

"CRISPR technology challenge facing the numerical Integrity of whole Human Genome DNA"

Author

Jean-claude PEREZ * £

* retired interdisciplinary researcher (IBM),

£ 7 avenue de terre-rouge F33127 Martignas Bordeaux metropole France jeanclaudeperez2@gmail.com

Keywords : Human genome, CRISPR, Biomathematics, Evolution, Autopoiesis.

Abstract

Background : Global analysis of 3 human genomes of increasing levels of evolution (Neanderthal / Sapiens Build34 / Sapiens hg38) reveals 2 levels of numerical constraints controlling, structuring and optimizing these genome's DNA sequences. A global constraint - called "HGO" for "Human Genome Optimum" - optimizes the genome at its global scale. The same operator applied to each of the 24 individual chromosomes reveals a hierarchical structure of these 24 chromosomes.

Results : Then analysing the single strand DNA CG / TA proportions at whole chromosomes and genome scale reveals strong fine-tuned numerical ratios evidencing the "closure" nature (Varela's autopoiesis theory) of whole human genome.

Introduction

Thanks to the CRISPR (Clustered regularly interspaced short palindromic repeats) technology, it is now possible to locally modify the genomes, and particularly the human genome (1). Almost simultaneously, the fractal and global structures of the human genome were demonstrated (2). In such a context, apart from ethical questions, can a local technology as powerful as CRISPR be applied, ignoring its possible effect on the possible global and long-range equilibria and balancing at the chromosome scale or even the entire genome scale? For more than 25 years, we have been looking for possible global, even numerical, structures that would organize DNA, genes, chromosomes and even whole genomes (3, 4, 5, 6, 7, 8, 9).

We have already demonstrated a numerical structure at the scale of each human chromosome as well as on the whole genome (10, 11, 12, 13, 14, 15). In (10) we have already highlighted this numerical value of 0.6909830056, the HGO in this article : it controls the population of triplets codons analysing single stranded DNA sequence from the whole human genome.

Materials and Methods

Analysed whole human genomes :

We analyzed completely and systematically each of the 24 chromosomes of each of the following three reference genomes:

Neanderthal genome : (2014) ref (16)

<http://www.nature.com/nature/journal/v505/n7481/full/nature12886.html>

Sapiens Build34 (2003) human reference genome ref (17)

<http://www.nature.com/nature/journal/v431/n7011/full/nature03001.html>

Sapiens hg38 (2013) human reference genome ref (18) <https://www.ncbi.nlm.nih.gov/grc/human>

Computing the HGOs : Let us now distinguish the two types of HGO that will be discussed:

1/ Theoretical HGO (tHGO)

tHGO = $(3-\Phi)^{-2}$ = 0.6909830056, where Φ is the Golden Ratio $\Phi = 1.618033989$

2/ Reference female HGO (rwHGO) : rwHGO = 0.6913477936

error (tHGO – rwHGO) = 0.6909830056 - 0.6913477936 = -0.0003647879784

and

Reference male HGO (rmHGO) : rmHGO = 0.6922864236

error (tHGO – rmHGO) = 0.6909830056 - 0.6922864236 = -0.001303417973

Details :

HGOwoman = [(sum C+G single strand 1 to 22 chromosomes) + (sum C+G chrX)

+ (sum C+G single strand 1 to 22 chromosomes) + (sum C+G chrX)]

/ [(sum T+A single strand 1 to 22 chromosomes) + (sum T+A chrX)

+ (sum T+A single strand 1 to 22 chromosomes) + (sum T+A chrX)]

HGOman = [(sum C+G single strand 1 to 22 chromosomes) + (sum C+G chrX)

+ (sum C+G single strand 1 to 22 chromosomes) + (sum C+G chrY)]

/ [(sum T+A single strand 1 to 22 chromosomes) + (sum T+A chrX)

+ (sum T+A single strand 1 to 22 chromosomes) + (sum T+A chrY)]

Results and Discussion

In all that follows, the general methodology will be as follows: we calculate, for the 46 chromosomes constituting each genome studied, only the single-stranded DNA sequences. In these sequences, we count the relative populations of bases T + A on the one hand, and C + G on the other hand.

1/ GENOME UNITY:

HGO of the 3 whole genomes : Neanderthal, Sapiens Build34 and Sapiens HG38

The three genomes we compare here are differentiated on the one hand by their respective evolution levels, on the other hand by the sample of individual genomes of which they form the syntheses, and finally by the precision of the sequencing of DNA.

The detailed analysis related to the 3 whole genomes shows the various distances and errors between real computed HGOs for each genome and theoretical HGO optimum value = 0.6909830055.

Particularly, it is found that the 3 HGOs calculated for the respective 3 genomes of Neanderthal, Sapiens (2003 Build34 and 2013 hg38 Sapiens) are very close to the ideal theoretical optimal HGO = 0.6909830056 (99.67% for the least optimal genome).

It is also observed that female genomes (XX) are more optimal than male genomes (XY).

On the other hand, the genomes of Neanderthal and Sapiens (Build34 of 2003) have very close optimization levels. We believe this results from the fact that the precisions of their respective DNA sequencing are similar.

On the contrary, the hg38 genomes of 2013 show the most optimal levels, this is most certainly due to the deeper quality of their DNA sequencing. Fig. 1 summarises HGO results for these 3 human genomes of varying levels of evolution.

Comparing HGO (Human Genome Optimum) for 3 Human Genomes

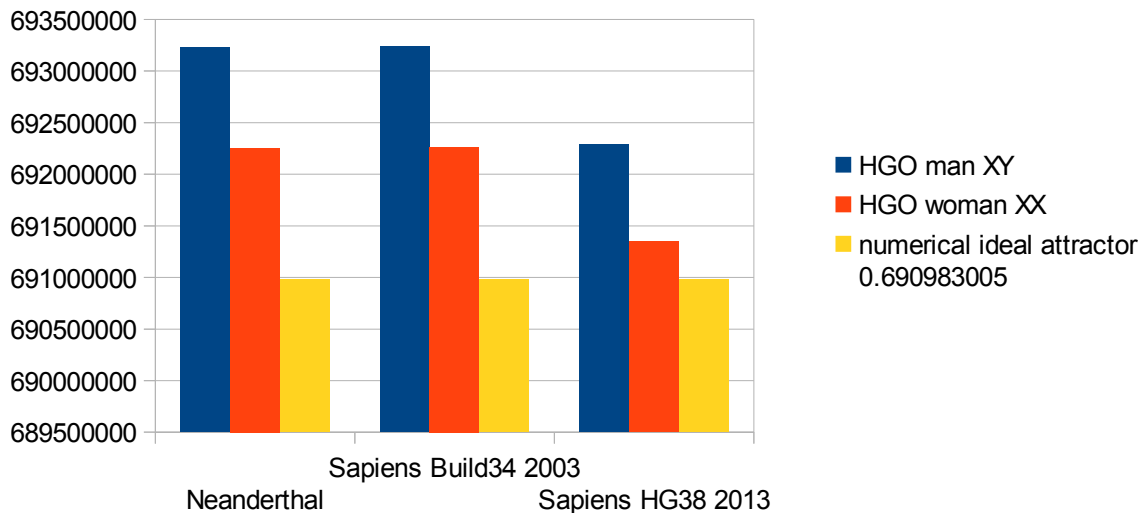


Fig. 1. The respective HGOs of 3 human genomes of varying levels of evolution are shown here.

Considerations on this theoretical Human Genetic Optimum (HGO) of $(3 - \Phi) / 2$:

This formula is particularly simple. We can even make it more "beautiful", indeed:

Since $1 + \Phi = \Phi^2$, we can write:

$$(3 - \Phi) / 2 = C+G / T+A = (4 - (1+\Phi)) / 2 = (4 - (\Phi^2)) / 2 = (2^2 - \Phi^2) / 2 = C+G / T+A$$

This new equivalent formula contains only the numbers "2" and "Phi".

This omnipresence of the number "2" in this formula has a strong analogy with the predictive formula of the periodic table of the Mendeleiev elements, also built around the "2" (19).

A second track to be studied could consist in replacing this writing by:

$$(3 - \Phi) / 2 = (3 - \Phi) / (5 - 3) = C+G / T+A$$

By this artifice of writing, we thus make the "3" appear in the numerator and the denominator (!)

The formula then becomes:

$$(3-\Phi) \times (T+A) = 2 \times (C+G) = (5-3) \times (C+G)$$

$$3(T+A) + 3(C+G) = 5(C+G) + \Phi(T+A)$$

$$3(T+A+C+G) = 5(C+G) + \Phi(T+A)$$

Therefore, if we consider that the single copy (single strand DNA) of the 24 chromosomes whole genomes XX or XY all lead to the same attractor $HGO = (3-\Phi) / 2$, to write :

Considering the cumulative population of 24 chromosomes of the single human genome (single strand DNA),

We check the following PERFECT BALANCE:

"THREE times the whole genome (T + A + C + G) = FIVE times (C + G) PLUS Phi times (T + A)"

Verification on 24 hg38 chromosomes single strand DNA:

$$CG = 1200551672$$

$$TA = 1737087441$$

$$3 \times (CG+TA) = 8812917339$$

$$(5 \times CG) + (\Phi \times TA) = 8813424881$$

$$8812917339+8813424881 = 0.9999424126$$

$$8812917339-8813424881 = \bar{5}07542$$

Finally, it is remarkable that this formula is based on integers 3 or 5. In fact, these numbers are very small integers and they are Fibonacci numbers. It will therefore be interesting to postpone the error calculations on the accuracy of these two integers 3 and 5:

$$(5 \times CG) + (\Phi \times TA) = 8813424881$$

/

$$(CG+TA) = 2937639113$$

$$8813424881 / 2937639113 = \mathbf{3.000172772}$$

and

$$3 \times (CG+TA) = 8812917339$$

-

$$(\Phi \times TA) = 2810666521$$

$$8812917339 - 2810666521 = 6002250818$$

$$CG = 1200551672$$

$$6002250818 \div CG = \mathbf{4.999577243}$$

The exact formula can then be written:

$$\mathbf{3.000172772} (T+A+C+G) = 5(C+G) + \Phi(T+A)$$

or

$$3(T+A+C+G) = \mathbf{4.999577243} (C+G) + \Phi(T+A)$$

2/ CHROMOSOMES HIERARCHY:

HGO spectral hierarchy of the 24 Human chromosomes :

The following 2 figures Fig. 2 and Fig. 3 illustrate the hierarchical spectrum of the individual HGOs of each of the 24 chromosomes for each of the three genomes analyzed. It should be noted that the upstream / downstream tipping point lies between chromosomes 14 and 21, which is closely related to the probable mechanisms explaining trisomy21 (whose disorders involve precisely these two chromosomes).

Finally, we note that it is the downstream region (Fig. 3) that contributes the most to the superiority of optimality of sapiens hg38 compared to sapiens Build34.

We have sorted the 24 chromosomes by increasing values of CG/TA ratios in the 3 cases of compared genomes.

It then reveals a hierarchical classification scale of 24 chromosomes ranging from 1 / Phi (chromosome4) to 3/2 Phi (chromosome 19).

Table 1 - the respective populations and ratios of each of the 24 chromosomes of the genome HG38

chr	C+G	T+A	CG/TA
UP			
4	72568001	117184666	0.6192619178
13	37772797	60210328	0.627347471
5	71611274	109654104	0.6530651511
0	61221521	93671508	0.6535767632
6	67360020	102718502	0.6557729979

3 78577742 119522393 0.6574311309
 18 31856106 48233499 0.6604560453
 0 10572683 15842360 0.66736793
 8 58133960 86634176 0.6710280248
 2 96769083 143779145 0.6730397722
 7 64696843 94273288 0.686269084
 12 54275482 78862334 0.6882307338
 14 36982791 53585358 0.6901659778
 DOWN
 21 16411625 23676994 0.693146478
 9 50270473 71520077 0.70288617
 11 55885058 78648684 0.7105657102
 10 55359481 77903481 0.7106162689
 1 96166571 134314441 0.7159808751
 15 35578844 49062481 0.7251741713
 20 28010605 35933652 0.7795089962
 16 36472718 45333225 0.8045471726
 17 37575444 45344760 0.8286612169
 22 18406838 20752939 0.8869509037
 19 28015712 30425046 0.9208108346

Diversity of Chromosomes from 3 Human Genomes

The Chromosomes UPSTREAM the HGO theoretical point = 0.6909830056

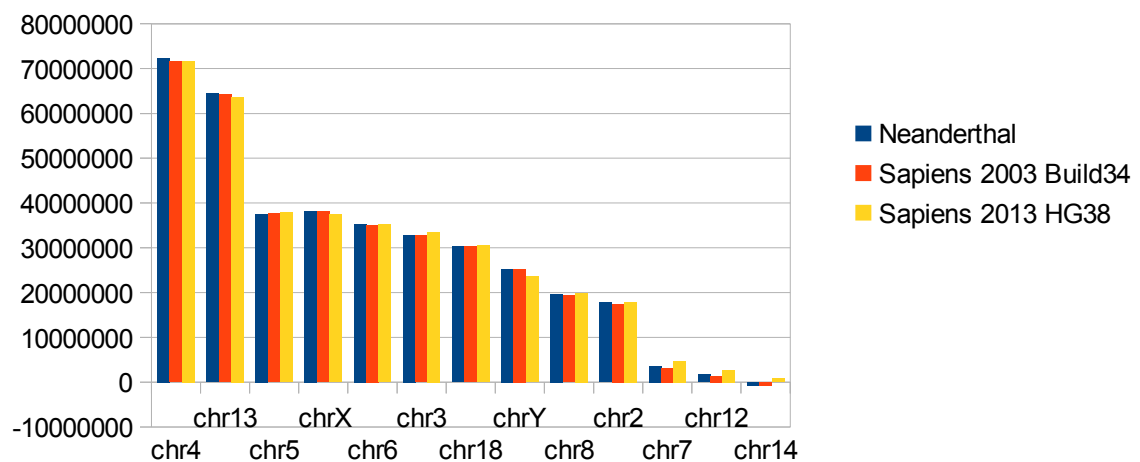


Fig. 2. « UP » chromosomes : HGO diversity of human chromosomes UPSTREAM of the numerical attractor

HGO = 0.6909830056.

Diversity of Chromosomes from 3 Human Genomes

The Chromosomes DOWNSTREAM the HGO theoretical point = 0.6909830056

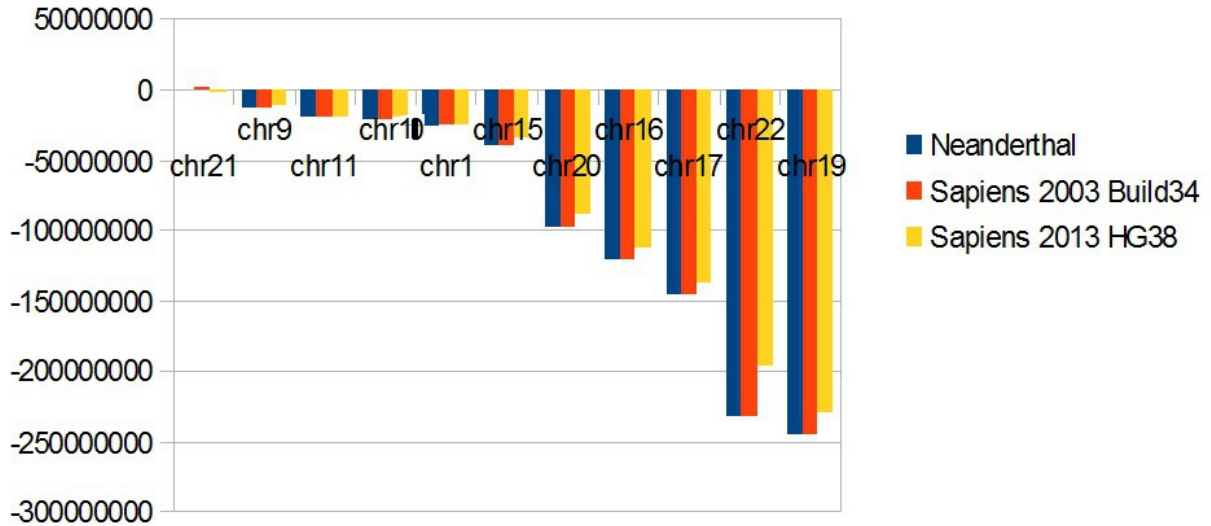


Fig. 3. « Down » chromosomes : Diversity of HGOs of human chromosomes DOWNSTREAM of the numerical attractor HGO = 0.6909830056.

3/ COHESION CHROMOSOMES / GENOME:

About the hierarchical classification of 24 single stranded chromosomes

In the following, we demonstrate a real interaction, a kind of "dialogue" with feedback between the equilibrium of the whole genome and the part of each of the individual chromosomes.

We must now regulate this high level of remarkable numerical constraints which seem to "frame" the CG and TA populations of each of the 24 human chromosomes on the one hand and of the entire genome on the other hand.

This will be verified for the human HG38 reference genome, but - as illustrated in Table 2 below - these remarkable properties will be extended to other higher primates in (26).

Genome	Extremum Top CG/TA Chr4		Extremum Down CG/TA chr19		Spectral Limits (CG/TA chr19) - (CG/TA Chr4)	
	value	Error CG/TA Chr4 vs 1/Phi	value	Error CG/TA chr19 vs 3/2 Phi	value	Error (3/2 Phi) – Spectral Limits
Sapiens HG38	0.6192619178	-0.0012279291	0.9208108346	0.0062401484	0.3015489168	0.0074680776
Sapiens BUILD34	0.6193778165	-0.0013438278	0.9364951603	-0.0094441773	0.3171173438	-0.0081003495

neanderthal	0.6185900969	-0.0005561082	0.9366477274	-0.0095967444	0.3180576305	-0.0090406362
chimp	0.6152388655	0.0027951232	0.9279395824	-0.0008885994	0.3127007169	-0.0036837226
Orangutang	0.6143645844	0.0036694043	0.9252214497	0.0018295333	0.3108568653	-0.001839871
gorilla	0.6177456029	0.0002883858	0.9299418695	-0.0028908865	0.3121962666	-0.0031792723
macaque	0.6536608193	-0.0356268306	0.929993709	-0.002942726	0.2763328897	0.0326841046

Table 2 - Evidence of strong numerical constraints surrounding the relative populations C+G / T+A constituting the hierarchical metastructure of the 24 chromosomes in humans and large primates

First, is there a simple relationship between HGO (P2), the numerical constraint at the scale of the entire genome, and the two extreme extremes of chromosome 4 (P1) and chromosome 19 (P3)? Then :

$$P1 = 1 \div \text{PHI} = 0.6180339887$$

$$P3 = 3 \div (2 \times \text{PHI}) = 0.927050983$$

$$P2 = (3 - \text{PHI}) \div 2 = 0.6909830055$$

We could compute :

$$P2 - P1 = 0.07294901685$$

$$P3 - P2 = 0.2360679775$$

$$\text{Then, } (P3 - P2) \div (P2 - P1) = 3.236067979$$

$$\text{Given that } 2 \times \text{PHI} = 3.236067978$$

$$\text{Then, } (P3 - P2) \div (P2 - P1) = 2 \times \text{PHI} = 3.236067979$$

$$\text{In other hand, } P3 - P1 = 1 \div (2 \times \text{PHI}) = 0.3090169943$$

Then finally, the high level of strong numerical constraints applied simultaneously to the 2 extrema chromosomes and to the whole human genome :

P1	P2	P3
chr4 1/Phi	genome (3-Phi)/2	chr19 3/2 Phi

$$P2 - P1 = 0.07294901685 \quad P3 - P2 = 0.2360679775$$

$$(P3 - P2) \div (P2 - P1) \ 3.236067979 = \mathbf{2 \times \text{PHI}} = 3.236067978$$

$$P3 - P1 \ 0.3090169943 \quad \mathbf{1 \div (2 \times \text{PHI})} \ 0.3090169943$$

4/ CLOSURE

We will now demonstrate a very strong property of the human genome very close to the theory of the autopoiesis of my friend franco-chilian biologist Francisco Varela (28, 29). In this theory, the coherence, consistency and integrity of living systems are modeled: the DNA of the human genome is a wonderful illustration of this.

Let us now look at the two UP chromosome populations (chr4 to chr14) and DOWN (chr21 to chr19). Would there exist particular constraints or remarkable relations on these 2 populations of chromosomes which determine the law described here?

Let us recall in table 1 the respective populations and ratios of each of the 24 chromosomes of the genome HG38:

Then cumulating in Table1 he populations C+G and T+A in each subclass UP and DOWN :

UP = 742398303 1124171661

DOWN = 458153369 612915780

DOWN / UP : C+G T+A

0.6171260995 0.5452154695

Or UP / DOWN : C+G T+A

1.62041437 1.834137246

This result is remarkable since it means that: on the one hand, the CG / TA ratio of chromosome4, a sort of leader or "semaphore", is equal to 1 / Phi.

On the other hand, the ratio of the C+G ratios between the 11 DOWN chromosomes to the 13 UP chromosomes is also equal to 1 / Phi.

Closure Varela's theory :

Distance amplitudes. $CG/TA \text{ Down/Up} = (P3-P2)/(P2-P1) = 2 \text{ Phi}$

Distance populations $CG \text{ Down/Up} = CG \text{ Down} / CG \text{ Up} = 1/Phi$

Then, distance amplitudes $Up/Down \text{ CG/TA} = 1/2Phi$

populations $CG \text{ Down} / CG \text{ Up} = \underline{2 \text{ times}}$ Distance amplitudes $CG/TA \text{ Up} / \text{Down}$.

It is remarkable to obtain this relation between AMPLITUDES on the one hand, and POPULATIONS (C + G) on the other hand.

We thus find again this number "2", symbol of the doubling of frequency such as the octave shift in music ... suggesting the possible wave nature of the DNA (23).

We still have a lot to discover on this fascinating CODE that is DNA (20-27)...

Finally, our approach may be related to these hundreds of unpredictable mutations resulting from manipulation of genomes by CRISPR revolutionary technology (30). Effectively in their 2017 article, authors note that « .../...They found that the technique had successfully corrected a gene that causes blindness in the mice, but the two mice that had undergone CRISPR gene-editing had sustained more than 1,500 unintended single-nucleotide mutations, and more than 100 larger deletions and insertions .../...».

Acknowledgements

We especially thank Dr. Robert Friedman M.D. practiced nutritional and preventive medicine in Santa Fe, New Mexico, which suggested new research interest on mtDNA genome. We also thank the mathematician Pr. Diego Lucio Rapoport (Buenos aires), the french biologist Pr. François Gros (Pasteur institute, co-discoverer of RNA messenger with James Watson and Walter Gilbert) and Pr. Luc Montagnier, medicine Nobel prizewinner for their interest in my research of biomathematical laws of genomes.

References

- 1 – Jun Wu et al, (2016), *In vivo* genome editing via CRISPR/Cas9 mediated homology-independent targeted integration, *Nature* (16 November 2016) | doi:10.1038/nature20565
- 2 - Lieberman-Aiden et al. (2009), Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science*, 2009; DOI: 10.1126/science.1181369 .
- 3 - Perez JC (1991) [Chaos DNA and neuro-computers: a golden link. *Speculations in Science and Technology* 14: 336-346.](#)
- 4 - Marcer PJ (1992) [Order and chaos in DNA – the Denis Guichard Prizewinner: Jean-Claude Perez. *Kybernetes* 21: 60-61.](#)
- 5 - Perez JC (1997) [L'adn Décrypté. Resurgence publisher Liege, Belgium.](#)
- 6 - Perez JC (2009) [Codex Biogenesis. Resurgence, Liege Belgium.](#)
- 7- Perez JC (2017) Sapiens Mitochondrial DNA Genome Circular Long Range Numerical Meta Structures are Highly Correlated with Cancers and Genetic Diseases mtDNA Mutations. *J Cancer Sci Ther* 9:512-527. doi: 10.4172/1948-5956.1000469

- <https://www.omicsonline.org/open-access/sapiens-mitochondrial-dna-genome-circular-long-range-numericalmeta-structures-are-highly-correlated-with-cancers-and-genetic-disea-1948-5956-1000469.php?aid=90737>
- 8- perez, DUF1220 Homo Sapiens and Neanderthal fractal periods architectures breakthrough(2017)SDRP Journal of Cellular and Molecular Physiology 1(1) <http://www.siftdesk.org/article-details/DUF1220%20Homo%20Sapiens%20and%20Neanderthal%20%20fractal%20periods%20architectures%20breakthrough/184?>
- 9- Perez J.C., (2017) "Global and long range fractal differences between sapiens and neanderthal genomes" , to be published in 2017.
- 10 - Perez JC (2010) [Codon Populations in Single-Stranded Whole Human Genome DNA Are Fractal and Fine-Tuned by the Golden Ratio 1.618. Interdisciplinary Sciences: Computational Life Sciences 2: 1-13.](#)
- 11 - Perez JC (2011) [Caminos Interdisciplinarios. Seminario CLAVE_INTER, Espacio Interdisciplinario, Universidad de la Republica Montevideo, Uruguay.](#)
- 12 - Perez JC (2011) [Decoding Non-Coding DNA Codes: Human Genome Meta-Chromosomes Architecture. BIT Life Sciences' 3rd Annual World Vaccine Congress, Beijing, China.](#)
- 13 - Perez JC (2013) [The "3 Genomic Numbers" Discovery: How Our Genome Single-Stranded DNA Sequence Is "Self-Designed" as a Numerical Whole. Applied Mathematics 4: 37-53.](#)
- 14 - Perez J (2015) Deciphering Hidden DNA Meta-Codes -The Great Unification & Master Code of Biology. J Glycomics Lipidomics 5:131. doi: 10.4172/2153-0637.1000131
- 15 – Perez J.C, Fractal Self-similarity, Scale Invariance and Stationary waves Codes Architecture Human Chromosomes DNA sequences, (submitted journal « chaos, fractals and solitons », <https://www.journals.elsevier.com/chaos-solitons-and-fractals/>).
- 16 - Neanderthal genome (2014) <http://www.nature.com/nature/journal/v505/n7481/full/nature12886.html>
- 17 - Sapiens Build34 (2003) genome <http://www.nature.com/nature/journal/v431/n7011/full/nature03001.html>
- 18 - Sapiens HG38 (2013) genome <https://www.ncbi.nlm.nih.gov/grc/human>
- 19 - Perez, J.C, (2017), Symmetry and Asymmetry in the MENDELEEV's Periodic Table Predictive EQUATION, SDRP Journal of Computational Chemistry & Molecular Modelling (accepted).
- 20 - Friedman R, Cross M (2013) [The Golden Ratio & Fibonacci Sequence: Golden Keys to Your Genius, Health, Wealth & Excellence. 1st \(Edn\), Hoshin Media, USA.](#)
- 21 - Pellionisz AJ, Graham R, Pellionisz PA , Perez JC (2012) [Recursive Genome Function of the Cerebellum: Geometric Unification of Neuroscience and Genomics. In: Manto M, DL, et al. \(Eds\) Handbook of the Cerebellum and Cerebellar Disorders. 1st \(Edn\), Springer, USA.](#)
- 22 - Rapoport DL (2015) Möbius strip and Klein Bottle Genomic Topologies, Self reference, Harmonics and Evolution, Life. <http://www.quantumbiosystems.org/admin/files/OBS%202016.%207.%20107-173.pdf> and https://www.academia.edu/30518156/Klein_Bottle_Logophysics_Self-reference_Heterarchies_Genomic_Topologies_Harmonics_and_Evolution_Part_III_The_Klein_Bottle_Logic_of_Genomics_and_its_Dynamics_Quantum_Information_Complexity_and_Palindromic_Repeats_in_Evolution
- 23– *Montagnier L. et al, (2017) Water Bridging Dynamics of Polymerase Chain Reaction in the Gauge Theory Paradigm of Quantum Fields, Water 2017, 9(5), 339; doi:10.3390/w9050339* <http://www.mdpi.com/2073-4441/9/5/339/pdf>
- 24 - Persaud, Dharam and O'Leary, James P., (2015) "Fibonacci Series, Golden Proportions, and the Human Biology" HWCOCOM Faculty Publications. Paper 27. http://digitalcommons.fiu.edu/com_facpub/27
- 25 - Roman V Yampolskiy, (2017) On the origin of synthetic life: attribution of output to a particular algorithm, The Royal Swedish Academy of Sciences , Phys. Scr. 92 (2017), DOI : [10.1088/0031-8949/92/1/013002](https://doi.org/10.1088/0031-8949/92/1/013002)
- 26- Perez J.C., (2017) "Humans and Primates Chromosomes4 Fractal CODES: periodic stationary waveforms characterizing and differentiating Neanderthal and Sapiens whole chromosomes DNA sequences" , (paper to be published, 2017).
- 27 - Perez, J.C., Deciphering "the MASTER CODE ®" Structure and Discovery of a Periodic Invariant Unifying 160 HIV1/HIV2/SIV Isolates Genomes. Biomed J Sci & Tech Res 1(2)-2017. BJSTR. MS.ID.000209. <http://biomedres.us/pdfs/BJSTR.MS.ID.000209.pdf>
- 28 - Varela F. J. (1981) Autonomy and autopoiesis. In: Roth G. & Schwegler H. (eds.) Self-organizing systems: An interdisciplinary approach. Campus Verlag, Frankfurt/New York: 14–24. Available at <http://cepa.info/2063>
- 29 - H.R. Maturana; F.J. Varela (1980). "The cognitive process". *Autopoiesis and cognition: The realization of the living*. Springer Science & Business Media. p. 13. ISBN 978-90-277-1016-1.
- 30 - **Kellie A Schaefer et al. ,(2017) Unexpected mutations after CRISPR-Cas9 editing in vivo, Nature Methods (2017). Nature Methods 14, 547–548 (2017) doi:10.1038/nmeth.4293** Published online 30 May 2017

31 - Perez JC (2017) The “Master Code of DNA”: Towards the Discovery of the SNPs Function (Single-Nucleotide Polymorphism). J Clin Epigenet. Vol. 3 No. 3:26 <http://clinical-epigenetics.imedpub.com/the-master-code-of-dna-towards-the-discovery-of-the-snps-function-singlenucleotide-polymorphism.pdf>