archaeal chromosomes provided they are long enough. Some our works were also devoted to study frequencies of nucleotides and hydrogen bonds in long DNA sequences [Petoukhov, 2018a,b, 2019; Petoukhov, Petukhova, Svirin, 2018; Petoukhov, Svirin, 2018; Darvas, 2018].

This article presents, first of all, the author's results of calculation of frequencies (or probabilities) of nucleotides located in algorithmically sparsed sequences inside long DNA of many eukaryotic and prokaryotic organisms. These numeric results testify in favor of existence of some general rules of nucleotide frequencies in long DNA sequences and moreover they lead to a new important notion of "DNA epi-chains" defined below. These rules of long DNA epi-chains belong to the field of quantum biology and they testify on existence of quantum-mechanical long-range coherence in long DNA and remind the Fröhlich's theory of quantum long-range coherence or collective quantum effects in biological systems. The presented results reveal new connections of organizational principles of the genetic system with formalisms of quantum mechanics and quantum informatics dealing with vectors of probabilities, unitary matrices, etc. Since quantum mechanics operates with notions of probabilities, the term "probabilities" is used in the title of the article though inside the article the author mainly uses similar terms "frequencies" and "percentages" as more familiar to a wide range of readers.

The presented results also reveal that the arrangement of nucleotides in the DNA double-helix model [Watson, Crick, 1953] is not arbitrary but are connected with multi-helices of epi-chains: any long DNA double-helix can be considered as a superposition (or a hybrid) of many helical epi-chains and so the DNA double helix has regular multi-helical character from the standpoint of results described in the article. The

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 May 2019 doi:10.20944/preprints201904.0011.v2

author believes that the described numeric rules of epi-chains are related to epigenetics and some other biological themes and model approaches, which are briefly discussed below.

2. The notion of DNA epi-chains of nucleotides

The DNA double helix is a long discrete polyatomic construction, whose parts are connected into the whole stabilized structure. In this construction, each of two complementary strands contains a discrete chain of nucleotides, which carry genetic information and are located along helix-like trajectorie (Fig. 1, left).

The article considers additional types of DNA nucleotide chains, which are termed as “single-stranded DNA epi-chains” and “double-stranded DNA epi-chains” of different orders n (n = 1, 2, 3, 4, … are natural numbers). The attention paid to these types of nucleotide chains in DNA is caused by the author's results on the approximate equality of frequencies of nucleotides in these different chains.

In the described study, a chain of nucleotides, which continuously follow each other on one strand of DNA, will be called the single-stranded DNA epi-chain of the first order. It is precisely these nucleotide epi-chains are used in the mentioned second Chargaff's rule. The consecutive nucleotides in a single-stranded DNA (that is in the single-stranded DNA epi-chain of the first order) can be sequentially numerated with the natural numbers 1, 2, 3, 4, 5, ..., starting with any randomly chosen nucleotide (Fig. 1, II). Below we will deal with numerated nucleotides, each of which carries conditionally its sequence number along the nucleotide sequence in a considered DNA strand.

In contrast to the traditional attention of researchers to two strands of the DNA double helix, this article focuses on epi-chains, the concept of which is new for molecular genetics. We will consider two following types of DNA nucleotide epi-chains: the first type is connected with single-stranded DNA whose epi-chains are termed conditionally «single-stranded epi-chains» (Fig. 1, II-VII); the second type is connected with double-stranded DNA whose epi-chains are termed «double-stranded epi-chains» (Fig. 2).

By definition, a nucleotide single-stranded epi-chain is such single-stranded DNA nucleotide sequence, in which the numerations of adjacent nucleotides differ by natural numbers n = 1, 2, 3, 4,..., where n defines the "order of epi-chains". For example single-stranded epi-chains of the second order are two nucleotide sequences $N_{2/1}$ and $N_{2/2}$ whose nucleotides have their sequence numbers differing by n = 2: the epi-chain $N_{2/1}$ contains nucleotides with odd numerations 1, 3, 5, 7,… (Fig. 1, III) and the epi-chain $N_{2/2}$ contains nucleotides with even numerations 2, 4, 6, 8, ... (Fig. 1, IV). By analogy, single-stranded DNA epi-chains of the third order are those three nucleotide sequences $N_{3/1}$, $N_{3/2}$ and $N_{3/3}$, each of which has sequence numbers differing by n = 3: these single-stranded epi-chains contain nucleotides with numerations 1, 4, 7, 10, ... or 2, 5, 8, 11, ... or 3, 6, 9, 12, ... , respectively (Fig. 1, V-VII). Any single-stranded epi-chain of the first order coincides with the nucleotide sequence of the appropriate separate strand in the DNA double helix (in contrast to the case of double-stranded epi-chains defined below).

The nucleotide quantity (or density) on a single-stranded DNA epi-chain of n-th order (if n = 2, 3, 4, ...) is reduced by "n" times in comparison with the nucleotide quantity in the corresponding DNA strand (that is in the single-stranded epi-chain of the order n = 1). In other words, the DNA epi-chains are sparsed DNA strand in relation to its nucleotide sequence. The term "epi-chain" uses the Ancient Greek prefix epi- (ἐπι- "over, outside of, around"), which implies features that are "on top of" or "in addition to" the traditional DNA double helix. In any DNA strand, each nucleotide belongs to many epi-chains having different orders n. The used symbol N in the designation of DNA epi-chains corresponds to the first letter of the word “nucleotides”. In the symbolic designation $N_{n/m}$ of single-stranded DNA epi-chains, the numerator "n" in the chain index indicates the order of the epi-chain, and the denominator "m" is the numeration of the first nucleotide of this epi-chain along the initial DNA strand (Fig. 1). For example, the symbol $N_{3/2}$ refers to a single-stranded epi-chain of the third order with the first nucleotide having the number 2 in the nucleotide numeration along the corresponding DNA strand: 2-5-8-11-14-… (Fig. 1, VI).

In the considered DNA epi-chains, their orders n are closely related to angles of rotation between adjacent nucleotides (in the B-form of the DNA double helix, the angle α of rotation between adjacent nucleotides is approximately 36^0). More precisely, in these epi-chains of the order n, the angle of rotation between adjacent nucleotides is equal to $n\alpha$. For example, in the epi-chains of orders 2 and 3 their angles of rotation are equal to 2α and 3α respectively (Fig. 1). These angular values $n\alpha$ can be used for definitions of orders of the epi-chains. Now let us turn to double-stranded DNA epi-chains.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 May 2019 doi:10.20944/preprints201904.0011.v2

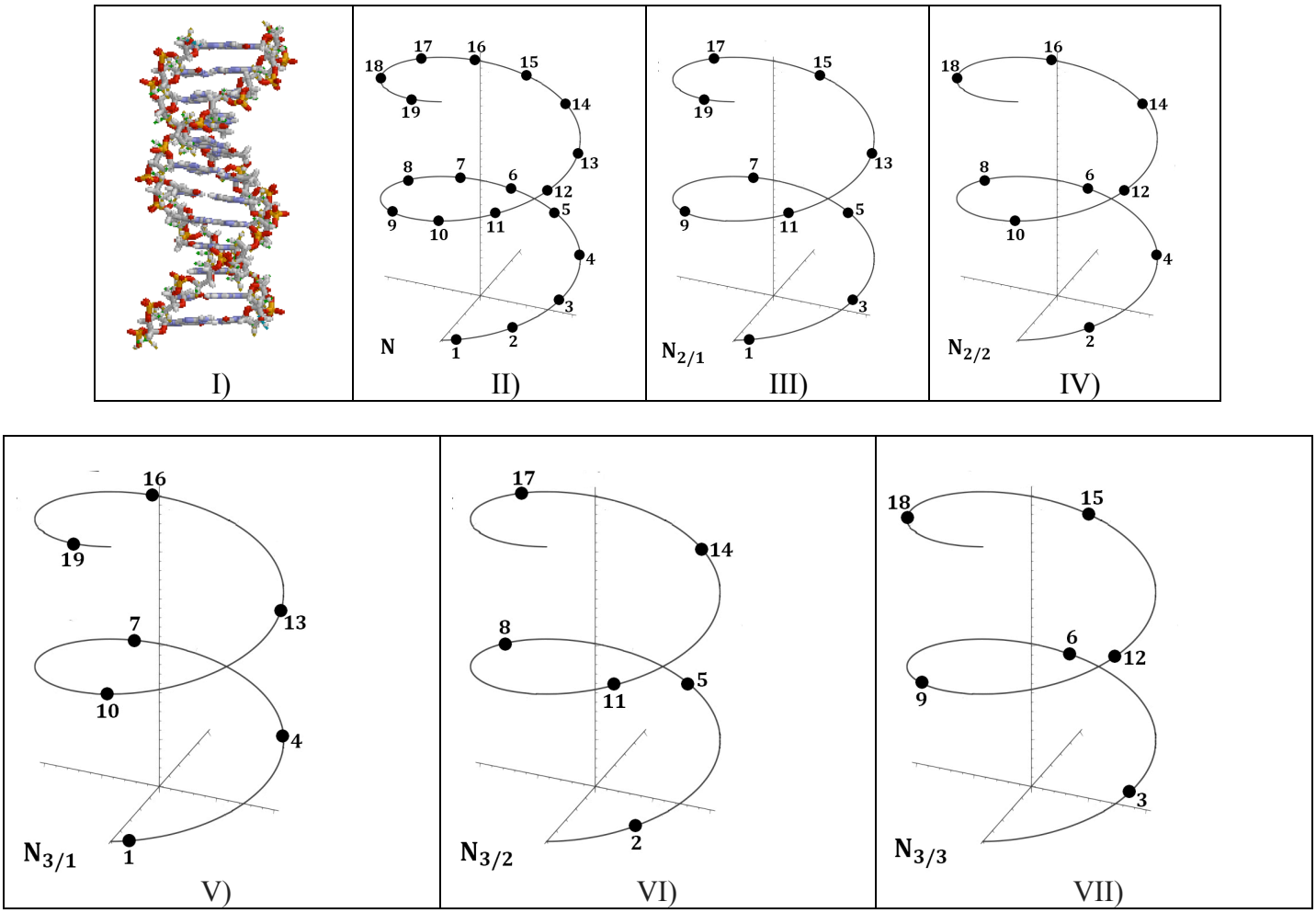


Fig. 1. The schematic representation of DNA and a few epi-chains of nucleotides on the separate DNA strand. I) The DNA double helix (from https://ru.wikipedia.org/wiki/Дезоксирибонуклеиновая_кислота). II) A sequence N of numerated nucleotides of single-stranded DNA termed as a DNA code-chain; III-IV) Sequences $N_{2/1}$ and $N_{2/2}$ represent two single-stranded epi-chains of the second order regarding the same code-chain N; V-VII) Sequences $N_{3/1}$, $N_{3/2}$ and $N_{3/3}$ represent three single-stranded epi-chains of the third order regarding the same code-chain N (explanations in the text). Black circles at helices represent conditionally nucleotides.

By definition, a double-stranded DNA epi-chain is a modified single-stranded epi-chain, in which every second nucleotide is replaced by its complementary belonged to the opposite DNA strand (systematic jumps from one DNA strand to another are realised) (Fig. 2). For example, under the replacement of every second nucleotide in the single-stranded epi-chains CGAATGCTAGG… by its complementary (C↔G and A↔T), the following double-stranded DNA epi-chain appears: CCATTCCAACG… . If two DNA strands are denoted by N_a and N_b and their nucleotides are denoted by 1a-2a-3a-4a-… and 1b-2b-3b-4b-… respetively, than their double-stranded epi-chains of the first order contain, for example, the following sequences of nucleotides: 1a-2b-3a-4b-5a-6b-… . The index «a» in the symbol N_a refers to nucleotide sequences of single-stranded DNA presented in the Genbank. A double-stranded epi-chain of the *n*-th order is denoted by the symbol $N_{n/ma}$ where the numerator "*n*" in the index indicates the order of the epi-chain, and the denominator "*m*a" (or "*m*b") is the numeration *m* of the first nucleotide of this epi-chain along the DNA strand N_a (or N_b). For example the symbol $N_{3/2b}$ refers to a double-stranded DNA epi-chain of the third order with the first nucleotide having the number 2 in the nucleotide numeration along the DNA strand N_b: 2b-5a-8b-11a-14b-…. (Fig. 2, right). The symbol N_{1ab} (or N_{1ba}) refers to a double-stranded DNA epi-chain of the first order.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 May 2019 doi:10.20944/preprints201904.0011.v2

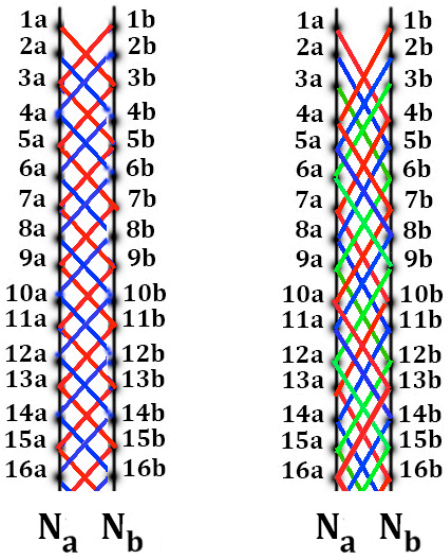


Fig. 2. The schematical representation of two separate DNA strands N_a and N_b, which are parts of the DNA double helix, in a form of straight lines carrying numerated nucleotides. Left: the system of double-stranded DNA epi-chains of the second order is shown, which begin from nucleotides 1a and 1b (red lines) and from nucleotides 2a and 2b (blue lines). Right: the system of double-stranded DNA epi-chains of the third order is shown, which begin from nucleotides 1a and 1b (red lines), from nucleotides 2a and 2b (blue lines) and from 3a and 3b (green lines).

This article presents data of the author's analysis of frequencies of nucleotides A, T, C, and G in single-stranded DNA epi-chains and also in double-stranded DNA epi-chains of long DNA in a wide set of eukaryotic and prokaryotic genomes.

3. Nucleotide frequencies in epi-chains of complete sets of chromosomes

This Section (jointly with Appendixes 1-4) presents the author's study results that frequencies of nucleotides in epi-chains of the considered orders in long DNA sequences of eukariotic genomes unexpectedly turn out to be almost identical independently on strong content differences of these nucleotide chains. These frequencies (or percentages, or probabilities) of nucleotides A, T, C and G in the studied DNA sequence will be denoted by the symbols %A, %T, %C and %G, respectively.

Let us describe, first of all, data about the human genome with its 22 autosomes and 2 sex chromosomes. These 24 nuclear chromosomes contain long DNA molecules, the lengths of texts in which lie in the range from 50 to 250 million letters approximately. These chromosomes differ greatly in their molecular dimensions, their arrangement of letters A, T, C and G, kinds and quantities of genes in them, cytogenetic bands (which shows biochemical specifity of different parts of chromosomes), etc. (Fig. 3).

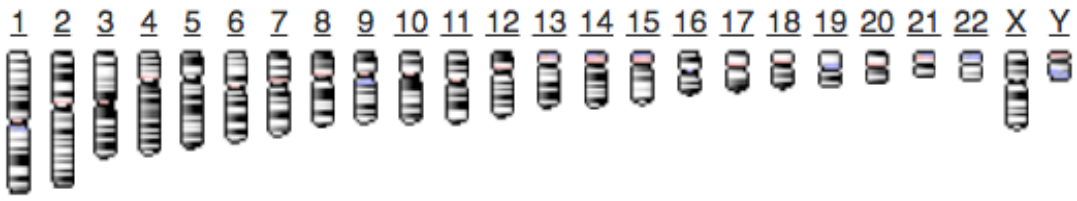


Fig. 3. Human chromosomes (from https://www.ncbi.nlm.nih.gov/genome/51)

Taking into account these great differences among human chromosomes, it was very unexpected for the author to reveal that these 24 human chromosomes are very similar each to other from the standpoint of nucleotide frequencies in their different single-stranded DNA epi-chains (Table 1) and also in their different double-stranded DNA epi-chains (Table 3).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 May 2019 doi:10.20944/preprints201904.0011.v2

Table 1. Frequencies %A, %T, %C and %G of nucleotides A, T, C and G in single-stranded DNA epi-chains of initial orders in all human nuclear chromosomes. Data relating single-stranded DNA epi-chains of the first order N_1, the second order ($N_{2/1}$, $N_{2/2}$) and the third order ($N_{3/1}$, $N_{3/2}$, $N_{3/3}$) are presented. Two left columns show ordinary symbols of chromosomes and their length. Initial data relating to these chromosomes were accessed from https://www.ncbi.nlm.nih.gov/genome/?term=Homo+sapiens+genome (the column RefSeq).

Chr	DNA length (bp)		N_1	$N_{2/1}$	$N_{2/2}$	$N_{3/1}$	$N_{3/2}$	$N_{3/3}$
1	248956422	%A	0.2910	0.2910	0.2910	0.2910	0.2910	0.2910
		%T	0.2918	0.2917	0.2918	0.2917	0.2917	0.2918
		%C	0.2085	0.2085	0.2085	0.2084	0.2085	0.2085
		%G	0.2087	0.2088	0.2087	0.2088	0.2088	0.2087
2	242193529	%A	0.2984	0.2985	0.2984	0.2983	0.2985	0.2985
		%T	0.2993	0.2992	0.2993	0.2993	0.2992	0.2993
		%C	0.2009	0.2009	0.2009	0.2009	0.2009	0.2008
		%G	0.2014	0.2014	0.2014	0.2014	0.2014	0.2014
3	198295559	%A	0.3013	0.3013	0.3013	0.3013	0.3013	0.3014
		%T	0.3020	0.3020	0.3020	0.3022	0.3020	0.3019
		%C	0.1980	0.1981	0.1980	0.1979	0.1981	0.1981
		%G	0.1986	0.1986	0.1986	0.1986	0.1986	0.1986
4	190214555	%A	0.3086	0.3086	0.3086	0.3087	0.3086	0.3086
		%T	0.3089	0.3090	0.3089	0.3089	0.3090	0.3090
		%C	0.1910	0.1909	0.1910	0.1909	0.1910	0.1910
		%G	0.1915	0.1915	0.1914	0.1915	0.1914	0.1914
5	181538259	%A	0.3018	0.3018	0.3018	0.3018	0.3017	0.3018
		%T	0.3032	0.3032	0.3032	0.3031	0.3032	0.3032
		%C	0.1971	0.1971	0.1971	0.1971	0.1972	0.1971
		%G	0.1979	0.1979	0.1980	0.1980	0.1979	0.1979
6	171718000	%A	0.3021	0.3021	0.3021	0.3021	0.3021	0.3022
		%T	0.3020	0.3020	0.3021	0.3022	0.3020	0.302
		%C	0.1979	0.1979	0.1979	0.1979	0.1980	0.1979
		%G	0.1979	0.1979	0.1979	0.1979	0.1979	0.1979
7	159345973	%A	0.2960	0.2960	0.2960	0.2960	0.2960	0.296
		%T	0.2970	0.2970	0.2970	0.2971	0.2970	0.297
		%C	0.2033	0.2033	0.2033	0.2032	0.2034	0.2033
		%G	0.2037	0.2037	0.2037	0.2037	0.2037	0.2037
8	146399655	%A	0.2994	0.2994	0.2994	0.2994	0.2994	0.2994
		%T	0.2990	0.2990	0.2989	0.2990	0.2990	0.2990
		%C	0.2008	0.2008	0.2008	0.2008	0.2008	0.2008
		%G	0.2008	0.2008	0.2009	0.2008	0.2008	0.2008
9	111583154	%A	0.2928	0.2928	0.2928	0.2927	0.2928	0.2929
		%T	0.2926	0.2927	0.2925	0.2926	0.2926	0.2927
		%C	0.2074	0.2073	0.2074	0.2074	0.2074	0.2073
		%G	0.2072	0.2072	0.2073	0.2073	0.2072	0.2072
10	133797422	%A	0.2917	0.2918	0.2916	0.2917	0.2917	0.2918
		%T	0.2929	0.2928	0.2929	0.2929	0.2928	0.2929
		%C	0.2074	0.2073	0.2075	0.2074	0.2075	0.2073
		%G	0.2080	0.2080	0.2080	0.2080	0.2080	0.2080

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 May 2019 doi:10.20944/preprints201904.0011.v2

11	135086622	%A	0.2920	0.2920	0.2921	0.2920	0.2921	0.2920
		%T	0.2926	0.2926	0.2926	0.2925	0.2926	0.2927
		%C	0.2074	0.2074	0.2074	0.2074	0.2074	0.2073
		%G	0.2080	0.2080	0.2080	0.2081	0.2079	0.2080
12	133275309	%A	0.2957	0.2957	0.2958	0.2958	0.2956	0.2958
		%T	0.2966	0.2966	0.2966	0.2966	0.2967	0.2966
		%C	0.2035	0.2035	0.2035	0.2035	0.2035	0.2035
		%G	0.2042	0.2042	0.2041	0.2041	0.2042	0.2042
13	95789532	%A	0.3069	0.3068	0.3069	0.3069	0.3069	0.3068
		%T	0.3079	0.3078	0.3079	0.3079	0.3079	0.3078
		%C	0.1926	0.1927	0.1925	0.1926	0.1925	0.1927
		%G	0.1926	0.1927	0.1926	0.1926	0.1926	0.1927
14	87316725	%A	0.2945	0.2944	0.2946	0.2946	0.2946	0.2944
		%T	0.2970	0.2970	0.2970	0.2970	0.2970	0.2970
		%C	0.2040	0.2041	0.2039	0.2040	0.2039	0.2040
		%G	0.2045	0.2045	0.2045	0.2045	0.2045	0.2046
15	101991189	%A	0.2896	0.2896	0.2895	0.2897	0.2894	0.2896
		%T	0.2901	0.2901	0.2901	0.2900	0.2902	0.2900
		%C	0.2097	0.2097	0.2098	0.2098	0.2097	0.2097
		%G	0.2106	0.2106	0.2106	0.2105	0.2106	0.210
16	90338345	%A	0.2758	0.2757	0.2758	0.2758	0.2757	0.2758
		%T	0.2784	0.2785	0.2783	0.2784	0.2784	0.2784
		%C	0.2221	0.2222	0.2221	0.2221	0.2222	0.2222
		%G	0.2237	0.2236	0.2238	0.2237	0.2237	0.2237
17	83257441	%A	0.2730	0.2730	0.2730	0.2730	0.2730	0.2730
		%T	0.2738	0.2739	0.2737	0.2739	0.2738	0.2737
		%C	0.2258	0.2257	0.2259	0.2258	0.2257	0.2259
		%G	0.2273	0.2274	0.2273	0.2272	0.2274	0.2274
18	74792881	%A	0.3011	0.3012	0.3009	0.3011	0.3011	0.3010
		%T	0.3014	0.3013	0.3014	0.3014	0.3013	0.3014
		%C	0.1987	0.1985	0.1988	0.1987	0.1987	0.1986
		%G	0.1989	0.1989	0.1989	0.1988	0.1989	0.1990
19	58617616	%A	0.2591	0.2592	0.2590	0.2592	0.2591	0.2591
		%T	0.2615	0.2615	0.2615	0.2616	0.2616	0.2613
		%C	0.2388	0.2388	0.2388	0.2387	0.2388	0.2388
		%G	0.2406	0.2405	0.2407	0.2405	0.2405	0.2408
20	59605541	%A	0.2778	0.2779	0.2777	0.2778	0.2777	0.2779
		%T	0.2810	0.2808	0.2811	0.2810	0.2810	0.2809
		%C	0.2201	0.2202	0.2200	0.2201	0.2201	0.2201
		%G	0.2211	0.2212	0.2211	0.2211	0.2212	0.2211
21	33543332	%A	0.2964	0.2964	0.2965	0.2963	0.2964	0.2966
		%T	0.2946	0.2947	0.2945	0.2946	0.2946	0.2947
		%C	0.2047	0.2046	0.2047	0.2046	0.2049	0.2044
		%G	0.2043	0.2043	0.2043	0.2045	0.2041	0.2043
22	50818468	%A	0.2651	0.2651	0.2651	0.2651	0.2652	0.2651
		%T	0.2648	0.2648	0.2648	0.2648	0.2647	0.2650
		%C	0.2339	0.2339	0.2340	0.234	0.2341	0.2337
		%G	0.2361	0.2362	0.2360	0.2362	0.2360	0.2362
X	156040895	%A	0.3019	0.3019	0.3018	0.3019	0.3018	0.3018
		%T	0.3029	0.3029	0.3029	0.3029	0.3029	0.3029

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 May 2019 doi:10.20944/preprints201904.0011.v2

		%C	0.1971	0.1970	0.1971	0.1971	0.1971	0.1971
		%G	0.1982	0.1982	0.1982	0.1982	0.1982	0.1982
Y	57227415	%A	0.2985	0.2986	0.2985	0.2985	0.2987	0.2984
		%T	0.3012	0.3013	0.3011	0.3013	0.3011	0.3012
		%C	0.2001	0.2001	0.2001	0.2001	0.2001	0.2001
		%G	0.2001	0.2000	0.2003	0.2001	0.2001	0.2003

It is surprising but for all 24 human chromosomes, each of rows of frequencies in Table 1 contains approximately the same value of frequencies of nucleotides in its 7 cells corresponding to the appropriate epi-chains N_1, $N_{2/1}$, $N_{2/2}$, $N_{3/1}$, $N_{3/2}$, $N_{3/3}$ in single-stranded DNA: small differences in the values of frequencies appear only in the fourth decimal place, though these epi-chains strongly differ by its length and by its nucleotide arrangement. It means that the single-stranded DNA epi-chains of the higher orders are practically copies (or duplicates) of nucleotide chains N_1 of DNA strands regarding frequencies of nucleotides A, T, C and G. Different human chromosomes differ little from each other in terms of their nucleotide frequencies %A, %T, %C and %G in single-stranded DNA epi-chains. The analogue of the second Chargaff's rule for single-stranded DNA is executed in the epi-chains of the higher orders: %A ≈ %T and %C ≈ %G.

But what one can say about frequencies of nucleotides in single-stranded DNA epi-chains of significantly higher orders? Whether such equalities of probabilties also exist there? The author's study gives a positive answer on this question. Table 2 shows frequencies %A, %T, %C and %G in examples of single-stranded DNA epi-chains of the 10th order, the 50th order, the 100th order in the human chromosome № 1. Tabular data testify that in all shown epi-chains these frequencies have practically the same values like in the basic nucleotide sequence N_1 of the single-stranded DNA. Other single-stranded epi-chains of the indicated orders n = 10, 50 and 100, which are not shown in the small Table 2, obey the same rule of approximate equality of nucleotide frequencies in epi-chains.

Table 2. Frequencies of nucleotides %A, %T, %C and %G in the human chromosome №1 for the following examples of randomly selected epi-chains: $N_{10/1}$ and $N_{10/9}$ refer to two epi-chains of the 10^{th} order; $N_{50/1}$ and $N_{50/49}$ – two epi-chains of the 50^{th} order; $N_{100/1}$ and $N_{100/99}$ – two epi-chains of the 100^{th} order. N_1 refers to the basic nucleotide sequence of the DNA strand of this chromosome (by analogy with Table 1).

	$\mathbf{N_1}$	$\mathbf{N_{10/1}}$	$\mathbf{N_{10/9}}$	$\mathbf{N_{50/1}}$	$\mathbf{N_{50/49}}$	$\mathbf{N_{100/1}}$	$\mathbf{N_{100/99}}$
%A	0.2910	0.2910	0.2911	0.2910	0.2915	0.2910	0.2915
%T	0.2918	0.2918	0.2919	0.2917	0.2916	0.2917	0.2920
%C	0.2085	0.2084	0.2084	0.2086	0.2085	0.2086	0.2080
%G	0.2087	0.2088	0.2087	0.2088	0.2085	0.2087	0.2086

But what situation exists in double-stranded DNA epi-chains of all human nuclear chromosomes? Table 3 gives answer on this question showing a very similar situation in double-stranded DNA epi-chains of all 24 human nuclear chromosomes from the standpoint of nucleotide frequencies %A, %T, %C and %G.

Table 3. Frequencies %A, %T, %C and %G of nucleotides A, T, C and G in double-stranded DNA epi-chains of initial orders in all human nuclear chromosomes. Data relating double-stranded DNA epi-chains of the first order (N_{1ab}), the second order ($N_{2/1}$, $N_{2/2}$) and the third order ($N_{3/1}$, $N_{3/2}$, $N_{3/3}$) are presented. Two left columns show ordinary symbols of chromosomes and their length. Initial data relating to these chromosomes were accessed from https://www.ncbi.nlm.nih.gov/genome/?term=Homo+sapiens+genome (the column RefSeq).

Chr	DNA length (bp)		$\mathbf{N_{1ab}}$	$\mathbf{N_{2/1a}}$	$\mathbf{N_{2/2a}}$	$\mathbf{N_{3/1a}}$	$\mathbf{N_{3/2a}}$	$\mathbf{N_{3/3a}}$
1	248956422	%A	0.2914	0.2914	0.2914	0.2914	0.2913	0.2914
		%T	0.2914	0.2914	0.2914	0.2914	0.2914	0.2914
		%C	0.2086	0.2086	0.2086	0.2086	0.2087	0.2086
		%G	0.2086	0.2086	0.2086	0.2086	0.2086	0.2086

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 May 2019 doi:10.20944/preprints201904.0011.v2

2	242193529	%A	0.2989	0.2989	0.2989	0.2990	0.2988	0.2989
		%T	0.2988	0.2988	0.2988	0.2987	0.2989	0.2989
		%C	0.2011	0.2011	0.2012	0.2012	0.2011	0.2010
		%G	0.2012	0.2012	0.2011	0.2012	0.2012	0.2012
3	198295559	%A	0.3017	0.3017	0.3016	0.3017	0.3016	0.3016
		%T	0.3017	0.3016	0.3017	0.3017	0.3016	0.3017
		%C	0.1983	0.1984	0.1983	0.1982	0.1983	0.1984
		%G	0.1983	0.1983	0.1983	0.1983	0.1984	0.1983
4	190214555	%A	0.3088	0.3088	0.3088	0.3088	0.3087	0.3087
		%T	0.3088	0.3088	0.3087	0.3088	0.3088	0.3088
		%C	0.1912	0.1912	0.1912	0.1912	0.1912	0.1911
		%G	0.1913	0.1912	0.1912	0.1912	0.1912	0.1913
5	181538259	%A	0.3025	0.3025	0.3025	0.3024	0.3025	0.3025
		%T	0.3025	0.3025	0.3025	0.3025	0.3025	0.3025
		%C	0.1976	0.1975	0.1975	0.1976	0.1975	0.1975
		%G	0.1975	0.1975	0.1975	0.1975	0.1976	0.1975
6	171718000	%A	0.3021	0.3021	0.3022	0.3021	0.3020	0.3021
		%T	0.3021	0.3021	0.3020	0.3021	0.3020	0.3021
		%C	0.1979	0.1980	0.1979	0.1978	0.1980	0.1980
		%G	0.1979	0.1979	0.1980	0.1979	0.1979	0.1979
7	159345973	%A	0.2965	0.2965	0.2965	0.2965	0.2964	0.2965
		%T	0.2965	0.2965	0.2966	0.2966	0.2965	0.2965
		%C	0.2035	0.2034	0.2035	0.2034	0.2035	0.2035
		%G	0.2035	0.2035	0.2035	0.2035	0.2035	0.2035
8	146399655	%A	0.2992	0.2992	0.2991	0.2991	0.2992	0.2991
		%T	0.2992	0.2993	0.2992	0.2993	0.2992	0.2992
		%C	0.2008	0.2008	0.2009	0.2008	0.2008	0.2009
		%G	0.2008	0.2008	0.2008	0.2008	0.2008	0.2007
9	111583154	%A	0.2927	0.2927	0.2926	0.2926	0.2927	0.2929
		%T	0.2927	0.2928	0.2927	0.2928	0.2926	0.2927
		%C	0.2073	0.2072	0.2073	0.2073	0.2073	0.2072
		%G	0.2073	0.2073	0.2074	0.2073	0.2074	0.2073
10	133797422	%A	0.2923	0.2923	0.2922	0.2924	0.2923	0.2924
		%T	0.2922	0.2923	0.2923	0.2922	0.2922	0.2923
		%C	0.2077	0.2077	0.2078	0.2077	0.2078	0.2076
		%G	0.2077	0.2076	0.2077	0.2077	0.2078	0.2077
11	135086622	%A	0.2923	0.2922	0.2923	0.2923	0.2923	0.2922
		%T	0.2923	0.2923	0.2924	0.2922	0.2924	0.2924
		%C	0.2077	0.2077	0.2077	0.2077	0.2076	0.2077
		%G	0.2077	0.2077	0.2077	0.2078	0.2077	0.2076
12	133275309	%A	0.2961	0.2962	0.2962	0.2961	0.2961	0.2962
		%T	0.2962	0.2961	0.2962	0.2963	0.2961	0.2962
		%C	0.2038	0.2039	0.2038	0.2038	0.2039	0.2039
		%G	0.2038	0.2038	0.2038	0.2038	0.2038	0.2038
13	95789532	%A	0.3074	0.3073	0.3074	0.3074	0.3074	0.3073
		%T	0.3074	0.3073	0.3075	0.3074	0.3074	0.3073
		%C	0.1926	0.1926	0.1925	0.1926	0.1926	0.1927
		%G	0.1926	0.1927	0.1926	0.1926	0.1926	0.1926
14	87316725	%A	0.2957	0.2956	0.2958	0.2958	0.2958	0.2956
		%T	0.2958	0.2958	0.2958	0.2957	0.2957	0.2958

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 May 2019 doi:10.20944/preprints201904.0011.v2

		%C	0.2043	0.2043	0.2042	0.2043	0.2041	0.2044
		%G	0.2042	0.2043	0.2042	0.2042	0.2043	0.2042
15	101991189	%A	0.2899	0.2898	0.2898	0.2898	0.2897	0.2898
		%T	0.2899	0.2899	0.2898	0.2898	0.2900	0.2898
		%C	0.2102	0.2102	0.2102	0.2101	0.2102	0.2102
		%G	0.2102	0.2100	0.2102	0.2102	0.2102	0.2101
16	90338345	%A	0.2770	0.2771	0.2771	0.2770	0.2772	0.2770
		%T	0.2772	0.2771	0.2770	0.2772	0.2769	0.2772
		%C	0.2230	0.2229	0.2230	0.2230	0.2229	0.2230
		%G	0.2229	0.2229	0.2229	0.2228	0.2230	0.2228
17	83257441	%A	0.2734	0.2734	0.2735	0.2735	0.2735	0.2733
		%T	0.2735	0.2735	0.2733	0.2735	0.2733	0.2734
		%C	0.2265	0.2265	0.2266	0.2264	0.2266	0.2265
		%G	0.2266	0.2266	0.2266	0.2266	0.2266	0.2267
18	74792881	%A	0.3013	0.3012	0.3011	0.3014	0.3012	0.3013
		%T	0.3011	0.3013	0.3012	0.3011	0.3012	0.3011
		%C	0.1987	0.1988	0.1988	0.1987	0.1988	0.1987
		%G	0.1988	0.1987	0.1989	0.1988	0.1988	0.1989
19	58617616	%A	0.2603	0.2603	0.2601	0.2604	0.2603	0.2603
		%T	0.2603	0.2604	0.2603	0.2603	0.2603	0.2601
		%C	0.2397	0.2397	0.2398	0.2397	0.2396	0.2398
		%G	0.2396	0.2396	0.2397	0.2395	0.2397	0.2398
20	59605541	%A	0.2795	0.2793	0.2793	0.2795	0.2792	0.2795
		%T	0.2793	0.2794	0.2796	0.2793	0.2795	0.2793
		%C	0.2206	0.2207	0.2205	0.2206	0.2206	0.2206
		%G	0.2206	0.2206	0.2206	0.2206	0.2207	0.2206
21	33543332	%A	0.2955	0.2958	0.2956	0.2955	0.2957	0.2955
		%T	0.2956	0.2954	0.2954	0.2954	0.2953	0.2957
		%C	0.2044	0.2045	0.2045	0.2046	0.2046	0.2043
		%G	0.2045	0.2044	0.2045	0.2044	0.2044	0.2045
22	50818468	%A	0.2650	0.2650	0.2650	0.2651	0.2652	0.2651
		%T	0.2650	0.2650	0.2650	0.2647	0.2647	0.2650
		%C	0.2350	0.2351	0.2349	0.2349	0.2350	0.2349
		%G	0.2351	0.2350	0.2351	0.2352	0.2351	0.2350
X	156040895	%A	0.3024	0.3024	0.3024	0.3025	0.3024	0.3023
		%T	0.3024	0.3023	0.3023	0.3023	0.3024	0.3024
		%C	0.1976	0.1976	0.1976	0.1976	0.1976	0.1976
		%G	0.1976	0.1976	0.1977	0.1976	0.1977	0.1977
Y	57227415	%A	0.2999	0.2999	0.2998	0.2998	0.2999	0.2996
		%T	0.2999	0.3000	0.2998	0.3000	0.2999	0.3000
		%C	0.2002	0.2000	0.2003	0.2001	0.2000	0.2003
		%G	0.2000	0.2001	0.2001	0.2001	0.2002	0.2001

Table 3 shows that each of its rows of frequencies contains approximately the same value of frequencies of nucleotides in its 7 cells corresponding to the appropriate double-stranded DNA epi-chains N_{1ab}, $N_{2/1a}$, $N_{2/2a}$, $N_{3/1a}$, $N_{3/2a}$, $N_{3/3a}$: small differences in the values of frequencies appear only in the fourth decimal place, though these epi-chains strongly differ by its length and by its nucleotide arrangement. It means that the double-stranded DNA epi-chains of the higher orders are practically copies (or duplicates) of nucleotide chains N_{1ab} of DNA strands regarding frequencies of nucleotides A, T, C and G. Different human chromosomes differ little from each other in terms of their nucleotide frequencies %A, %T, %C and %G in double-stranded DNA epi-chains. It should be specially emphasised that the proposed by the author

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 May 2019 doi:10.20944/preprints201904.0011.v2

analogue of the second Chargaff's for epi-chains (%A ≈ %T and %C ≈ %G) is executed with higher level of accuracy in the case of double-stranded DNA epi-chains than in the case of single-stranded DNA epi-chains (compare Table 1 and Table 3). Comparing Tables 1 and 3 shows that cases of single-stranded and double-stranded DNA epi-chains are approximately identical from the standpoint of their frequencies %A, %T, %C and %G.

But what one can say about frequencies of nucleotides in double-stranded DNA epi-chains of significantly higher orders? Whether such equalities of probabilties also exist there? Yes, this question has a positive answer. Table 4 shows frequencies %A, %T, %C and %G in examples of double-stranded DNA epi-chains of the 10th order, the 50th order, the 100th order in the human chromosome № 1. Tabular data testify that in all shown epi-chains these frequencies have practically the same values like in the not sparse nucleotide sequence N_{1ab} of the first order. Other double-stranded epi-chains of the indicated orders n = 10, 50 and 100, which are not shown in this small Table 4, obey the same of approximate equality of nucleotide frequencies in their set. The cases of single-stranded and double-stranded DNA epi-chains are very similar from the standpoint of nucleotide frequencies %A, %T, %C and %G.

Table 4. Frequencies of nucleotides %A, %T, %C and %G in the human chromosome №1 for the following examples of randomly selected double-stranded DNA epi-chains: $N_{10/1a}$ and $N_{10/9a}$ refer to two double-stranded epi-chains of the 10th order; $N_{50/1a}$ and $N_{50/49a}$ – two double-stranded epi-chains of the 50th order; $N_{100/1a}$ and $N_{100/99a}$ – two double-stranded epi-chains of the 100th order. N_{1ab} refers to the double-stranded DNA epi-chain of the first order in this chromosome (by analogy with Tables 2 and 3).

	N_{1ab}	$N_{10/1a}$	$N_{10/9a}$	$N_{50/1a}$	$N_{50/49a}$	$N_{100/1a}$	$N_{100/99a}$
%A	0.2914	0.2914	0.2915	0.2914	0.2917	0.2918	0.2919
%T	0.2914	0.2913	0.2915	0.2916	0.2913	0.2914	0.2915
%C	0.2086	0.2087	0.2086	0.2085	0.2088	0.2082	0.2083
%G	0.2086	0.2087	0.2085	0.2086	0.2082	0.2086	0.2082

Till now only nuclear chromosomes were under consideration in the article. But what about frequencies of nucleotides in epi-chains of mitochondrions, whose length much more shorter that length of nuclear chromosomes? Table 5 shows these frequencies for the human mitochondrion whose length is only 16569 bp.

Table 5. Frequencies %A, %T, %C and %G of nucleotides A, T, C and G in single-stranded DNA epi-chains of initial orders in the human mitochondrion. Data relating single-stranded epi-chains of the first order (N_1), the second order ($N_{2/1}$, $N_{2/2}$) and the third order ($N_{3/1}$, $N_{3/2}$, $N_{3/3}$) are presented. Initial data relating to the human mitochondrion were accessed from https://www.ncbi.nlm.nih.gov/nuccore/NC_012920.1.

DNA length (bp)		N_1	$N_{2/1}$	$N_{2/2}$	$N_{3/1}$	$N_{3/2}$	$N_{3/3}$
16569	%A	0.3093	0.3095	0.3090	0.3147	0.3126	0.3006
	%T	0.2471	0.2496	0.2446	0.2433	0.2486	0.2493
	%C	0.3127	0.3087	0.3168	0.3111	0.3131	0.3140
	%G	0.1309	0.1322	0.1296	0.1309	0.1257	0.1362

These tabular data testify that - in human mitochondrion - frequencies of its nucleotides in the not sparse single-stranded DNA epi-chain N_1 and in other considered epi-chains are also approximately equal to each other but with a lower level of accuracy than in the case of nuclear chromosomes. In contrast to this, the second Chargaff's and its extensions for single-stranded DNA epi-chains are not executed at all in this short mitochondrion: %A differ significantly from %T and %C differ significantly from %G in all these chains.

The case of double-stranded DNA epi-chains in human mitochondrion is represented in Table 6. Comparing Table 5 and Table 6, one can see that values %A, %T, %C and %G have noticeable differences in them. But frequencies of nucleotides in the not sparse double-stranded DNA epi-chain N_{1ab} and in other considered epi-chains in Table 6 are also approximately equal to each other. Unlike the Table 5, Table 6

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 May 2019 doi:10.20944/preprints201904.0011.v2

shows that the second Chargaff's and its extensions for double-stranded DNA epi-chains are approximately executed in this short mitochondrion for all considered epi-chains: %A ≈ %T and %C ≈ %G.

Table 6. Frequencies %A, %T, %C and %G of nucleotides A, T, C and G in double-stranded DNA epi-chains of initial orders in the human mitochondrion. Data relating double-stranded epi-chains of the first order (N_{1ab}), the second order ($N_{2/1a}$, $N_{2/2a}$) and the third order ($N_{3/1a}$, $N_{3/2a}$, $N_{3/3a}$) are presented. Initial data relating to the human mitochondrion were accessed from https://www.ncbi.nlm.nih.gov/nuccore/NC_012920.1.

DNA length (bp)		N_{1ab}	$N_{2/1a}$	$N_{2/2a}$	$N_{3/1a}$	$N_{3/2a}$	$N_{3/3a}$
16569	%A	0.2770	0.2775	0.2766	0.2815	0.2823	0.2707
	%T	0.2793	0.2816	0.2770	0.2765	0.2789	0.2792
	%C	0.2192	0.2150	0.2244	0.2187	0.2275	0.2274
	%G	0.2245	0.2259	0.2220	0.2232	0.2113	0.2227

The author has received similar results of approximate equality of frequencies of nucleotides A, T, C and G in single-stranded and double-stranded DNA epi-chains of different orders in complete sets of nuclear chromosomes of a few model organisms, which are used long ago in the study of genetics, development and disease: a house mouse *Mus musculus*, a fruit fly *Drosophila melanogaster*, a nematode *Caenorhabditis elegans* and a plant *Arabidopsis thaliana*. Received data on frequencies in single-stranded and double-stranded DNA epi-chains of genomes of these eukariots are presented in Appendixes 1-4.

4. Nucleotide frequencies in DNA epi-chains of prokaryotic genomes

What one can say about nucleotide frequencies in epi-chains of prokaryotic genomes? This Section represents author's results testified that in prokariotic genomes the same situation relating these frequencies exists like in eukaryotic genomes (but with a slightly lower level of accuracy in equalities). Tables 7 and 8 show frequencies of nucleotides A, T, C and G in single-stranded and double-stranded DNA for all 19 genomes of bacteria and archaea from their full list in the article [Rapoport, Trifonov, 2012]. After these Tables, additional data on these genomes are given.

Table 7. Frequencies %A, %T, %C and %G of nucleotides A, T, C and G in single-stranded epi-chains of initial orders in 19 prokaryotic genomes. Data relating epi-chains of the first order (N_1), the second order ($N_{2/1}$, $N_{2/2}$) and the third order ($N_{3/1}$, $N_{3/2}$, $N_{3/3}$) are presented.

№	DNA length		N_1	$N_{2/1}$	$N_{2/2}$	$N_{3/1}$	$N_{3/2}$	$N_{3/3}$
1	Aquifex aeolicus, 1551335 bp	%A	0.2841	0.2834	0.2849	0.2820	0.2859	0.2845
		%T	0.2811	0.2818	0.2805	0.2815	0.2827	0.2791
		%C	0.2168	0.2173	0.2163	0.2167	0.2156	0.2181
		%G	0.2179	0.2176	0.2183	0.2198	0.2157	0.2183
2	Acidobacteria bacterium, 4996384 bp	%A	0.2155	0.2156	0.2153	0.2165	0.2154	0.2146
		%T	0.2171	0.2169	0.2173	0.2167	0.2186	0.2161
		%C	0.2855	0.2855	0.2855	0.2836	0.2863	0.2868
		%G	0.2819	0.2819	0.2818	0.2833	0.2798	0.2825
3	Bradyrhizobium japonicum, 9224208 bp	%A	0.1819	0.1820	0.1818	0.1806	0.1836	0.1815
		%T	0.1815	0.1814	0.1816	0.1797	0.1825	0.1824
		%C	0.3184	0.3185	0.3183	0.3197	0.3166	0.3189
		%G	0.3182	0.3182	0.3182	0.3201	0.3173	0.3173
4	Bacillus subtilis, 4025326 bp	%A	0.2805	0.2801	0.2809	0.2813	0.2795	0.2807
		%T	0.2806	0.2806	0.2805	0.2809	0.2803	0.2805
		%C	0.2192	0.2196	0.2189	0.2181	0.2196	0.2200
		%G	0.2197	0.2197	0.2197	0.2196	0.2206	0.2189

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 May 2019 doi:10.20944/preprints201904.0011.v2

5	Chlamydia trachomatis, 1025839 bp	%A	0.2942	0.2939	0.2945	0.2927	0.2963	0.2936
		%T	0.2930	0.2931	0.2930	0.2921	0.2935	0.2935
		%C	0.2067	0.2066	0.2067	0.2074	0.2049	0.2077
		%G	0.2061	0.2063	0.2058	0.2078	0.2053	0.2051
6	Chromobacterium violaceum, 127377 bp	%A	0.1670	0.1658	0.1682	0.1758	0.1651	0.1602
		%T	0.1819	0.1812	0.1827	0.1883	0.1808	0.1766
		%C	0.3049	0.3044	0.3055	0.3009	0.3112	0.3027
		%G	0.3461	0.3486	0.3437	0.3350	0.3429	0.3606
7	Dehalococcoides mccartyi, 1521287 bp	%A	0.2664	0.2665	0.2663	0.2706	0.2617	0.2668
		%T	0.2645	0.2647	0.2643	0.2653	0.2686	0.2596
		%C	0.2338	0.2337	0.2339	0.2295	0.2376	0.2343
		%G	0.2353	0.2351	0.2355	0.2346	0.2321	0.2394
8	Escherichia coli, 5231428 bp	%A	0.2480	0.2481	0.248	0.2479	0.2468	0.2494
		%T	0.2472	0.2470	0.2474	0.2489	0.2461	0.2466
		%C	0.2526	0.2523	0.2528	0.2519	0.2544	0.2515
		%G	0.2522	0.2526	0.2518	0.2514	0.2527	0.2525
9	Flavobacterium psychrophilum, 2860382 bp	%A	0.3326	0.3326	0.3326	0.3335	0.3306	0.3338
		%T	0.342	0.3416	0.3425	0.3411	0.3423	0.3427
		%C	0.1640	0.1643	0.1637	0.1616	0.1665	0.1638
		%G	0.1614	0.1615	0.1612	0.1638	0.1606	0.1597
10	Gloeobacter violaceus, 4659019 bp	%A	0.1906	0.1905	0.1907	0.1892	0.1925	0.1901
		%T	0.1894	0.1899	0.1890	0.1881	0.1899	0.1903
		%C	0.3101	0.3098	0.3104	0.312	0.3074	0.3108
		%G	0.3099	0.3099	0.3099	0.3107	0.3103	0.3088
11	Helicobacter pylori, 1643831 bp	%A	0.3033	0.3034	0.3032	0.3033	0.3042	0.3022
		%T	0.3048	0.3049	0.3048	0.3037	0.3028	0.3080
		%C	0.1970	0.1969	0.1970	0.1989	0.1958	0.1962
		%G	0.1949	0.1949	0.1950	0.1941	0.1972	0.1936
12	Methanosarcina acetivorans, 5751492 bp	%A	0.2848	0.2848	0.2848	0.2830	0.2851	0.2863
		%T	0.2884	0.2881	0.2887	0.2890	0.2875	0.2887
		%C	0.2136	0.2139	0.2132	0.2148	0.2133	0.2126
		%G	0.2132	0.2132	0.2133	0.2132	0.2141	0.2124
13	Nanoarchaeum equitans, 490885 bp	%A	0.3422	0.3424	0.3420	0.3413	0.3423	0.3431
		%T	0.3422	0.3419	0.3425	0.342	0.3413	0.3433
		%C	0.1576	0.1581	0.1571	0.1593	0.1571	0.1563
		%G	0.1580	0.1577	0.1583	0.1573	0.1594	0.1573
14	Syntrophus aciditrophicus, 3179300 bp	%A	0.2431	0.2431	0.2431	0.2441	0.2418	0.2433
		%T	0.2423	0.2423	0.2424	0.2446	0.2413	0.2412
		%C	0.2556	0.2556	0.2557	0.2546	0.2573	0.2551
		%G	0.2590	0.2590	0.2589	0.2567	0.2597	0.2604
15	Streptomyces coelicolor, 8667507 bp	%A	0.1389	0.1386	0.1391	0.1366	0.1408	0.1392
		%T	0.1400	0.1400	0.1399	0.1381	0.1414	0.1403
		%C	0.3601	0.3600	0.3602	0.3614	0.3580	0.3609
		%G	0.3611	0.3613	0.3608	0.3638	0.3598	0.3596
16	Sulfolobus solfataricus, 2727337 bp	%A	0.3181	0.3179	0.3184	0.3182	0.3185	0.3177
		%T	0.3233	0.3231	0.3234	0.3225	0.3230	0.3243
		%C	0.1798	0.1797	0.1799	0.1798	0.1782	0.1814
		%G	0.1788	0.1793	0.1783	0.1795	0.1803	0.1765
17	Treponema denticola,	%A	0.3095	0.3095	0.3096	0.3079	0.3098	0.3109
		%T	0.3106	0.3104	0.3107	0.3120	0.3089	0.3108

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 May 2019 doi:10.20944/preprints201904.0011.v2

	1850823 bp	%C	0.188	0.1877	0.1882	0.1882	0.1879	0.1877
		%G	0.1919	0.1923	0.1915	0.1918	0.1934	0.1906
18	Thermotoga maritima, 1859582 bp	%A	0.2695	0.2700	0.2689	0.2703	0.2697	0.2684
		%T	0.2678	0.2681	0.2675	0.2689	0.2694	0.2651
		%C	0.2281	0.2277	0.2284	0.2271	0.2289	0.2282
		%G	0.2347	0.2341	0.2352	0.2337	0.2320	0.2383
19	Thermus thermophilus, 2121526 bp	%A	0.1543	0.1543	0.1542	0.1515	0.1542	0.1571
		%T	0.1557	0.1556	0.1558	0.1556	0.1529	0.1586
		%C	0.3461	0.3458	0.3464	0.3490	0.3465	0.3428
		%G	0.3439	0.3443	0.3435	0.3439	0.3464	0.3414

Table 8. Frequencies %A, %T, %C and %G of nucleotides A, T, C and G in double-stranded epi-chains of initial orders in 19 prokaryotic genomes. Data relating epi-chains of the first order (N_{1ab}), the second order ($N_{2/1}$, $N_{2/2}$) and the third order ($N_{3/1}$, $N_{3/2}$, $N_{3/3}$) are presented.

№	DNA length		N_{1ab}	$N_{2/1a}$	$N_{2/2a}$	$N_{3/1a}$	$N_{3/2a}$	$N_{3/3a}$
1	Aquifex aeolicus, 1551335 bp	%A	0.2819	0.2825	0.2823	0.2802	0.2850	0.2819
		%T	0.2833	0.2826	0.2831	0.2833	0.2837	0.2802
		%C	0.2178	0.2173	0.2170	0.2188	0.2153	0.2185
		%G	0.2170	0.2175	0.2177	0.2177	0.2161	0.2179
2	Acidobacteria bacterium, 4996384 bp	%A	0.2165	0.2164	0.2165	0.2164	0.2164	0.2154
		%T	0.2161	0.2161	0.2161	0.2167	0.2175	0.2153
		%C	0.2837	0.2838	0.2834	0.2829	0.2832	0.2842
		%G	0.2837	0.2837	0.2839	0.2840	0.2828	0.2851
3	Bradyrhizobium japonicum, 9224208 bp	%A	0.1818	0.1814	0.1817	0.1801	0.1826	0.1819
		%T	0.1816	0.1820	0.1817	0.1801	0.1835	0.1820
		%C	0.3184	0.3182	0.3182	0.3200	0.3174	0.3185
		%G	0.3183	0.3184	0.3184	0.3198	0.3166	0.3177
4	Bacillus subtilis, 4025326 bp	%A	0.2803	0.2803	0.2811	0.2809	0.2802	0.2805
		%T	0.2807	0.2804	0.2803	0.2814	0.2797	0.2807
		%C	0.2196	0.2195	0.2193	0.2187	0.2201	0.2201
		%G	0.2193	0.2197	0.2193	0.2190	0.2201	0.2188
5	Chlamydia trachomatis, 1025839 bp	%A	0.2935	0.2935	0.2940	0.2920	0.2953	0.2939
		%T	0.2938	0.2936	0.2935	0.2928	0.2944	0.2932
		%C	0.2062	0.2059	0.2066	0.2070	0.2053	0.2066
		%G	0.2065	0.2071	0.2060	0.2081	0.2050	0.2062
6	Chromobacterium violaceum, 127377 bp	%A	0.1743	0.1727	0.1766	0.1834	0.1771	0.1705
		%T	0.1747	0.1743	0.1743	0.1807	0.1689	0.1663
		%C	0.3240	0.3252	0.3243	0.3140	0.3264	0.3305
		%G	0.3270	0.3278	0.3248	0.3219	0.3277	0.3328
7	Dehalococcoides mccartyi, 1521287 bp	%A	0.2654	0.2656	0.2652	0.2677	0.2650	0.2632
		%T	0.2655	0.2655	0.2654	0.2682	0.2653	0.2632
		%C	0.2346	0.2347	0.2341	0.2322	0.2347	0.2367
		%G	0.2345	0.2342	0.2353	0.2319	0.2350	0.2369
8	Escherichia coli, 5231428 bp	%A	0.2477	0.2473	0.2473	0.2488	0.2466	0.2481
		%T	0.2475	0.2477	0.2481	0.2480	0.2463	0.2479
		%C	0.2520	0.2524	0.2523	0.2514	0.2542	0.2519
		%G	0.2527	0.2526	0.2523	0.2519	0.2529	0.2521

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 May 2019 doi:10.20944/preprints201904.0011.v2

9	Flavobacterium psychrophilum, 2860382 bp	%A	0.3375	0.3368	0.3385	0.3371	0.3355	0.3383
		%T	0.3371	0.3374	0.3366	0.3376	0.3373	0.3383
		%C	0.1627	0.1631	0.1624	0.1628	0.1633	0.1615
		%G	0.1626	0.1626	0.1625	0.1625	0.1639	0.1620
10	Gloeobacter violaceus, 4659019 bp	%A	0.1898	0.1904	0.1897	0.1882	0.1913	0.1901
		%T	0.1903	0.1898	0.1900	0.1891	0.1910	0.1903
		%C	0.3099	0.3099	0.3103	0.3114	0.3092	0.3097
		%G	0.3101	0.3097	0.3100	0.3113	0.3084	0.3099
11	Helicobacter pylori, 1643831 bp	%A	0.3041	0.3040	0.3041	0.3032	0.3032	0.3053
		%T	0.3040	0.3042	0.3039	0.3039	0.3039	0.3050
		%C	0.1959	0.1959	0.1961	0.1971	0.1966	0.1944
		%G	0.1960	0.1959	0.1959	0.1959	0.1963	0.1954
12	Methanosarcina acetivorans, 5751492 bp	%A	0.2868	0.2865	0.2867	0.2862	0.2862	0.2877
		%T	0.2864	0.2864	0.2868	0.2858	0.2865	0.2873
		%C	0.2136	0.2132	0.2132	0.2143	0.2134	0.2125
		%G	0.2132	0.2139	0.2133	0.2137	0.2140	0.2126
13	Nanoarchaeum equitans, 490885 bp	%A	0.3424	0.3419	0.3417	0.3422	0.3414	0.3430
		%T	0.3420	0.3423	0.3429	0.3411	0.3421	0.3434
		%C	0.1582	0.1571	0.1571	0.1585	0.1577	0.1574
		%G	0.1574	0.1587	0.1583	0.1582	0.1585	0.1563
14	Syntrophus aciditrophicus, 3179300 bp	%A	0.2427	0.2429	0.2429	0.2447	0.2416	0.2424
		%T	0.2427	0.2425	0.2426	0.2440	0.2414	0.2421
		%C	0.2573	0.2576	0.2573	0.2557	0.2585	0.2576
		%G	0.2573	0.2570	0.2572	0.2556	0.2584	0.2579
15	Streptomyces coelicolor, 8667507 bp	%A	0.1393	0.1393	0.1392	0.1370	0.1411	0.1397
		%T	0.1395	0.1393	0.1398	0.1377	0.1411	0.1398
		%C	0.3604	0.3611	0.3604	0.3625	0.3590	0.3600
		%G	0.3608	0.3603	0.3606	0.3627	0.3587	0.3605
16	Sulfolobus solfataricus, 2727337 bp	%A	0.3207	0.3203	0.3207	0.3201	0.3207	0.3211
		%T	0.3207	0.3207	0.3211	0.3206	0.3207	0.3209
		%C	0.1790	0.1790	0.1794	0.1793	0.1795	0.1787
		%G	0.1796	0.1800	0.1788	0.1800	0.1790	0.1793
17	Treponema denticola, 1850823 bp	%A	0.3101	0.3099	0.3105	0.3101	0.3094	0.3111
		%T	0.3100	0.3101	0.3098	0.3099	0.3092	0.3106
		%C	0.1896	0.1895	0.1899	0.1901	0.1914	0.1888
		%G	0.1902	0.1906	0.1898	0.1899	0.1900	0.1985
18	Thermotoga maritima, 1859582 bp	%A	0.2688	0.2691	0.2684	0.2695	0.2697	0.2670
		%T	0.2685	0.2690	0.2680	0.2697	0.2695	0.2665
		%C	0.2315	0.2306	0.2327	0.2298	0.2302	0.2328
		%G	0.2312	0.2312	0.2309	0.2310	0.2307	0.2337
19	Thermus thermophilus, 2121526 bp	%A	0.1550	0.1552	0.1552	0.1537	0.1536	0.1580
		%T	0.1549	0.1547	0.1549	0.1534	0.1535	0.1577
		%C	0.3447	0.3448	0.3451	0.3459	0.3464	0.3416
		%G	0.3454	0.3453	0.3449	0.3470	0.3465	0.3426

The following prokaryotic genomes are shown in Tables 7 and 8 under numbers 1-19:

1) Aquifex aeolicus VF5, complete genome, 1551335 bp, accession AE000657, version AE000657.1, https://www.ncbi.nlm.nih.gov/nuccore/AE000657.1?report=genbank ;
2) Acidobacteria bacterium KBS 146 M015DRAFT_scf7180000000004_quiver.1_C, whole genome shotgun sequence, 4996384 bp, accession JHVA01000001,

https://www.ncbi.nlm.nih.gov/nuccore/JHVA01000001.1?report=genbank;
3) Bradyrhizobium japonicum strain E109, complete genome, 9224208 bp, accession CP010313, https://www.ncbi.nlm.nih.gov/nuccore/CP010313.1?report=genbank ;
4) Bacillus subtilis strain UD1022, complete genome, 4025326 bp, accession CP011534, https://www.ncbi.nlm.nih.gov/nuccore/CP011534.1?report=genbank;
5) Chlamydia trachomatis strain QH111L, complete genome, 1025839 bp, accession CP018052, https://www.ncbi.nlm.nih.gov/nuccore/CP018052.1?report=genbank;
6) Chromobacterium violaceum strain LK30 1, whole genome shotgun sequence, 127377 bp, accession LDUX01000001 version LDUX01000001.1, https://www.ncbi.nlm.nih.gov/nuccore/LDUX01000001.1?report=genbank;
7) Dehalococcoides mccartyi strain CG3, complete genome, NCBI Reference Sequence: NZ_CP013074.1, 1521287 bp, https://www.ncbi.nlm.nih.gov/nuccore/NZ_CP013074.1?report=genbank;
8) Escherichia coli CFT073, complete genome, GenBank: AE014075.1, 5231428 bp, https://www.ncbi.nlm.nih.gov/nuccore/AE014075.1?report=genbank;
9) Flavobacterium psychrophilum JIP02/86, complete genome, 2860382 bp, accession NC_009613, https://www.ncbi.nlm.nih.gov/nuccore/NC_009613.3;
10) Gloeobacter violaceus PCC 7421 DNA, complete genome, GenBank: BA000045.2, 4659019 bp, accession BA000045 AP006568-AP006583 version BA000045.2, https://www.ncbi.nlm.nih.gov/nuccore/BA000045.2?report=genbank;
11) Helicobacter pilory, NCBI Reference Sequence: NC_000921.1, complete genome, 1643831 bp, accession NC_000921 NZ_AE001440-NZ_AE001571 version NC_000921.1, https://www.ncbi.nlm.nih.gov/nuccore/NC_000921.1 ;
12) Methanosarcina acetivorans str. C2A, complete genome, 5751492 bp, accession AE010299 AE010656-AE011189 version AE010299.1, https://www.ncbi.nlm.nih.gov/nuccore/AE01029;
13) Nanoarchaeum equitans Kin4-M, complete genome, 490885 bp, accession AE017199 AACL01000000 AACL01000001 version AE017199.1, https://www.ncbi.nlm.nih.gov/nuccore/AE017199.1?report=genbank;
14) Syntrophus aciditrophicus SB, complete genome, 3179300 bp, accession CP000252, https://www.ncbi.nlm.nih.gov/nuccore/CP000252.1?report=genbank;
15) Streptomyces coelicolor A3(2) complete genome, 8667507 bp, accession AL645882, https://www.ncbi.nlm.nih.gov/nuccore/AL645882.2?report=genbank;
16) Sulfolobus solfataricus strain SULA, complete genome, 2727337 bp, accession CP011057, https://www.ncbi.nlm.nih.gov/nuccore/CP011057.1?report=genbank;
17) Treponema denticola SP33 supercont1.1, whole genome shotgun sequence, NCBI Reference Sequence: NZ_KB442453.1, 1850823 bp, accession NZ_KB442453 NZ_AGDZ01000000 version NZ_KB442453.1, https://www.ncbi.nlm.nih.gov/nuccore/NZ_KB442453.1?report=genbank
18) Thermotoga maritima strain Tma200, complete genome, 1859582 bp, accession CP010967, https://www.ncbi.nlm.nih.gov/nuccore/CP010967.1?report=genbank;
19) Thermus thermophilus DNA, complete genome, strain: TMY, 2121526 bp, accession AP017920, https://www.ncbi.nlm.nih.gov/nuccore/AP017920.1?report=genbank.

Data in Tables 7 and 8 testify that, in the prokariotic genomes, frequencies of each of kinds of nucleotides A, T, C and G have practically constant values in all considered single-stranded and double-stranded DNA epi-chains. Corresponding frequencies of nucleotides in single-stranded and double-stranded DNA epi-chains are approximately the same. All this for the prokariotic genomes is similar to what was obtained for described eukariotic genomes.

5. The formulation of rules of nucleotide frequencies in DNA epi-chains

This article presents results of analysing mainly epi-chains in long DNA sequences whose length is more than 1000000 (one million) bp. For such long sequences the above presented results testify in favor of existence of the following four universal genetic rules for eukariotic and prokariotic genomes relating to their epi-chains of n-orders (n = 1, 2, 3, 4, 5, … but not extremely large):

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 May 2019 doi:10.20944/preprints201904.0011.v2

- **The rule № 1**: for each of eukariotic and prokariotic genomes, frequencies %A, %T, %C and %G of each kind of nucleotides A, T, C and G are approximately constant in all long single-stranded and double-stranded DNA epi-chains of *n*-orders (*n* = 1, 2, 3, 4, 5, …… but not extremely large);
- **The rule № 2**: for each of eukariotic and prokariotic genomes, in all its long single-stranded and double-stranded DNA epi-chains of *n*-orders, frequencies of nucleotides A and T are approximately equal (%A ≈ %T) and frequencies of nucleotides C and G are also approximately equal (%C ≈ %G) (this rule for long DNA epi-chains is analogous to the second Chargaff's rule for long DNA strands);
- **The rule № 3**: for each of eukariotic and prokariotic genomes, its single-stranded and double-stranded DNA epi-chains of *n*-orders little differ each other from the standpoint of values of their nucleotide frequencies %A, %T, %C and %G;
- **The rule № 4**: in each of eukariotic genomes, all its nuclear chromosomes little differ each other from the standpoint of values of their nucleotide frequencies %A, %T, %C and %G in single-stranded and double-stranded DNA epi-chains.

Of course, at this stage of research, these rules are only candidates for the role of the universal genetic rules; further research of wider sets of genomes is needed to establish the degree of their universality. In the described study, the author met no one example of genomes with a violation of these rules till now. The question of the minimal length of long DNA sequences, from which value these rules of epi-chains begin to be true, is open and it needs additional study.

These rules belong to the field of quantum biology and support the Jordan's idea that «*life's missing laws were the rules of chance and probability (the indeterminism) of the quantum world that were somehow scaled up inside living organisms*» [McFadden, Al-Khalili, 2018]. These rules tell about probabilities (or frequencies), which is the main term in the language of quantum mechanics and quantum informatics.

The presented results demonstrate that long DNA sequences are not ramdom at all but they are regular constructions with their rich nets (or lattices) of epi-chains interrelated with each other. Integer numbers n = 1, 2, 3, 4,… play essential role in these regular constructions, which can be termed as harmonical in this relation. Long DNA sequences demonstrate themselves as quantum-mechanical entities, in which there is a long-range interaction among nucleotides regarding their arrangements.

The arrangement of nucleotides in long DNA is not at all arbitrary, but organized in such a way that the noted rules for nucleotide frequencies are implemented on its wide set of DNA epi-chains. This is directly related to the fundamental issue of the structural regularities in long DNA sequences, which are known today due to achivements of molecular biology. The modern situation in the theoretic field of genetic informatics, where many millions of nucleotide sequences are described, can be characterized by the following citation: "*We are in the position of Johann Kepler when he first began looking for patterns in the volumes of data that Tycho Brahe had spent his life accumulating. We have the program that runs the cellular machinery, but we know very little about how to read it.*" [Fickett & Burks, 1989].

6. Regarding frequencies of triplets in DNA epi-chains of different orders

This Section presents some author's results on studying frequencies of each of 64 triplets in long DNA epi-chains of various orders. These results testify in favor that, in relation to triplets, all epi-chains of each concrete order have practically identical frequencies of anindividual type, which can be used as an individual characteristic of the order of epi-chains (in contrast to the above described situation with frequencies of separate kinds of nucleotides, which are practically identical in epi-chains of different orders). Such differences in triplets frequencies allow additional speculating on participation of epi-chains of various orders in genetic processes.

Table 8 shows frequencies of each of 64 triplets in epi-chains of various orders in the human chromosome № 1.

Table 8. Frequencies of 64 triplets in single-stranded DNA epi-chains of the human chromosome №1, which contains 248956422 bp. The cases of epi-chains of the first order (N_1), the second order ($N_{2/1}$, $N_{2/2}$), the third order ($N_{3/1}$,$N_{3/2}$, $N_{3/3}$) and the fourth order ($N_{4/1}$, $N_{4/2}$, $N_{4/3}$, $N_{4/4}$) are shown. All values are rounded to four decimal places. Epi-chains of each of orders are marked by an individual color. Initial data relating to the chromosome were accessed from https://www.ncbi.nlm.nih.gov/nuccore/NC_000001.11.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 May 2019 doi:10.20944/preprints201904.0011.v2

	N_1	$N_{2/1}$	$N_{2/2}$	$N_{3/1}$	$N_{3/2}$	$N_{3/3}$	$N_{4/1}$	$N_{4/2}$	$N_{4/3}$	$N_{4/4}$
CCC	0.0138	0.0122	0.0122	0.0138	0.0138	0.0138	0.0124	0.0124	0.0124	0.0124
CCG	0.0029	0.0123	0.0123	0.0103	0.0103	0.0103	0.0109	0.0110	0.0110	0.0110
CCT	0.0185	0.0131	0.0131	0.0126	0.0127	0.0127	0.0135	0.0134	0.0134	0.0134
CCA	0.0188	0.0116	0.0116	0.0121	0.0121	0.0121	0.0119	0.0119	0.0119	0.0119
CGC	0.0025	0.0113	0.0113	0.0107	0.0107	0.0107	0.0117	0.0117	0.0117	0.0117
CGG	0.0029	0.0122	0.0122	0.0102	0.0103	0.0103	0.0109	0.0109	0.0109	0.0109
CGT	0.0026	0.0111	0.0111	0.0133	0.0133	0.0133	0.0123	0.0124	0.0123	0.0124
CGA	0.0023	0.0132	0.0132	0.0122	0.0121	0.0122	0.0125	0.0125	0.0125	0.0125
CTC	0.0176	0.0139	0.0140	0.0137	0.0137	0.0137	0.0130	0.0130	0.0130	0.0129
CTG	0.0209	0.0116	0.0116	0.0110	0.0110	0.0110	0.0103	0.0104	0.0103	0.0103
CTT	0.0201	0.0186	0.0186	0.0175	0.0175	0.0175	0.0180	0.0180	0.0180	0.0180
CTA	0.0127	0.0149	0.0150	0.0149	0.0149	0.0149	0.0153	0.0152	0.0152	0.0152
CAC	0.0152	0.0112	0.0113	0.0127	0.0127	0.0127	0.0130	0.0130	0.0130	0.0130
CAG	0.0210	0.0114	0.0114	0.0108	0.0109	0.0109	0.0102	0.0102	0.0103	0.0102
CAT	0.0179	0.0137	0.0137	0.0164	0.0164	0.0164	0.0161	0.0162	0.0161	0.0162
CAA	0.0186	0.0161	0.0161	0.0161	0.0162	0.0162	0.0165	0.0165	0.0164	0.0164
GCC	0.0125	0.0109	0.0109	0.0094	0.0094	0.0094	0.0104	0.0104	0.0104	0.0104
GCG	0.0025	0.0114	0.0114	0.0107	0.0107	0.0107	0.0116	0.0116	0.0116	0.0116
GCT	0.0144	0.0116	0.0116	0.0106	0.0105	0.0106	0.0108	0.0108	0.0108	0.0108
GCA	0.0146	0.0104	0.0104	0.0105	0.0105	0.0105	0.0116	0.0116	0.0117	0.0117
GGC	0.0126	0.0109	0.011	0.0093	0.0094	0.0093	0.0104	0.0104	0.0103	0.0103
GGG	0.0138	0.0121	0.0121	0.0137	0.0137	0.0137	0.0124	0.0123	0.0124	0.0124
GGT	0.0119	0.0129	0.0129	0.012	0.0119	0.0119	0.0129	0.0129	0.0129	0.0129
GGA	0.016	0.0133	0.0133	0.0139	0.0138	0.0138	0.0131	0.0131	0.0131	0.0131
GTC	0.0096	0.0112	0.0112	0.0122	0.0122	0.0122	0.0116	0.0116	0.0117	0.0116
GTG	0.0153	0.0114	0.0114	0.0128	0.0128	0.0127	0.0132	0.0132	0.0131	0.0132
GTT	0.0145	0.0165	0.0165	0.0162	0.0162	0.0162	0.0165	0.0165	0.0165	0.0165
GTA	0.0112	0.0149	0.0148	0.0167	0.0166	0.0167	0.0165	0.0166	0.0165	0.0165
GAC	0.0096	0.0114	0.0114	0.0123	0.0123	0.0123	0.0116	0.0115	0.0115	0.0116
GAG	0.0176	0.0138	0.0138	0.0139	0.0138	0.0138	0.013	0.0131	0.0131	0.0130
GAT	0.0133	0.0168	0.0168	0.0162	0.0162	0.0162	0.0160	0.0160	0.0160	0.0160
GAA	0.0196	0.0193	0.0193	0.0186	0.0186	0.0186	0.0174	0.0173	0.0173	0.0173
TCC	0.0159	0.0133	0.0132	0.0138	0.0138	0.0138	0.0131	0.0131	0.0131	0.0131
TCG	0.0023	0.0130	0.0130	0.0122	0.0122	0.0122	0.0124	0.0124	0.0125	0.0124
TCT	0.0223	0.0190	0.0191	0.0183	0.0183	0.0183	0.0170	0.0170	0.0170	0.0171
TCA	0.0196	0.0158	0.0159	0.0164	0.0164	0.0164	0.0154	0.0154	0.0154	0.0154
TGC	0.0146	0.0104	0.0104	0.0107	0.0107	0.0107	0.0117	0.0117	0.0117	0.0116
TGG	0.019	0.0118	0.0118	0.0122	0.0122	0.0122	0.0120	0.0120	0.0120	0.0120
TGT	0.0199	0.0147	0.0147	0.0171	0.0171	0.0171	0.0171	0.0171	0.0170	0.0171
TGA	0.0195	0.0160	0.0160	0.0166	0.0166	0.0166	0.0154	0.0154	0.0154	0.0153
TTC	0.0197	0.0193	0.0193	0.0185	0.0184	0.0185	0.0174	0.0174	0.0174	0.0174
TTG	0.0188	0.0162	0.0162	0.0163	0.0163	0.0163	0.0166	0.0166	0.0166	0.0166
TTT	0.0372	0.0374	0.0374	0.0321	0.0321	0.0322	0.0340	0.0341	0.0340	0.0341
TTA	0.0198	0.0258	0.0258	0.0249	0.0249	0.025	0.0263	0.0263	0.0263	0.0263
TAC	0.011	0.0149	0.0149	0.0167	0.0167	0.0167	0.0166	0.0166	0.0166	0.0166
TAG	0.0128	0.0150	0.0150	0.0149	0.0149	0.0149	0.0152	0.0152	0.0152	0.0152
TAT	0.0194	0.0231	0.0231	0.0262	0.0263	0.0263	0.0253	0.0253	0.0254	0.0254
TAA	0.0199	0.0257	0.0257	0.0248	0.0249	0.0248	0.0262	0.0263	0.0262	0.0262
ACC	0.0118	0.0128	0.0128	0.0118	0.0118	0.0118	0.0129	0.0129	0.0128	0.0129
ACG	0.0025	0.0111	0.0111	0.0133	0.0133	0.0133	0.0125	0.0125	0.0125	0.0124
ACT	0.0162	0.0153	0.0153	0.0155	0.0156	0.0156	0.0153	0.0153	0.0153	0.0152
ACA	0.0198	0.0146	0.0146	0.0172	0.0172	0.0172	0.0168	0.0169	0.0169	0.0168

AGC	0.0144	0.0116	0.0115	0.0104	0.0104	0.0104	0.0108	0.0108	0.0107	0.0108
AGG	0.0185	0.0131	0.0131	0.0127	0.0127	0.0126	0.0134	0.0133	0.0134	0.0133
AGT	0.0161	0.0152	0.0152	0.0155	0.0156	0.0155	0.0154	0.0154	0.0154	0.0155
AGA	0.0224	0.0188	0.0188	0.0183	0.0183	0.0183	0.0169	0.0170	0.0169	0.0169
ATC	0.0132	0.0168	0.0167	0.0163	0.0163	0.0163	0.0159	0.0159	0.0159	0.0159
ATG	0.0178	0.0138	0.0137	0.0165	0.0165	0.0165	0.0160	0.0160	0.0161	0.0160
ATT	0.0239	0.0263	0.0262	0.0259	0.0260	0.0261	0.0258	0.0258	0.0259	0.0259
ATA	0.0194	0.0232	0.0231	0.0262	0.0262	0.0262	0.0254	0.0253	0.0254	0.0254
AAC	0.0145	0.0163	0.0164	0.0162	0.0162	0.0162	0.0163	0.0163	0.0163	0.0163
AAG	0.0199	0.0185	0.0185	0.0173	0.0173	0.0172	0.0180	0.0180	0.0180	0.0180
AAT	0.0238	0.0264	0.0263	0.0260	0.0260	0.0260	0.0257	0.0257	0.0257	0.0257
AAA	0.0369	0.0373	0.0373	0.0318	0.0317	0.0318	0.0339	0.0339	0.0339	0.0339

These tabular data show, for example, that frequencies P(CCC) of the triplet CCC are different in the single-stranded DNA epi-chains of various orders but practically identical in all epi-chains of the same order:

- In the epi-chain of the first order N_1, P(CCC) = 0,0138;
- In both epi-chains of the second order $N_{2/1}$ and $N_{2/2}$, P(CCC) = 0,0122;
- In all three epi-chains of the third order $N_{3/1}$,$N_{3/2}$ and $N_{3/3}$, P(CCC) = 0,0138;
- In all four epi-chains of the fourth order $N_{4/1}$, $N_{4/2}$, $N_{4/3}$ and $N_{4/4}$, P(CCC) = 0,0124.

Frequencies of all other triplets in Table 8 show a similar strong connection with orders of epi-chains; inside the considered epi-chains of the same order, these frequencies have only small fluctuations in the fourth decimal place.

But what a situation is realized in double-stranded DNA epi-chains of different orders for frequencies of 64 triplets? Corresponding data in Table 9 show that the case of double-stranded epi-chains has similar regularities for triplets frequencies like the case of single-stranded epi-chains (though values of frequencies of triplets differ in both cases).

Table 9. Frequencies of 64 triplets in double-stranded DNA epi-chains of the human chromosome №1, which contains 248956422 bp. The cases of epi-chains of the first order (N_1), the second order ($N_{2/1}$, $N_{2/2}$), the third order ($N_{3/1}$,$N_{3/2}$, $N_{3/3}$) and the fourth order ($N_{4/1}$, $N_{4/2}$, $N_{4/3}$, $N_{4/4}$) are shown. All values are rounded to four decimal places. Epi-chains of each of orders are marked by an individual color. Initial data relating to the chromosome were accessed from https://www.ncbi.nlm.nih.gov/nuccore/NC_000001.11.

	N_{1ab}	$N_{2/1a}$	$N_{2/2a}$	$N_{3/1a}$	$N_{3/2a}$	$N_{3/3a}$	$N_{4/1a}$	$N_{4/2a}$	$N_{4/3a}$	$N_{4/4a}$
CCC	0.0025	0.0113	0.0113	0.0107	0.0107	0.0107	0.0116	0.0116	0.0117	0.0117
CCG	0.0077	0.0116	0.0115	0.0098	0.0098	0.0098	0.0107	0.0106	0.0106	0.0107
CCT	0.0086	0.0108	0.0108	0.0119	0.0119	0.0119	0.012	0.012	0.012	0.012
CCA	0.0083	0.0124	0.0124	0.0114	0.0113	0.0114	0.0116	0.0117	0.0116	0.0117
CGC	0.0138	0.0121	0.0121	0.0137	0.0138	0.0137	0.0124	0.0124	0.0124	0.0124
CGG	0.0077	0.0116	0.0117	0.0098	0.0098	0.0098	0.0106	0.0107	0.0107	0.0106
CGT	0.0173	0.0132	0.0132	0.0133	0.0133	0.0133	0.0132	0.0133	0.0133	0.0133
CGA	0.0153	0.0122	0.0122	0.012	0.012	0.012	0.0124	0.0124	0.0124	0.0125
CTC	0.0153	0.0113	0.0113	0.0127	0.0127	0.0127	0.0131	0.0131	0.0131	0.0131
CTG	0.0153	0.0113	0.0113	0.0115	0.0115	0.0116	0.0109	0.0109	0.011	0.0109
CTT	0.0145	0.0143	0.0143	0.0166	0.0165	0.0166	0.0163	0.0163	0.0163	0.0163
CTA	0.0165	0.0163	0.0163	0.0162	0.0162	0.0162	0.0165	0.0165	0.0164	0.0164
CAC	0.0176	0.0139	0.0139	0.0138	0.0137	0.0138	0.013	0.013	0.013	0.013
CAG	0.0152	0.0115	0.0115	0.0117	0.0117	0.0117	0.0109	0.0109	0.0109	0.0109
CAT	0.0198	0.019	0.019	0.0181	0.018	0.018	0.0177	0.0177	0.0176	0.0176
CAA	0.013	0.0158	0.0159	0.0155	0.0155	0.0155	0.0156	0.0156	0.0156	0.0156
GCC	0.0077	0.0116	0.0116	0.0098	0.0099	0.0098	0.0106	0.0107	0.0106	0.0106
GCG	0.0138	0.0122	0.0122	0.0137	0.0137	0.0137	0.0124	0.0123	0.0124	0.0124
GCT	0.0153	0.0122	0.0123	0.012	0.012	0.012	0.0124	0.0124	0.0124	0.0124

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 May 2019 doi:10.20944/preprints201904.0011.v2

GCA	0.0173	0.0132	0.0132	0.0133	0.0132	0.0133	0.0133	0.0132	0.0132	0.0133
GGC	0.0077	0.0116	0.0116	0.0098	0.0098	0.0098	0.0107	0.0107	0.0107	0.0107
GGG	0.0025	0.0113	0.0113	0.0107	0.0107	0.0107	0.0117	0.0117	0.0116	0.0116
GGT	0.0083	0.0124	0.0124	0.0114	0.0113	0.0114	0.0116	0.0117	0.0116	0.0116
GGA	0.0086	0.0107	0.0107	0.0119	0.0119	0.0119	0.012	0.012	0.012	0.012
GTC	0.0153	0.0115	0.0115	0.0117	0.0117	0.0117	0.0109	0.0109	0.0109	0.0109
GTG	0.0176	0.0139	0.0139	0.0138	0.0138	0.0138	0.013	0.013	0.013	0.013
GTT	0.013	0.0159	0.0159	0.0156	0.0155	0.0155	0.0157	0.0156	0.0156	0.0156
GTA	0.0199	0.019	0.019	0.018	0.018	0.018	0.0177	0.0176	0.0177	0.0176
GAC	0.0153	0.0113	0.0113	0.0115	0.0116	0.0116	0.0109	0.0109	0.011	0.0109
GAG	0.0153	0.0113	0.0113	0.0127	0.0127	0.0127	0.0131	0.0131	0.0131	0.0131
GAT	0.0165	0.0163	0.0163	0.0162	0.0162	0.0162	0.0165	0.0165	0.0165	0.0165
GAA	0.0145	0.0143	0.0143	0.0165	0.0165	0.0165	0.0163	0.0164	0.0163	0.0164
TCC	0.0086	0.0108	0.0108	0.012	0.012	0.012	0.0121	0.012	0.0121	0.012
TCG	0.0154	0.0123	0.0123	0.012	0.012	0.012	0.0124	0.0124	0.0124	0.0124
TCT	0.0198	0.0146	0.0147	0.0172	0.0172	0.0171	0.017	0.017	0.017	0.0169
TCA	0.0178	0.0157	0.0156	0.0161	0.0161	0.0161	0.0153	0.0153	0.0153	0.0153
TGC	0.0172	0.0132	0.0132	0.0133	0.0133	0.0132	0.0132	0.0132	0.0132	0.0132
TGG	0.0084	0.0123	0.0123	0.0113	0.0113	0.0113	0.0116	0.0116	0.0116	0.0116
TGT	0.0223	0.0189	0.0189	0.0183	0.0183	0.0183	0.0169	0.017	0.017	0.017
TGA	0.0179	0.0155	0.0155	0.0159	0.016	0.016	0.0154	0.0154	0.0154	0.0154
TTC	0.0144	0.0143	0.0143	0.0166	0.0166	0.0166	0.0163	0.0163	0.0163	0.0163
TTG	0.013	0.0159	0.0159	0.0156	0.0157	0.0156	0.0156	0.0156	0.0156	0.0155
TTT	0.0194	0.0232	0.0231	0.0262	0.0263	0.0262	0.0254	0.0253	0.0254	0.0254
TTA	0.0219	0.026	0.026	0.0254	0.0254	0.0255	0.026	0.026	0.026	0.026
TAC	0.0198	0.0189	0.0189	0.0179	0.0178	0.0178	0.0177	0.0177	0.0177	0.0177
TAG	0.0167	0.0163	0.0163	0.0163	0.0162	0.0162	0.0164	0.0165	0.0165	0.0164
TAT	0.0371	0.0373	0.0374	0.0319	0.0319	0.032	0.034	0.034	0.034	0.034
TAA	0.0218	0.0261	0.0261	0.0255	0.0254	0.0255	0.026	0.026	0.026	0.0259
ACC	0.0084	0.0123	0.0123	0.0113	0.0113	0.0113	0.0116	0.0116	0.0116	0.0116
ACG	0.0172	0.0132	0.0132	0.0133	0.0132	0.0132	0.0132	0.0132	0.0132	0.0132
ACT	0.0179	0.0155	0.0156	0.0159	0.016	0.0159	0.0154	0.0154	0.0154	0.0154
ACA	0.0223	0.0189	0.019	0.0183	0.0183	0.0183	0.017	0.017	0.0169	0.017
AGC	0.0154	0.0123	0.0123	0.012	0.012	0.012	0.0124	0.0124	0.0124	0.0125
AGG	0.0086	0.0108	0.0108	0.012	0.012	0.012	0.0121	0.0121	0.0121	0.0121
AGT	0.0179	0.0157	0.0157	0.0161	0.0161	0.016	0.0153	0.0153	0.0153	0.0153
AGA	0.0198	0.0147	0.0146	0.0172	0.0171	0.0172	0.0169	0.017	0.0169	0.017
ATC	0.0166	0.0163	0.0163	0.0162	0.0162	0.0162	0.0164	0.0164	0.0165	0.0165
ATG	0.0198	0.0189	0.0189	0.0179	0.0179	0.0178	0.0177	0.0177	0.0177	0.0177
ATT	0.0218	0.0261	0.026	0.0255	0.0255	0.0255	0.026	0.026	0.0261	0.026
ATA	0.0371	0.0374	0.0374	0.032	0.032	0.032	0.034	0.034	0.0339	0.034
AAC	0.013	0.0159	0.0159	0.0156	0.0156	0.0156	0.0156	0.0155	0.0155	0.0156
AAG	0.0144	0.0143	0.0143	0.0166	0.0166	0.0166	0.0163	0.0163	0.0163	0.0163
AAT	0.0219	0.026	0.026	0.0254	0.0254	0.0254	0.0260	0.0261	0.0261	0.0261
AAA	0.0194	0.0231	0.0232	0.0262	0.0262	0.0262	0.0254	0.0254	0.0254	0.0254

7. Resonances and some scientific analogies refered to DNA epi-chains

Each DNA epi-chain is a complex vibrational system having many degrees of freedom and correspondingly having many resonance frequencies. The author believes that genetic role of epi-chains are connected with their physical essence as vibrational systems having many resonance frequencies. Concerning importance of resonances in quantum mechanical systems, E. Schrodinger wrote: "*The one thing which one has to accept and which is the inalienable consequence of the wave-equation as it is used in*

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 May 2019 doi:10.20944/preprints201904.0011.v2

every problem, under the most various forms, is this: that the interaction between two microscopic physical systems is controlled by a peculiar law of resonance» [Schrodinger, 1952, p.115]. Structures of molecular genetics are deeply connected with mathematical formalisms of the theory of resonances of vibrational systems having many degrees of freedom [Petoukhov, 2015a,b,c, 2016a; Petoukhov, Petukhova, 2017]. From the point of view of quantum mechanics, the interaction of molecules is based on the emission and absorption of photons with the participation of resonance correspondences. The energy state of parts of molecular-genetic systems depends on the emission and the absorption of photons, which are the force carriers for electromagnetic field (see some details in [Petoukhov, 2018b; Hu, Petoukhov, Petukhova, 2018]).

The creator of the theory of resonances in structural chemistry L. Pauling wrote about an important meaning of resonances in organization of living matter [Pauling, 1940]. Pauling's theory uses the fundamental principle of a minimal energy because – in resonant combining of parts into a single unit – each of members of the ensemble requires less energy for performing own work than when working individually. Of course, this fundamental principle can be used in many other cases of resonances in different systems as the physical base. The principle of energetic minimum in resonance processes has some correlations with the principle of relaxation in morphogenetic processes proposed in [Igamberdiev, 2012]. From the point of view of quantum mechanics, the interaction of molecules is based on the emission and the absorption of photons with the participation of resonance correspondences.

An organism during its life on genetic basis should solve algorithmic problems of two types: 1) informational, providing coordinated energy processes; 2) energetic, providing information processes. Systems of resonances can be used as a common basis of such "two-faced" algorithms since resonances are associated both with oscillatory energy and with informatics of communications among objects. The notion “resonance” was introduced into quantum mechanics by W. Heisenberg in 1926 year in connection with analyzes of multi-body systems. He emphasized that in quantum mechanics the phenomenon of resonances has much more general character than in classical physics. In classic theory, two periodic oscillating systems come into their own resonance only in the case when a frequency of a separate sub-system doesn’t depend on energy of the system and when this frequency is approximately equal in both sub-systems. In quantum mechanics, two atomic systems come into their resonance only in the case when a frequency of absorption of one system coincides with a frequency of emitting another system, or vice versa [Heisenberg, 1926, §2]. Quantized electromagnetic field is represented as a set of oscillators. The book [Ji, 2017] contains new data about a role of resonances in biological structures.

Let us consider some possible approaches for analising and understanding a genetic role of DNA epi-chains taking into account analogies with known data from different branches of science.

- 7.1. DNA epi-chains and helical and fractal antennas

Many authors have suggested that DNA double helices can serve as electromagnetic wave antennas for communication (see for example a discussion in [Blank, Goodman, 2011; Foster, 2011]). Regarding this topic, it is necessary to recall the important role of helical antennas in communication technology for space communications, radar, cellular telephony and much more. Helixal antenna theory and applications are described in many books, for example, in [Kraus, 1988; Kraus, Marhefka, 2002; Stutzman, Thiele, 1998]). Fig. 4 shows examples of photographs of technical helical antennas.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 May 2019 doi:10.20944/preprints201904.0011.v2

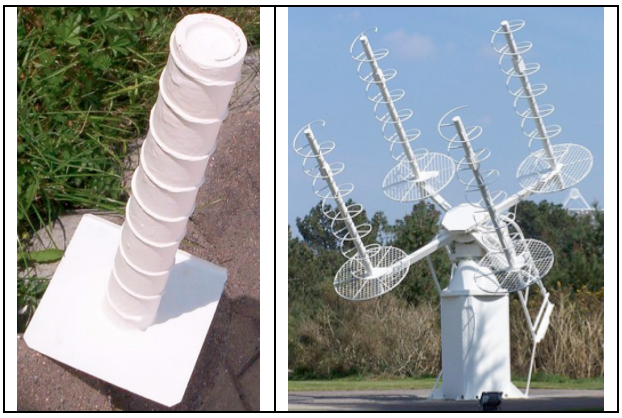


Fig. 4. Examples of technical helical antennas (photos are taken from web sites https://ru.wikipedia.org/wiki/Спиральная_антенна and https://en.wikipedia.org/wiki/Helical_antenna).

Let us remind a few known facts about helical antennas, which can relate to the topic of DNA helices as possible helical antennas. Helical antennas are structurally insensitive to manufacturing errors in essential degree and they can operate in one of two principal modes - axial mode or normal mode. In the *axial mode* helical antenna functions as a directional antenna radiating a beam off the ends of the helix, along the antenna's axis (in this mode the diameter and pitch of the helix are comparable to a wavelength). It radiates circularly polarized radio waves, which distinguishes it from other antennas with directional radiation. In radio transmission, circular polarisation is often used where the relative orientation of the transmitting and receiving antennas cannot be easily controlled, such as in animal tracking and spacecraft communications (that is, for example, a spacecraft rotation does not influence on the communication). At present, intensive study of quantum polarization states is being conducted all over the world for successful mastering of quantum information technologies, especially in the field of quantum communication.

The direction of rotation of the circular polarization of the space receiving-transmitting antenna must coincide with the direction of rotation of the ground receiving and transmitting antenna operating from the space antenna. In space communication, polarization isolation is used, that is, antennas of opposite directions of rotation of polarization operate at the same frequency. In other words, regarding these helical antennas, the factor of chirality (left or right polarization) is very important for communication. But the problem of chirality (or biological dissymmetry) also quite important for molecular biological systems as known beginning from famous experiments by Louis Pasteur in 1848 (see for example [Darvas, 2007; Hargittai, Hargittai, 2003; Kizel, 1980]). Helical chiral biomolecules emit and absorb the electromagnetic waves of the corresponding circular polarization. This provides the opportunity for helical chiral biomolecules to exchange radio waves of the corresponding circular polarization selectively with helical biomolecules of the same kind of chirality. The author believes that the principle of the chiral stereochemical organisation of biomolecules in living nature is deeply related with the informational principle of communication among biomolecules on the basis of electromagnetic waves of appropriate circular polarization inside biological bodies. Existence of chirality in biomolecular structures and also in their electromagnetic waves of appropriate circular polarization are very significant for farmacology and some methods of physiotherapy. It should be added that many species have the useful informational ability to perceive the polarization light by their organs of sight (some birds, insects, etc.).

Here it should be recalled that not only DNA molecule has its helical configuration but many proteins have helical substructures, known as alpha helices. Moreover helices are termed long ago as «*curves of life*» due to the multiple implementation of inherited helical structures and processes in living bodies on various lines and levels of biological evolution [Cook, 1914; Petoukhov, Svirin, Khazina, 2015]. For example in the human body, helical structures are genetically inherited from generation to generation in muscles, heart, vessels, bones, tendons, ligaments, nerves, organ of hearing (cochlea ear), etc. Science reveals new and new facts about helices in biological structures. The recent work [Zabeo et al., 2018] describes a discovery of left-handed helices in tips of human spermatozoon tails; its authors speculate that these helical structures in particularly can «*play a role in controlling the swimming direction of spermatozoa*» that is play a role of antennas for communication with environment.

The discovery of the described phenomenon of DNA epi-chains generates new thoughts and questions to the theme of DNA molecules as helical antennas of electromagnetic waves of appropriate

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 May 2019 doi:10.20944/preprints201904.0011.v2

circular polarizations (as known, the A and B forms of DNA are right-handed helices and the Z form of DNA is left-handed). This phenomenon attracts additional attention to a cooperative quantum essence of long DNA sequences in eukariotic and prokarioric genomes. By contrast to simple ordinary helical antennas known in technology, each long DNA helix is characterized by a great system of its receiving and transmitting elements in form of nitrogenous bases A, T, C and G, which are located along the DNA helix and each of which can be considered conditionally as an emmiter and an absorber of certain electromagnetic waves.

The distance between adjacent nucleotides in DNA epi-chains depends on the order *n* of epi-chains. Can individual DNA epi-chains of different orders function as separate helical antennas in certain circumstances? Can electromagnetic waves of circular polarizations and different wave-frequencies, arriving from outside in DNA, interact in various ways with nucleotide sets on epi-chains of different orders *n*? Are epi-chains of different orders *n* intended for individual or cooperative work with certain types of electromagnetic waves of circular polarization? These are only a few examples of many new open questions in this theme for future study.

When considering DNA helices as analogues of helical antennas, it should be taken into account that DNA is a dynamic structure capable of greatly changing its spatial configuration in different conditions. By this reason they should be considered as dynamic antennas with parameters varied in time. Because of this dynamism, it seems that DNA is in constant searching and production of wave information with circular polarizations.

According to described data on epi-chains in long DNA sequences, all nucleotides are united by common regularities of frequencies (or percentages) of their spatial distribution along DNA on different epi-chains and, apparently, can be united by a common cooperative state. It seems probable that in some situation such a cooperative system can operate in a pulsed mode like a laser, including with alternating impulse work of epi-chains.

The described data on DNA epi-chains show that long helical DNA nucleotide sequences have fractal-like features. In this regard, the attention of researchers seeking to understand the functioning of DNA as an antenna structure should also be additionally drawn to fractal antennas, which are another important type of antenna with many applications in technology. There are many publications on the theory and applications of fractal antennas and on fractal methods of information transfer (see, for example, [Potapov, 2000, 2005]).

- 7.2. DNA epi-chains and the Fröhlich theory of long-range coherence in biological systems

Described numerical characteristics of DNA nucleotide epi-chains give new evidences in favor of long-range coherence in biological systems. This Section discusses possible relations between the collective features of DNA nucleotide epi-chains and the known Fröhlich's theory of long-range coherence or collective quantum effects in biological systems. The foundations of this theory were presented in works [Fröhlich, 1968, 1969,1970, 1980, 1988; Fröhlich, Kremer, 1983]. Many other works consider the Fröhlich's theory and its applications, for example [Grundler, Keilmann, 1983; Hyland, 1998, 2009; Kadantsev, Goltsov, 2018; Lockwood, 1989; Lundholm et al, 2015; Marshall, 1989; Mesquita, Vasconcellos, Luzzi, 1993; Penrose, 1994; Vasconcellos et al, 2012].

Let me remind about this Fröhlich's theory to those readers, who are not familiar with it, using data and quotations from these works beginning from the book [Penrose, 1994, p. 352]: «*the distinguished physicist Herbert Frohlich (who, in the 1930s had made one of the fundamental breakthroughs in the understanding of 'ordinary' low-temperature superconductivity) suggested a possible role for collective quantum effects in biological systems. ... Frohlich was led to propose, in 1968, that there should be vibrational effects within active cells, which would resonate with microwave electromagnetic radiation, at* 10^{11} *Hz, as a result of a biological quantum coherence phenomenon. Instead of needing a low temperature, the effects arise from the existence of a large energy of metabolic drive. There is now some respectable observational evidence, in many biological systems, for precisely the kind of effect that Frohlich had predicted in 1968*".

According to Fröhlich, under appropriate conditions a phenomenon quite similar to a Bose condensation may occur in substances which possess longitudinal electric modes. If energy is fed into these modes and thence transferred to other degrees of freedom of the substance, then a stationary state will be reached in which the energy content of the electric modes is larger than in the thermal equilibrium. This excess energy

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 May 2019 doi:10.20944/preprints201904.0011.v2

is found to be channelled into a single mode - exactly as in Bose condensation - provided the energy supply exceeds a critical value. Under these circumstances a random supply of energy is thus not completely thermalized but partly used in maintaining a coherent electric wave in the substance [Fröhlich, 1968; Vasconcellos et al, 2012].

"*Fröhlich's results are based on the idea that alive biological systems are open and very far from equilibrium and have considerable amounts of energy available, through metabolic processes, that cause non-linear changes in molecules and larger biological subsystems. ... In Fröhlich model vibrational-polar modes are excited by a continuous supply of energy pumped by an external source, while these modes interact with the surrounding medium acting as a thermal bath. The interplay of these two effects—pumping of energy subtracting entropy from the system and dissipative internal effects adding entropy to the system—may lead to the emergence of complex behaviour in the system consisting in what can be called Fröhlich effect: Provided the energy supply is sufficiently large compared with the energy loss, the system attains a stationary state in which the energy that feeds the polar modes is channelled into the modes with the lowest frequencies. The latter largely increase their populations at the expenses of the other higher-in-frequency modes, in a way reminiscent of a Bose-Einstein condensation. Fröhlich's synchronous large-scale collective oscillations imply inter-cellular microwave emissions, which would constitute a non-chemical and non-thermal interaction between cells. These oscillations could therefore be revealed by detection of emissions of GHz or THz radiation»* [Vasconcellos et al, 2012].

Some evidence of a non-thermal influence of coherent microwave radiation on the genome conformational state in *E. coli* has been reported [Hyland, 1998], which may indicate that «*chromosomal DNA could be the target of mm microwave irradiation within this system. Also low intensity microwave irradiation of leukocytes results in a significant increase in bio-photon emission in the optical range, the origin of which is thought to involve DNA. Also it is worth noticing the possible influence of the concept of bio-coherence on the very particular dipolar system, which is water. It can be considered the possibility that biological water might itself support coherent dipolar excitations extending over mesoscopic regions; thus water, instead of passive space-filling solvent, would be raised to an important singular position whose full significance has yet to be elucidated»* [Vasconcellos et al, 2012].

The work [Lundholm et al, 2015] presents the experimental observation of Frohlich condensation in a protein structure. It describes that 0.4 THz electromagnetic radiation induces non-thermal changes in electron density. In particular, the authors observed a local increase of electron density in a long α-helix motif consistent with a subtle longitudinal compression of the helix. The analysis shows that the received experimental results "*can only be explained by Frohlich condensation, a phenomenon predicted almost half a century ago*".

The Fröhlich theory of long-range coherence in biological systems and its consequences have relations to medical diagnosis [Hyland, 1998]. The book [Penrose, 1994] considers the Frohlich theory as related to brain functioning and artificial intelligence. Taking into account the Hameroff's hypothesis [Hameroff, 1974, 1987] that microtubules in the cytoskeletons of neurons might act as "*dielectric wavequides*", Penrose supposed a subordination of microtubules to similar quantum-coherent behavior: "*Recall Frohlich's idea that large-scale collective quantum phenomenon – perhaps of the nature of a Bose-Einstein condensate – is a definite biological possibility, even within the 'hot' brain. Here we envisage that not only must single microtubules be involved in a relatively large-scale quantum-coherent state, but that such a state must extend from one microtubule to the next. Thus, not only must this quantum coherence stretch to the entire microtubule (and we recall that microtubules can extend to considerable length), but a good many of the different microtubules in the cytoskeleton within a neuron, if not all of them, must together take part in this same quantum-coherent state. Not only this, but the quantum coherence must leap the synaptic barrier between neuron and neuron. It is not much of a globality if it iinvolves only individual cells! The unity of a single mind can arise, in such a description, only if there is some form of quantum coherence extending across at least an appreciable part of the entire brain"*.

F. Fröhlich (the son of Herbert Fröhlich) wrote the article "Genetic code as Language" regarding quantum coherence states and long-range communication in genomes [Fröhlich F., 1988]. He noted: "*Beyond the chromosome, the genome as a whole must contain some sort of long-distance communication in order not to produce on one section antigens produced on other parts. So there is a complexity of entities which which use genetic language in different ways. There might be said to be a logic of cells"* [Fröhlich F., 1988, p. 194]. The H. Fröhlich's hypothesis *"suggests that long-range coherent vibrations will lead to resonance between a differentiated cell with its own characteristic vibrations and the chromosome such that*

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 May 2019 doi:10.20944/preprints201904.0011.v2

the chromosome-particular region responding to this characteristic frequency will be activated or opened up so it can produce the appropriate proteins. Such a resonance could transport the embryological, already partially induced cells to their target and there they would be further fixed into producing the correct proteins for this organ through superimposed resonance and why if it were too mature, it would not so adapt. There might be degrees of resonance" [ibid, p. 199]. *"This have been done more specifically in the case of cell division during mitosis, where it has been hypothesized that corresponding chromosomes line up through resonance having the same frequency, finding each other from the apparenr confusion of the genome by long-range communication [Holland, 1972]. Resonance oscillations draw like to like. More generally there might be some kind of coherent long-range interaction among the members of the genome creating the self-marking. If this broke down through subsequent muataions, auto-immune diseases would arise"* [ibid, 203].

Works [Petoukhov, 2019a,b] should be added to the topic of quantum coherence and language. These works show an amizing mathematical relationship between two long binary sequences from quite different levels of biological organization: 1) binary sequences of two kinds of hydrogen bonds 2 and 3 of complementary nitrogenous bases in long DNA sequences of different species; 2) binary sequences of letters from two phonetic groups of the Russian alphabet in novels by L.Tolstoy, F.Dostoevsky, A.Pushkin, etc (in this case all phonetic significant letters of the Russian alphabet are separated on two equivalence classes and denoted by symbols 0 or 1: the first class contains all short (iotated) vowels and all deaf consonants (е, ё, ю, я, п, ф, к, т, ш, с, х, ц, ч, щ); the second class contains all long vowels and all voiced consonants (а, и, о, у, ы, э, б, в, г, д, ж, з, й, л, м, н, р). This close relationship of both binary sequences draws attention to their hidden rules, which are modelled on the basis of formalisms from quantum mechanics (more precisely by means of the tensor product of vectors of frequencies).

Here one can remind that nitrogenous bases A and T are linked by 2 hydrogen bonds and C and G are linked by 3 hydrogen bond in DNA double helices. Correspondingly DNA can be formally considered as numeric sequences of these two digits like 32232233323…. . It is obvious that all data described above in this article about frequencies of nucleotides A, T, C and G in nucleotide epi-chains of long DNA can be reformulated for "hydrogen bonds epi-chains" in long DNA sequences of hydrogen bonds.

The mentioned analogies between systems of genetic and linguistic information are of wide scientific interest. Some authoritative authors in fields of linguistics and genetics proposed long ago that genetic language is the structural basis of linguistic languages [Jacob et al., 1968; Jakobson, 1985]. According to Jakobson, all relations among linguistic phonemes are decomposed into a series of binary oppositions of elementary differential attributes (or traits). As Jakobson wrote, the genetic code system is the basic simulator, which underlies all verbal codes of human languages. "*The heredity in itself is the fundamental form of communications … Perhaps, the bases of language structures, which are imposed on molecular communications, have been constructed by its structural principles directly*" [Jakobson, 1985, p. 396]. These questions had arisen to Jakobson as consequence of his long-term research into the connections between linguistics, biology and physics. Such connections were considered at a united seminar of physicists and linguists, organized by Niels Bohr and Roman Jakobson, jointly, at the Massachusetts Institute of Technology. The book "Linguistic Genetics" [Makovsky, 1992] noted: "*A look at language as a living organism, subject to the natural laws of nature, ascends to a deep antiquity … Research of a nature, of disposition and of reasons of isomorphism between genetic and linguistic regularities is one of the most important fundamental problems for linguistics of our time*". One can think that mentioned relationships of genetic and linguistic languages are consequences of quantum coherences in biological organisms related with the Fröhlich's theory.

Since long DNA sequences have quantum long-range interaction between elements, quantum entanglement in this quantum genetic system can be essential for its surprisingly organized construction. The thought about an important role of quantum entanglement in DNA organization is not new: see, for example, the article «Quantum entanglement between the electron clouds of nucleic acids in DNA» about the entanglement as a glue for DNA constructions where lattice vibrations or phonons are significant [Rieper, Anders, Vedral, 2011].

Returning to the Jordan's thoughts on quantum biology, the difference between biological and inanimate objects should be explained. Jordan correctly pointed out that inanimate objects were governed by the average random motion of millions of particles, such that the motion of a single molecule has no influence whatsoever on the whole object. This insight is usually credited to Erwin Schrödinger, who later claimed

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 May 2019 doi:10.20944/preprints201904.0011.v2

that life was different from inorganic chemistry because of its dependence on the dynamics of a small number of molecules. Jordan similarly argued that the few molecules that control the dynamics of living cells within the control center have a dictatorial influence, such that quantum-level events that govern their motion, such as Heisenberg's uncertainty principle, are amplified to influence the entire organism. Jordan believed that living organisms were uniquely able to carry out this amplification in a way that was conspicuously different from inanimate matter. Jordan was convinced he could extend quantum indeterminism from the subatomic world to macroscopic biology. He even made a connection with free will by suggesting a link between quantum mechanics and psychology [McFadden, Al-Khalili, 2018].

- **7.3. DNA epi-chains, epigenetics and transformations of DNA configurations**

This paragraph is devoted to initial author's thoughts and suppositions on the possible meaning of DNA nucleotide epi-chains for problems of epi-genetics and changes of DNA configurations.

Since DNA strands carry genetic information, the phenomenon of numerical equalities in the frequencies of nucleotides between DNA strands and their epi-chains of higher orders provokes the following thought: not only DNA strands but also epi-chains and their nucleotide sequences are participants of genetic and epigenetic processes and play an informational genetic and epigenetic role.

Taking into account the theory of Fröhlich condensation described in the previous paragraph, energy status of DNA nucleotide epi-chains may change depending on the current external electromagnetic actions. These epi-chains can be selective accumulators of energy from external electromagnetic waves and their energy level determines a lot in epigenetic phenomena and in DNA transformations. Energetic states and sets of resonance wave-frequencies of DNA epi-chains can be changed also under influence of other external physical and chemical factors, for example, under influence of joining new chemical elements or bonds to certain sites of DNA (the term “wave-frequencies” is used here to distinguish wave frequencies, wave measured in Hertz, from frequencies (or percentages) of nucleotides in DNA chains). DNA epi-chains with their helix configurations and changeable energetic states can be essential for DNA supercoiling. Supercoiling, which needs an additional energy, is important in a number of biological processes, such as compacting DNA, DNA metabolism and possibly gene expression [Irobalieva, Zechiedrich et al., 2015; Kim, Ganji, Kim, et al, 2018; Singer, 2016]. Changing energetic states of epi-chains, which is connected with changing their sets of resonance wave-frequencies, can serve as a manage factor for DNA interaction with proteins and for DNA supercoiling.

In various cells, energetic states and sets of resonance wave-frequencies of their DNA epi-chains can be mechanically depended on a location of separate cells inside a holistic ensemble of cells. Such differences in energetic states and resonance wave-frequencies of DNA epi-chains in various cells can explain the phenomen of a different fate of various cells in ontogenesis, although all cells contain DNA of the same nucleotide composition. This also relates the well-known embryological rule: “*The fate of an embryo part is a function of its position within a whole*” [Beloussov, 2015, p. 4; Driesch, 1921].

The additional argument in favor of a participation of DNA epi-chains in epigenetic phenomena is given by the results described in the Section 6 (Tables 8 and 9) about regular frequencies of triplets in long DNA epi-chains of various orders. The author believe that knowledge about DNA epi-chains and about DNA telomeres allows developing new model versions of senescence. For example, one of possible model approaches considers epi-chains starting at the most distal nucleotides of telomeres. The model assumes that at each life stage a predominant epigenetic activity belongs to the DNA epi-chain, which begins from this most distal nucleotide №1 and has a fixed order among all epi-chains containing this telomeric nucleotide №1. Under cell division, the telomeric ends become shorter and the nucleotide №1 is removed, giving the most distal place in DNA to another nucleotide №2. At this new stage of cell life, the predominant epigenetic activity is transmitted to the following epi-chain, which contains this telomeric nucleotide №2 and has the same fixed order among all epi-chains containing this telomeric nucleotide. This relay process of transmitting activity to new and new epi-chains, which have specifical energetic states and resonance frequencies, repeats with each cell division. Such model approach combines knowledge about DNA telomeres and DNA epi-chains.

The author's study of DNA epi-chains provokes many new questions for further investigations, for example, the following: do epi-chains with their own nucleotide sequences have their own epi-codes or not? Are there nucleotide sequences on DNA epi-chains that correspond to genes on the DNA strand? What will

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 May 2019 doi:10.20944/preprints201904.0011.v2

the comparative analysis of the characteristics of epi-chains in genomes of different organisms for the tasks of evolutionary biology?

The described studies conducted by the author use computer programs developed by his graduate student V.I. Svirin according to the author's technical task.

- 7.4. DNA double helix as a numerical tree of epi-chains with binary representations of their nucleotides

This paragraph shows a natural possibility of binary representations of DNA nucleotides and represents DNA molecules as numerical trees of DNA epi-chains having a form of appropriate binary sequences.

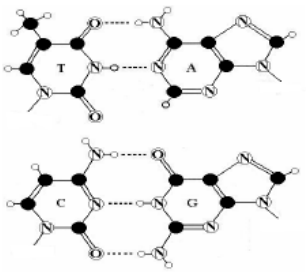
Science does not know why the nucleotide alphabet of DNA has been created by the Nature of just from four letters, and why just these very simple molecules were chosen for DNA-alphabet (out of millions of possible molecules). But science knows (Petoukhov, 2008; Petoukhov and He, 2010; Stambuk, 1999; Fimmel et al., 2013) that these four molecules are interrelated by means of their symmetrical peculiarities into the united molecular ensemble with its three pairs of binary-oppositional traits (Fig. 5):

1) two letters are purines (A and G), and other two are pyrimidines (C and T). From the standpoint of these binary-oppositional traits one can denote C=T=1, A=G=0. From the standpoint of these traits any of nucleotide sequences is represented by a corresponding binary sequence. For example, GCATGAAGT is represented by 010100001;

2) two letters are amino-molecules (A and C) and the other two are keto-molecules (G and T). From the standpoint of these traits one can designate A=C=1, G=T=0. Correspondingly the same sequence GCATGAAGT is represented by another binary sequence 011001100;

3) the pairs of complementary letters A-T and C-G are linked by 2 and 3 hydrogen bonds respectively. From the standpoint of these binary traits one can designate C=G=0, A=T=1. Correspondingly the same sequence GCATGAAGT is read as 001101101.

Accordingly, each of DNA nucleotide sequences is the carrier of three parallel messages on three different binary languages. At the same time, these three types of binary representations form a common logical set on the basis of logical operation of modulo-2 addition ⊕: modulo-2 addition of any two such binary representations of the DNA-sequence coincides with the third binary representation of the same DNA-sequence. For example the mentioned three representations of the sequence GCATGAAGT form a common logical set: modulo-2 addition of any its two binary representations gives its third binary representation , for example, 010100001 ⊕ 011001100 = 001101101. The operation of modulo-2 addition is very important in quantum computing [Nielsen, Chuang, 2010] and this opeartion serves as the group operation in dyadic groups of binary numbers noted below in relation to DNA alphabets of *n*-plets.



№	Binary symbols	C	A	G	T
1	0_1 – purines 1_1 – pyrimidines	1_1	0_1	0_1	1_1
2	0_2 – keto 1_2 – amino	1_2	1_2	0_2	0_2
3	0_3 – three hydrogen bonds; 1_3 – two hydrogen bonds	0_3	1_3	0_3	1_3

	1	0
1	C 11	A 10
0	T 01	G 00

Fig. 5. Left: the four nitrogenous bases of DNA: adenine A, guanine G, cytosine C and thymine T. Middle: three binary sub-alphabets of the nucleotide alphabet on the basis of three pairs of binary-oppositional traits. Right: the definition of DNA-letters C, A, T and G as 2-bit binary numbers from the standpoint of their molecular binary-oppositional traits (the explanation in the text).

Taking into account the phenomenological fact that each of DNA-letters C, A, T and G in DNA *n*-plets is uniquely defined by any two kinds of mentioned molecular binary-oppositional traits or indicators (Fig. 5, middle), these genetic letters can be represented by means of corresponding pairs of binary symbols, for example, from the standpoint of two first binary-oppositional indicators. For example, one can use at the first position of each of DNA-letters its binary symbol from the second pair of binary-oppositional indicators

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 May 2019 doi:10.20944/preprints201904.0011.v2

(the indicators "amino or keto": C=A=0, T=G=1) and at the second positions of each of letters its binary symbol from the first pair of binary-oppositional indicators (the indicators "pyrimidine or purine": C=T=0, A=G=1). In this case the letter G is represented by the binary symbol $0_2 0_1$ (that is as 2-bit binary number), T – by the symbol $0_2 1_1$, A – by the symbol $1_2 0_1$, C – by the symbol $1_2 1_1$ (Fig. 5, right). Using these representations of separate DNA-letters, each of 16 genetic doublets is represented as the concatenation of the binary symbols of its letters (that is as 4-bit binary number): for example, the doublet CC is represented as 4-bit binary number $1_2 1_1 1_2 1_1$, the doublet CA – as 4-bit binary number $0_2 0_1 1_2 0_1$, etc. By analogy, each of 64 triplets is represented as the concatenation of the binary symbols of its letters (that is as 6-bit binary number): for example, the triplet CCC is represented as 6-bit binary number $1_2 1_1 1_2 1_1 1_2 1_1$, the triplet CCA – as 6-bit binary number $1_2 1_1 1_2 1_1 1_2 0_1$, etc. In general, each of *n*-plets is represented as the concatenation of the binary symbols of its letters (below we will not show these indexes 2 and 1 of separate letters in binary representations of *n*-plets but will remember that each of positions corresponds to its own kind of indicators from the first or from the second set of indicators in Fig. 5, middle).

It should be noted that the set of the binary representations of nucleotide *n*-plets forms appropriate dyadic groups of binary numbers, for example:

- the binary representation of 4 nucleotides G, T, A and C corresponds to the dyadic group of 2-bit binary numbers 00, 01, 10, 11;

- the binary representation of 16 doublets GG, GT, …, CC corresponds to the dyadic group of 4-bit binary numbers 0000, 0001, 0010, 0011, 0100, 0101, 0110, 0111, 1000, 1001, 1010, 1011, 1100, 1101, 1110, 1111;

- the binary representation of 64 triplets GGG, GGT, …, CCC corresponds to the dyadic group of 6-bit binary numbers 000000, 000001, …, 111111.

Now it is obvious that each epi-chain of nucleotides can be represented as an appropriate numerical sequence of 2-bit binary numbers. But one can consider each DNA epi-chain not only as the chain of separate nucleotides but also as the appropriate epi-chain of doublets or as the appropriate epi-chain of triplets, etc. In these cases, each DNA epi-chain can be represented and analysed as an appropriate sequence of 4-bit binary numbers, or of 6-bit binary numbers, etc. Various represenations of the DNA double helix are possible as 3D-trees (or hybrids) of binary numbers from different dyadic groups used in computer technologies.

These binary representations of the wide set of DNA epi-chains allow discovering new hidden rules in structural organization of DNA nucleotide sequences. In addition, effective mathematical methods of dyadic analysis, Walsh functions and sequency analysis by Harmuth [Harmuth, 1970, 1977, 1981, 1989] can be apply to study DNA with using notions of multidimensional dyadic spaces, dyadic shifts, dyadic derivative, Hadamard matrices, modulo-2 addition, etc. These mathematical methods have long ago important applications in radio-engineering, acoustics, optics, etc. In comparison with spectral analysis by means of sine waves, which is applicable to linear time-invariant systems, the sequency analysis is based on non-sinusoidal waves and it is used to study systems, which are changed in time (biological systems belong to such systems). In particular, the problem of absorption of radio waves and acoustic waves, which is important for biological systems, is bypassed by means of the sequency analysis. Here one should mention that structural alphabets of doublets and triplets of DNA nucleotides are mathematically connected with Hadamard matrices, Walsh functions, unitary matrices, dyadic shifts, Boolean algebra and united-hypercomplex numerical systems as it was previously shown by the author in a series of publications on matrix genetics, geno-logic coding and algebraical biology [Petoukhov, 2008, 2011a,b, 2013, 2015a-c, 2016a,b, 2017, 2018a,b; Petoukhov, Petukhova, 2017a,b; Petoukhov, Petukhova, Svirin, 2018; Hu, Petoukhov, Petukhova, 2017b; etc.]

In the whole the noted binary representations of DNA epi-chains contibute to identify patents of living nature for storing and processing genetic information; to develop new methods of comparative analysis of genomes for studying biological evolution; to facilitate developing algebraical biology; to create new methods in medicine, biotechnology, quantum computing and artificial intelligence.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 May 2019 doi:10.20944/preprints201904.0011.v2

8. Some concluding remarks

Described numerical and analytical materials testify in favor of the following:

- Any long DNA double helix in eukariotic and prokariotic genomes is a superposition (or a hybride) of many helical nucleotide epi-chains (or nucleotide epi-helices) of different *n*-orders, where numbers *n* belong to the beginning of the series of natural numbers (*n* =1, 2, 3, 4, ... is not too large). These epi-chains in each single-stranded and double-stranded long DNA are regular related with each other since they practically do not differ in values of frequencies (or percentages, or probabilities) of nitrogenous bases A, T, C and G in them. These frequencies values play a role of approximate invariants of each concrete genome, but they differ for genomes of various species in a general case. From the standpoint of existence of these epi-chains in long DNA, the long DNA double helix is a multi-helical design, briefly speaking.
- Four rules of long DNA epi-chains, formulated in the Section 5, are candidates for the role of the universal genetic rules. The further research of wider set of genomes of various species is needed to establish the degree of their universality;
- The content of these rules supports ideas on quantum long-range coherence in DNA and in biological systems declared by the theory of H. Fröhlich and his followers about biological quantum coherence phenomena; in other words, the fact that the frequency characteristics of many long DNA epi-chains in a concrete genome are practically identical indicates quantum coherence;
- Described author's results about long DNA epi-chains of helical configurations give new opportunities to develop theoretical genetics and quantum biology using various achievements in different fields of science and technology including, for example, knowledges about helical antennas with their circular polarizations of electromagnetic waves, which remind on chirality of biological objects (in other words, on the phemomenon of biological dissymmetry);
- New opportunities are opened for apllications of effective algebraic methods (such as dyadic analysis, sequency analysis by Harmuth, etc.) to study structural organisation of long DNA nucleotide sequences represented as numerical trees (or hybrids) of helical epi-chains of binary numbers.

The author hopes that presented results will be useful for deeper understanding genetic principles of storing and processing genetic information; for developing new methods of comparative analysis of genomes for studying biological evolution; for developing algebraic and quantum biology; for creating new methods in medicine, biotechnology and artificial intelligence.

Appendix 1. Nucleotide frequencies in DNA epi-chains of the genome of a house mouse *Mus musculus*

This appendix presents frequencies %A, %T, %C and %G of nucleotides in single-stranded and double-stranded DNA epi-chains of all nuclear chromosomes of the genome of a house mouse *Mus musculus*. All initial data were accessed from the CenBank https://www.ncbi.nlm.nih.gov/genome?term=mus%20musculus.

The presented tabular values confirm the rules formulated above in the Section 5 about nucleotide probabilties %A, %T, %C and %G in single-stranded and double-stranded DNA epi-chains.

Table A1/1. Frequencies %A, %T, %C and %G of nucleotides A, T, C and G in single-stranded DNA epi-chains of initial orders in all nuclear chromosomes of a house mouse *Mus musculus*. Data relating single-stranded DNA epi-chains of the first order N_1, the second order ($N_{2/1}$, $N_{2/2}$) and the third order ($N_{3/1}$, $N_{3/2}$, $N_{3/3}$) are presented. Two left columns show ordinary symbols of chromosomes and their length.

Chr	DNA length (bp)		N_1	$N_{2/1}$	$N_{2/2}$	$N_{3/1}$	$N_{3/2}$	$N_{3/3}$
1	195471971	%A	0.2946	0.2946	0.2945	0.2946	0.2946	0.2945
		%T	0.2940	0.2940	0.2939	0.2940	0.2940	0.294
		%C	0.2058	0.2058	0.2058	0.2058	0.2058	0.2058
		%G	0.2057	0.2056	0.2057	0.2057	0.2056	0.2057

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 May 2019 doi:10.20944/preprints201904.0011.v2

2	182113224	%A	0.2894	0.2894	0.2893	0.2894	0.2894	0.2893
		%T	0.2898	0.2898	0.2898	0.2898	0.2898	0.2899
		%C	0.2103	0.2102	0.2104	0.2103	0.2103	0.2104
		%G	0.2105	0.2105	0.2105	0.2105	0.2106	0.2104
3	160039680	%A	0.2973	0.2973	0.2973	0.2973	0.2974	0.2973
		%T	0.2982	0.2982	0.2982	0.2980	0.2983	0.2982
		%C	0.2021	0.2021	0.2021	0.2022	0.2019	0.2021
		%G	0.2024	0.2024	0.2024	0.2025	0.2024	0.2025
4	156508116	%A	0.2882	0.2882	0.2882	0.2883	0.2881	0.2883
		%T	0.2889	0.2888	0.2889	0.2888	0.2889	0.2889
		%C	0.2114	0.2114	0.2114	0.2114	0.2114	0.2114
		%G	0.2115	0.2116	0.2115	0.2115	0.2116	0.2114
5	151834684	%A	0.2872	0.2873	0.2872	0.2872	0.2873	0.2873
		%T	0.2874	0.2873	0.2875	0.2874	0.2874	0.2875
		%C	0.2127	0.2127	0.2126	0.2127	0.2127	0.2126
		%G	0.2127	0.2127	0.2127	0.2127	0.2127	0.2126
6	149736546	%A	0.2928	0.2928	0.2928	0.2927	0.2928	0.2929
		%T	0.2931	0.2931	0.2930	0.2932	0.2930	0.2931
		%C	0.2072	0.2071	0.2072	0.2072	0.2071	0.2072
		%G	0.2070	0.2069	0.2071	0.207	0.2071	0.2069
7	145441459	%A	0.2839	0.2840	0.2838	0.2840	0.2839	0.2839
		%T	0.2856	0.2855	0.2856	0.2856	0.2855	0.2857
		%C	0.2154	0.2154	0.2154	0.2154	0.2155	0.2153
		%G	0.2151	0.2151	0.2152	0.2151	0.2152	0.2151
8	129401213	%A	0.2884	0.2883	0.2884	0.2883	0.2885	0.2884
		%T	0.2879	0.2879	0.2880	0.2879	0.2878	0.288
		%C	0.2119	0.2119	0.2119	0.2119	0.2119	0.2119
		%G	0.2118	0.2118	0.2117	0.2119	0.2118	0.2117
9	124595110	%A	0.2866	0.2866	0.2866	0.2867	0.2866	0.2865
		%T	0.2864	0.2864	0.2863	0.2864	0.2863	0.2864
		%C	0.2136	0.2136	0.2136	0.2135	0.2136	0.2137
		%G	0.2134	0.2134	0.2135	0.2134	0.2135	0.2134
10	130694993	%A	0.2926	0.2926	0.2927	0.2927	0.2927	0.2926
		%T	0.2933	0.2933	0.2934	0.2932	0.2935	0.2933
		%C	0.2068	0.2069	0.2068	0.2068	0.2068	0.2068
		%G	0.2072	0.2073	0.2072	0.2073	0.2071	0.2073
11	122082543	%A	0.2813	0.2812	0.2813	0.2812	0.2813	0.2814
		%T	0.2806	0.2806	0.2805	0.2807	0.2805	0.2806
		%C	0.2191	0.2192	0.2191	0.2192	0.2192	0.2191
		%G	0.2190	0.2190	0.2190	0.2189	0.2191	0.2189
12	120129022	%A	0.2899	0.2900	0.2899	0.2898	0.2900	0.2900
		%T	0.2927	0.2928	0.2927	0.2927	0.2927	0.2928
		%C	0.2085	0.2084	0.2085	0.2086	0.2084	0.2084
		%G	0.2089	0.2088	0.2090	0.2090	0.2089	0.2088
13	120421639	%A	0.2925	0.2925	0.2925	0.2924	0.2925	0.2926
		%T	0.2913	0.2913	0.2913	0.2914	0.2913	0.2911
		%C	0.2081	0.2081	0.2082	0.2080	0.2083	0.2081
		%G	0.2081	0.2081	0.2081	0.2082	0.2080	0.2081
14	124902244	%A	0.2939	0.2939	0.2940	0.2940	0.2939	0.2939
		%T	0.2945	0.2945	0.2944	0.2944	0.2945	0.2944
		%C	0.2057	0.2057	0.2057	0.2055	0.2058	0.2058
		%G	0.2059	0.2059	0.2059	0.2060	0.2059	0.2058
15	104043685	%A	0.2899	0.2900	0.2898	0.2900	0.2899	0.2898
		%T	0.2905	0.2904	0.2907	0.2906	0.2906	0.2905
		%C	0.2098	0.2098	0.2098	0.2098	0.2099	0.2099

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 May 2019 doi:10.20944/preprints201904.0011.v2

		%G	0.2097	0.2098	0.2097	0.2097	0.2096	0.2098
16	98207768	%A	0.2950	0.2950	0.2951	0.2950	0.2950	0.2952
		%T	0.2956	0.2956	0.2956	0.2955	0.2957	0.2954
		%C	0.2046	0.2046	0.2045	0.2046	0.2046	0.2046
		%G	0.2048	0.2048	0.2048	0.2049	0.2047	0.2048
17	94987271	%A	0.2862	0.2864	0.2861	0.2861	0.2864	0.2863
		%T	0.2868	0.2867	0.2869	0.2870	0.2867	0.2868
		%C	0.2136	0.2136	0.2136	0.2136	0.2136	0.2136
		%G	0.2134	0.2133	0.2134	0.2134	0.2134	0.2133
18	90702639	%A	0.2929	0.2930	0.2928	0.2930	0.2928	0.2929
		%T	0.2927	0.2928	0.2926	0.2927	0.2927	0.2927
		%C	0.2069	0.2068	0.2070	0.2068	0.2069	0.2070
		%G	0.2075	0.2074	0.2075	0.2074	0.2076	0.2074
19	61431566	%A	0.2875	0.2874	0.2875	0.2875	0.2875	0.2874
		%T	0.2852	0.2853	0.2852	0.2854	0.2852	0.2852
		%C	0.2139	0.2139	0.2139	0.2139	0.2138	0.2140
		%G	0.2134	0.2134	0.2134	0.2132	0.2135	0.2135
X	171031299	%A	0.3038	0.3038	0.3037	0.3039	0.3037	0.3037
		%T	0.3037	0.3037	0.3037	0.3037	0.3038	0.3036
		%C	0.1962	0.1963	0.1962	0.1962	0.1962	0.1963
		%G	0.1963	0.1963	0.1963	0.1962	0.1963	0.1964
Y	91744698	%A	0.3046	0.3046	0.3046	0.3045	0.3047	0.3046
		%T	0.3065	0.3066	0.3065	0.3065	0.3066	0.3066
		%C	0.1949	0.1949	0.1949	0.1950	0.1948	0.1949
		%G	0.1940	0.1939	0.1940	0.1941	0.1939	0.1939

Table A1/2. Frequencies %A, %T, %C and %G of nucleotides A, T, C and G in double-stranded DNA epi-chains of initial orders in all nuclear chromosomes of a house mouse *Mus musculus*. Data relating double-stranded DNA epi-chains of the first order (N_{1ab}), the second order ($N_{2/1a}$, $N_{2/2a}$) and the third order ($N_{3/1a}$, $N_{3/2a}$, $N_{3/3a}$) are presented. Two left columns show ordinary symbols of chromosomes and their length.

Chr	DNA length (bp)		N_{1ab}	$N_{2/1a}$	$N_{2/2a}$	$N_{3/1a}$	$N_{3/2a}$	$N_{3/3a}$
1	195471971	%A	0.2943	0.2943	0.2943	0.2943	0.2943	0.2942
		%T	0.2943	0.2943	0.2942	0.2942	0.2943	0.2943
		%C	0.2057	0.2056	0.2057	0.2057	0.2057	0.2057
		%G	0.2057	0.2058	0.2058	0.2058	0.2057	0.2058
2	182113224	%A	0.2896	0.2896	0.2895	0.2897	0.2896	0.2896
		%T	0.2896	0.2896	0.2895	0.2895	0.2896	0.2896
		%C	0.2104	0.2104	0.2104	0.2104	0.2105	0.2104
		%G	0.2105	0.2104	0.2105	0.2104	0.2104	0.2105
3	160039680	%A	0.2978	0.2978	0.2978	0.2977	0.2979	0.2977
		%T	0.2977	0.2977	0.2977	0.2977	0.2979	0.2977
		%C	0.2023	0.2023	0.2022	0.2023	0.2021	0.2023
		%G	0.2022	0.2022	0.2023	0.2023	0.2021	0.2023
4	156508116	%A	0.2886	0.2885	0.2886	0.2886	0.2884	0.2886
		%T	0.2885	0.2885	0.2885	0.2885	0.2885	0.2886
		%C	0.2115	0.2115	0.2115	0.2115	0.2116	0.2115
		%G	0.2115	0.2115	0.2114	0.2115	0.2115	0.2114
5	151834684	%A	0.2874	0.2873	0.2874	0.2874	0.2872	0.2874
		%T	0.2873	0.2873	0.2873	0.2873	0.2874	0.2873
		%C	0.2127	0.2127	0.2126	0.2127	0.2126	0.2126
		%G	0.2127	0.2127	0.2127	0.2127	0.2128	0.2126

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 May 2019 doi:10.20944/preprints201904.0011.v2

6	149736546	%A	0.2929	0.2930	0.2928	0.2930	0.2930	0.2929
		%T	0.2929	0.2929	0.2929	0.2929	0.2928	0.2930
		%C	0.2071	0.2071	0.2072	0.2071	0.2070	0.2070
		%G	0.2071	0.2070	0.2071	0.2071	0.2072	0.2070
7	145441459	%A	0.2848	0.2848	0.2847	0.2848	0.2845	0.2848
		%T	0.2846	0.2848	0.2847	0.2847	0.2848	0.2847
		%C	0.2153	0.2153	0.2153	0.2153	0.2154	0.2153
		%G	0.2153	0.2151	0.2153	0.2152	0.2153	0.2152
8	129401213	%A	0.2882	0.2881	0.2882	0.2881	0.2881	0.2881
		%T	0.2882	0.2881	0.2881	0.2881	0.2882	0.2883
		%C	0.2118	0.2119	0.2118	0.2118	0.2117	0.2118
		%G	0.2118	0.2119	0.2118	0.2120	0.2119	0.2118
9	124595110	%A	0.2864	0.2864	0.2863	0.2865	0.2864	0.2863
		%T	0.2865	0.2864	0.2865	0.2866	0.2864	0.2865
		%C	0.2135	0.2134	0.2135	0.2135	0.2135	0.2136
		%G	0.2135	0.2135	0.2136	0.2134	0.2135	0.2135
10	130694993	%A	0.2930	0.2930	0.2931	0.2930	0.2931	0.2929
		%T	0.2930	0.2929	0.2930	0.2928	0.2930	0.2930
		%C	0.2070	0.2071	0.2070	0.2070	0.2069	0.2070
		%G	0.2070	0.2071	0.2069	0.2072	0.2070	0.2070
11	122082543	%A	0.2809	0.2809	0.2810	0.2809	0.2810	0.2810
		%T	0.2810	0.2809	0.2809	0.2810	0.2808	0.2810
		%C	0.2191	0.2190	0.2190	0.2190	0.2191	0.2190
		%G	0.2191	0.2191	0.2191	0.2191	0.2192	0.2190
12	120129022	%A	0.2913	0.2914	0.2912	0.2913	0.2914	0.2914
		%T	0.2913	0.2914	0.2913	0.2912	0.2913	0.2914
		%C	0.2087	0.2087	0.2088	0.2087	0.2086	0.2086
		%G	0.2087	0.2086	0.2087	0.2088	0.2087	0.2086
13	120421639	%A	0.2919	0.2920	0.2919	0.2919	0.2920	0.2919
		%T	0.2919	0.2918	0.2918	0.2919	0.2918	0.2918
		%C	0.2081	0.2081	0.2081	0.2081	0.2081	0.2081
		%G	0.2081	0.2082	0.2081	0.2081	0.2081	0.2082
14	124902244	%A	0.2941	0.2942	0.2942	0.2942	0.2943	0.2942
		%T	0.2942	0.2942	0.2942	0.2942	0.2941	0.2941
		%C	0.2058	0.2058	0.2058	0.2058	0.2058	0.2058
		%G	0.2058	0.2059	0.2058	0.2058	0.2059	0.2058
15	104043685	%A	0.2903	0.2902	0.2903	0.2904	0.2901	0.2901
		%T	0.2901	0.2902	0.2902	0.2901	0.2904	0.2902
		%C	0.2098	0.2098	0.2097	0.2096	0.2098	0.2099
		%G	0.2098	0.2098	0.2098	0.2099	0.2098	0.2098
16	98207768	%A	0.2953	0.2953	0.2953	0.2952	0.2954	0.2953
		%T	0.2953	0.2953	0.2953	0.2953	0.2953	0.2954
		%C	0.2047	0.2047	0.2046	0.2048	0.2046	0.2047
		%G	0.2047	0.2047	0.2047	0.2047	0.2047	0.2047
17	94987271	%A	0.2866	0.2865	0.2866	0.2866	0.2864	0.2866
		%T	0.2864	0.2866	0.2864	0.2864	0.2866	0.2865
		%C	0.2135	0.2134	0.2135	0.2136	0.2134	0.2134
		%G	0.2134	0.2134	0.2135	0.2134	0.2136	0.2135
18	90702639	%A	0.2928	0.2928	0.2927	0.2928	0.2927	0.2928
		%T	0.2928	0.2930	0.2928	0.2929	0.2928	0.2927
		%C	0.2072	0.2071	0.2074	0.2072	0.2072	0.2071
		%G	0.2072	0.2071	0.2072	0.2071	0.2073	0.2073
19	61431566	%A	0.2863	0.2863	0.2864	0.2864	0.2864	0.2861
		%T	0.2864	0.2864	0.2863	0.2865	0.2864	0.2864
		%C	0.2136	0.2136	0.2136	0.2136	0.2136	0.2136

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 May 2019 doi:10.20944/preprints201904.0011.v2

		%G	0.2136	0.2137	0.2136	0.2135	0.2137	0.2138
X	171031299	%A	0.3037	0.3037	0.3038	0.3038	0.3037	0.3037
		%T	0.3037	0.3038	0.3037	0.3038	0.3038	0.3036
		%C	0.1963	0.1963	0.1963	0.1962	0.1962	0.1964
		%G	0.1962	0.1962	0.1962	0.1962	0.1963	0.1963
Y	91744698	%A	0.3056	0.3055	0.3056	0.3054	0.3056	0.3056
		%T	0.3056	0.3056	0.3055	0.3055	0.3056	0.3056
		%C	0.1945	0.1944	0.1944	0.1946	0.1943	0.1944
		%G	0.1944	0.1945	0.1945	0.1945	0.1944	0.1944

Appendix 2. Nucleotide frequencies in DNA epi-chains of the genome of a fruit fly *Drosophila melanogaster*

This appendix presents frequencies %A, %T, %C and %G of nucleotides in single-stranded and double-stranded DNA epi-chains of all nuclear chromosomes of the genome of a fruit fly *Drosophila melanogaster.* All initial data were accessed from the CenBank (https://www.ncbi.nlm.nih.gov/genome/?term=drosophila+melanogaster).

The presented tabular values confirm the rules formulated above in the Section 5 about nucleotide probabilties %A, %T, %C and %G in single-stranded and double-stranded DNA epi-chains.

Table A2/1. Frequencies %A, %T, %C and %G of nucleotides A, T, C and G in single-stranded DNA epi-chains of initial orders in all nuclear chromosomes of a fruit fly *Drosophila melanogaster*. Data relating single-stranded DNA epi-chains of the first order N_1, the second order ($N_{2/1}$, $N_{2/2}$) and the third order ($N_{3/1}$, $N_{3/2}$, $N_{3/3}$) are presented. Two left columns show ordinary symbols of chromosomes and their length.

Chr	DNA length (bp)		N_1	$N_{2/1}$	$N_{2/2}$	$N_{3/1}$	$N_{3/2}$	$N_{3/3}$
2L	23513712	%A	0.2915	0.2914	0.2915	0.2915	0.2915	0.2913
		%T	0.2907	0.2908	0.2907	0.2911	0.2912	0.2899
		%C	0.2089	0.2089	0.2089	0.2089	0.2085	0.2092
		%G	0.2089	0.2090	0.2089	0.2085	0.2087	0.2096
2R	25286936	%A	0.2877	0.2877	0.2877	0.2877	0.2879	0.2875
		%T	0.2862	0.2861	0.2863	0.2859	0.2862	0.2865
		%C	0.2134	0.2134	0.2135	0.2136	0.2134	0.2133
		%G	0.2127	0.2129	0.2125	0.2128	0.2126	0.2127
3L	28110227	%A	0.2909	0.2909	0.2909	0.2909	0.2913	0.2906
		%T	0.2929	0.2928	0.2929	0.2924	0.2933	0.2929
		%C	0.2081	0.2082	0.2080	0.2084	0.2080	0.2080
		%G	0.2081	0.2080	0.2081	0.2083	0.2074	0.2085
3R	32079331	%A	0.2872	0.2873	0.2871	0.2874	0.2871	0.2871
		%T	0.2869	0.2869	0.2870	0.2871	0.2870	0.2867
		%C	0.2132	0.2131	0.2132	0.2131	0.2131	0.2135
		%G	0.2127	0.2127	0.2127	0.2125	0.2129	0.2127
4	1348131	%A	0.3195	0.3191	0.3198	0.3199	0.3197	0.3188
		%T	0.3280	0.3285	0.3276	0.329	0.3273	0.3279
		%C	0.1747	0.1750	0.1744	0.1734	0.1747	0.176
		%G	0.1778	0.1774	0.1781	0.1777	0.1784	0.1773
X	23542271	%A	0.2868	0.2867	0.2869	0.2871	0.2868	0.2864
		%T	0.2886	0.2885	0.2886	0.2885	0.2889	0.2883
		%C	0.2120	0.2120	0.2119	0.2119	0.2116	0.2123
		%G	0.2127	0.2127	0.2126	0.2124	0.2127	0.2129
Y	3667352	%A	0.3099	0.3098	0.3100	0.3104	0.3098	0.3096
		%T	0.2958	0.2960	0.2957	0.2955	0.2954	0.2965
		%C	0.2002	0.2003	0.2002	0.2002	0.2003	0.2001
		%G	0.1940	0.1939	0.1941	0.1938	0.1945	0.1938

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 May 2019 doi:10.20944/preprints201904.0011.v2

Table A2/2. Frequencies %A, %T, %C and %G of nucleotides A, T, C and G in double-stranded DNA epi-chains of initial orders in all nuclear chromosomes of a fruit fly *Drosophila melanogaster*. Data relating double-stranded DNA epi-chains of the first order (N_{1ab}), the second order ($N_{2/1a}$, $N_{2/2a}$) and the third order ($N_{3/1a}$, $N_{3/2a}$, $N_{3/3a}$) are presented. Two left columns show ordinary symbols of chromosomes and their length.

Chr	DNA length (bp)		N_{1ab}	$N_{2ab/1}$	$N_{2ab/2}$	$N_{3ab/1}$	$N_{3ab/2}$	$N_{3ab/3}$
2L	23513712	%A	0.2910	0.2909	0.2910	0.2911	0.2914	0.2906
		%T	0.2911	0.2912	0.2912	0.2915	0.2914	0.2906
		%C	0.2089	0.2090	0.2090	0.2086	0.2087	0.2094
		%G	0.2090	0.2089	0.2088	0.2087	0.2085	0.2094
2R	25286936	%A	0.2870	0.2867	0.2871	0.2869	0.2869	0.2870
		%T	0.2869	0.2871	0.2870	0.2867	0.2872	0.2871
		%C	0.2129	0.2131	0.2128	0.2131	0.2130	0.2127
		%G	0.2132	0.2132	0.2131	0.2133	0.2129	0.2132
3L	28110227	%A	0.2919	0.2918	0.2918	0.2917	0.2924	0.2919
		%T	0.2919	0.2920	0.2921	0.2916	0.2922	0.2917
		%C	0.2082	0.2082	0.2079	0.2083	0.2076	0.2084
		%G	0.2080	0.2081	0.2083	0.2084	0.2078	0.2081
3R	32079331	%A	0.2871	0.2870	0.2872	0.2873	0.2871	0.2871
		%T	0.2870	0.2871	0.2869	0.2871	0.2869	0.2867
		%C	0.2129	0.2130	0.2130	0.2127	0.2131	0.2132
		%G	0.2130	0.2129	0.2129	0.2128	0.2128	0.2130
4	1348131	%A	0.3234	0.3233	0.3233	0.3249	0.3244	0.3227
		%T	0.3241	0.3243	0.3242	0.3240	0.3225	0.3240
		%C	0.1766	0.1763	0.1763	0.1759	0.1763	0.1770
		%G	0.1759	0.1761	0.1763	0.1752	0.1768	0.1763
X	23542271	%A	0.2877	0.2877	0.2877	0.2878	0.2879	0.2875
		%T	0.2877	0.2875	0.2878	0.2879	0.2878	0.2873
		%C	0.2123	0.2123	0.2124	0.2120	0.2120	0.2127
		%G	0.2123	0.2124	0.2121	0.2123	0.2123	0.2125
Y	3667352	%A	0.3027	0.3032	0.3024	0.3030	0.3029	0.3029
		%T	0.3030	0.3026	0.3033	0.3029	0.3023	0.3032
		%C	0.1972	0.1971	0.1969	0.1971	0.1977	0.1973
		%G	0.1971	0.1971	0.1974	0.1969	0.1971	0.1966

Appendix 3. Nucleotide frequencies in DNA epi-chains of the genome of a nematode Caenorhabditis elegans

This appendix presents frequencies %A, %T, %C and %G of nucleotides in single-stranded and double-stranded DNA epi-chains of all nuclear chromosomes of the genome of a nematode *Caenorhabditis elegans*. All initial data were accessed from the CenBank (https://www.ncbi.nlm.nih.gov/genome?term=caenorhabditis%20elegans).

The presented tabular values confirm the rules formulated above in the Section 5 about nucleotide probabilties %A, %T, %C and %G in single-stranded and double-stranded DNA epi-chains.

Table A3/1. Frequencies %A, %T, %C and %G of nucleotides A, T, C and G in single-stranded DNA epi-chains of initial orders in all nuclear chromosomes of a nematode *Caenorhabditis elegans*. Data relating single-stranded DNA epi-chains of the first order N_1, the second order ($N_{2/1}$, $N_{2/2}$) and the third order ($N_{3/1}$, $N_{3/2}$, $N_{3/3}$) are presented. Two left columns show ordinary symbols of chromosomes and their length.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 May 2019 doi:10.20944/preprints201904.0011.v2

Chr	DNA length (bp)		N_1	$N_{2/1}$	$N_{2/2}$	$N_{3/1}$	$N_{3/2}$	$N_{3/3}$
I	15072434	%A	0.3208	0.3207	0.3210	0.3205	0.3212	0.3209
		%T	0.3217	0.3218	0.3216	0.3216	0.3216	0.3218
		%C	0.1789	0.1789	0.1788	0.1788	0.1793	0.1784
		%G	0.1786	0.1787	0.1786	0.1791	0.1779	0.1789
II	15279421	%A	0.3193	0.3193	0.3193	0.3190	0.3192	0.3196
		%T	0.3187	0.3187	0.3187	0.3189	0.318	0.3192
		%C	0.1812	0.1813	0.1812	0.1814	0.1814	0.1809
		%G	0.1808	0.1807	0.1809	0.1807	0.1814	0.1803
III	13783801	%A	0.3225	0.3226	0.3223	0.3223	0.3227	0.3223
		%T	0.3209	0.3209	0.3209	0.3206	0.3213	0.3209
		%C	0.1777	0.1778	0.1776	0.1780	0.1776	0.1775
		%G	0.1789	0.1787	0.1792	0.1791	0.1784	0.1792
IV	17493829	%A	0.3265	0.3263	0.3266	0.3264	0.3265	0.3265
		%T	0.3276	0.3277	0.3275	0.3273	0.3275	0.328
		%C	0.1735	0.1735	0.1734	0.1735	0.1737	0.1732
		%G	0.1725	0.1724	0.1725	0.1728	0.1724	0.1723
V	20924180	%A	0.3226	0.3225	0.3228	0.3227	0.3226	0.3225
		%T	0.3231	0.3232	0.3230	0.3229	0.3228	0.3235
		%C	0.1774	0.1775	0.1773	0.1772	0.1778	0.1773
		%G	0.1769	0.1768	0.1770	0.1772	0.1768	0.1767
X	17718942	%A	0.3244	0.3244	0.3243	0.3245	0.3240	0.3245
		%T	0.3236	0.3235	0.3237	0.3237	0.3234	0.3237
		%C	0.1761	0.1761	0.1761	0.1757	0.1765	0.1759
		%G	0.1760	0.1760	0.1759	0.1760	0.1760	0.1758

Table A3/2. Frequencies %A, %T, %C and %G of nucleotides A, T, C and G in double-stranded DNA epi-chains of initial orders in all nuclear chromosomes of a nematode *Caenorhabditis elegans*. Data relating double-stranded DNA epi-chains of the first order (N_{1ab}), the second order ($N_{2/1a}$, $N_{2/2a}$) and the third order ($N_{3/1a}$, $N_{3/2a}$, $N_{3/3a}$) are presented. Two left columns show ordinary symbols of chromosomes and their length.

Chr	DNA length (bp)		N_{1ab}	$N_{2ab/1}$	$N_{2ab/2}$	$N_{3ab/1}$	$N_{3ab/2}$	$N_{3ab/3}$
I	15072434	%A	0.3211	0.3212	0.3211	0.3209	0.3216	0.3211
		%T	0.3214	0.3212	0.3214	0.3211	0.3213	0.3216
		%C	0.1787	0.1789	0.1784	0.1791	0.1787	0.1787
		%G	0.1788	0.1786	0.1790	0.1789	0.1785	0.1786
II	15279421	%A	0.3190	0.3190	0.3190	0.3188	0.3186	0.3195
		%T	0.3190	0.3190	0.3190	0.3191	0.3186	0.3193
		%C	0.1811	0.1811	0.1808	0.1809	0.1811	0.1807
		%G	0.1809	0.1810	0.1812	0.1811	0.1817	0.1805
III	13783801	%A	0.3218	0.3217	0.3216	0.3216	0.3220	0.3217
		%T	0.3216	0.3218	0.3216	0.3213	0.3220	0.3215
		%C	0.1785	0.1781	0.1784	0.1787	0.1779	0.1787
		%G	0.1781	0.1783	0.1784	0.1784	0.1781	0.1780
IV	17493829	%A	0.3269	0.3272	0.3270	0.3269	0.3266	0.3272
		%T	0.3271	0.3269	0.3270	0.3271	0.3270	0.3268
		%C	0.1730	0.1730	0.1728	0.1733	0.1730	0.1728
		%G	0.1729	0.1729	0.1731	0.1730	0.1731	0.1727
V	20924180	%A	0.3227	0.3227	0.3230	0.3227	0.3227	0.3227
		%T	0.3230	0.3229	0.3228	0.3229	0.3228	0.3233
		%C	0.1772	0.1772	0.1771	0.1773	0.1773	0.1771
		%G	0.1771	0.1771	0.1771	0.1770	0.1773	0.1769

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 May 2019 doi:10.20944/preprints201904.0011.v2

X	17718942	%A	0.3241	0.3238	0.3240	0.3241	0.3235	0.3241
		%T	0.3239	0.3241	0.3240	0.3241	0.3240	0.3242
		%C	0.1760	0.1760	0.1760	0.1759	0.1763	0.1757
		%G	0.1760	0.1761	0.1760	0.1758	0.1763	0.1760

Appendix 4. Nucleotide frequencies in DNA epi-chains of the genome of a plant *Arabidopsis thaliana*

This appendix presents frequencies %A, %T, %C and %G of nucleotides in single-stranded and double-stranded DNA epi-chains of all nuclear chromosomes of the genome of a plant *Arabidopsis thaliana.* All initial data were accessed from the CenBank (https://www.ncbi.nlm.nih.gov/genome/4, the column RefSeq).

The presented tabular values confirm the rules formulated above in the Section 5 about nucleotide probabilties %A, %T, %C and %G in single-stranded and double-stranded DNA epi-chains.

Table A4/1. Frequencies %A, %T, %C and %G of nucleotides A, T, C and G in single-stranded DNA epi-chains of initial orders in all nuclear chromosomes of a plant *Arabidopsis thaliana.* Data relating single-stranded DNA epi-chains of the first order N_1, the second order ($N_{2/1}$, $N_{2/2}$) and the third order ($N_{3/1}$, $N_{3/2}$, $N_{3/3}$) are presented. Two left columns show ordinary symbols of chromosomes and their length.

Chr	DNA length (bp)		N_1	$N_{2/1}$	$N_{2/2}$	$N_{3/1}$	$N_{3/2}$	$N_{3/3}$
1	30427671	%A	0.3208	0.3207	0.3209	0.3212	0.3206	0.3207
		%T	0.3204	0.3204	0.3204	0.3204	0.3207	0.3202
		%C	0.1796	0.1796	0.1796	0.1794	0.1798	0.1796
		%G	0.1791	0.1792	0.1791	0.1790	0.1789	0.1795
2	19698289	%A	0.3207	0.3205	0.3208	0.3205	0.3204	0.3210
		%T	0.3207	0.3208	0.3206	0.3210	0.3208	0.3203
		%C	0.1799	0.1800	0.1798	0.1797	0.1802	0.1797
		%G	0.1788	0.1787	0.1788	0.1788	0.1785	0.1790
3	23459830	%A	0.3191	0.3194	0.3189	0.3195	0.3192	0.3187
		%T	0.3176	0.3175	0.3177	0.3171	0.3179	0.3176
		%C	0.1816	0.1815	0.1816	0.1816	0.1813	0.1818
		%G	0.1817	0.1817	0.1818	0.1818	0.1816	0.1818
4	18585056	%A	0.3197	0.3197	0.3196	0.3198	0.3195	0.3197
		%T	0.3183	0.3182	0.3183	0.318	0.3183	0.3185
		%C	0.1814	0.1814	0.1815	0.1812	0.1817	0.1814
		%G	0.1806	0.1807	0.1805	0.1810	0.1805	0.1803
5	26975502	%A	0.3197	0.3197	0.3198	0.3196	0.3199	0.3198
		%T	0.3209	0.3208	0.3209	0.321	0.3206	0.3210
		%C	0.1792	0.1793	0.1791	0.1796	0.1790	0.1790
		%G	0.1802	0.1802	0.1802	0.1797	0.1805	0.1803

Table A4/2. Frequencies %A, %T, %C and %G of nucleotides A, T, C and G in double-stranded DNA epi-chains of initial orders in all nuclear chromosomes of a plant *Arabidopsis thaliana.* Data relating double-stranded DNA epi-chains of the first order (N_{1ab}), the second order ($N_{2/1a}$, $N_{2/2a}$) and the third order ($N_{3/1a}$, $N_{3/2a}$, $N_{3/3a}$) are presented. Two left columns show ordinary symbols of chromosomes and their length.

Chr	DNA length (bp)		N_{1ab}	$N_{2ab/1}$	$N_{2ab/2}$	$N_{3ab/1}$	$N_{3ab/2}$	$N_{3ab/3}$
1	30427671	%A	0.3206	0.3204	0.3207	0.3207	0.3206	0.3203
		%T	0.3207	0.3208	0.3206	0.3209	0.3207	0.3206
		%C	0.1794	0.1795	0.1794	0.1791	0.1792	0.1796
		%G	0.1794	0.1793	0.1793	0.1793	0.1794	0.1796

2	19698289	%A	0.3206	0.3207	0.3207	0.3205	0.3207	0.3207
		%T	0.3208	0.3206	0.3207	0.3210	0.3206	0.3206
		%C	0.1794	0.1792	0.1793	0.1792	0.1794	0.1796
		%G	0.1793	0.1795	0.1793	0.1793	0.1793	0.1791
3	23459830	%A	0.3185	0.3182	0.3184	0.3184	0.3184	0.3185
		%T	0.3182	0.3186	0.3182	0.3182	0.3187	0.3179
		%C	0.1817	0.1816	0.1817	0.1818	0.1814	0.1819
		%G	0.1816	0.1816	0.1817	0.1816	0.1814	0.1818
4	18585056	%A	0.3189	0.3188	0.3190	0.3191	0.3187	0.3191
		%T	0.319	0.3191	0.3189	0.3188	0.3191	0.3192
		%C	0.1810	0.1811	0.1810	0.1811	0.1810	0.1806
		%G	0.1811	0.1810	0.1810	0.1811	0.1812	0.1811
5	26975502	%A	0.3203	0.3201	0.3204	0.3203	0.3205	0.3206
		%T	0.3203	0.3204	0.3203	0.3203	0.3199	0.3202
		%C	0.1797	0.1798	0.1797	0.1798	0.1797	0.1796
		%G	0.1797	0.1797	0.1796	0.1796	0.1798	0.1797

Acknowledgments. Some results of this paper have been possible due to a long-term cooperation between Russian and Hungarian Academies of Sciences on the topic "Non-linear models and symmetrologic analysis in biomechanics, bioinformatics, and the theory of self-organizing systems", where S.V. Petoukhov was a scientific chief from the Russian Academy of Sciences. Special thanks to V.I. Svirin who created computer programs for described researches on the basis of author's technical tasks. The author is grateful to G. Darvas, E. Fimmel, M. He, Z.B. Hu, Yu.I. Manin and I.V. Stepanyan for their collaboration. Special thanks also to the German Academic Exchange Service (DAAD) for providing the very useful internship for the author in autumn 2017 at the Institute of Mathematical Biology of the Mannheim University of Applied Sciences (Germany) where the host was Prof. E. Fimmel.

References

Abbott D., Davies P.C.W., Pati A.K. (Eds.) (2008). *Quantum Aspects of Life*, foreword by Sir Roger Penrose, 2008 / ISBN-13: 978-1-84816-253-2.

Albrecht-Buehler G. (2006). Asymptotically increasing compliance of genomes with Chargaff's second parity rules through inversions and inverted transpositions. - *Proc Natl Acad Sci USA*. 103 (47): p. 17828–17833, doi:10.1073/pnas.0605553103.

Altaisky M.V., Filatov F.P., 2001. Genetic information and quantum gas. - arXiv:quant-ph/0106123v1, submitted on 22.06.2001.

Belousov L.V. (2015). *Morphomechanics of development*. Springer International Publishing Switzerland, ISBN 978-3-319-13989-0, DOI 10.1007/978-3-319-13990-6.

Blank M., Goodman R. (2011). DNA is a fractal antenna in electromagnetic fields. *International Journal of Radiation Biology*, 87:4, 409-415, DOI: 10.3109/09553002.2011.538130. http://dx.doi.org/10.3109/09553002.2011.538130

Chargaff E., Lipshitz R., Green C. (1952). Composition of the deoxypentose nucleic acids of four genera of seaurchin. - *J Biol Chem*. 195 (1), p. 155–160.

Chargaff E. (1971). Preface to a Grammar of Biology: A hundred years of nucleic acid research, Science, 172, p. 637-642.

Cook T.A. (1914). *The Curves of Life*. London: Constable and Co., 490 p.

Darvas G. (2007). *Symmetry*. Basel: Birkhauser, 508 p.

Darvas G. (2018). Petoukhov's rules on symmetries in long DNA-texts. *Symmetry: Culture and Science*, v. 29, # 2, pages 318-320 , https://doi.org/10.26830/symmetry_2018_2_318, http://journal-scs.symmetry.hu/abstract/?pid=673

Driesch H. (1921). *Philosophie des Organischen*. Engelmann, Leipzig.

Fimmel E., Petoukhov S., 2019. Genetic Code Modeling from the Perspective of Quantum Informatics. – In: *Advances in Artificial Systems for Medicine and Education II.* / Hu Z.B., He M., Petoukhov S.V. (Eds). Springer (in print), 978-3-030-12081-8, 478911_1_En.

Foster K.R. (2011). Comments on DNA as a fractal antenna. - International Journal of Radiation Biology**,** v. 87, 2011 - Issue 12**,** 1208–1209 https://www.tandfonline.com/doi/full/10.3109/09553002.2011.626490?src=recsys

Fröhlich F. (1973). Notes and articles in honor of Herbert Fröhlich. In *Cooperative Phenomena*; Haken H., Wagner M., Eds.; Springer: Berlin, Germany.

Fröhlich, H. (1968). Long-range coherence and energy storage in biological systems. *Int. J. Quantum Chem.* 1968, *2*, 641–649.

Fröhlich F. (1988). Genetic code as Language. – In: Fröhlich H. Editor. *Biological Coherence and Response to External Stimuli*, Springer, p. 192-204. ISBN 978-0-387-18739-6.

Fröhlich, H. (1969). Quantum mechanical concepts in biology. In *From Theoretical Physics to Biology*; Marois, M., Ed.; North-Holland: Amsterdam, the Netherlands, 1969; pp. 13–22.

Fröhlich, H. (1970). Long range coherence and the action of enzymes. *Nature* 1970.

Fröhlich H. (1978). Coherent electric vibrations in biological systems and cancer problem. // *IEEE Trans. Microwave Theory Tech.*, v. 26, p. 613–617.

Fröhlich, H. (1980). The biological effects of microwaves and related questions. In *Advances in Electronics and Electron Physics*, v. 53, p. 85–152. New York, USA: Academic Press.

Fröhlich H. (1988). Introduction. Theoretical Physics and Biology. – In: Herbert Fröhlich editor, *Biological Coherence and Response to External Stimuli* (Springer, 1988) ISBN 978-0-387-18739-6, p. 3-24.

Fröhlich H., Kremer F. (1983). *Coherent Excitations in Biological Systems.* Springer-Verlag, ISBN 978-3-642-69186-7.

Grundler W., Keilmann F. (1983). Sharp resonances in yeast growth proves nonthermal sensitivity to microvaves. *Phys. Rev. Letts*., v. 51, p. 1214-1216.

Hameroff S. R. (1974). Chi: a neural hologram? *Am. J. Clin. Med*., 2(2), p. 163-70.

Hameroff S. R. (1987). *Ultimate computing. Biomolecular consciousness and nano-technology*. North-Holland, Amsterdam.

Hameroff S.R., Penrose R. (1996). Conscious events as orchestrated spacetime selections. *J. Conscious. Stud.*, 3, p. 36–53.

Hameroff S., Penrose R. (2014). Consciousness in the universe: A review of the"Orch OR" theory**.** *Phys. Life Rev.*, 11, p. 39–78.

Harmuth H.F. (1970). Transmission of information by orthogonal functions. Berlin: Springer.

Harmuth H.F. (1977). *Sequency Theory. Foundations and Applications*; Academic Press: New York, NY, USA, 1977.

Harmuth H.F. (1981). Nonsinusoidal waves for radar and radio communication. N.-Y.: Academic Press.

Harmuth H.F. (1989). *Information theory applied to space-time physics*. - The Catholic University of America, DC.: Washington, USA.

Hargittai M., Hargittai I. (2003). Eternal dissymmetry. - *Mendeleev Communications Electronic Version*, Issue 3, p. 1-2, http://www.niic.nsc.ru/files/institute/history/iycr2014/papers/Dissymmetry.pdf

Holland B.W. (1972). Dynamic specificity and the pairing of homologous chromosomes in meiosis. – J. Theor. Biol., 35(2), p. 395-397.

Hu Z.B., Petoukhov S.V., **Petukhova E.S**. (2017a). Generalized crystallography, the genetic system and biochemical esthetics. - *Structural Chemistry*, v. 28, №1, pp. 354-368. doi:10.1007/s11224-016-0880-0, http://link.springer.com/journal/11224/28/1/page/2

Hu Z.B., Petoukhov S.V., Petukhova E.S. (2017b). I-Ching, dyadic groups of binary numbers and the geno-logic coding in living bodies. - *Progress in Biophysics and Molecular Biology,* v. 131, December 2017, pp. 354-368. (https://doi.org/10.1016/j.pbiomolbio.2017.08.018).

Hu Z.B., Petoukhov S.V., Petukhova E.S. (2018). On symmetries, resonances and photonic crystals in morphogenesis. *Biosystems*, online 14.09.2018, https://doi.org/10.1016/j.biosystems.2018.09.004

Hyland G.J. (1998). Coherent GHz and THz excitations in active biosystems, and their implications. In The Future of Medical Diagnostics? Proceeding of Matra Marconi UK, Directorate of Science, Internal Report, Portsmouth, UK, 25 June 1998; pp. 14–27.

Igamberdiev A.U. (1993). Quantum mechanical properties of biosystems: a framework for complexity, structural stability, and transformations.- *Biosystems*, 31 (1), pp. 65–73.

Igamberdiev A.I., 2004. Quantum computation, non-demolition measurements, and reflective control in living systems. – *BioSystems*, v. 77, p. 47–56.

Igamberdiev A.I., 2007. Physical limits of computation and emergence of life. – *BioSystems*, v. 90, p. 340–

349.
Igamberdiev A.I., 2008. Objective patterns in the evolving network of non-equivalent observers. – *BioSystems*, v. 92, p. 122–131.
Igamberdiev A.I., Shklovskiy-Kordi N.E., 2016. Computational power and generative capacity of genetic systems. – *BioSystems*, v. 142–143, p. 1–8.
Igamberdiev A.I., Shklovskiy-Kordi N.E., 2017. The quantum basis of spatiotemporality in perception and consciousness. - *Progress in Biophysics and Molecular Biology*, v. 130, p. 15-25.
Irobalieva R.N, Fogg J.M., Catanese D.J, Sutthibutpong T, Muyuan Chen, Barker A.K, Ludtke S.J., Harris S.A., Schmid M.F., Chiu W., Zechiedrich L. (12 October 2015). "Structural diversity of supercoiled DNA". Nature Communications. 6 (8440): 8440. doi:10.1038/ncomms9440. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4608029/
Ji S. (2017). *The Cell Language Theory: Connecting Mind and Matter*. - World Scientific Publishing Europe Ltd. , 450 p. ISBN-13: 978-1848166608.
Jordan P. (1932). Die Quantenmechanik und die Grundprobleme der Biologie und Psychologie. Naturwissenschaften 20, 815–821. (doi:10.1007/BF01494844).
Josephson B.D. (2018). The Physics of Mind and Thought. Preprint, DOI: 10.13140/RG.2.2.36516.32640/2 (https://www.researchgate.net/publication/328968105).
Kadantsev V.N., Goltsov A.N. (2018). Collective excitations in alpha-helical protein molecule. - *Russian technological journal*, v. 6, No. 2, p. 32-45.
Kelley S.O. (1999). Electron transfer through the DNA double helix: spectroscopic and electrochemical studies. Thesis (Dissertation (Ph.D.)). California Institute of Technology.
Kim S.H., Ganji, M., Kim, E., van der Torre J., Abbondanzieri E., Dekker C. (2018). Laub M.T, Barkai N., eds. "DNA sequence encodes the position of DNA supercoils". eLife. 7: e36557. doi:10.7554/eLife.36557. ISSN 2050-084X. PMID 30523779. - https://elifesciences.org/articles/36557.
Kizel V.A. (1980). Optical activity and dissymmetry in living systems. - Soviet Physics Uspekhi, v.23, 6, p. 277-295.
Kraus J.D. (1988). Antennas. 2nd Ed, MacGraw Hill, 1988
Kraus J.D., Marhefka R.J. (2002). Antennas: For All Applications. Third Edition, McGraw-Hill Higher Education.
Lockwood M. (1989). *Mind, brain and the quantum*. Blackwell, Oxford.
Lundholm I.V., Rodilla H., Wahlgren W., Duelli A., Bourenkov G., Vukusic J., Friedman R., Stake J., Schneider T. (2015). Terahertz radiation induces non-thermal structural changes, associated with Fröhlich condensation in a protein crystal // Structural Dynamics, v. 2(5), 054702, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4711649/.
Marshall I. N. (1989). Consciousness and Bose-Einstein condensates. *New Ideas in Psychology*, v.1.
Matsuno K. (1999). Cell motility as an entangled quantum coherence. – *BioSystems*, 51, pp. 15–19.
Matsuno K. (2003). Quantum mechanics in first, second and third person descriptions. – *BioSystems*, 68, pp. 107-118.
Matsuno K., Paton R.C. (2000). Is there a biology of quantum information? – *BioSystems*, 55, p.39–46.
McFadden J., Al-Khalili J. (2018). The origins of quantum biology. *Proceedings of the Royal Society A*: *Mathematical, Physical and Engineering Sciences*. 12 December 2018, p. 1-13. https://royalsocietypublishing.org/doi/full/10.1098/rspa.2018.0674.
Mesquita M.V., **Vasconcellos A.R., Luzzi R**. (1993). Selective amplification of coherent polar vibrations in biopolymers. *Phys. Rev. E*, *48*, p. 4049–4059.
Mikheenko P. (2018). Possible superconductivity in brain. https://arxiv.org/abs/1812.05602 , submitted on 13.12.2018.
Nielsen M.A., Chuang I.L. (2010). *Quantum Computation and Quantum Information*. New York: Cambridge University Press.
Patel A. (2001a). Quantum algorithms and the genetic code. *Pramana – Journal of Physics*, v. 56, 2-3, pp. 367-381, arXiv:quant-ph/0002037.
Patel A. (2001b). Testing quantum dynamics in genetic information processing. – *Journal of Genetics*, v. 80, 1, pp. 39-43.
Patel A. (2001c). Why genetic information processing could have a quantum basis. – *Journal of Biosciences*, v. 26, 2, pp. 145-151.
Pauling L. (1940). *The Nature of the Chemical Bond and the Structure of Molecules and Crystals: An I*

introduction to Modern Structural Chemistry, 2nd ed. Oxford University Press, London, 664 p.

Penrose R. (1994). *Shadows of the Mind: a Search for the Missing Science of Consciousness*. New York: Oxford University Press Inc., ISBN 0 19 853978 9.

Penrose R. (2019). Your eyes are not meant for seeing. – *Community*, February 21, 2019, https://thriveglobal.com/stories/your-eyes-are-not-meant-for-seeing/ .

Petoukhov S.V. (2008). Matrix genetics, algebras of the genetic code, noise-immunity. Moscow: RCD, 316 p. ISBN 978-5-93972-643-6 http://petoukhov.com/matrix-genetics-petoukhov-2008.pdf (in Russian).

Petoukhov S.V. (2011a). Matrix genetics, 8-dimensional hypercomplex numbers and dyadic shifts. – *Proceedings of the 2nd International Workshop on Genomic Signal Processing* (GSP 2011), 27-28 June 2011, Bucharest, Romania (Edited by Paul Dan Cristea), Biomedical Engineering Center, Bucharest, Romania, p. 27-34.

Petoukhov S.V. (2011b). Matrix Genetics and Algebraic Properties of the Multi-Level System of Genetic Alphabets**. -** *Neuroquantology*, Vol 9, No 4, p. 60-81. http://www.neuroquantology.com/index.php/journal/article/view/501

Petoukhov S.V. (2013). Matrix genetics: algebra of projection operators, cyclic groups and inherited ensembles of biological cycles. - *Proceedings of the 2nd International Conference "Theoretical Approaches to BioInformation Systems*" (TABIS.2013, September 17 – 22, 2013, Belgrade, Serbia). Editors B. Dragovich, R. Panajotović, D. Timotijević. - Institute of Physics, Belgrade, 2014, Serbia, p. 189-204. http://www.tabis2013.ipb.ac.rs/tabis2013.pdf

Petoukhov S.V. (2015a). Resonances and genetic biomechanics. - *Symmetry: Culture and Science*, т. 26, №3, p. 379-397, http://petoukhov.com/PETOUKHOV_IN_SCS_2015.pdf

Petoukhov S.V. (2015b). The concept of resonances in genetics ("resonant bioinformatics"). - Selected Works of the VII International Congress "Weak and Superweak Fields and Radiation in Biology and Medicine", September 07-11, 2015, St. Petersburg, pp. 1-41, http://www.biophys.ru/lib/sci/resonance/460-resonance-00001 (in Russian).

Petoukhov S. V. (2015c). Music and the Modeling Approach to Genetic Systems of Biological Resonances. *Presentation at the conference «IS4IS Summit Vienna 2015. The information society at the crossroads*», 3-7 June 2015, Vienna, Austria, http://sciforum.net/conference/70/paper/2812

Petoukhov S.V. (2016a). The system-resonance approach in modeling genetic structures. – *Biosystems*, January 2016, v. 139, p. 1-11, http://petoukhov.com/PETOUKHOV_ARTICLE_IN_BIOSYSTEMS.pdf .

Petoukhov S.V. (2016b). Symmetries of the genetic code, Walsh functions and the theory of genetic logical holography. - Symmetry: Culture and Science, 27, 95-98.

Petoukhov S.V. (2017). Genetic coding and united-hypercomplex systems in the models of algebraic biology. - Biosystems, 158, 31–46.

Petoukhov S.V. (2018a). The Genetic Coding System and Unitary Matrices. *Preprints* 2018, 2018040131 (doi: 10.20944/preprints201804.0131.v2).

Petoukhov S.V. (2018b). The rules of long DNA-sequences and tetra-groups of oligonucleotides. - arXiv:1709.04943v5, 5th version (8 October 2018), 159 pages.

Petoukhov S.V. (2019a). Structural Connections between Long Genetic and Literary Texts. *Preprints* **2018**, 2018120142 , online 15 02 2019. DOI: 10.20944/preprints201812.0142.v2, https://www.preprints.org/manuscript/201812.0142/v2.

Petoukhov S.V. (2019b). Connections Between Long Genetic and Literary Texts. The Quantum-Algorithmic Modelling. In: Hu Z., Petoukhov S., Dychka I., He M. (eds) *Advances in Computer Science for Engineering and Education II*. pp 534-543, ICCSEEA 2019. Advances in Intelligent Systems and Computing, vol 938. Springer, Cham. Online publication on 29 March 2019. Springer copyright – 2020. DOI https://doi.org/10.1007/978-3-030-16621 -2_50, https://link.springer.com/chapter/10.1007/978-3-030-16621-2_50#citeas (online).

Petoukhov S.V., Petukhova E.S. (2017a). Symmetries in genetic systems and the concept of geno-logical coding. - *Information*, *8*(1), 2; doi:10.3390/info8010002 (registering DOI), http://www.mdpi.com/2078-2489/8/1/2/htm; PDF Version: http://www.mdpi.com/2078-2489/8/1/2/pdf .

Petoukhov S.V., Petukhova E.S. (2017b). Resonances and the quest for transdisciplinarity. – *Information Studies and the Quest for Transdisciplinarity*. Editors M. Burgin, W.Hofkirchner, World

Scientific, p. p. 467-487.

Petoukhov S.V., Petukhova E.S., Svirin V.I. (2018). New Symmetries and Fractal-Like Structures in the Genetic Coding System. In: Hu Z., Petoukhov S., Dychka I., He M. (eds). *Advances in Computer Science for Engineering and Education*. ICCSEEA 2018. P. 588-600. Advances in Intelligent Systems and Computing, vol 754. Springer, Cham, DOI: https://doi.org/10.1007/978-3-319-91008-6_59

Petoukhov S.V., Svirin V.I. (2012). Fractal genetic nets and the rules of long genetic sequences. – *The Integrated Scientific Journal*, № 8-9, p. 50-52 (in Russian).

Petoukhov S.V., Svirin V.I. (2018). The New Wide Class of Symmetries in Long DNA-Texts. Elements of Quantum-Algorithmic Genetics. - *Biologia Serbica*, v. 40, №1, p. 51, Special Edition, ISSN 2334 6590, UDK 57(051). Book of Abstracts, Belgrade Bioinformatics Conference 2018, 18-22 June, 2018, Belgrade, Serbia, page 51.

Petoukhov S.V., Svirin V.I. (2015). Bionics of spiral structures. *Journal of Machinery Manufacture and Reliability*, Vol. 44, No. 3, pp. 249–253.

Potapov A.A. (2000). Fractals in Radiophysics and Radar. Elements of the Theory of Fractals: A Review. - *Journal of Communications Technology and Electronics*, 45(11):1157-1164.

Potapov A.A. (2005). *Fractals in radiophysics and radiolocation*. – 2nd ed., Moscow, Universitatskaia kniga, 848 p. (in Russian).

Rapoport A.E., Trifonov E.N. (2012): Compensatory nature of Chargaff's second parity rule. *Journal of Biomolecular Structure and Dynamics*, November, 1-13, DOI:10.1080/07391102.2012.736757

Rieper E., Anders J., Vedral V. (2011). Quantum entanglement between the electron clouds of nucleic acids in DNA. - arXiv:1006.4053v2.

Schrödinger E. (1952). Are there quantum jumps? Part I. - *The British Journal for the Philosophy of Science*, Vol. 3, No. 10, pp. 109-123. Published by: Oxford University Press on behalf of The British Society for the Philosophy of Science Stable, URL: http://www.jstor.org/stable/685552.

Singer E. (2016). "How Strange Twists in DNA Orchestrate Life". Quanta Magazine. https://www.quantamagazine.org/dna-supercoils-change-the-way-that-cells-work-20160105

Stutzman W., Thiele G. (1998). Antenna Theory and Design. 2nd Ed., John Wiley and Sons.

Vasconcellos A.R., Vannucchi F.S., Mascarenhas S., Luzzi R. (2012). Frohlich Condensate: Emergence of Synergetic Dissipative Structures in Information Processing Biological and Condensed Matter Systems. – Information, v.3(4), p. 601-620, doi:10.3390/info3040601. https://www.mdpi.com/2078-2489/3/4/601/htm

Watson J.D., Crick F.H.C. (1953). A Structure for Deoxyribose Nucleic Acid. – *Nature*, 171, p. 737-738.

Zabeo D., Heumann J.M., Schwartz C.L., Suzuki-Shinjo A., Morgan G., Per O. Widlund P.O, Höög J.L. (2018). A lumenal interrupted helix in human sperm tail microtubules. - *Sci Rep*., 8: 2727. Published online 2018 Feb 9. doi: 10.1038/s41598-018-21165-8, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5807425/