Fractional dynamics in DNA

J.A. Tenreiro Machado, António C. Costa, Maria Dulce Quelhas

A B S T R A C T

This paper addresses the DNA code analysis in the perspective of dynamics and fractional calculus. Several mathematical tools are selected to establish a quantitative method without distorting the alphabet represented by the sequence of DNA bases. The association of Gray code, Fourier transform and fractional calculus leads to a categorical representation of species and chromosomes.

Keywords:
DNA
Chromosome
Phylogenetics
Dynamics
Fourier transform
Fractional calculus

1. Introduction

Phylogenetics is the study of the evolutionary relations between groups of organisms. With the advent of genome sequencing and genome databases [1,2], considerable new information is available for computational processing, allowing decoding and understanding the informational structure present on DNA sequences.

Fractional calculus (FC) goes back to the beginning of the theory of differential calculus and deals with the generalization of standard integrals and derivatives to a non-integer or even complex order [3–16]. FC was somehow considered an "exotic" mathematical tool, but during the last decade fractional dynamics was found to play a crucial role in many phenomena. In fact, a large range of potential application fields are possible by bringing to a broader paradigm the concepts of physics, chemistry and engineering. Consequently, fractional dynamics emerged as the concept of adopting FC in the study of dynam-ical systems by tacking advantage of the long memory properties of the fractional operators.

This paper studies the deoxyribonucleic acid (DNA) code [17,18] in the perspective of system dynamics and fractional calculus (FC). A close observation of the DNA structure leads to the conclusion that "dynamic tools" may prove to be powerful allies in this endeavor. It is believed that, besides the information about the "structural construction" of a given species, DNA also includes the history of evolution towards the particular species and the instructions for the growth of each individual during its lifetime. These two different time scales show that we are in the presence of a complex system with a complicated dynamics, and that the systems analysis tools may be helpful. This observation motivated the association of logical and

mathematical concepts namely, Gray coding, Fourier transform (FT) and fractional calculus for the analysis of the DNA data of twenty species. The results reveal important relationships between chromosomes and species, pointing to the goodness of the proposed methodology, and motivating further research with the usual formalisms of system dynamics.

Within this mindset, this paper is organized as follows. Section 2 briefly presents the main biological concepts and mathematical tools, and formulates their application in the framework of the DNA sequence decoding. Section 3 analyzes the relationship between chromosomes and species. Finally, Section 4 outlines the main conclusions.

2. Mathematical tools and DNA decoding

DNA is made up of two polymers connected by the bonding of hydrogen atoms, leading to a double helix structure [19]. Each polymer contains nucleotides that can be classified into three types: deoxyribose, a phosphate group, and a nitrogenous base. There are four different nitrogenous bases: thymine, cytosine, adenine, and guanine, represented as "T", "C", "A", and "G". Each type of base on one strand forms a bond with just one type of base on the other strand. This arrangement is called "base pairing", with A bonding only to T, and C bonding only to G. For example, in a human being, each cell holds 23 pairs of separate DNA–protein complexes (chromosomes), each containing, on average, 160 million nucleotide pairs. This massive amount of information is being collected and decoded during the last years, as the result of a large collaborative effort among many individuals and at research institutions around the world, and is available [20–27] for scientific research.

From the available DNA sequences a substantial part is organized into chromosomes and has been used in this study. For converting the DNA code into a numerical value it is observed that we are handing an alphabet with symbols {T, C, A, G}. The available data includes a fifth symbol, represented by "N", which has no practical meaning for the DNA coding and, therefore, this symbol was considered as "zero" during the calculations.

We have different values when considering DNA sequences with length ranging from $n = 1$, representing a counting of $m = 4^1$ states, up to $n = 5$, representing the dynamics of a system with $m = 4^5$ states. It must be noted that we are handling

Table 1
Main characteristics of species and their chromosomes.

Specie	Tag	Group	Chromosomes
Human	Hu	Mammal	Hu1, Hu2, Hu3, Hu4, Hu5, Hu6, Hu7, Hu8, Hu9, Hu10, Hu11, Hu12, Hu13, Hu14, Hu15, Hu16, Hu17, Hu18, Hu19, Hu20, Hu21, Hu22, HuX, HuY
Chimpanzee	Ch	Mammal	Ch1, Ch2a, Ch2b, Ch3, Ch4, Ch5, Ch6, Ch7, Ch8, Ch9, Ch10, Ch11, Ch12, Ch1, Ch14, Ch15, Ch16, Ch17, Ch18, Ch19, Ch20, Ch21, Ch22, ChX, ChY
Orangutan	Or	Mammal	Or1, Or2a, Or2b, Or3, Or4, Or5, Or6, Or7, Or8, Or9, Or10, Or11, Or12, Or13, Or14, Or15, Or16, Or17, Or18, Or19, Or20, Or21, Or22, OrX
Rhesus	Rm	Mammal	Rm1, Rm2, Rm3, Rm4, Rm5, Rm6, Rm7, Rm8, Rm9, Rm10, Rm11, Rm12, Rm13, Rm14, Rm15, Rm16, Rm17, Rm18, Rm19, Rm20, RmX
Pig	Pi	Mammal	Pi1, Pi2, Pi3, Pi4, Pi5, Pi6, Pi7, Pi8, Pi9, Pi10, Pi11, Pi12, Pi13, Pi14, Pi15, Pi16, Pi17, Pi18, PiX
Opossum	Op	Mammal	Op1, Op2, Op3, Op4, Op5, Op6, Op7, Op8, OpX
Mouse	Mm	Mammal	Mm1, Mm2, Mm3, Mm4, Mm5, Mm6, Mm7, Mm8, Mm9, Mm10, Mm11, Mm12, Mm13, Mm14, Mm15, Mm16, Mm17, Mm18, Mm19, MmX, MmY
Rat	Rn	Mammal	Rn1, Rn2, Rn3, Rn4, Rn5, Rn6, Rn7, Rn8, Rn9, Rn10, Rn11, Rn12, Rn13, Rn14, Rn15, Rn16, Rn17, Rn18, Rn19, Rn20, RnX
Dog	Do	Mammal	Do1, Do2, Do3, Do4, Do5, Do6, Do7, Do8, Do9, Do10, Do11, Do12, Do13, Do14, Do15 Do16, Do17, Do18, Do19, Do20, Do21, Do22, Do23, Do24, Do25, Do26, Do27, Do28, Do29, Do30, Do31, Do32, Do33, Do34, Do35, Do36, Do37, Do38
Cow	Co	Mammal	Co1, Co2, Co3, Co4, Co5, Co6, Co7, Co8, Co9, Co10, Co11, Co12, Co13, Co14, Co15, Co16, Co17, Co18, Co19, Co20, Co21, Co22, Co23, Co24, Co25, Co26, Co27, Co28, Co29, CoX
Horse	Eq	Mammal	Eq1, Eq2, Eq3, Eq4, Eq5, Eq6, Eq7, Eq8, Eq9, Eq10, Eq11, Eq12, Eq13, Eq14, Eq15, Eq16, Eq17 Eq18, Eq19, Eq20, Eq21, Eq22, Eq23, Eq24, Eq25, Eq26, Eq27, Eq28, Eq29, Eq30, Eq31, EqX
Chicken	Ck	Bird	Ck1, Ck2, Ck3, Ck4, Ck5, Ck6,Ck7, Ck8, Ck9, Ck10, Ck11, Ck12, Ck13, Ck14, Ck15, Ck16, Ck17, Ck18, Ck19, Ck20, Ck21, Ck22, Ck23, Ck24, Ck25, Ck26, Ck27, Ck28, CkW, CkZ
Zebra Finch	Tg	Bird	Tg1a, Tg1b, Tg1, Tg2, Tg3, Tg4, Tg4a, Tg5, Tg6, Tg7, Tg8, Tg9, Tg10, Tg11, Tg12, Tg13, Tg14, Tg15, Tg17, Tg18, Tg19, Tg20, Tg21, Tg22, Tg23, Tg24, Tg25, Tg26, Tg27, Tg28, TgZ
Zebrafish	Zf	Fish	Zf1, Zf2, Zf3, Zf4, Zf5, Zf6, Zf7, Zf8, Zf9, Zf10, Zf11, Zf12, Zf13, Zf14, Zf15, Zf16, Zf17, Zf18, Zf19, Zf20, Zf21, Zf22, Zf23, Zf24, Zf25
Tetraodon	Tn	Fish	Tn1, Tn2, Tn3, Tn4, Tn5, Tn6, Tn7, Tn8, Tn9, Tn10, Tn11, Tn12, Tn13, Tn14, Tn15, Tn16, Tn17, Tn18, Tn19, Tn20, Tn21
Mosquito (*Anopheles gambiae*)	Ag	Insect	Ag2l, Ag2r, Ag3l, Ag3r, AgU, AgX
Honeybee (*Apis mellifera*)	Am	Insect	Am1, Am2, Am3, Am4, Am5, Am6, Am7, Am8, Am9, Am10, Am11, Am12, Am13, Am14, Am15, Am16
Caenorhabditis elegans	Ce	Nematode	Ce1, Ce2, Ce3, Ce4, Ce5, CeX
Caenorhabditis briggsae	Cb	Nematode	Cb1, Cb2, Cb3, Cb4, Cb5, CbX
Yeast (*Saccharomyces cerevisiae*)	Sc	Fungus	Sc1, Sc2, Sc3, Sc4, Sc5, Sc6, Sc7, Sc8, Sc9, Sc10, Sc11, Sc12, Sc13, Sc14, Sc15, Sc16

non-numerical quantities. Therefore, in order to prevent inserting a numerical order, it was decided to adopt numerical values according to the binary Gray encoding [28] applied to the DNA alphabet. Since the standard Gray binary code changes only one symbol between adjacent states, in our case we implemented a novel encoding scheme, corresponding to a generalization to base four, keeping one base change per state. For example, we get the sequences {A} {C} {G} {T}, and {AA} {AC} {AG} {AT} {CT} {CG} {CC} {CA} {GA} {GC} {GG} {GT} {TT} {TG} {TC} {TA} for $n = 1$ and $n = 2$, respectively. Furthermore, for the Gray code sequence conversion windows were adopted with an overlapping of $n - 1$ consecutive bases. Once performed the code to state translation, it is considered a circular function, leading to a numerical value capable of being processed by a dynamic tool such as the FT. In other words, the numerical output of the DNA encoding is given by $y = \sin(2\pi x/m)$ where $x = 0, 1, \ldots, m - 1$ for the consecutive sequences of n symbols in the Gray encoding.

Once defined the mathematical tool for studying the DNA dynamics, we decided to analyze eleven mammals, two birds, two fishes, two insects, two nematodes and one fungus, namely, Human (Hu), Common Chimpanzee (Ch), Orangutan (Or), Rhesus monkey (Rm), Pig (Pi), Opossum (Op), Mouse (Mm), Rat (Rn), Dog (Do), Cow (Co), Horse (Eq), Chicken (Ck), Zebra Finch (Tg), Zebrafish (Zf), Tetraodon (Tn), Gambiae mosquito (Ag), Honeybee (Am), *Caenorhabditis elegans* (Ce), *Caenorhabditis briggsae* (Cb), and Yeast {Sc}. The chromosomes characteristics of each DNA species are presented in Table 1.

3. Fourier analysis of DNA

The combination of Gray encoding and trigonometric circular function calculation was applied to the chromosomes of the twenty species and its FT was calculated. It was observed that the real and imaginary components depict considerable noise and that the FT amplitude can be considered more reliable for proceeding with the analysis. For all cases it was verified that the amplitude (A) of FT versus the frequency (ω) could be approximated by a power function $A \approx a\,\omega^b$, with the parameters (a, b) to be determined by a least square fit procedure. For example, Fig. 1 shows the amplitude and the power law trend line for Human chromosome 1.

We now discuss more deeply the power law fitting (1) of data generated by the Fourier Transform. In fact, we can not assert that this empirical dependence is single and unique for the "noisy" distribution of amplitudes represented in Fig. 1, since other hypothesis may also be acceptable. It must be noted that a considerable research effort was devoted [29–39] to the application of signal tools for DNA analysis, but it has been verified that considerable "noise" occurs. We can only speculate about the source of the "noise", a word that we are using in the absence of a more appropriate term. One possibility is, simply, that the signal conversion of DNA alphabet is not the most adequate, leading to an intrinsic "numerical deformation"

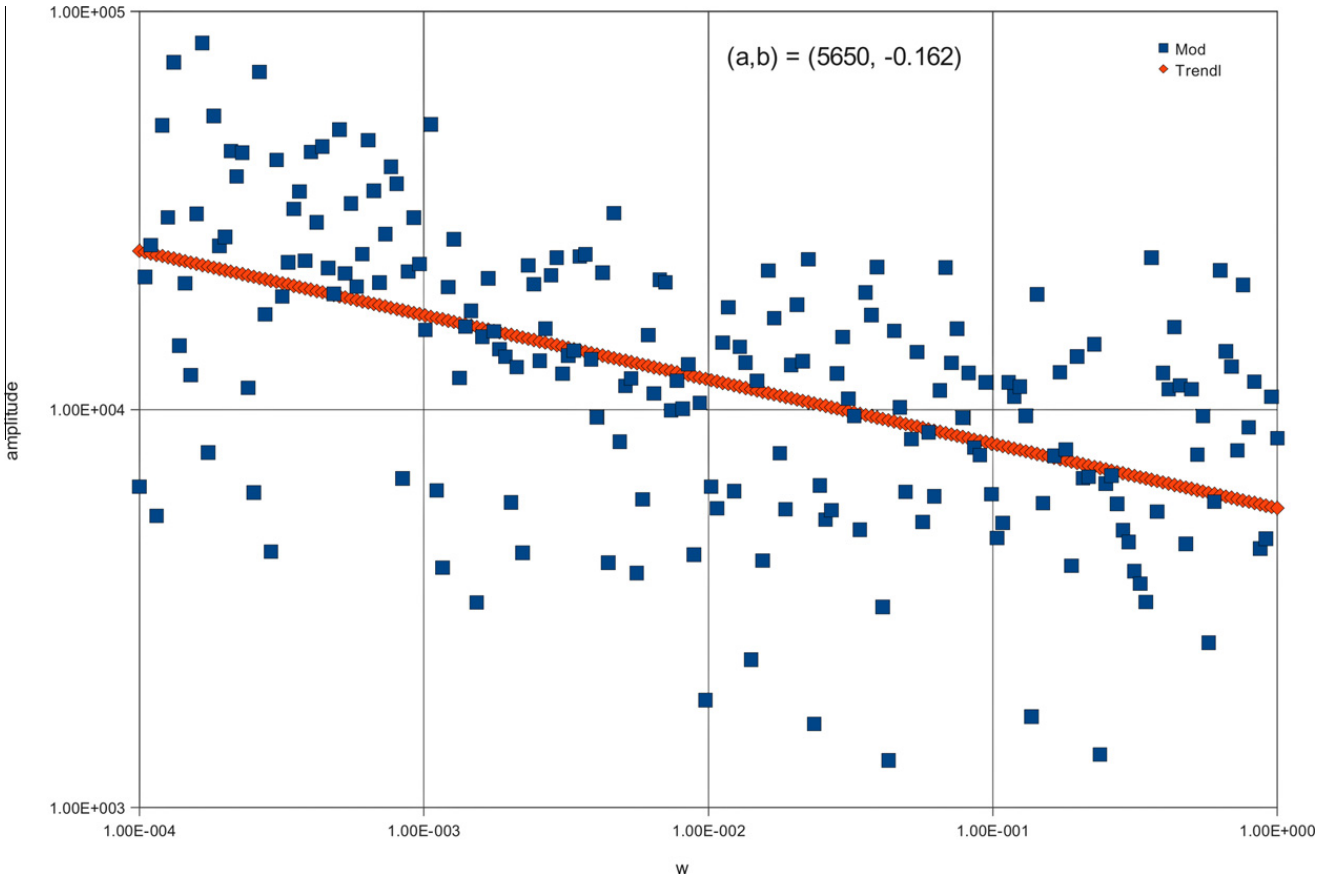

Fig. 1. Fourier transform of the signal for the Human chromosome 1 when $n = 3$: amplitude versus ω and power law approximation $A \approx a\,\omega^b$.

responsible for the "noise". The authors experimented several previously proposed schemes for yielding numerical values, but the proposed method showed to generate lower "noise" levels. Therefore, the adoption of the power law fitting (1) may also be considered as a filter that compensates errors introduced by the so called "numerical deformation".

The first test to consider in the FT of DNA code consists in evaluating the effect of the value of n upon the amplitude. Therefore, FTs were evaluated for all species with $n = \{1,2,3,4,5\}$ and the corresponding power law trend lines were obtained. Fig. 2 depicts the locus of the parameters (a,b) for the Human 24 chromosomes. It can be observed that the results "converge" and that for $n \geqslant 3$ there is almost no significant variation. In general each trace moves from bottom to top and from left to right when varying from $n = 1$ up to $n = 3$, while for $n = \{4,5\}$ there the trace remains essentially in the same location. Due to this "property", and in order to limit the computational load, in the sequel was considered solely the $n = 3$ encoding.

Again, some further discussion about the quality of the curve fitting is needed. In fact, these channels of uncertainty can considerably distort the process of information processing and, consequently, change the basic conclusions. The amplitude of the Fourier Transform reveals a considerable randomness in all cases and, consequently, the description of the dynamical properties through the power law fitting (1) can be interpreted as a filtering in the perspective of FC, but does not preclude the adoption of other descriptors. For example, Table 2 presents the parameters (a,b) and the correlation coefficient R^2 (also

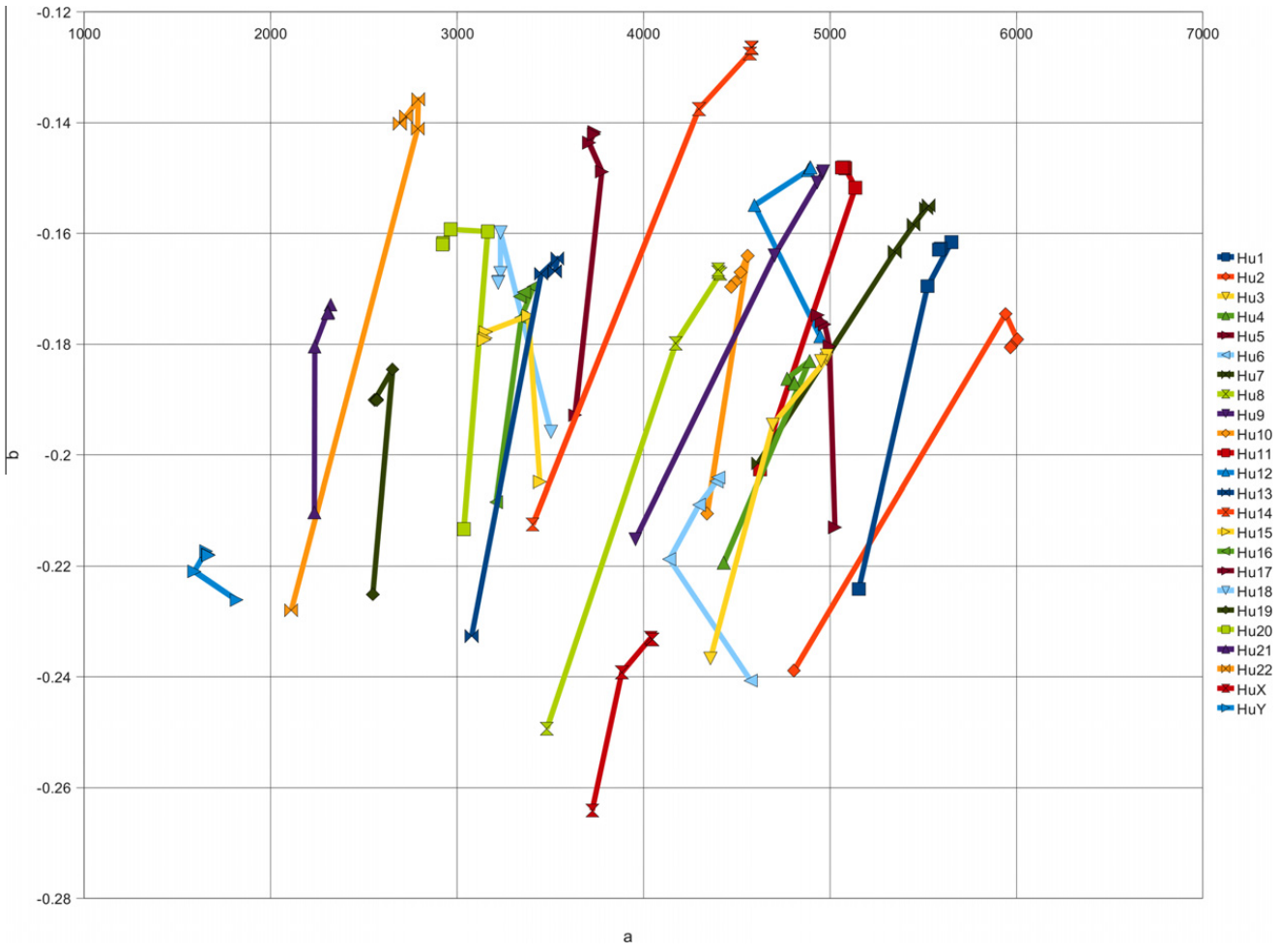

Fig. 2. Locus of the parameters (a,b) for the Human 24 chromosomes when $n = \{1,2,3,4,5\}$.

Table 2
Parameters (a,b) of the power law trend line correlation coefficient R^2 for the case of Human chromosome 1.

n	a	b	R^2
1	5154.7	−0.2241	0.4502
2	5523.6	−0.1695	0.3048
3	5651.4	−0.1616	0.2682
4	5593.2	−0.1628	0.2633
5	5583.3	−0.1629	0.2627

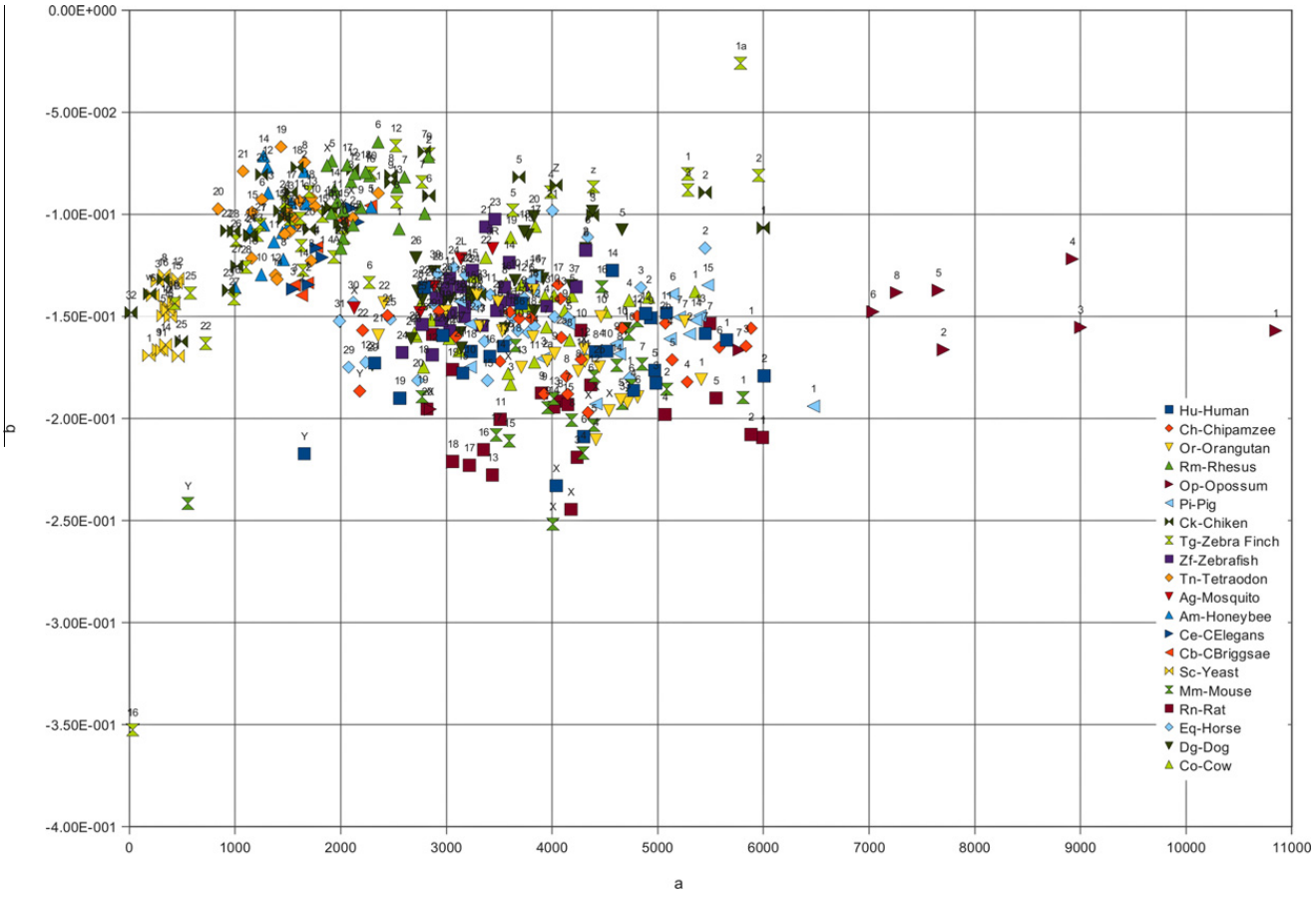

Fig. 3. Locus of the parameters (a,b) for all 415 chromosomes of the twenty species when $n = 3$.

known as *R*-squared value of correlation coefficient) when $n = \{1,2,3,4,5\}$ for the case of Human chromosome 1. We verify the stabilization of the numerical values for $n \geqslant 3$ and, on the other hand, the low values R^2 corresponding to higher noise and smaller slope. A more complex curve fitting could hardly lead to significantly better values of R^2, an observation that supports the speculation on the "filtering action" of (1) for the "numerical deformation" within the numerical conversion of DNA code.

In all cases the power law trend line reveals the fractional dynamics of the code and the intrinsic long memory dependence of the "signal" [40–42].

Having established the conceptual and numerical framework for the study, the final phase consists in performing the FT calculation and power law approximation for all chromosomes of the twenty species. Fig. 3 shows the *locus* of (a,b) parameters for all 415 chromosomes. The parameter a is related to the "energy of the signal" which reflects partially the size of the chromosome. Therefore, we observe a tendency for smaller/larger values of the point labels in the right/left of the *locus* of (a,b). The parameter b is related with the information content, being more close/apart to/from zero as the DNA "signal" is more random/correlated along the sequence. We verify that mammals have more negative values of b.

We note also a separation both in the perspective of species and chromosomes. In terms of species, we observe at the top left side a cluster constituted by the Sc, followed by the group Rm, Tn, Am, Ce, and Cb. Somewhat lower to the right we have the main part of the mammals namely the Hu, Ch, Or, Pi, Mm, Rn, Do, Co, and Eq. Somewhat peculiar is the place of Op separated to the right from the rest of the mammals. The birds Ck and Tg are in the middle of the two groups covering all range from left to right. The Zf and the Ag superimpose partially in the mammals.

In terms of chromosomes, we observe particularly in mammals that, in general, chromosomes with the same numbering are relatively close. For example chromosomes 1 and 3 are very similar for Human, Chimpanzee and Orangutan. Nevertheless, chromosomes 2, X and Y (when it exists) reveal a remarkable difference. Obviously, much more can be extracted from the *locus* of (a,b) with 415 points and a more detailed analysis will be developed in the future.

4. Conclusions

After verifying that chromosomes have a code based on a four symbol alphabet, in this paper it was adopted a Gray-like encoding and a sinusoidal numerical conversion of the DNA code. This information can be analyzed as a "signal" representative of a system dynamics. For that purpose it was adopted the Fourier transform, with the resulting amplitude versus

frequency charts approximated by power law trend lines. The locus of power law parameters reveals species' representative clusters. Furthermore, it is also observed a second level of grouping according to the type of chromosome. The results are in agreement with what is currently known in phylogenetics and this research opens new research directions to pursuit.

Acknowledgments

The authors thank

– The reviewers for the careful reading and for the suggestions that improved this paper.
– The following organizations for allowing access to genome data:

- Human – Genome Reference Consortium, http://www.ncbi.nlm.nih.gov/projects/genome/assembly/grc/
- Common Chimpanzee – Chimpanzee Genome Sequencing Consortium
- Orangutan – Genome Sequencing Center at WUSTL, http://genome.wustl.edu/genome.cgiGENOME=Pongo%20abelii
- Rhesus – Macaque Genome Sequencing Consortium, http://www.hgsc.bcm.tmc.edu/projects/rmacaque/
- Pig – The Swine Genome Sequencing Consortium, http://piggenome.org/
- Cow – The Baylor College of Medicine Human Genome Sequencing Center, http://www.hgsc.bcm.tmc.edu/projects/bovine/
- Dog Genome Sequencing Project – http://www.broad.mit.edu/mammals/dog/, Lindblad-Toh K, et al.Genome sequence, comparative analysis and haplotype structure of the domestic dog.Nature.2005 Dec 8;438:803-19
- Horse – The Broad Institute, http://www.broad.mit.edu/mammals/horse/
- Mouse – Mouse Genome Sequencing Consortium.Initial sequencing and comparative analysis of the mouse genome.-Nature, 420, 520-562 (2002), http://www.hgsc.bcm.tmc.edu/projects/mouse/
- Rat – The Baylor College of Medicine Human Genome Sequencing Center, http://www.hgsc.bcm.tmc.edu/projects/rat/, Rat Genome Sequencing Project Consortium.Genome sequence of the Brown Norway rat yields insights into mammalian evolution.Nature 428(6982), 493–521 (2004)
- Opossum – The Broad Institute,http://www.broad.mit.edu/mammals/opossum/
- Chicken – International Chicken Genome Sequencing Consortium, Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. Nature. 2004 Dec 9; 432(7018): 695–716. PMID: 15592404
- Zebra Finch – Genome Sequencing Center at Washington University St. Louis School of Medicine
- Zebrafish – The Wellcome Trust Sanger Institute, http://www.sanger.ac.uk/Projects/D_rerio/
- Tetraodon – Genoscope, http://www.genoscope.cns.fr/
- Honeybee – The Baylor College of Medicine Human Genome Sequencing Center, http://www.hgsc.bcm.tmc.edu/projects/honeybee/
- Gambiae Mosquito – The International Anopheles Genome Project
- Elegans nematode – Wormbase, http://www.wormbase.org/
- Briggsae nematode – Genome Sequencing Center at Washington University in St. Louis School of Medicine
- Yeast – Sacchromyces Genome Database, http://www.yeastgenome.org/

References

[1] Schuh RT, Brower AVZ. Biological Systematics: principles and applications. 2nd ed. Cornell University Press; 2009.
[2] The tree of life project – <http://tolweb.org/tree/home.pages/abouttol.html>.
[3] Oldham Keith B, Spanier Jerome. The fractional calculus: theory and application of differentiation and integration to arbitrary order. Academic Press; 1974.
[4] Bagley RL, Torvik PJ. Fractional calculus – a different approach to the analysis of viscoelastically damped structures. AIAA J 1983;21(5):741–8.
[5] Oustaloup A. La commande CRONE: commande robuste d'ordre non entier. Hermes; 1991.
[6] Samko Stefan G, Kilbas Anatoly A, Marichev Oleg I. Fractional integrals and derivatives: theory and applications. Gordon and Breach Science Publishers; 1993.
[7] Miller Kenneth S, Ross Bertram. An introduction to the fractional calculus and fractional differential equations. John Wiley and Sons; 1993.
[8] Anastasio Thomas J. The fractional-order dynamics of brainstem vestibulo-oculomotor neurons. Biol Cybernet 1994;72:69–79.
[9] Tenreiro Machado JA. Analysis and design of fractional-order digital control systems. J Syst Anal Model Simul 1997;27:107–22.
[10] Podlubny I. Fractional differential equations. San Diego: Academic Press; 1999.
[11] Zaslavsky George M. Hamiltonian chaos and fractional dynamics. Oxford University Press; 2005.
[12] Kilbas Anatoly A, Srivastava Hari M, Trujillo Juan J. Theory and applications of fractional differential equations. Elsevier; 2006.
[13] Tarasov Vasily E. Fractional dynamics: applications of fractional calculus to dynamics of particles, fields and media. Springer; 2010.
[14] Mainardi Francesco. Fractional calculus and waves in linear viscoelasticity: an introduction to mathematical models. Imperial College Press; 2010.
[15] Ionescu Clara, Tenreiro Machado J. Mechanical properties and impedance model for the branching network of the seiva system in the leaf of *Hydrangea macrophylla*. Nonlinear Dyn 2010;60(1-2):207–16.
[16] Tenreiro Machado J, Kiryakova Virginia, Mainardi Francesco. Recent history of fractional calculus. Communications in nonlinear science and numerical simulations, vol. 16. Elsevier; 2011.
[17] The Official Web Site of the Nobel Prize, <http://nobelprize.org/nobel_prizes/medicine/laureates/1968/>.
[18] Seitz Harald, editor. Analytics of protein–DNA interactions. Advances in biochemical engineering biotechnology. Springer; 2007.

[19] Pearson H. Genetics: what is a gene? Nature 2006;441(7092):398–401.
[20] UCSC genome bioinformatics – <http://hgdownload.cse.ucsc.edu/downloads.html>.
[21] Sims Gregory E, Jun Se-Ran, Wu Guohong A, Kim Sung-Hou. Alignment-free genome comparison with feature frequency profiles (FFP) and optimal resolutions. Proc Nat Acad Sci United States of America 2009;106(8):2677–82.
[22] Murphy William J, Pringle Thomas H, Crider Tess A, Springer Mark S, Miller Webb. Using genomic data to unravel the root of the placental mammal phylogeny. Genome Res 2007;17:413–21.
[23] Zhao Hao, Bourque Guillaume. Recovering genome rearrangements in the mammalian phylogeny. Genome Res 2009;19:934–42.
[24] Prasad Arjun B, Allard Marc W. Confirming the phylogeny of mammals by use of large comparative sequence data sets. Mol Biol Evolution 2008;25(9):1795–808.
[25] Ebersberger Ingo, Galgoczy Petra, Taudien Stefan, Taenzer Simone, Platzer Matthias, Haeseler Arndt von. Mapping human genetic ancestry. Mol Biol Evolution 2007;24(10):2266–76.
[26] Dunn Casey W et al. Broad phylogenomic sampling improves resolution of the animal tree of life. Nature 2008;452:745–50.
[27] Hillier LaDeana W et al. International chicken genome sequencing consortium. Nature 2004;432:695–716.
[28] Paul E Black. Gray code. In: Paul E Black, editor. Dictionary of algorithms and data structures [online], U.S. National Institute of Standards and Technology, 31 August 2009.
[29] Tiwari Shrish, Ramachandran S, Bhattacharya Alok, Bhattacharya Sudha, Ramaswamy Ramakrishna. Prediction of probable genes by Fourier analysis of genomic sequences. Comput Appl Biosci 1997;13(3):263–70.
[30] Dodin Guy, Vandergheynst Pierre, Levoir Patrick, Cordier Christine, Marcourt Laurence. Fourier and Wavelet transform analysis, a tool for visualizing regular patterns in DNA sequences. J Theor Biol 2000;206(3):323–6.
[31] Afreixo Vera, Ferreira Paulo JSG, Santos Dorabella. Fourier analysis of symbolic data: a brief review. Digital Signal Process 2004;14:523–30.
[32] Afreixo Vera, Ferreira Paulo JSG, Santos Dorabella. Spectrum and symbol distribution of nucleotide sequences. Phys Rev E 2004;70(3):031910.
[33] Yin Changchuan, Yau Stephen S-T. A Fourier characteristic of coding sequences: origins and a non-fourier approximation. J Comput Biol 2005;12(9):1153–65.
[34] Vincent A Emanuele II, Thao T Tran, Tong Zhou G. A Fourier product method for detecting approximate TANDEM repeats in DNA. In: 2005 IEEE/SP 13th workshop on statistical signal processing, Bordeaux, France, 17–20 July; 2005, p. 1390–95.
[35] Leitão Helena Cristina G, Pessôa Luciana S, Stolfi Jorge. Mutual information content of homologous DNA sequences. Genet Mol Res 2005;4(3):553–62.
[36] Jeng Cheng-Chang, Yang I-Ching, Hsieh Kun-Lin, Lin Chun-Nan. Clustering analysis for bacillus genus using Fourier transform and self-organizing map. In: King I et al., editors. ICONIP 2006, Part III, LNCS, 4234. Springer-Verlag; 2006. p. 48–57.
[37] Yu Zhou, Li-Qian Zhou, Zu-Guo Yu, Vo Anh. Distinguish coding and noncoding sequences in a complete genome using Fourier transform. In: IEEE third international conference on natural computation, Haikou, China; 2007, p. 295–299.
[38] Yin Changchuan, Yau Stephen S-T. Numerical representation of DNA sequences based on genetic code context and its applications in periodicity analysis of genomes. IEEE symposium on computational intelligence in bioinformatics and computational biology. Idaho: Sun Valley; 2008.
[39] Arniker Swarna Bai, Kwan Hon Keung. Graphical representation of DNA sequences. 2009 IEEE international conference on electro/information technology. Ontario, Canada: Windsor; 2009.
[40] Eduardo Roman H, Porto Markus. Fractional derivatives of random walks: time series with long-time memory. Phys Rev E 2008;78(3):031127.
[41] Tenreiro Machado JA. Fractional derivatives: probability interpretation and frequency response of rational approximations. Communications in nonlinear science and numerical simulations, vol. 14. Elsevier; 2009. p. 9–10.
[42] Nigmatullin RR, Baleanu D. Is it possible to derive Newtonian equations of motion with memory? Int J Theor Phys 2010;49(4):701–8.