SDRP Journal of Cellular and Molecular Physiology(ISSN: 2574-4046)

WHY THE GENOMIC LOCATION OF INDIVIDUAL SNPS IS FUNCTIONAL?

DOI: 10.25177/JCMP.2.1.4

Research

AUTHOR: **Jean-claude PEREZ**

Sepember 2017

Received Date: 14th June 2017
Accepted Date: 25th Aug 2017
Published Date:10th Sep 2017

Jean-claude PEREZ, PhD Maths & Computer Science, retired interdisciplinary researcher (IBM),
7 avenue de terre-rouge F33127 Martignas Bordeaux metropole France, phone 33 0781181112

CORRESPONDING AUTHOR:
Jean-claude PEREZ
Email: jeanclaudeperez2@gmail.com

CONFLICTS OF INTEREST
There are no conflicts of interest for any of the authors.

ABSTRACT

Background: Every human individual is differentiated from all other humans by the few million SNPs. We are interested in the immediate neighborhood of each SNP. Would the SNP point have particular properties with respect to the surrounding nucleotides at short or medium distance?

Methods: The regions encompassing each SNP are analyzed by the biomathematical method of the "Master Code of DNA" published elsewhere. In particular, we will use a progressive approach of dichotomy type.

Results: There are then two types of remarkable results. Evidence of fractal properties such as self-similarity and scale invariance. On the other hand, the position of each SNP seems to play a functional role of the "active site" type as it is found in genes and proteins.

Conclusions: Each SNP would be more important by its precise location in the genome than by the value of its local mutation T / C or A / G for example. Consequently, SNPs would play a major functional role. We demonstrate here this property on 5 SNPs located in intron functional regions.

Keywords : Human genome, SNP, Fractals, self similarity, scale invariance

INTRODUCTION

Facing the new CRISPR genome editing technology [1] it is urgent to do research on possible hidden codes of DNA, chromosomes and genomes [1]. Since 1990, we are looking for possible hidden numerical codes that would control and structure the DNA of genes and genomes [2, 3, 4]. On the one hand, all the genetic difference between two human beings comes from some 1% of TCAG bases of the whole human genome constituting the SNPs [5, 6]. On the other hand, we have discovered and published a law and a method of analysis: the "master code of DNA" [7, 8], which UNIFIES all the information of the 3 types of CODES of Biology: DNA, RNA and amino acid sequences. Other numerical codes of DNA and genomes were discovered during these 25 years [8, 9, 10, 11]. Some of these codes, as will be the case here, are fractal [12, 13, 14, 15, 16, 17, 18].

Figure 1: The conventional representation of the Mendeleev's periodic table.

AUTHOR: **Jean-claude PEREZ**

METHODS

Master code summary:

Starting from the atomic masses constituting nucleotides and amino acids, a numerical scale of integers characterizing each bioatom, each TCAG DNA base, each UCAG RNA base, or each amino acid, an integer numbers scale code is obtained. Then, for each sequence of double - stranded DNA to be analyzed, the sequence of integers that characterizes it (genomics) is constructed as well as the sequence of amino acids that would encode this double strand if each of the strands was a potential protein (proteomics). The remarkable fact is that this proteomics image still exists, even for regions not translated into proteins (junk dna). The computational methodology of the Master code (7, 8) then produces 2 patterned images (2D curves, see Figure 1) which are very strongly correlated. This would mean that beyond the visible sequence of DNA there would be a kind of MASTER CODE being manifested by two supports of biological information: the sequences of DNA and of amino acids, the RNA image constituting a kind of neutral element like the zero of the mathematics. Our thousands of genes and genomes Master Code analyses (viruses, archaeas, bacteria, eucharyotes) demonstrated [7, 8] that the extremums (max and min) signify functional regions like proteins active sites, fragility points like chromosomes breakpoints). The main discovery of the paper will be the fact that each SNP is located precisely in such extremums. This allows us to consider the probable functionality of each SNP.

Figure 1 below illustrates Master Code computing showing strong correlation between the 2 images Genomics and Proteomics.

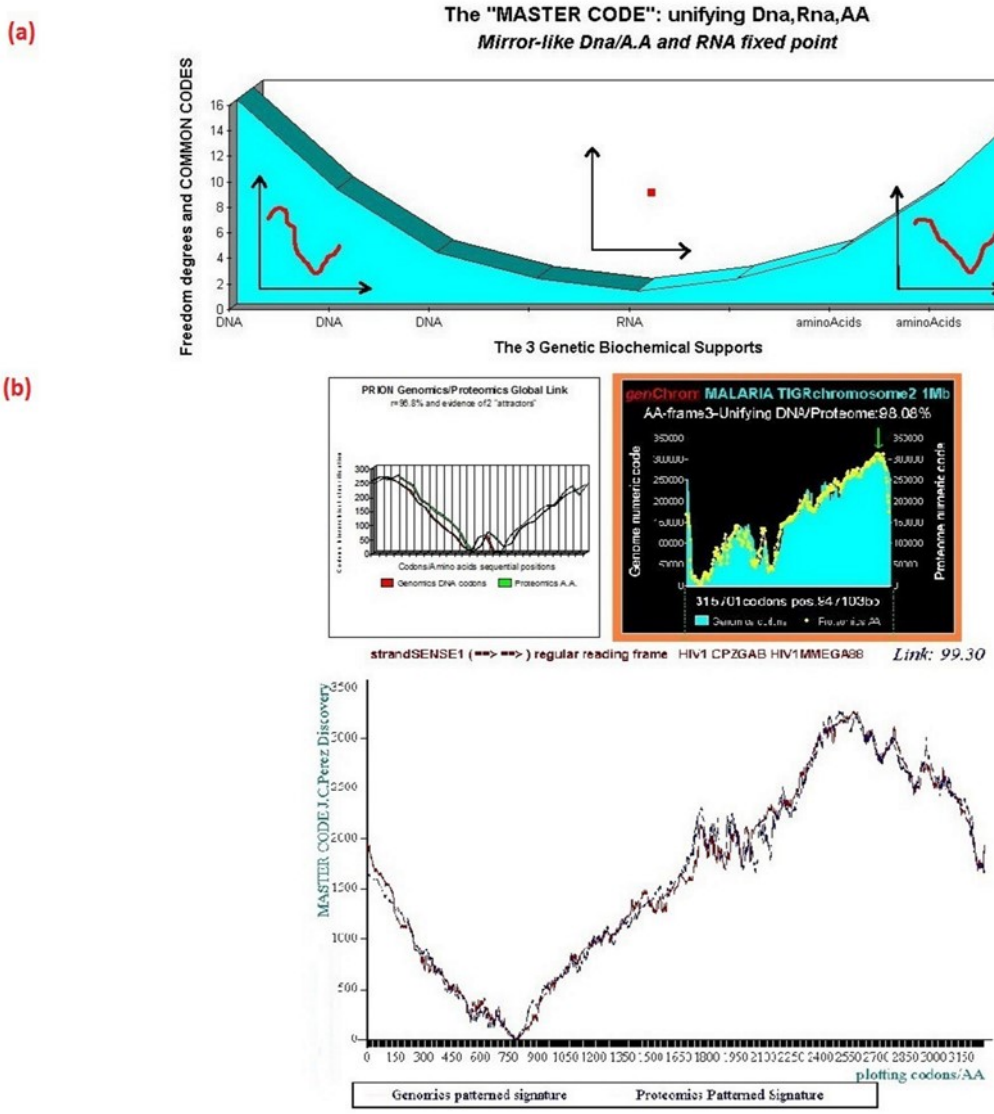


Figure 1 • (a) “Master code of biology” and Great Unification shows an equivalence of both Genomics (DNA) and Proteomics (amino acids) signatures while the RNA signature is a neutral area like a “zero”. (b) A typical correlation between Genomics and Proteomics signatures related to the Prion protein, the whole Malaria chromosome 2, and the whole HIV1 genome.

We work with the public SNPs WEB database from “CSHL”:
https://www.ncbi.nlm.nih.gov/pubmed/12519964

SNP		dbSNP #	Chrom	GenBank Ver	Golden Path	SNP Pos Relative to:		
reference number	identifier				Contig	GB	Contig	Chrom (kb)
		…/...						
Example 2	TSC1004969	2827238	Chr1	AL391359.6	ctg12483	71095	47949	75163.3
		…/...						
Example 3	TSC0080881	2076091	Chr1	AL391359.6	ctg12483	114011	70061	75185.5
		…/...						
Example 4	TSC1004962	2827231	Chr1	AL391359.6	ctg12483	130072	80327	75195.7
		…/...						
Example 5	TSC1004975	2827244	Chr1	AL391359.6	ctg12483	151041	93765	75209.2
		…/...						
Example 1	TSC1270753		Chr1	AL391359.5	ctg12483	58750	119856	75235.3
	…/…	…/...						

Table 1 - Five SNPs from human chromosome 1 in the region referenced AL391359.6

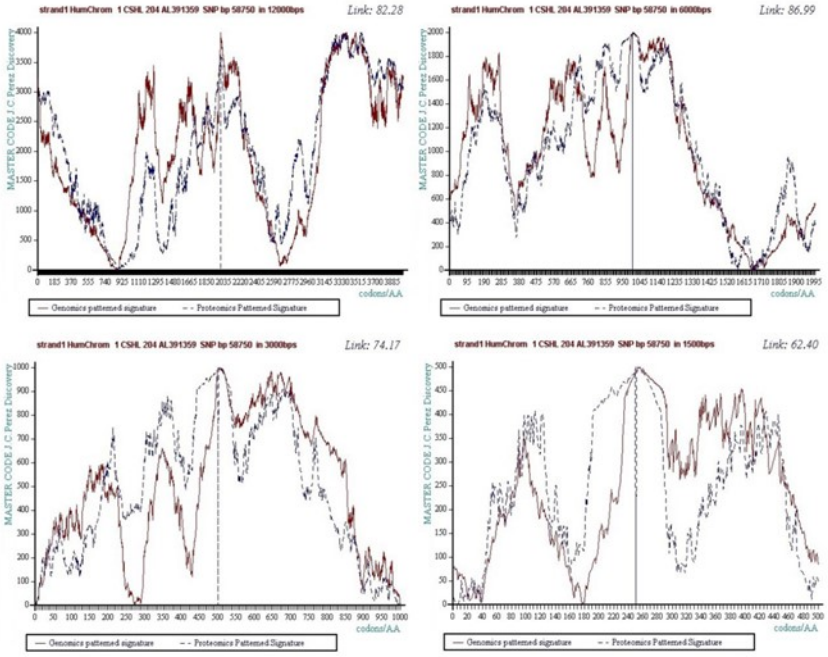
We proceed by successive analyzes of dichotomous type (embedded zooms): a sequence of 12000 bases is defined on each side of the SNP, then a second half-reduced sequence (6000 bases) is analyzed, then 3000 bases, then 1500 bases, and finally 750 bases.

Results

We analyzed several tens of SNPs according to this Master Code of DNA method. We have included only the five most representative types in this synthesis document, knowing however that ALL the cases analyzed lead to results of this nature. Here we present results for 5 examples of SNPs from the human chromosome1. They were choosen randomly within this long chromosome1 SNPs region.

Example 1 https://www.ncbi.nlm.nih.gov/projects/SNP/snp_ref.cgi?rs=2211603 Function class:
rs2211603 is located in the intron region of NM_001037341.1. Homo sapiens phosphodiesterase 4B (PDE4B), transcript variant d, mRNA
rs2211603 is located in the intron region of NM_001297440.1. Homo sapiens phosphodiesterase 4B (PDE4B), transcript variant e, mRNA
rs2211603 is located in the intron region of NM_001297441.1. Homo sapiens phosphodiesterase 4B (PDE4B), transcript variant f, mRNA
rs2211603 is located in the intron region of NM_002600.3. Homo sapiens phosphodiesterase 4B (PDE4B),

transcript variant a, mRNA
rs2211603 is located in the intron region of XM_011541566.1. PREDICTED: Homo sapiens phosphodiesterase 4B (PDE4B), transcript variant X3, mRNA



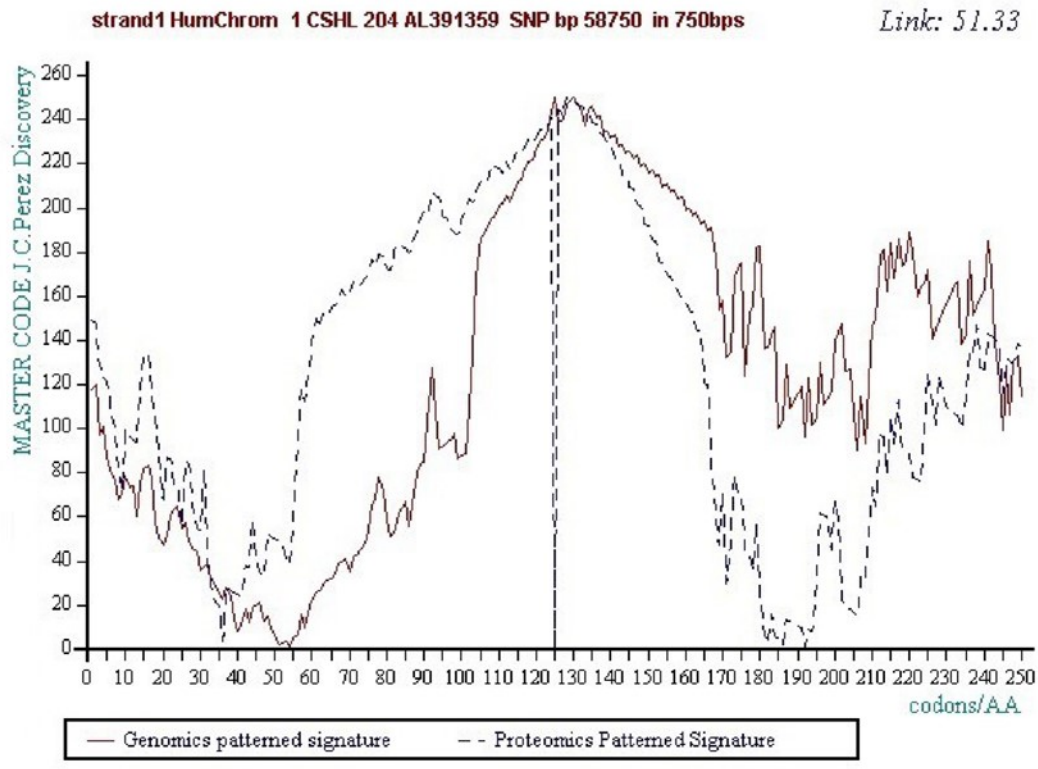
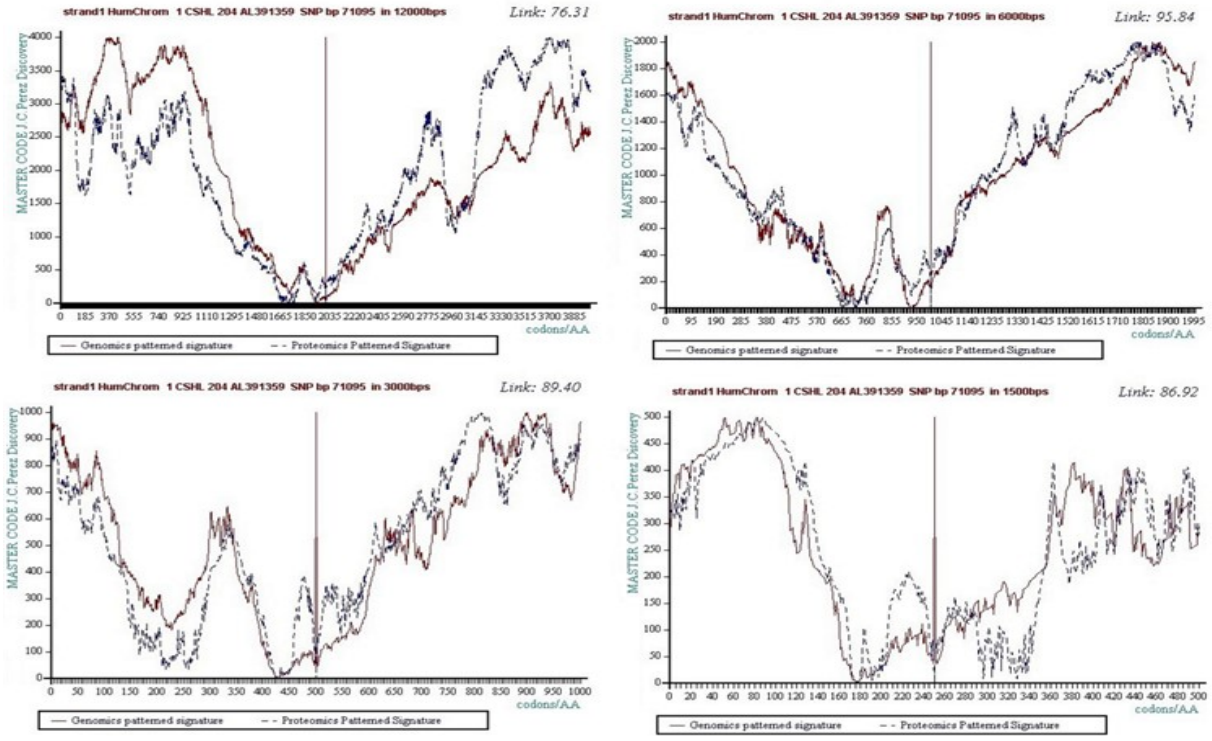


Figure 2 – Example 1, the SNP TSC1270753.

Example 2
https://www.ncbi.nlm.nih.gov/snp/?term=TSC1004969 intron variant in PDE4B **gene (enzyme cAMP-specific 3',5'-cyclic phosphodiesterase 4B)**



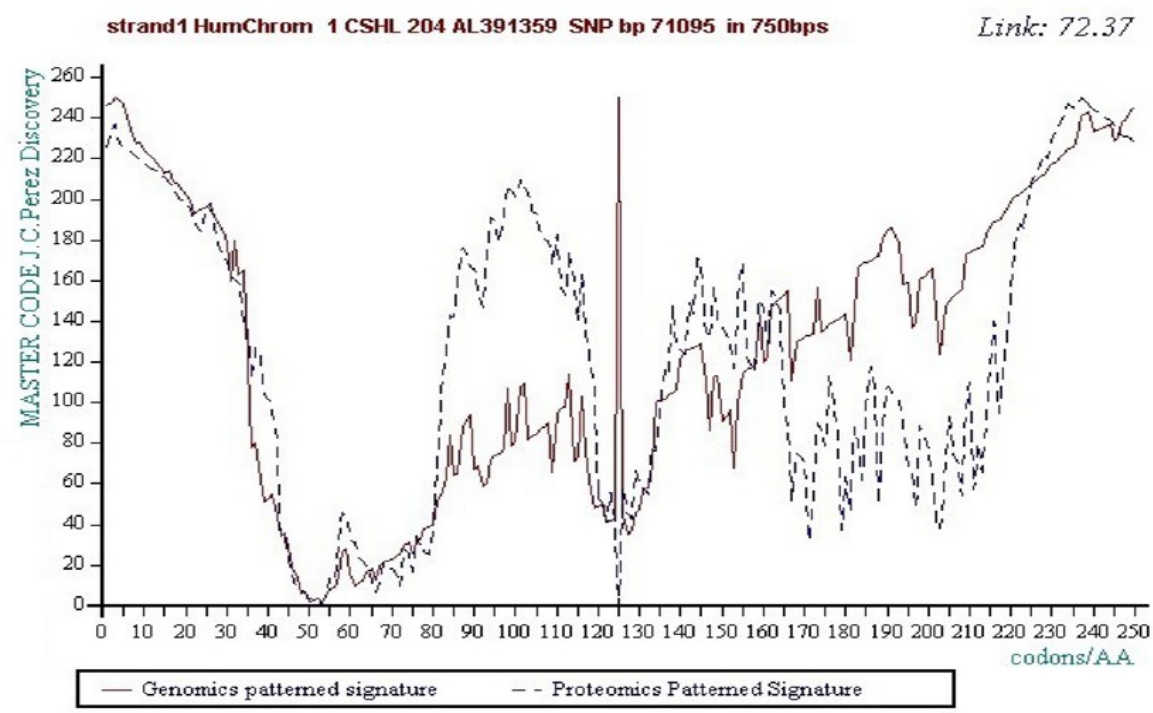
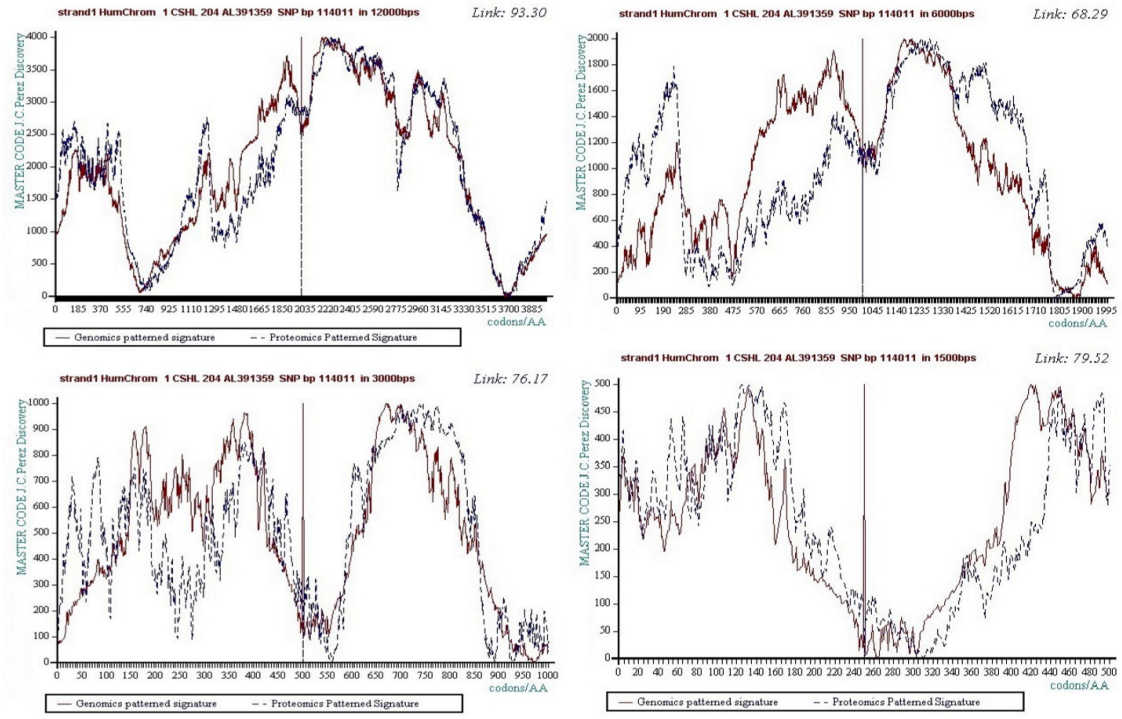


Figure 3 – Example 2, the SNP TSC1004969
https://www.ncbi.nlm.nih.gov/snp/?term=TSC0080881
intron variant in PDE4B **gene (enzyme cAMP-specific 3',5'-cyclic phosphodiesterase 4B)**

Example3:



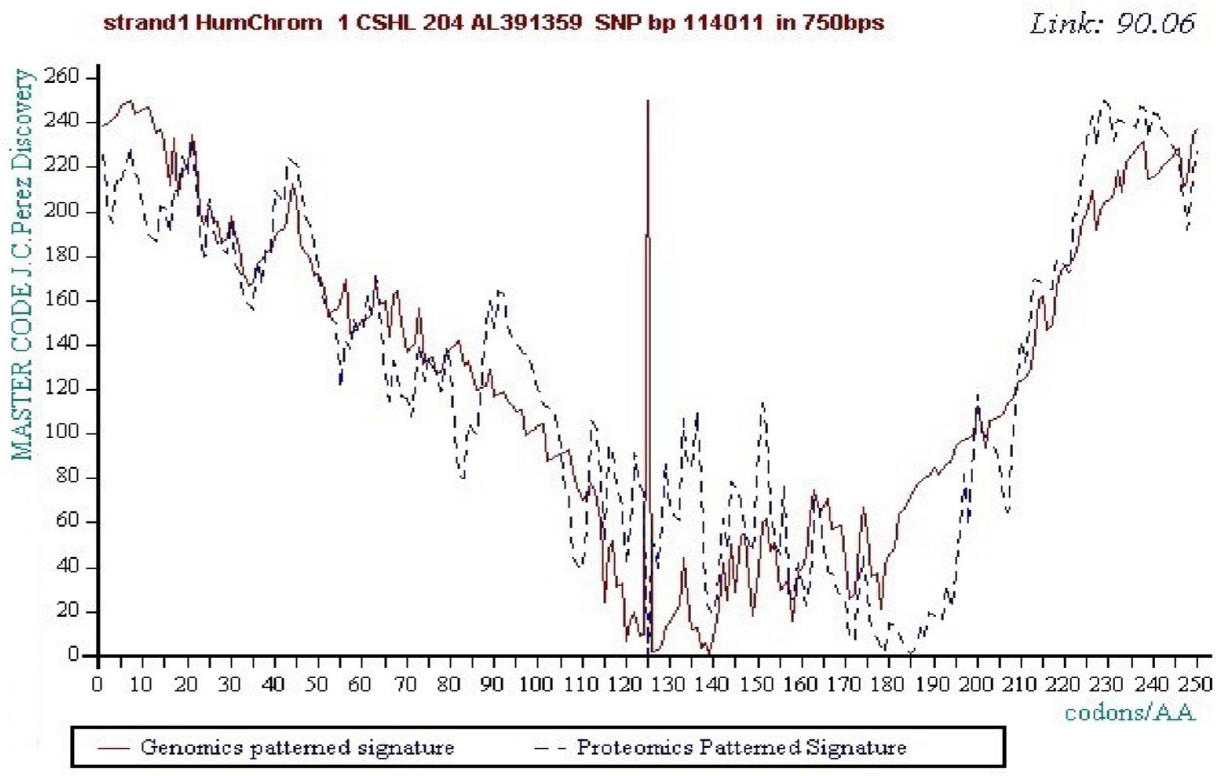


Figure 4 – Example 3, the SNP TSC008088 1

Example 4

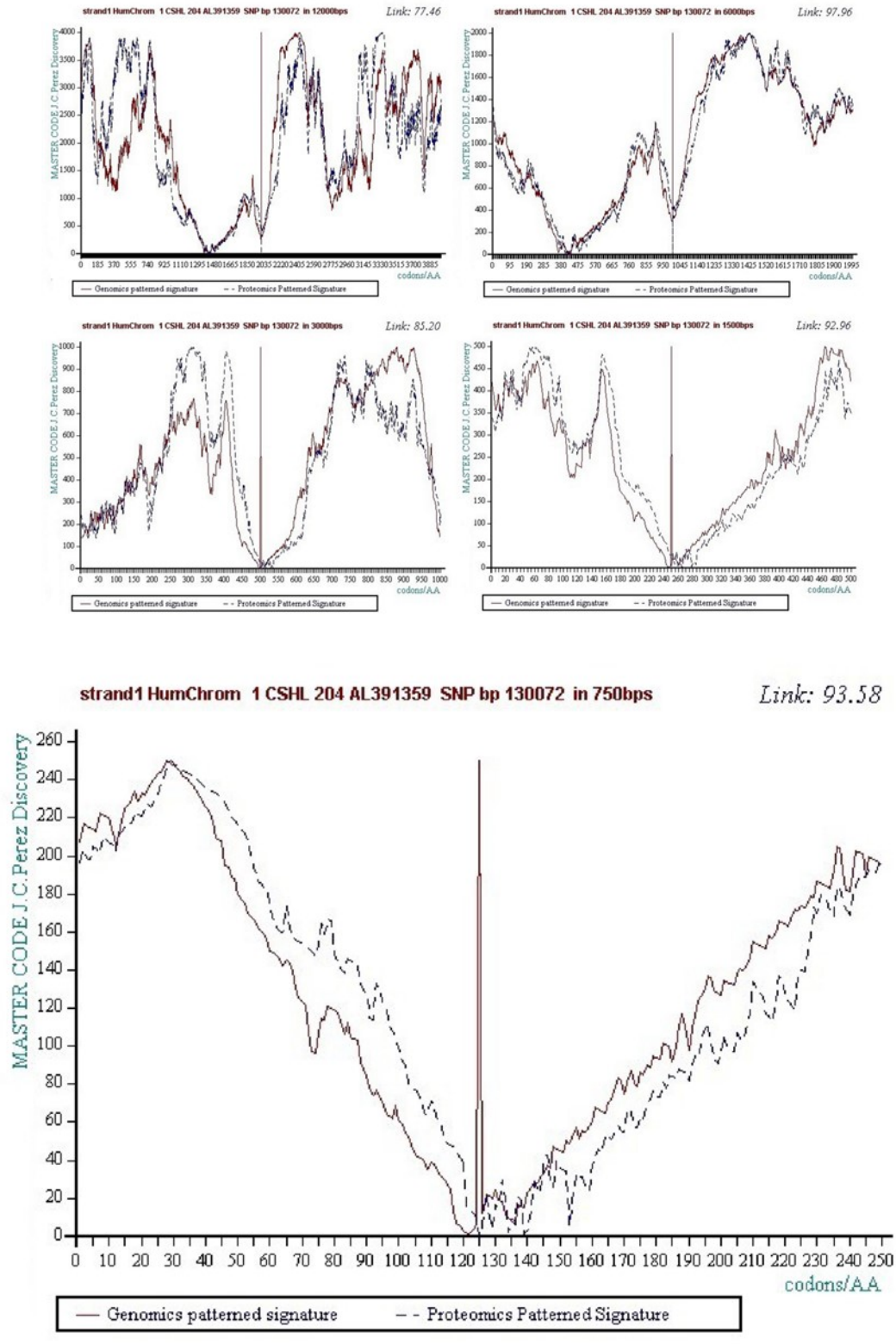
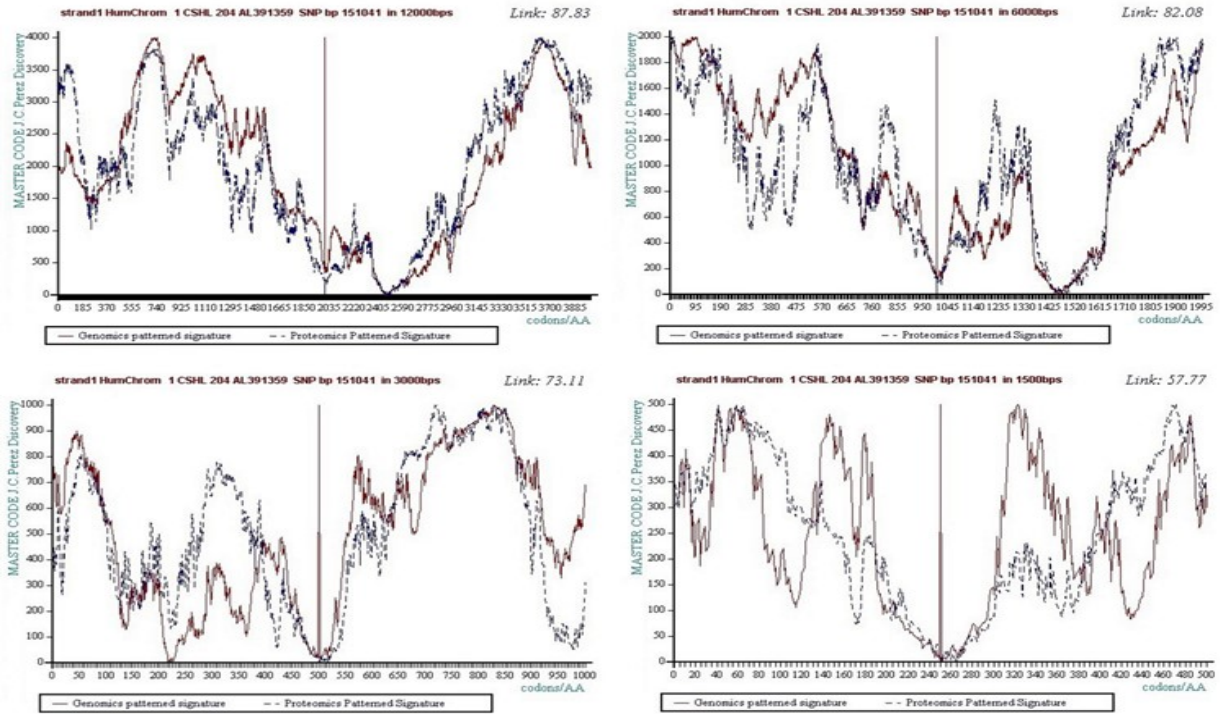


Figure 5 – Example 4, the SNP TSC1004962 https://www.ncbi.nlm.nih.gov/snp/?term=TSC1004975 intron variant in PDE4B gene (enzyme **cAMP-specific 3',5'-cyclic phosphodiesterase 4B**)

Example 5



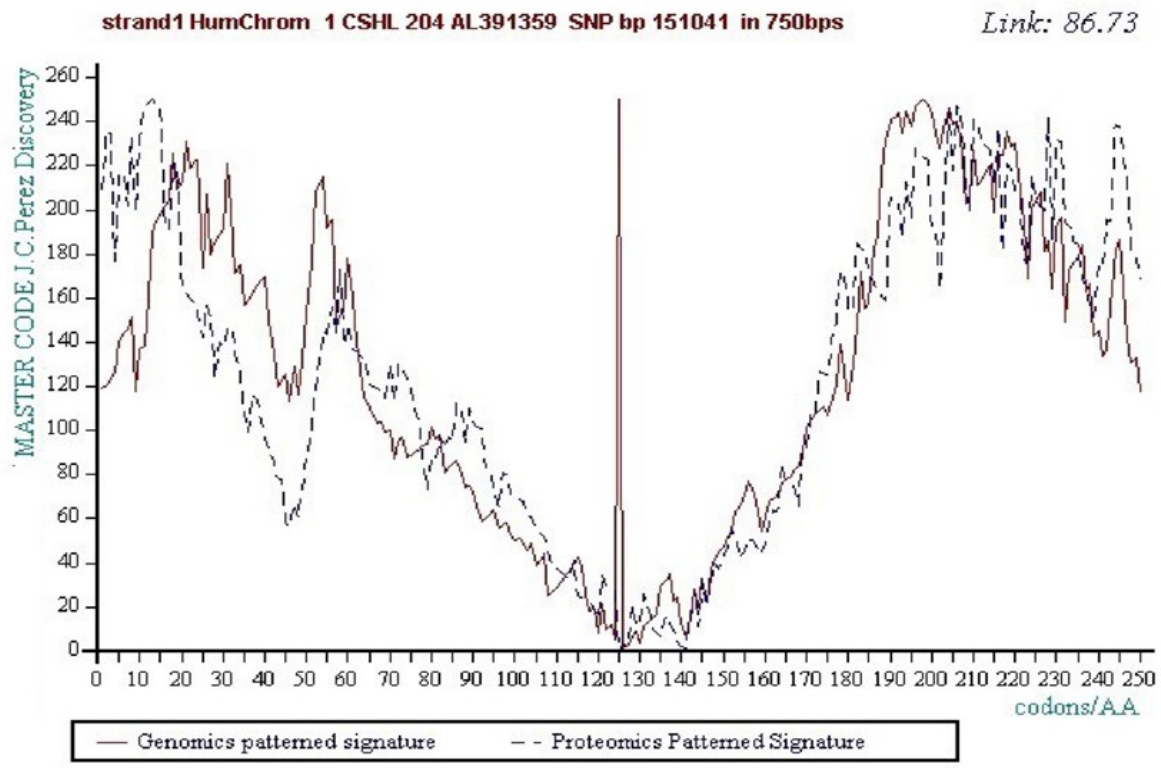


Figure 6 – Example 5, the SNP TSC1004975

Discussion

The 5 examples of SNPs analyzed here were chosen at random. On the one hand, it is noted that these 5 SNPs are located in functional regions (introns). On the other hand, in all the cases analyzed here, these SNPs are located in precise regions revealed FUNCTIONAL according to the MASTER CODE (peak or trough). We deduce that the location of SNPs is a remarkable and very important point at the functional level.

Conclusions

We will retain here the point most remarkable:

" the SNPs are LOCATED in strategic and functional sites areas within the genome. Beyond the nucleotide value of a SNP (mutation/polymorphism), It IS THUS MUCH MORE ITS ADDRESS WHICH IS VERY IMPORTANT ".

Finally, analysing the SNP locations within genomic contigs establishes the evidence of a strong link between the "FUNCTIONAL SITES" proposed by MASTER CODE law and the exact LOCATIONS of SNPs within the genome.

In other hand: SNPs LOCATIONS are FUNCTIONAL...

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 prize-winner for their interest in my research of biomathematical laws of genomes.

References

1. Baltimore D. et al, (2015) A prudent path forward for genomic engineering and germline gene modification Science. 2015 Apr 3; 348(6230): 36–38. Published online 2015 Mar 19. doi: 10.1126/science.aab1028 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4394183/ https://doi.org/10.1126/science.aab1028
2. Perez JC (1991) Chaos DNA and neuro-computers: a golden link. Speculations in Science and Technology 14: 336-346.
3. Marcer PJ (1992) Order and chaos in DNA – the Denis Guichard Prizewinner: Jean-Claude Perez. Kybernetes 21: 60-61. https://doi.org/10.1108/eb005922
4. Perez JC (1997) L'adn Décrypté, Resurgence publisher Liege, Belgium.
5. Genetics Home Reference, What are single nucleotide polymorphisms (SNPs)? , https://ghr.nlm.nih.gov/primer/genomicresearch/snp
6. Human Genome Variation Society, http://www.hgvs.org/central-mutation-snp-databases
7. Perez JC (2009) Codex Biogenesis. Resurgence, Liege Belgium. https://www.amazon.fr/Codex-Biogenesis13-codes-lADN/dp/2874340448
8. 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 https://www.omicsonline.org/openaccess/deciphering-hidden-dna-metacodes-the-great-unification--mastercode-of-biology-2153-0637-1000131.php?aid=55261# https://doi.org/10.4172/2153-0637.1000131
9. Perez JC (2011) Decoding Non-Coding DNA Codes: Human Genome Meta-Chromosomes Architecture. BIT Life Sciences' 3rd Annual World Vaccine Congress, Beijing, China.
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 (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. https://doi.org/10.4236/am.2013.410A2004
12. 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.
13. Perez J.C, (2017) Fractal Self-similarity, Scale Invariance and Stationary waves Codes Architecture Human Chromosomes DNA sequences, (submitted journal « chaos, fractals and soloitons », https://www.journals.elsevier.com/chaos-solitons-and-fractals/).

14. Perez JC (2011) Caminos Interdisciplinaios. Seminario CLAVE_INTER, Espacio Interdisciplinario, Universidad de la Republica Montevideo, Uruguay.
15. Perez J.C., (2017) DUF1220 Homo Sapiens and Neanderthal fractal periods architectures breakthrough (2017)SDRP Journal of Cellular and Molecular Physiology 1(1) http://www.siftdesk.org/articledetails/DUF1220%20Homo%20Sapiens%20and%20Neanderthal%20%20fractal%20periods%20architectures%20breakthrough/184
16. 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-genomecircular-long-range-numericalmeta-structures-are-highly-correlated-with-cancers-and-genetic-disea-1948-59561000469.php?aid=90737 https://doi.org/10.4172/1948-5956.1000469
17. Perez, J.C., (2017) Decyphering "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
18. perez. J.C., (2017) Sapiens mtDNA circular long-range numerical meta-structures are highly correlated with mtDNA diseases mutations(2017)SDRP Journal of Cellular and Molecular Physiology 1(1)

Contact Us:
SIFT DESK
Deerpark Dr, #75, Fullerton,CA,92831,United States. E-mail: helpdesk@siftdesk.org
Visit us on the web at: **www.siftdesk.org**