**Intriguing Humans And Primates Chromosomes 4**

You have free access to this content

**Supplementary Materials**

Dr Jean-claude PEREZ, PhD Maths § Computer Science, retired interdisciplinary researcher (IBM Emeritus),

7 avenue de terre-rouge F33127 Martignas Bordeaux metropole France, phone 33 0781181112

[jeanclaudeperez2@gmail.com](mailto:jeanclaudeperez2@gmail.com) [jeanclaudeperez3@free.fr](mailto:jeanclaudeperez3@free.fr)

*The basic question to which this article will answer is:*

*"Of the 24 human chromosomes besides the 2 chromosomes X and Y is there a chromosome that would be radically different from all the other chromosomes"?*

*The answer is YES. This chromosome exists: it is the chromosome4 ...*

**Part 1/3 : Methods ( 5 codes of biological Life)**

We describe here the 5 embedded steps of 'Fractal Life Codes”:

-I-Atomic mass code.

-II-Master code.

-III-Binary codes.

-IV-Undulatory code.

-V-²Standing waves meta-code

Their nature is fractal, each new code is based on entries on the previous code (s).

-I-Atomic mass code.

FUNCTION: Transform any atomic mass (real number) into an integer number between -3 and +7, corresponding to multiples of Pi / 10 (-3Pi / 10 ... / ... + 7Pi / 10).

INPUTS: a real number corresponding to any atomic mass mono-isotopic, average, or composed of one or more atoms

OUTPUTS: the « Pi-mass », an integer number between -3 and +7, corresponding to multiples of Pi / 10 (-3Pi / 10 ... / ... + 7Pi / 10).

SUMMARY :

A quick presentation of the formula for life: In [5,6,14,15] we introduced the law we call Formula for Life. This law unifies all of the components of living including bio-atoms, CONHSP and their various isotopes, to genes, RNA, DNA, amino acids, chromosomes and whole genomes. This law is the result of a simple non-linear projection formula of the atomic masses. The result of this projection is then organized in a linear scale of integer number based codes (e.g., -2, -1, 0, 1, 2, 3...) coding multiples Pi/10 regular values. These codes are called Pi-masses.



Figure 1 - The numerical projection formula of the atomic masses of any biological component.



Figure 2 - geometric meaning of the formula for life numerical projection

PROCESS :

Computing the “Formula for Life” associated with any atomic mass of Life components:

For atomic mass of any biological compound, we operate the “projection” of the atomic mass numerical value using the following operator:

where

then P = 0.742340663...

Now, consider the “v” value, where v is always a negative or zero real number.

Then consider the function:

Where Abs (v) is the absolute value of v, and « remainder » or « residue » the decimal remainder of the numerical projection

For example: remainder (-27.85) = 28-27.85 = 0.15

We then defined PPI (m) such that:

Note that (1-P.PI) is always negative because m is always positive, and (1-P.PI) is always negative.

As an example, consider the amino acid GLY:

We defined the average mass of GLY as: GLY = 75.067542

Then: (1-P.PI) . GLY = -99.99987286

Thus, PPI (GLY) = remainder [(1-P.PI) . GLY] = 0.0001271351803

Then finally, the result is a real number which we retain only the residues (decimal remainder),

PPI (GLY) = 0.0001271351803

Although no longer considered the decimal part, we note that, if we were interested in the set

(1-P.PI) . GLY = -99.99987286, this value is substantially equal to 100 = 10\*2 ... which is not “just any number” ... So then, what is the geometric reality of this projection? As Fig 2 summarizes above, everything happens as if the atomic mass was “filtered” through the competitive interference of two projections: one through a cube of side = 1 and the second through of a sphere of radius = φ × 7/4.

More precisely, let's take an example, by extending the example already presented relating to Glycine (GLY):

We calculated PPI (GLY) = 0.0001271351803.

We can then calculate the 21 PPI (GLY) -R (N.PI / 10) deviations where R (N.PI / 10) is the rightmost column in the previous table of N.PI / 10.

The following 21 values ​​are then obtained:

0.8582802112 0.1724394766 0.4865987419 0.8007580073 0.1149172727 0.429076538 0.7432358034

0.05739506874 0.3715543341 0.6857135995 0.0001271351803 0.3140321302 0.6281913955 0.9423506609

0.2565099263 0.5706691916 0.884828457 0.1989877223 0.5131469877 0.8273062531 0.1414655184

It can be seen that the minimum difference (underlined) corresponds to an angle of 0 °, so to N = 0.

So we'll say that "PI-MASS (GLY) = 0"

The successive stages characteristic having transformed the mass GLY in PI-MASS = 0 are:

Atomic mass GLY = 75.067542

PPI projection (GLY) = 0.0001271351803

Angle N.PI / 10 the nearest: 0.000000000 (N = 0)

Approximate error: EPI (GLY, 0) = 0.0001271351803

PI-MASS (GLY) = 0 either 0 ° or else 0.PI / 10

As a documentary, we will calculate the PI-MASSES relating to 10 sgnificant different genetic materials. Consider any atomic mass « m », which may be that of a bio-atom, of a nucleotide, a codon or an amino acid or any other genetic compound based on bio-atoms or even, any atoms from Mendeleiëv periodic table.

This process will work especially on the average masses (mix of various isotopes % proportions). But it may also be applied to a particular isotope or any derivative of specific atomic mass proportions of the various isotopes.

Table 1 - A set of Pi-mass projections for some main Life compounds.

| Nature | Molecule or bioatom | Average atomic mass | Projection PPI(m) | Pi-mass NPI(m) = N.Pi/10 | Angle | Error EPI(m,N) |
| --- | --- | --- | --- | --- | --- | --- |
| Bioatom | C12 Carbon isotope 12 | 12.000000 | 0.01441631887 | 0 PI/10  (0°) | 0° | 0.01441631887 |
| Bioatom | C (Carbon average mass) | 12.0111 | 0.0003703460363 | 0 PI/10  (0°) | 0° | 0.0003703460363 |
| Nucleotide | G (G nucleotide) | 150.120453 | 0.01974469326 | 0 PI/10  (0°) | 0° | 0.01974469326 |
| Codon | Codon TCA | 369.324471 | 0.01106361166 | 0 PI/10  (0°) | 0° | 0.01106361166 |
| Codon | Codon UCA | 355.297477 | 0.6968708101 | -1 PI/10  (-18°) | -18° | 0.0110300755 |
| Codon | Codon AGT (TCA complement) | 409.349065 | 0.6930222208 | -1 PI/10  (-18°) | -18° | 0.0071814862 |
| double-stranded DNA | DNA double strand : TCA+AGT | 778.673536 | 0.7040858325 | -1 PI/10  (-18°) | -18° | 0.0182450978 |
| Amino acid | PRO (Proline amino acid) | 115.13263 | 0.6281423922 | +2 PI/10  (+36°) | +36° | 0.0001761385 |
| Amino acid | LYS (Lysine amino acid) | 146.190212 | 0.2553443926 | +4 PI/10  (+72°) | +72° | 0.0012926688 |
| Peptide link | CONH Peptidic link | 43.025224 | 0.6847234457 | -1 PI/10  (-18°) | -18° | 0.0011172889 |
| Notes: Projections PPI(m) are multiples of Pi:10. Example: 0.314... = 1Pi/10, 0.628... = 2Pi/10, etc... But, symmetrically vs. 0Pi/10,  it appears another regular scale of attractors in the negative region of Pi/10: -1Pi/10 = 1-0.314 = 0.685..., -2Pi/10 = 1-0.628... | | | | | | |

Table 2 - Synoptic of PI-masses of different components of genetics.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | -3 PI/10  and less | -2  PI/10 | -1 PI/10 | 0 PI/10 | +1 PI/10 | +2 PI/10 | +3 PI/10 | +4 PI/10 | +5 and +7 PI/10 |
| Bioatoms | P(-4pi/10) |  | H O | C | N |  |  | S |  |
| Nucleotides |  |  |  | U G I | T C A |  |  |  |  |
| Others compounds |  | Ph/  sugar  RNA | CONH | H2O | CH2  Ph/  sugar DNA |  |  |  |  |
| Amino acids |  |  | Asp | Asn Glu  Gly Ser | Ala Gln  His Thr | Pro Tyr  Cys (+2) | Arg Phe  Trp Val | Ile Leu  Lys  Met (+4) | Cys (+5)  Met (+7) |
| Codons DNA | ggg | gtg gcg gag tgg  cgg agg  ggt ggc  gga | ttg ctg  atg gtt  gtc gta  tcg ccg  acg gct  gcc gca  tag cag  aag gat  gac gaa  tgt tgc  tga cgt  cgc cga  agt agc  aga | ttt ttc  tta ctt  ctc cta  att atc  ata tct  tcc tca  cct ccc  cca act  acc aca  tat tac  taa cat  cac caa  aat aac  aaa |  |  |  |  |  |
| Codons RNA | uuu uug  guu gug  ugu ugg  ggu ggg | uuc uua  cuu cug  auu aug  guc gua  ucu ucg  gcu gcg  uau uag  gau gag  ugc uga  cgu cgg  agu agg  ggc gga | cuc cua  auc aua  ucc uca  ccu ccg  acu acg  gcc gca  uac uaa  cau cag  aau aag  gac gaa  cgc cga  agc aga | ccc cca  acc aca  cac caa  aac aaa |  |  |  |  |  |

Table 3 – Sensitivity of “Formula for Life” vs atomic mass fine tuning.

|  |  |  |  |
| --- | --- | --- | --- |
| Genetic compounds | Atomic mass m | PI-mass N PI/10 | Error EPI(m,N) |
| Regular GLYCINE Amino acid.  GLY=NH2-CH2-COOH  HYDROGEN atom mass H=1.007947 | 75.067542 | 0 PI/10 (0°) | 0.0001271351803 |
| GLYCINE modified by the atomic mass of only one of the HYDROGEN atoms that becomes H\*=1.0080424374  (the other H remain unchanged).  GLY=NH2-CHH\*-COOH | 75.06763744 | 0 PI/10 (0°) | 3.173283858 10\*-11  soit 0.0000000000317… |
| Electron (à titre indicatif) | 0.000549 | 0 PI/10 (0°) | 0.0007313405 |

A startling observation opens the door to enormous opportunities in astrobiology: Table 4 and Figure 3 shows a very curious fact: the Pi-mass projection formula seems optimal only for the atomic masses of average atomic weights of basic life bioatoms C O N H. Instead of tiny perturbations on these atomic masses and atomic masses of the individual isotopes (example O16) of each of these atoms “destroy” the optimality and fine-tuning of these projections then, also, consequently all resulting master code perfect tuning. Example here (Table 4) for the Pi-mass projection of Oxygen isotopes and % average weighted atom mass. As shown in Figure 3, isotopes of oxygen lightest and heaviest O16 O18 both produce an error on the projections Pi-mass much higher than that of the average atomic mass of that atom of oxygen consisting of: 99,757% + 0.04% O16 O17 + 0.2% O18.

Table 4 – Example of Pi-mass fine-tuned projection selectivity for Oxygen average mass vs. individual isotopes

| Atom | Isotope | Relative atomic mass | % isotopic composition | Pi projection residue and Pi mass value | Pi-mass NPI(m) = N.Pi/10 | Error EPI(m,N) |
| --- | --- | --- | --- | --- | --- | --- |
| Oxygen | Average % balance | 15.9994(3) | - | 0.686647751  0.685840735 | -1 | 0.000807016 |
|  | O16 | 15.994 914 619 56(16) | 0.997 57(16) | 0.692662834  0.685840735 | -1 | 0.006822099 |
|  | O17 | 16.999 131 70(12) | 0.000 38(1) | 0.354913152  0.371681469 | -2 | 0.016768318 |
|  | O18 | 17.999 161 0(7) | 0.002 05(14) | 0.022742056  0.000000000 | 0 | 0.022742056 |
| Notes: 0.685840735 = 1 – Pi/10, 0.371681469 = 1 – 2Pi/10 | | | | | | |

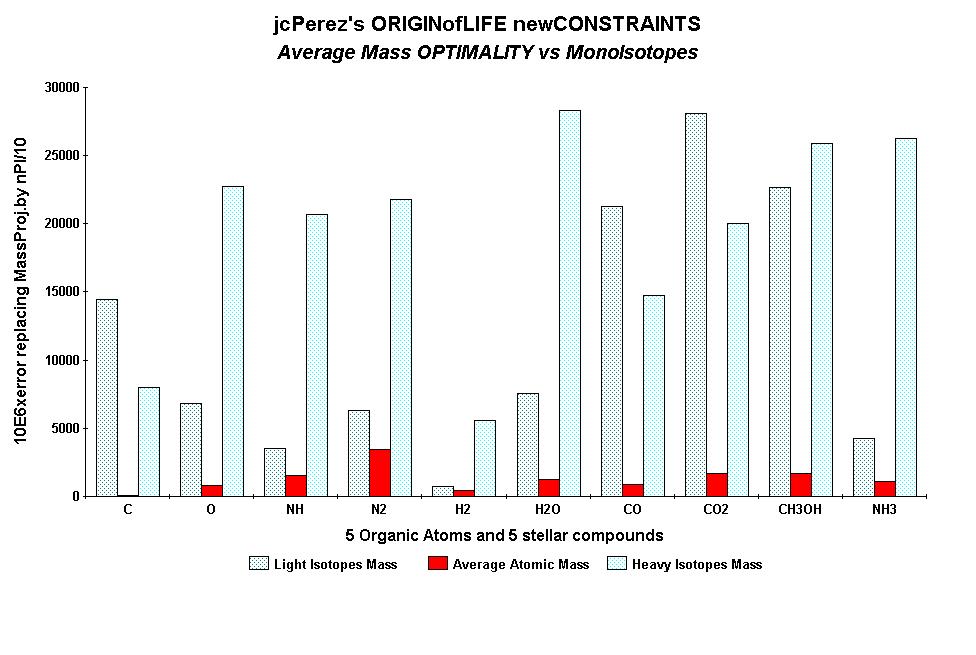


Figure 3 - The OPTIMALITY of the average atomic mass is proved here on the PI-masses of ten primordial organic components.

-II-Master code.

FUNCTION: Global integration of Genomics and Proteomics Pi-mass codes at the whole sequence level.

INPUTS: DNA double stranded sequence by codons pairs producing Genomics (DNA) Pi-mass code and Proteomics (coresponding potential amino acid) Pi-mass code (an integer number for each codons pair)

OUTPUTS: 2 Genomic and Proteomic numerical vectors generating 2 patterned 2-D Genomic and Proteomic Images signatures.

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 (3, 4) then produces 2 patterned images (2D curves, see Figures 4, 5, 6) 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 (Figure 4). Our thousands of genes and genomes Master Code analyses (viruses, archaeas, bacteria, eucharyotes) demonstrated that the extremums (max and min) signify functional regions like proteins active sites, fragility points like chromosomes breakpoints).

PROCESS:

An overviev on the “Biology Master Code” Great Unification of DNA, RNA and Amino acids : this process run 3 sequential substeps :

a) The coding step

b) The globalization and integration step

c) The great Unification between Genomics and Proteomics Master Code images

It may seem surprising that such a fine tuned process like biology of Life requires the use of three languages as diverse and heterogeneous as DNA with its alphabet of four bases TCAG; RNA with its alphabet of four bases UCAG; and proteins with their language of 20 amino acids. Obviously, the main discoveries in biology were made by those who managed to unearth the respective areas and “bridges” between these three languages. However, any “aesthete” researcher will think the table of the universal genetic code seems rather “ad hoc” and heterogeneous.

Starting only from the double-stranded DNA sequence data, the “Master Code” is a digital language unifying DNA, RNA and proteins that provide a common alphabet (Pi-mass scale) to the three fundamental languages of Genetics, Biology and Genomics.

The construction method of “the Master Code” will be now fully described below. It will highlight a significant discovery we summarize as follows: “Above the 3 languages of Biology - DNA, RNA and amino acids, there is a universal common code that unifies, connects and contains all these three languages”. We call this code the “Master Code of Biology.”

Here is a brief description of our process for computing the Master Code:

a) The coding step: First, we apply it to any DNA sequence encoding a gene or any non-coding sequence (formerly mislabelled as junk DNA). So it may be either a gene, a contig of DNA, or an entire chromosome or genome. In this sequence, we always consider double-stranded DNA as we explore the following three codon reading frames and following the two possible directions of strand reading (3’ ==> 5’ or 5’ ==> 3’). The base unit will always be the triplet codon consisting of three bases.

As shown in above sample, we calculate the Pi-mass related to double stranded triplets DNA bases, double stranded triplets RNA bases, and double-stranded pseudo amino acids. In fact, for each DNA single triplet codon, we deduce the complementary Crick Watson law bases pairing. We do the same work for RNA pseudo triplet codon pairs, then, similarly for amino acids translation of these DNA codon couples using the Universal Genetic Code table. Then we obtain 3 samples of pairs codes: DNA, RNA and amino acids and this, systematically even when this DNA region is gene-coding or junk-DNA.

A simple example: the starting region of Prion gene:

DNA image coding:

ATG CTG GTT CTC TTT...

-1 -1 -1 0 0...

Complement:

TAC GAC CAA GAG AAA...

0 -1 0 -2 0...

RNA image coding:

AUG CUG CUU CUC UUU...

-2 -2 -3 -1 -3...

Complement:

UAC GAC GAA GAG AAA...

-1 -1 0 -2 0...

Proteomics image coding:

MET LEU VAL LEU PHE

4 3 3 4 3...

Complement:

TYR ASP GLN GLU LYS

2 -1 1 0 4...

Pi-masses corresponding to two strands are then added for each triplet:

Double strand DNA image coding: -1 -2 -1 -2 0...

Double strand RNA image coding: -3 -3 -3 -3 -3...

Double strand Proteomics image coding: 6 2 4 4 7...

This produces three digital vectors relating to each of the 3 DNA, RNA, and proteomics coded images. At this point we already reach an absolutely remarkable result, as symbolized in Figure 1.

We will focus now – exclusively - on the DNA code (genomics) and amino acids code (proteomics).

b) The globalization and integration step: To these two numeric vectors we apply a simple globalization or integration linear operator. It will “spread” the code for each position triplet across a short, medium or long distance, producing an impact or “resonance” for each position and also on the most distant positions, reciprocally by feedback. This gives a new digital image where we retain not the values but the rankings by sorting them.

We run this process for each codon triplet position, for each of the three codon reading frames and for the two sequence reading directions (3’ ==> 5’ and 5’ ==> 3’).

For example, to summarize this method: on starting area of the GENOMICS (DNA) code of Prion above, the “radiation” of triplet codon number 1 would propagate well:

-1 -2 -1 -2 0... ==>

-1 -3 -4 -6 -6...then, we cumulate these values: -20

So we made a gradual accumulation of values.

The same operation from the codon number 2 produces:

-1 -2 -1 -2 0... ==>

-2 -3 -5 -5...then, we cumulate these values: -15

etc.

Similarly, the same process on starting area of the PROTEOMICS code of Prion above, the “radiation” of triplet codon number 1 would propagate well:

6 2 4 4 7... ==>

6 8 12 16 23...then, we cumulates these values: 65

So we made a gradual accumulation of values.

The same operation from the codon number 2 produces:

6 2 4 4 7... ==>

2 6 10 17...then, we cumulate these values: 35

etc.

Finally, after computing by this method these “global signatures” for each codon position at Genomics and Proteomics levels, we sort each genomic and proteomic vector to obtain the codon positions ranking: example: as illustrated bellow, the Genomics ranking patterned signature is 2 1 4 3 5 for this Prion starting 5 codons mini subset sequence of 5 codons positions (arbitrary values). Then, to summarize the Master Code computing method on these 5 codon positions starting Prion protein sequence:

Genomics signature:

Codon 1:

Codon / Basic codes / Potentials (with circular closure) / circular complements:

!

-1 -2 -1 -2 0

-1 -3 -4 -6 -6

0

Cumulates: -20

Codon 2:

Codon / Basic codes / Potentials (with circular closure) / circular complements:

!

-1 -2 -1 -2 0

-2 -3 -5 -5

-6

Cumulates: -21

Codon 3:

Codon / Basic codes / Potentials (with circular closure) / circular complements:

!

-1 -2 -1 -2 0

-1 -3 -3

-4 -6

Cumulates: -17

Codon 4:

Codon / Basic codes / Potentials (with circular closure) / circular complements:

!

-1 -2 -1 -2 0

-2 -2

-3 -5 -6

Cumulates: -18

Codon 5:

Codon / Basic codes / Potentials (with circular closure) / circular complements:

!

-1 -2 -1 -2 0

0

-1 -3 -4 -6

Cumulates: -14

Final rankings:

Codon positions: 1 2 3 4 5

Potentials: -20 -21 -17 -18 -14

Rankings: 2 1 4 3 5

Then we run similar computing for Proteomics...

Codon 1:

Codon / Basic codes / Potentials (with circular closure) / circular complements:

!

6 2 4 4 7

6 8 12 16 23

0

Cumulates: 65

Codon 2:

Codon / Basic codes / Potentials (with circular closure) / circular complements:

!

6 2 4 4 7

2 6 10 17

23

Cumulates: 58

Codon 3:

Codon / Basic codes / Potentials (with circular closure) / circular complements:

!

6 2 4 4 7

4 8 15

21 23

Cumulates: 71

Codon 4:

Codon / Basic codes / Potentials (with circular closure) / circular complements:

!

6 2 4 4 7

4 11

17 19 23

Cumulates: 74

Codon 5:

Codon / Basic codes / Potentials (with circular closure) / circular complements:

!

6 2 4 4 7

7

13 15 19 23

Cumulates: 77

Final rankings:

Codon positions: 1 2 3 4 5

Potentials: 65 58 71 74 77

Rankings: 2 1 3 4 5

Then finally:

Codon position: 1 2 3 4 5

Genomics vector: 2 1 4 3 5

Proteomics vector: 2 1 3 4 5

To complete, the same work must be also operate on each codon reading frame...

Meanwhile, a more synthetic means to compute these “long range potentials” for each codon position is the following formula:

Cumulate potential of codon location “i”

Then finally,

Example for Genomics image of codon “i”

The initial computing method described above provides:

-1 -2 -1 -2 0... ==>

-1 -3 -4 -6 -6...then, we cumulate these values: -20

becomes, using this new generic formula:

(-1)x5 + (-2)x4 + (-1)x3 +(-2)x2 +(0)x1 = (-5) + (-8) + (-3) + (-4) + 0 = -20

c) The great Unification between Genomics and Proteomics Master Code images: When applying the process described above in any sequence – gene coding, DNA contig, junk-DNA, whole chromosome or genome - a second surprise appears just as stunning as that of RNA neutral element. We find that for one of the three reading frames of the codons given, the Genomics patterned signature and the Proteomics patterned signature are highly correlated.

Contrary to the three genomics signatures which are correlated in all cases, the proteomics signatures are correlated with genomics signatures only for one codon reading frame, and generally in dissonance for the two remaining codon reading frames. Also, there are perfect local areas matching’s focusing on functional sites of proteins, hot-spots, chromosomes breaking points, etc.

In this global correlation, specific codon positions were a perfect match. This is remarkable when regions correspond to biologically functional areas: hot-spots, the active sites of proteins, breakpoints and chromosome fragility regions (i.e., Fragile X genetic disease), etc.

EXAMPLES:

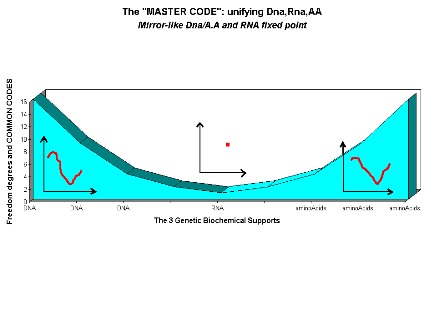
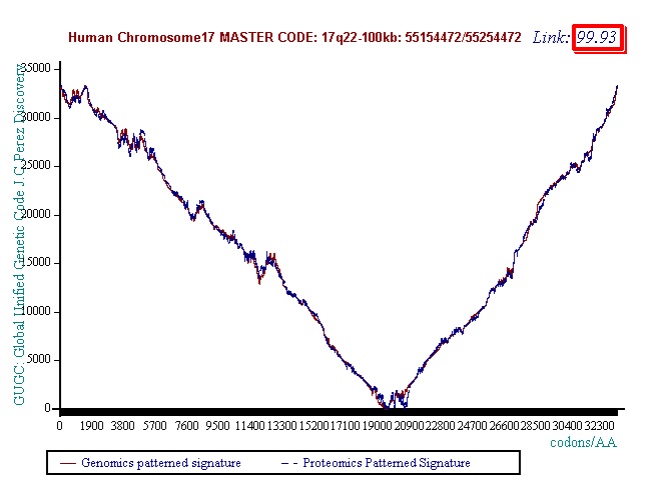


Figure 4 – A symbolic representation of the 3 worlds of double stranded DNA (Genomics) highly corelated with potential double stranded amino acids (Proteomics) while RNA double stranded image is like a neutral element.

Figure 5 - illustration of the high correlation coupling between genomics and proteomics images of a 100kb stretch of chromosome 7 (99.93% correlation).

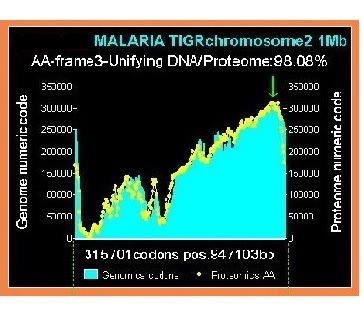


Figure 6 - illustration of the high correlation coupling between genomics and proteomics images of a complete MALARIA chromosome (98.08% correlation)..

-III-Binary codes.

FUNCTION:

We compute first-order differential texture and roughness analysis (Leibniz) on Master Code Genomics and Proteomics images.

INPUTS: Genomics and Proteomics images data from step -II- (2 numerical integer numbers vectors).

OUTPUTS: 2 binary vectors (0/1) related Genomics and Proteomics textures analysis.

SUMMARY :

If this work is carried out for Genomic patterned pictures, we see that if this trend seems self-organized around one attractor for DNA double strand (Genomics), it shows two levels, two “attractors” for the second (Proteomics). A curious fact then emerges: although two genomics and proteomics curves are still highly correlated in their respective forms and shapes, we discover that their textures are radically different.

Thus the population of Genomics curves will be relatively dispersed around one single withdrawing attractor in a kind of Gaussian dispersion, while the population of Proteomic curves will be distributed around two attractors, bringing out a kind of binary frequency modulation.

We are witnessing the emergence, the “birth” of a Binary Code as demonstrated by Figures 8 , 9, 10, and particularly 3D!

Let us not forget that the initial information was the atomic mass of each bio-atom, which is... a real (decimal) number! Then it is transformed into a code which is an integer number... and it now emerges Binary Code, then 0/1 bits which are binary numbers!

Preliminary analysis shows that the average levels of these two attractors are around 0.61 (61%) and 0.30 (30%) then appear to be in a ratio of two. We will return to these two values bellow...



Figure 7 – Evidence of first-order differential texture and roughness analysis (Leibniz) on Master Code Genomics and Proteomics images.

Here is a small example of a sequence of 312 bases where genomics (red) and proteomics (blue) signature (amino acids) are studied. Note the beauty of these mathematical structures which always increase and that some compare to artistic works by M.C. Escher or J.S. Bach.

PROCESS :

Towards discrete Waveforms and logic Biobits overlapping whole chromosomes and genomes

Here we analyze the texture, that is to say, the “roughness” of genomic and proteomic signatures provided by the Master Code. For this, we need only to analyze the slopes or mathematical differentiations from these patterned curves: slopes and gradients - in the sense of LEIBNIZ? - of order 1.

The curves of the Master Code are discontinuous (each point represents a position of triplet codon).

If we note M (i) the Master Code function as defined in -II-, then we agree that:

slope = 1 = ”growing” i.e., “increase” if M (i + 1) > M (i)

and slope = 0 = “decreasing” i.e., “decrease” if M (i + 1) < M (i).

Biobits: The emergence of a “binary language” from the Proteomics Master Code of any DNA sequence:

A detailed analysis of the texture of Genomics and Proteomics curves reveals a strange phenomenon: as shown in Figures 8, 9and 10, a curious roughness or “sawtooth” usually characterizes these images. This somehow amounts to a search for the “derivative of order 1”, that is to say the slope between two successive points. It becomes apparent that these slopes are mostly in the same direction: always growing or always decreasing.

EXAMPLES:

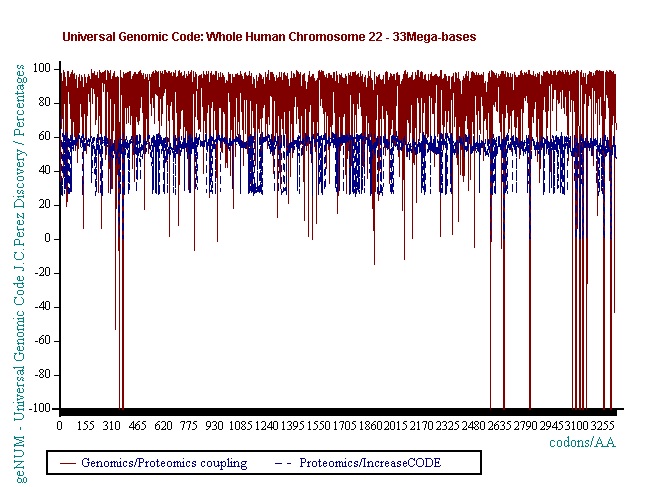


Figure 8 – Clouds of points - In whole chromosome 22, population of Genomics curves will be relatively dispersed around one single withdrawing attractor in a kind of Gaussian dispersion (red), while the population of Proteomic curves will be distributed around two binary attractors (blue).

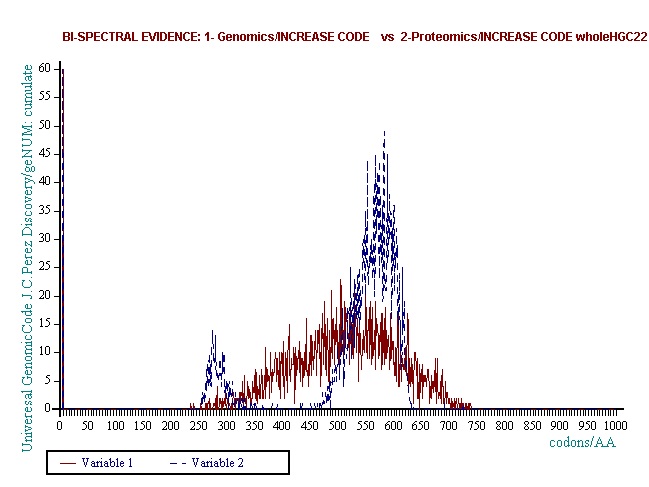
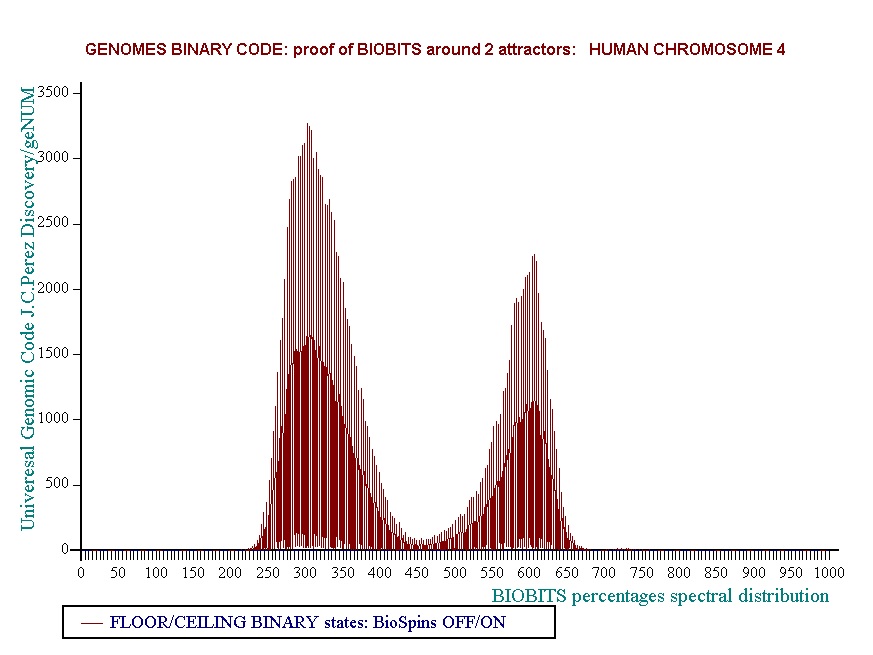


Figure 9 - Frequency distribution - In whole chromosome 22, population of Genomics curves will be relatively dispersed around one single withdrawing attractor in a kind of Gaussian dispersion (red), while the population of Proteomic curves will be distributed around two binary attractors (blue).

Figure 10 - whole chromosome 4 – evidence of the perfect Proteomics Binary Code.

-IV-Undulatory code.

FUNCTION: Building discrete waveforms from Genomic Master Code (-II-).

INPUTS: Genomics and Proteomics images data from step -II- (2 numerical integer numbers vectors).

OUTPUTS: descrete waveforms related Genomics textures analysis.

SUMMARY :

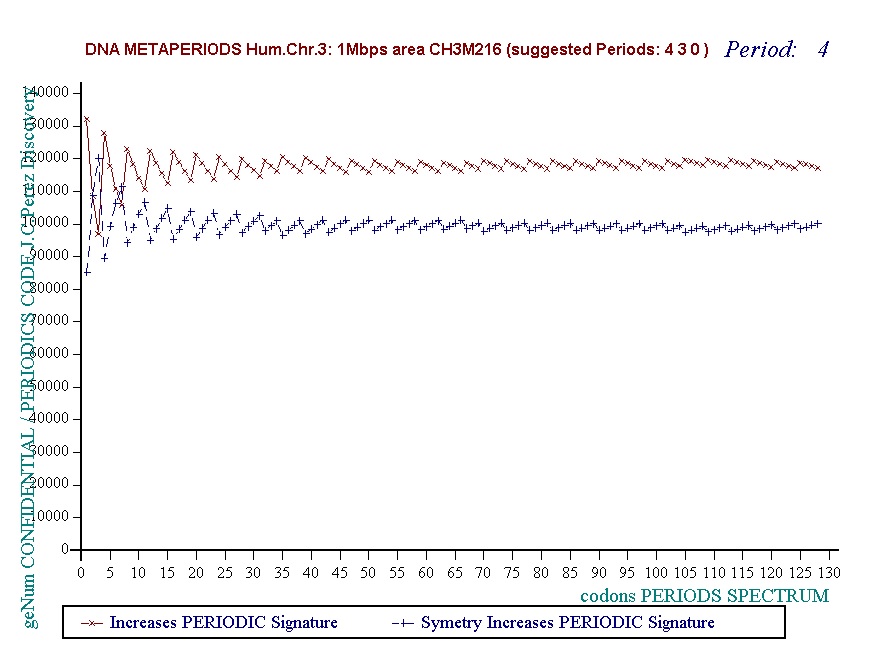
Discrete Waveforms: The emergence of “a modulated waveform code” from the Genomics Master Code of any DNA sequence: the generalization of previous gradient differentiations from second, third or nth gradient differentiation order now highlight “bits”… But waveforms, more precisely discrete waveforms of which we will measure periods: period of short-wave or 2 or 3 or even medium-wave wavelengths (greater than 10 times).

PROCESS:

Thus, we calculate exhaustively all successive gradients or slopes: S(i, i + 1), and S(i, i + 2), S(i, i + 3), ... S(i, i + n). From all these successive gradients periodicities emerge.

Figure 11 shows shortwave period = 4 codons, then 12 bases pairs in one million base pairs within human chromosome 3. Figure 12 shows long wave period = 12 codons, then 36 base pairs in the first 300000 base pairs within one of the largest human genes, the gene for the genetic disease Duchenne DMD.

EXAMPLES:

Figure 11 – Short period 4 waves from Genomics master code images (1 million bases from chromosome 3)..

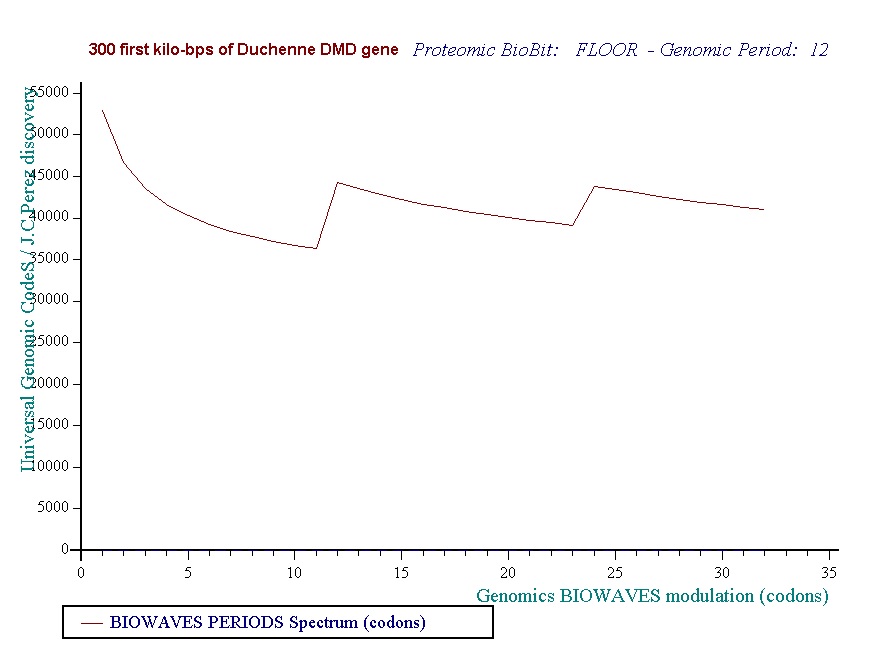


Figure 12 – Long waves from Genomics master code images (300000 bases from DMG gene (Duchenne Muscular Distrophy).

-V- Standing waves meta-code

FUNCTION:

The Genomics master code (-II-) is generalized to meta-codons that no longer have 3 nucleotides as a codon, but 4, 5, ... 377 nucleotides. Then we analyze the textures by the undulatory code (-IV-). It then appears dissonances and resonances that will reveal periods of discrete waves, resonances, and standing waves. The Genomics Binary code analysis (-III-) confirms these periods using a complementary independant method.

INPUTS: Double strand DNA sequence Pi-mass grouped by meta-codons (each Pi-mass is = -1 times number of « G » bases in meta-codon double strand or also = -1 times number of « C+G » bases in single strand meta-codon.

OUTPUTS: Peiod and resonance standing wave computed by two complementary methods.

SUMMARY :

We introduce here a method of global analysis of the roughness or fractal texture of the DNA sequences at the chromosome scale. To do this, we generalize the method of numerical analysis of the "Master Code" (-II-). Thus, we restructure the sequence into different generic sequences based on "meta codons", no longer triplets of 3 nucleotides, but values ranging from 17 to 377 nucleotides, ie 360 simulations. This method of analysis will then reveal, in most cases, discrete waves or interferences, most often dissonances (based on Genomics Undulatory waves described here in -IV-). However, sometimes there will emerge kinds of resonances where all scales of analysis appear to be in symbiosis.

PROCESS:

The discrete interferences fields resulting from the analysis of an entire chromosome are therefore a three- dimensional space: Dim y (vertical) restructuring in meta codons of lengths 17 to 377 nucleotides Dim x (horizontal) Leibnitz differentiations such that prmary 1/2 secondary 1/3... 1/4 ... 1 / n Dim z cumulated populations from the "Master code" operators. The + 1 / -1 derivatives will be of type increase, ie +1 if derivative increasing and will be of type decrease, ie -1 if derived decreasing. In this context we will explore these 3D spaces in 2 forms:

-Horizontally (IV- Undulatory code), meta codons dimension: curves for a given meta codon dimension, see in the example "resonances" below (see Figures 17 and 18).

-Vertically (-III- Genomics binary code), spectral differentiation: discrete series d2-d1 is +1 if increase and -1 if decrease (see Figure 19). We represent in top the +1 and in low the -1, (see Figure 19).

Table 5 – Computing the periodic standing waves and resonances for various metacodons Genomics Master code..

Dim x d1 d2 …/... d100

Dim y

…/... 377

Horizontal scan : exp. meta codons of 22 bases : 22 761233 774174 779102 783714 786854 …/...

(see Figure 13)

Vertical scan example derivations of first order: 1 if d2>d1 and -1 if d2<d1 then : -1 1 -1 1 -1 1 1 -1 1 -1 1 1 …/...

(see Figures 14, 15).

Figure 13 - zoom on vertical scan method revealing PERIOD = 22 from HG38 reference chromosome21.

These two independent methods lead in all the cases analyzed to the same period value: here, for example, the period "horizontal scan" is a resonance of 22bp (Figure 18) and the period "vertical scan" is a period of repeatability of 22bp also (Figure 15).

Figure 14 - Evidence of a resonance of 22bp period in the whole HG38 human reference chromosome21 (horizntal undulatory code -IV-).

Figure 15 - Confirmation of a 22bp period in the whole HG38 human reference chromosome21 -vertical genomic binary code -III-).

A third complementary method (Fig 16) is presented here: knowing the period determined and confirmed by the two previous methods, we segment the complete sequence of the chromosome by consecutive segments according to this period, for example here for the chromosome21, we will "cut" the entire sequence of the chromosome in successive sections of 22 bases, the length of the period discovered. Then we record for each segment the C + G populations on the one hand and T + A on the other hand. We then represent the cumulative distribution curve of these different CG and TA populations throughout the chromosome sequence.

Table 6 - This table shows a C+G top for 8 bases value within 22 bases segments distribution.

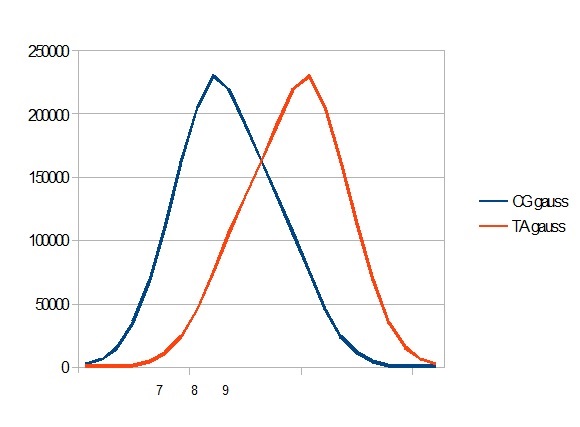
segmented by 22 bases periods. 

Figure 16 - Gauss like CG / TA distribution within the whole human HG38 chromosome21

**Part 2/3 : Then we compare now Sapiens Build34 and Neanderthal Chomosomes4 for these Fibonacci/Lucas harmonic periods.**

Now we analyze and compare according to the same methods the older genome of SAPIENS Build34 (2003) and the prehistoric genome of Neanderthal.

In this comparative figure we note the main 34 resonance but also the 29 Lucas number harmonic secondary resonance.

Then we compare now Sapiens Build34 and Neanderthal Chomosomes4 for these Fibonacci/Lucas harmonic periods.

**Figure 17** - Neanderthal chromosome4 harmonic resonance Lucas 18.

**Figure 18**  - Sapiens Build34 chromosome4 harmonic resonance Lucas 18.

**Figure 19** - Neanderthal chromosome4 harmonic resonance Fibonacci 21.

**Figure 20** - Sapiens Build34 chromosome4 harmonic resonance Fibonacci 21.

**Figure 21** - Neanderthal chromosome4 harmonic resonance Lucas 29.

**Figure 22** - Sapiens Build34 chromosome4 harmonic resonance Lucas 29.

**Figure 23** - Neanderthal chromosome4 MAIN resonance Fibonacci 34.

**Figure 24** - Neanderthal chromosome4 MAIN periods Fibonacci 34 and Lucas 123.

It will be noted that there is a 21-base DEPHASING between the sapiens barcodes (HG38 and Build34) on the one hand and the Neanderthal barcode above. It suffices to compare the positions of the first resonant bar34 in each of the 3 barcode patterns: Position 5 for Sapiens HG38 and Build34, Position 26 for Neanderthal. We deduce this phase shift of 26-5 = 21. Note that 21 is - also - a number of Fibonacci.

About possible relations between these two resonances of 34 and 123:

In the figure below we try to analyze possible links between these two resonances. We note first that 123/34 = 3.617647059 = 1 + Phi \* 2 = 2 + Phi = 3 + (1 / Phi). At the same time in the figure below we have divided the first period 123 of the Neandertal barcode. As shown by the bars "-4" of the barcode graph, we have delimited this first resonance 123 in 3 sub sections: 34 55 34. This structure is symmetrical vis-à-vis the middle of the pattern 55, the 2 patterns 34 being reversed head spades on both sides.

**Figure 25** - Sapiens Build34 chromosome4 MAIN resonance Fibonacci 34.

**Figure 26** - Sapiens Build34 chromosome4 MAIN period Fibonacci 34.

**Figure 27** - Neanderthal chromosome4 harmonic resonance Lucas 47.

**Figure 28** - Sapiens Build34 chromosome4 harmonic resonance Lucas 47.

**Figure 29** - Neanderthal chromosome4 harmonic resonance Fibnacci 55.

**Figure 30** - Sapiens Build34 chromosome4 harmonic resonance Fibonacci 55.

**Figure 31** - Neanderthal chromosome4 harmonic resonance Lucas 76.

**Figure 32** - Sapiens Build34 chromosome4 harmonic resonance Lucas 76.

**Figure 33** - Neanderthal chromosome4 harmonic resonance Fibonacci 89.

**Figure 34** - Sapiens Build34 chromosome4 harmonic resonance Fibonacci 89.

**Figure 35** - Neanderthal chromosome4 SECOND MAIN resonance Lucas 123.

**Figure 36** -Sapiens Build34 chromosome4 harmonic resonance Lucas 123.

**Figure 37** -Neanderthal chromosome4 harmonic resonance Fibonacci 144.

**Figure 38** -Sapiens Build34 chromosome4 harmonic resonance ibonacci 144.

**Figure 39** -Neanderthal chromosome4 harmonic resonance Lucas 199.

**Figure 40** -Sapiens Build34 chromosome4 harmonic resonance Lucas 199.

As in the case of Sapiens HG38, harmonic resonances are observed for the Fibonacci and Lucas periods upstream and downstream of 34.

However, the periods of Fibonacci (55 89) and the Lucas period (47 76) situated between the two main resonances (34 and 123) of Neanderthal have their harmonic waves longer and fewer than in the corresponding cases in Sapiens Build34 which, it contains only the major resonance 34.

***Part 3/3 : Does this phenomenon of harmonic resonances of Fibonacci and Lucas also extend to great apes?***

***The case of the chimpanzee:***

***Its major resonance is 21, does it have similar harmonic resonances of Fibonacci (34 55) or Lucas (29 47)?***

***Figure 41 -Chimp chromosome4 harmonic Lucas resonance 29***

***Figure 42  -Chimp chromosome4 harmonic Fibonacci resonance 34***

***Figure 43 - Chimp chromosome4 harmonic Lucas resonance 47***

***Figure 44 - Chimp chromosome4 harmonic Fibonacci resonance 55***

***The case of Orangutan:***

***Its major resonance is 21, does it have similar harmonic resonances of Fibonacci (34 55) or Lucas (29 47)?***

***Figure 45 - Orangutan chromosome4 harmonic Lucas resonance 29***

***Figure 46 - Orangutan chromosome4 harmonic Fibonacci resonance 34***

***Figure 47 - Orangutan chromosome4 harmonic Lucas resonance 47***

***Figure 48 - Orangutan chromosome4 harmonic Fibonacci resonance 55***

***The case of the Gorilla:***

***Its major resonance is 55, does it have harmonic resonances close to Fibonacci (21 34 89) or Lucas (29 47 76)?***

***Figure 49 - Gorilla chromosome4 harmonic Fibonacci resonance 21***

***Figure 50 - Gorilla chromosome4 harmonic Lucas resonance 29***

***Figure 51 - Gorilla chromosome4 harmonic Fibonacci resonance 34***

***Figure 52 - Gorilla chromosome4 harmonic Lucas resonance 47***

***Figure 53 - Gorilla chromosome4 harmonic Lucas resonance 76***

***Figure 54 - Gorilla chromosome4 harmonic Fibonacci resonance 89***

***The case of the Macaque:***

***Its major resonance is 43, which is neither a Fibonacci number nor a Lucas number, does it have similar harmonic resonances of Fibonacci (21 34 89) or Lucas (18 29 47 76)?***

***Figure 55 - Macaque chromosome4 Lucas resonance 18***

***Figure 56 - macaque chromosome4 quasi 21 Fibonacci resonance (21-1)***

***Figure 57 - macaque chromosome4 quasi 29 Lucas resonance (29-1)***

***Figure 58 - macaque chromosome4 quasi 34 Fibonacci resonance (34-1)***