ISSN 2278-3091
Volume 10, No.2, March - April 2021
International Journal of Advanced Trends in Computer Science and Engineering
Available Online at http://www.warse.org/IJATCSE/static/pdf/file/ijatcse751022021.pdf
https://doi.org/10.30534/ijatcse/2021/751022021

Involving FCGR images in studying fractality and multifractality in human chromosome 22 and bacteria complete genome

Zeineb Chebbi Babchia[1], Afef Elloumi Oueslati[2]
[1,2]University of Tunis El Manar, National School of Engineers of Tunis (ENIT), BP 37, le Belvédère, 1002, Tunis, Tunisia, Email: zeineb-chebbi@hotmail.com
[2] University of Carthage, National School of Engineers of Carthage (ENICarthage), Electrical Engineering Department, Tunisia, Email: afef.elloumi@enit.utm.tn

ABSTRACT

The frequency chaos game representation (FCGR) is a simple yet powerful visualization method of DNA sequences. It provides the possibility of representing genomes by images, revealing in such a way different fractal structures. In this paper, we perform a fractal and multifractal analysis of Human chromosome 22 and some complete genomes based on the FGCR image. We used the fractal dimension (FD) and the multifractality degree (ΔDq) to characterize and distinguish genomes. First, we construct the FCGR image with different orders of human chromosome 22. Next, we calculate the fractal dimension, the general dimension spectrum and the multifractal spectrum of each FCGR image using the box-counting method. Then, we examine the FCGR image fractal and multifractal characteristics impact on highlighting the existence of repetitive DNA sequences in human chromosome 22. We also observe the relationship between fractality and multifractality. After that, we apply this study to bacteria completes genomes and C.elegans chromosome I. The obtained results show that the multifractal spectra of all organisms studied are multifractal-like and chromosome 22 strong multifractality proves its richness of repetitive sequences. Also, we observed that with the increasing the FCGR order value, the multifractality grows and the fractal dimension lessens. Finally, by assigning to each sequence a point in two-dimensional space (FD, ΔDq), we obtained three classes of genomes. We can easily distinguish the human chromosome 22 from other genomes and Bacteria are almost close in the spaces (FD, ΔDq).

Keywords: Box-counting methode, DNA sequence, Frequency Chaos game representation, Fractal images, Multifractal analysis

1. INTRODUCTION

The genetic information of all living is encoded in a macromolecule called DNA. This DNA by its appearance, its composition and its complexity, contains preciously the elementary details of each organism. Microscopically, the DNA reveals chains of characters constituted by the bases adenine (A), thymine (T), cytosine (C) and guanine (G) [1].
DNA molecule structure is described by the double helix. This helix folds to fit into the nucleus of an organism cell, which is only about a hundredth of a millimetre in diameter. So it has a very particular fractal characteristic which is to have an infinite length, whereas it is contained only in surface a bounded and reduced. This helix particularity is shown that DNA is a fractal [2].
Fractal analysis is a nontraditional mathematical approach for studying objects which irregularity exclude fromEuclidean geometry [3], [4]. It has proven to be a useful tool in the analysis of medical signals [5] like electrocardiogram (ECG) [6] and electroencephalogram (EEG) [7] signals and DNA sequence [8].
The DNA sequences fractality is studied in the long-distance correlation [2], [9], [10]. In fact, the appearance of a nucleotide in a specific position depends on the previous nucleotides and the appearance of a small nucleotide segment depends on large-scale segments. Such long-range correlation is directly related to the power-law and fractal structure of the DNA sequence [11]. However, due to the DNA complexity, one exponent may not be enough for its characterization. The multifractal formalism allows using more exponents [12]. It is a powerful tool in both the theory and practice to describe the spatial heterogeneity of the fractal object systematically [13]. In this case, the analysis object is divided into several fractal sets, each generating a fractal dimension that is then translated into a continuous exponents spectrum.

Multifractal analysis has been useful in studying different problems at DNA sequence [14], [15]. It has been used to reconstruct phylogeny from mitochondria DNA [16] to study the proteins [17] and to distinguish coding and non-coding sequences in DNA sequences [18]. Also, this formalism was applied to study human chromosomes [19] and the C. elegans genome chromosomes [20] by using the chaos game representation (CGR). The (CGR) is a genomic sequence mapping. It has been proposed by Jeffrey [21]. The frequency chaos game representation (FCGR) is a nucleotide frequency

matrix extracted from the CGR [22], [23]. It's a way to represent the genome in a single image. This image shows the DNA fractal structure. In this work, we used the FCGR image to perform a fractal and multifractal analysis of human chromosome 22 and some complete genomes.

We choose to study chromosome 22 because it has a high gene density and it is repeat sequences rich [24], [25]. We studied also the possibility to distinguish genomes and classifying bacterias using fractal dimension and multifractality degree.

The paper is organized as follows. In section 2, we give a brief overview of the fractal dimension and multifractal analysis. In section 3, we present the image construction FCGR and we describe our methodology for the fractal and multifractal analysis. The results and discussions of the analysis are presented in section 4. Finally, section 5 gives the conclusion of this paper.

2. FRACTAL AND MULTIFRACTAL THEORY

2.1 Fractal dimension

The fractal dimension (FD) is a useful way for characterizing the self-similarity of an irregular fractal object. Indeed, it is a parameter which quantifies the complexity and the irregularity.

A rigorous mathematical definition of fractal dimension has been introduced by Hausdorff [24]–[26]. For any subset S of the n-dimensional Euclidean space, it is covered by a $\{S_i\}$, each of which has a diameter d_i=diam(F_i)< ϱ, ϱϵ[0,+∞[.
The Hausdorff measure $H^{\alpha}(S)$ is is given by :

$$H^{\alpha} = \lim_{\varrho \to 0} \inf_{\{S_i\}} \sum_i d_i^{\alpha}, \qquad \text{which} \alpha \epsilon [0, +\infty[, \qquad (1)$$

and thus, the fractal dimension FD can be written as:

$$\text{FD} = sup\{\alpha / \text{H}^{\alpha}(S) = \infty\} = inf\{\alpha / \text{H}^{\alpha}(S) = 0\} \qquad (2)$$

The FD is a real number whose value depends on the property of the object. It is a non-integer, less than the space dimension (SD) and greater than the topological dimension (TD) [3],[25].

Many methods exist to compute the FD like box-counting methods, fractional Brownian motion (fBm) methods and area measurement methods [26]. Each method estimates this dimension by using a different algorithm. The box-counting method is the most popular and suitable method for the object FD determination [28], [29]. Indeed, its simple and easy to develop. It consists on covering a binary image with boxes of size r and counting the boxes number $N_{box}(r)$ that contain pixels [30], [31]. This is repeated for different r size boxes. For a fractal image, the number of boxes $N_{box}(r)$ and r have the following power-law relationship :

$$N_{box}(r) \sim r^{FD} \quad (3)$$

By using the equation (1), the FD is obtained by:

$$FD = \lim_{r \to +\infty} \frac{\log (N_{box}(r))}{\log (r)} \qquad (4)$$

A log–log plot of $N_{box}(r)$ against r then yields a line of slope equal to FD. The fractal dimension describes fractals with a single scaling factor [12], [13], [32]. However, when the fractal object is more complex and several scaling factors are present, a more detailed description is required. In this case, the multifractal formalism is better suitable to use.

2.2 The multifractal theory

A multifractal object is a system of homogeneous fractal structure superposition that only one fractal dimension is insufficient to describe it. It can be analyzed by an interdependent fractal dimensions function or spectrum [33]. They exist many methods to approximate the multifractals spectrums. They are divided into two classes: the methods said box-counting and the methods based on wavelets [26]. In this study, we decided to use the box-counting method for our multifractal analysis [34] like for fractal analysis. In the multifractal analysis, The object is covered by a boxes grid $B_i(r)$ of normalized size r and the number of pixels m_i contained in each box is counted. Then the measurement of the ith box covering the object is defined as follows:

$$P_i(r) = \frac{m_i}{M} \quad (5)$$

Where M is the total number of image pixels [23]. A-weighted factor (mass exponent, qϵ] +∞, -∞ [) is applied to datasets extracted from the object giving more or less, importance to the high or low mass density areas. Then the partition function can be calculated by the following equation :

$$X(q, r) = \sum_{i=1}^{N(r)} (P_i(r))^q \quad (6)$$

where N(r) is the number of boxes covering the image. For a multifractal image this function has the following scaling properties:

$$X(q, r) \sim r^{\tau(q)} \qquad (7)$$

Where τ(q) the correlation exponent. For each q, τ(q) may be obtained as the slope of a log-log $X(q, r)$ against r. The τ(q) curve is a straight line for the monofractal object and it is nonlinear for the multifractal object. A generalized dimension function Dq is then derived as :

$$\begin{cases} D(q) = \frac{\tau(q)}{q-1} & q \neq 1 \\ D(1) = \frac{\sum_{i=1}^{N(r)} P_i(r) log P_i(r)}{\log (r)} & q = 1 \end{cases} \qquad (8)$$

The generalized dimension spectrum Dq is a monotonically decreasing function, with horizontal asymptotes at $D_{q_{max}} = \lim_{q \to +\infty} D_q$ and $D_{q_{min}} = \lim_{q \to -\infty} D_q$. Their values can be used to describe the heterogeneity, if $D_{q_{max}} \neq D_{q_{min}}$ the fractal is heterogeneous (multifractal), and homogeneous otherwise. The multifractality degree ΔDq is defined as [19], [20], [23]:

$$\Delta D(q) = |D_{q_{max}} - D_{q_{min}}| \qquad (9)$$

We use ΔDq to observe how the values of D(q) change along the spectrum. It can be regarded as a direct measure of the multifractality complexity degree [19], [20]. It shows the length to which the fractal exponent extent in the series, being an indicator of the signal structure richness. If ΔDq is

high, the multifractal spectrum is rich in information and the study object is very irregular, for a small ΔDq, the resulting dimension spectrum is poor in information.
From the correlation exponent τ(q), we can obtain the multifractal singularity spectrum $f(\alpha)$. It is a way to study the fractal dimensions (Holder exponents) distribution in a multifractal object [20], [23], [34]. It provides information about the structure scaling properties [19], [20]. The $f(\alpha)$ is obtained as follows:

$$f(\alpha) = q\alpha(q) - \tau(q) \quad (10)$$

where $\alpha(q) = \frac{d\tau(q)}{dq}$characterizes the singularity strength in the ith area. The $f(\alpha)$ describes the bigger probability subset property with smaller α. With bigger α, $f(\alpha)$ describes a smaller probability subset property. The spectrum $f(\alpha)$ is a single-humped function for a multifractal object. For a monofractal signal or image, the spectrum is reduced to a point.

3. MATERIALS AND METHODS

3.1 Materials: Coding DNA by the FCGR image

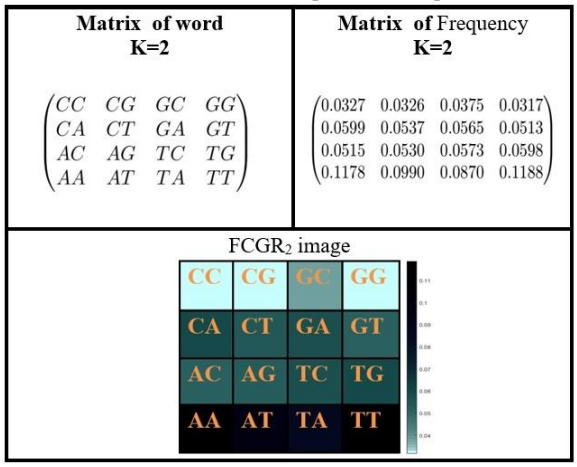
The input sequence (collected from the NCBI database [38]) is a long character string made up of four nucleotides: A, C, G and T. To be able to apply the fractal and multifractal analysis one must convert the sequence into an image. In this work, we choose to use the Frequency Chaos Game Representation (FCGR) [21], [22], [23]. In table I, we present the list of the genomes considered for analysis, which are Homo sapiens chromosome 22, C.elegans chromosome I, three archaea and seven bacteria.

Table 1: Thegenomesdescriptions

Species	Category	Data lenght
Homo sapies chromosome22	Human	31264301
C.elegans chromosome I	Nematode	15072434
Agrobacterium tumefaciens	Bacteria	6083998
Achromobacter	Bacteria	5876049
Bacillus Cereus	Bacteria	5221581
Lactobacillus	Bacteria	1026169
Bordetella bronchiseptica	Bacteria	5191712
Borrelia Garinii	Bacteria	904246
Eubacterium	Bacteria	2403485
Pyrococcus horikoshii	Archaebacteria	1738505
Archaeoglobus fulgidus	Archaebacteria	2178400
Aeropyrum pernix	Archaebacteria	1669696

The FCGR is a frequency matrix extracted of the chaos game representation (CGR) of genomic sequence [21], [22]. Using the FCGR, a given genomic sequence can be displayed as a square single image form in which each pixel intensity is associated with a specific word frequency (figure 1). The grayscale indicates the relative frequency of each word: the darker the pixel, the greater the frequency. As an illustrative example of FCGR procedure, we consider the sequence of Bacillus Cereus bacteria which we encode by $FCGR_2$.

Figure 1: Illustration of the $FCGR_2$ process to represent the



sequence ofBacillus Cereus bacteria

The genomic signatures obtained by $FCGR_8$ of allgenomes described in table 1 are shown in figure 2.

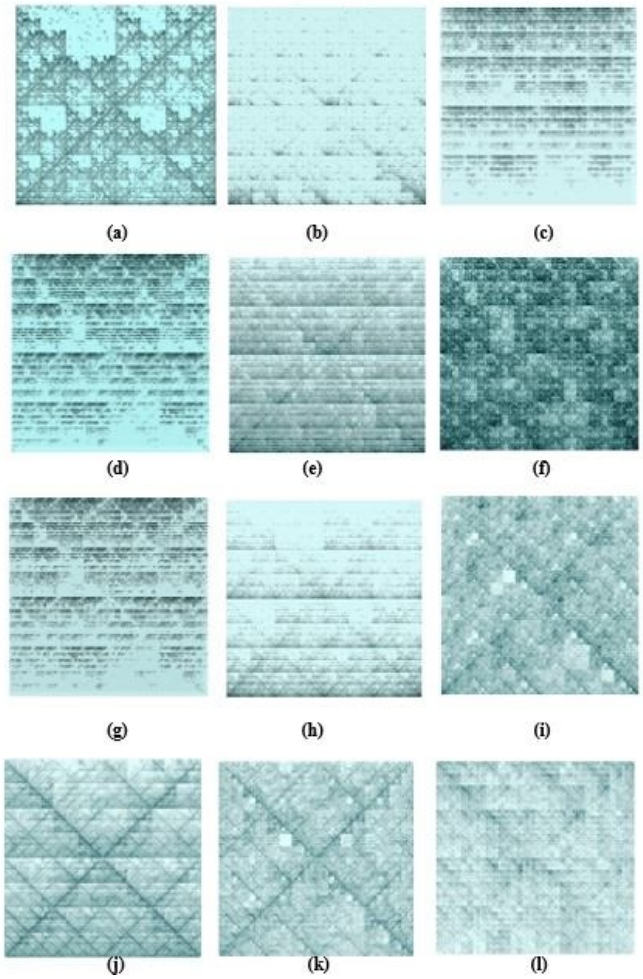


Figure 2: Genomic signatures obtained by FCGR8 from bacteria: (a) Homosapies chromosome 22, (b) C.elegans chromosome I, (c) Agrobacteriumtumefaciens, (d) Achromobacter, (e) BacillusCereus, (f) Lactobacillus, (g)Bordetella bronchiseptica, (h) Borrelia Garinii, (i) Eubacterium, (j) Pyrococcushorikoshii, (k) Archaeoglobus fulgidus, (l) Aeropyrum pernix

As shown in figures 2, all genomic signatures images reveal different fractal structures. All FCGR images have a structure specific to the study genomes. In this work, we applied a fractaland multifractal analysis on these FCGR images. These twoanalyzes allow us to characterize data that will be useful forclassifying the images.

3.2Methods

The goal of this paper is not only the fractal and multifractal analysis of genomes, but also, we want to examine the relationship between the fractality and multifractality and to provethat this analysis can be employed to distinguish and classifygenomes. To this end, we used the box-counting method fofractal and multifractal analysis of genome FCGR image.The box-counting method algorithms for fractal dimension andmultifractal spectrum calculation are given as follows:

Fractal dimension calculation algorithm :

1) Generate an $FCGR_i$ image to represent each genome (i is the scale of the FCGR image)
2) Binarization of $FCGR_i$ image
3) Superpose a cubic mesh with $r = 2^k$ on the surface of the image with k ϵ [0,11]
4) Calculate $N_{box}(r)$, which is the number of all the cubes containing at least one pixel of the image
5) The steps 3 and 4 are repeated until $r=2^{11}$
6) The FD is given directly by the slope of the graph (log(r), log($N_{box}(r)$))

Multifractal spectrums calculation algorithm:

1) Generate an $FCGR_i$ image to represent each genome (i is the scale of the FCGR image)
2) Divide the $FCGR_i$ to N boxes with $N=2^i \times 2^i$
3) Calculate the partition function $X(q)$ with q ϵ [-40,40]
4) Calulate $\tau(q), D(q)$, ΔDq and *f(α)* for all q
5) Represent $\tau(q), D(q)$and *f(α)* spectrums

Our methodology for fractal and multifractal analysis of all genomic images is composed of four phases as shown in figure 3.

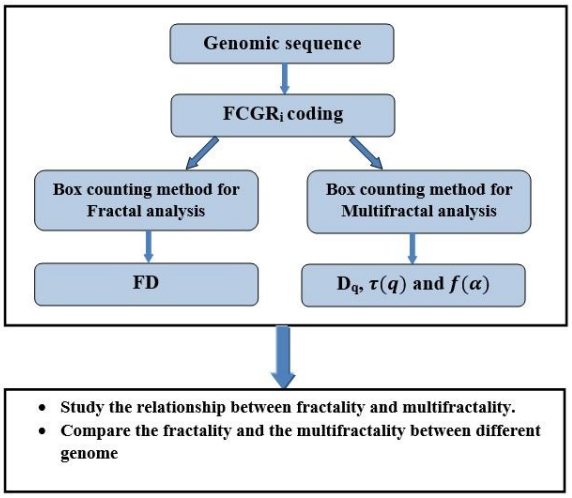


Figure 3: Fractal and multifractal analysis methodology flowchart

First, we collect the genomic sequences from the NCBI database. Second, we convert the sequences to the image by using the FCGR coding. Third, we perform the fractal and multifractal analysis by using the box-counting method and we obtain the FD and _Dq. Finally, we use the results to compare fractality and multifractality, distinguish and classify the studied genomes.

4. RESULT AND DISCUSSION

In this section, we study the fractality and multifractality of the genomes described in table 1. First, we examine theFCGR coding impact on highlighting these characteristics in
the Homosapies chromosome 22 and we also observe the relationship between fractality and multifractality. Next, we apply this study to C.elegans genome chromosome I and bacteria, then, we compare his results whit the human chromosome 22. Finally, we attribute to each genome a point in two dimensional space (FD, ΔDq) and we evaluate the possibility of employed this analysis to distinguish and classify genomes.

4.1 Fractal and multifractal analysis of the Homo sapienschromosome 22:

Before calculating the Homo sapiens chromosome 22fractal dimensions, we test our program performance by calculating the dimension of two deterministic fractals: the Koch snowflake and the Sierpinski triangle (figure 4).

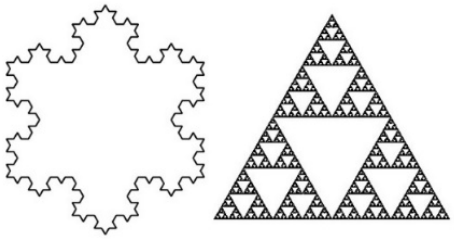


Figure 4: The Koch snow flake and the Sierpinski triangle

In the table 2, we summarize the FDcalculated values, the FD theoretical value and the error of calculation.

Table 2: Theboxcountingresultsthekochsnowflakeand thesierpinskitringle

	FD calculated	Theoretical FD	error
Koch snowflake	1,2572	1, 2619	0,0047
Sierpinski triangle	1,5797	1,5850	0,0053

The results illustrated in Tables 2 show that our programme gives good results for fractal dimension estimation.
We applied the box-counting method on each FCGR images of human chromosome 22. We summarized the fractal dimension for all FCGR images in Table 3.

Table 3: The fractal dimension of chromosome 22 FCGR images

Image	FD
$FCGR_1$	1.7957
$FCGR_2$	1.7613
$FCGR_3$	1.7545
$FCGR_4$	1.7292
$FCGR_5$	1.7289
$FCGR_6$	1.7223
$FCGR_7$	1.6775
$FCGR_8$	1.4569

The fractal dimensions calculated for the different FCGR images are between the topological dimension (TD = 1) and the space dimension (SD = 2), so all FCGR images of human chromosomes 22 are fractal [3], [4]. From the table 3, we observed that the FD minimum value is 1.4569 for $FCGR_8$ image. Its maximum value is 1.7957 for the $FCGR_1$ image, which proves that the $FCGR_1$ has a strong momofracatlity compared to the other FCGR images.

For multifractal analysis of human chromosomes 22, the general dimension spectrum D(q), the multifractal spectrum $f(\alpha)$and the qth mass exponent τ(q) of each FCGR images chromosome 22 are calculated and represented respectevly in figure 5, figure 6 and figure 7. We summarize in table 4 the D(q) result calculation with qϵ]-40, 40[.

Table 4: MultifractalparametresDq by FCGR images chromosome 22

	D_{-40}	D_1	D_2	D_{40}	ΔD_q
$FCGR_1$	2.2195	1.9819	1.9610	1.7760	**0.4436**
$FCGR_2$	2.6953	1.9582	1.9319	1.7422	**0.9531**
$FCGR_3$	2.8916	1.9720	1.9553	1.8711	**1.0205**
$FCGR_4$	2.7234	1.9492	1.9159	1.6233	**1.1000**
$FCGR_5$	2.7679	1.9414	1.8996	1.5028	**1.2651**
$FCGR_6$	2.7989	1.9331	1.8790	1.3813	**1.4175**
$FCGR_7$	2.8150	1.9230	1.8455	1.2528	**1.5622**
$FCGR_8$	2.8461	1.9133	1.7907	1.1393	**1.7068**

From the table 4, we observe that the D(q) value for all data sets depend on q values. The ΔDq minimum value is 0.4436 for the $FCGR_1$ image. Its maximum value is 1.7068 for the $FCGR_8$ image.

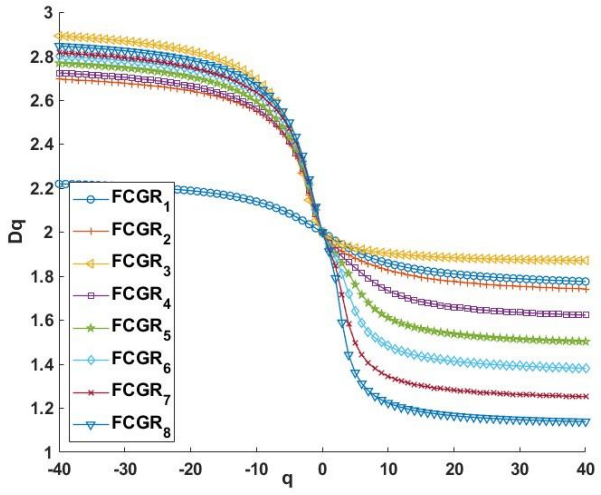


Figure 5: The general dimension spectrum D(q)

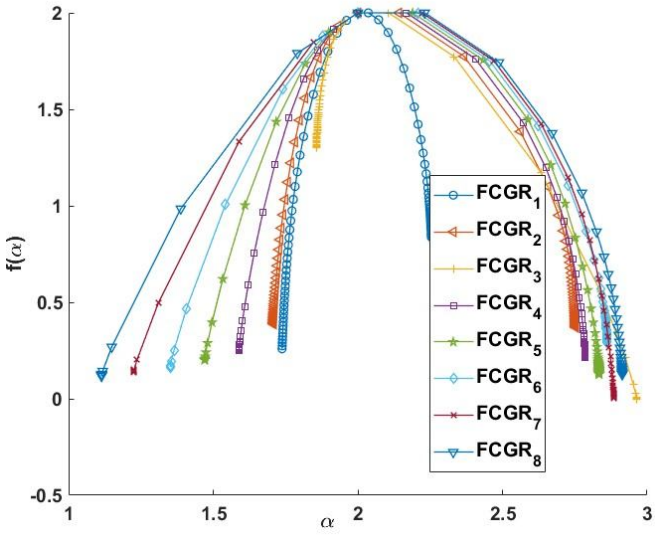


Figure 6: Multifractal spectrum$f(\alpha)$

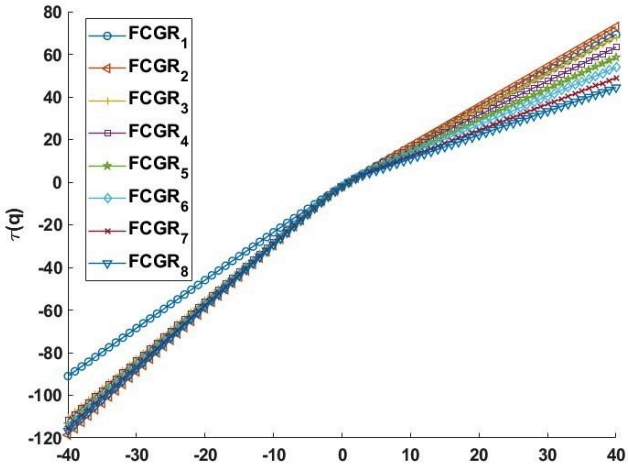


Figure 7: The qth mass exponentτ(q)

According to the figure 5, the figure 6 and the figure 7 we observe that : The D(q) values for all FCGR images,

decreasing with increasing q values and this evident in multifractal nature. The D(q) curve of FCGR1 is almost a horizontal line. All the $f(\alpha)$spectra present the typical single-humped shape, that characterizes multifractal signals. The differences among the spectra $f(\alpha)$ FCGR8 image spectrum $f(\alpha)$ is larger than the other FCGRimages spectrums. There is a difference between the τ(q)curves of $FCGR_1$ and $FCGR_8$. The $FCGR_8$ image τ(q)curve has the most nonlinear form than other FCGR imagesτ(q) curves.
From these results, we can deduce that all FCGR images are multifractal images. Also, these results prove that the $FCGR_8$ image has a strong multifractality compared to the other FCGR images while image $FCGR_1$ has a less multifractality.

To study the relationship between fractality and multifractality,We represent the fractal dimension and the multifractalitydegree for each image in the figure 8:

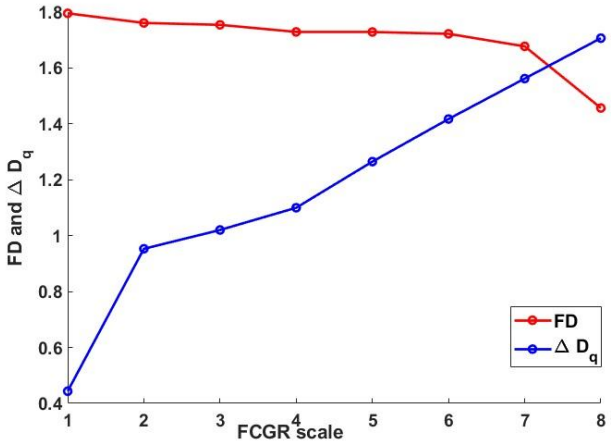


Figure 8: The ΔD(q) and the FD for each FCGR image Chromosome 22

According to figure 7, the curve which corresponds to the fractal dimension decreases when the FCGR images scale increases, unlike the multifractality degree curve. This shows that when the FCGR order k increases, the number of repeated sequences of size k increases and new fractal structures appearing. This explains the increased complexity and heterogeneity of the genomic signature. The FCGR image,therefore, becomes less monofractal and more multifractal.
The chromosome 22 low fractal dimension shows that there is not just one type of repeating sequence but several. The $FCGR_8$ image strong multifractality shows that the number of repetitive sequences of 8bp length is very large. Indeed the multifractality degree shows that several types of monofractal coexist in the same set.

4.2Compare fractality and multifractality between genomes:

In this section, we study the fractal and multifractal behaviour of each genome described in table 1 and we compare the fractality and multifractality between these genomes and Human Chromosome 22. We choose to usethe $FCGR_8$image. The general dimensions spectrums D(q) and themultifractal spectrums $f(\alpha)$are calculated and representedrespectively in figure 9 and figure 10. We summarize in table 5,the ΔD(q) and the FD result calculation for all genomes.

Table 5: MultifractaldegreeΔD(q) and FD for all genomes

From the values of ΔDq and FD, it is seen that there exists

Species	FD	ΔD
Homo sapies chromosome 22	1.4569	1.7068
C.elegans chromosome I	1.6163	1.3221
Agrobacterium tumefaciens	1,7072	1,4207
Achromobacter	1.7588	1.4147
Bacillus Cereus	1.7634	1.2216
Lactobacillus	1.8075	0.7864
Bordetella bronchiseptica	1.7017	1.4333
Borrelia Garinii	1.7250	1.1391
Eubacterium	1.8008	0.9340
Pyrococcus horikoshii	1.7875	0.9834
Archaeoglobus fulgidus	1.7864	1.0107
Aeropyrum pernix	1.7648	1.0461

a clear difference between the DNA sequences of all organismsconsidered. From table 5, we observe that the FD maximumvalue is 1.8075 for Lactobacillus. We also observed thatLactobacillus has a ΔDq minimum value 0.7864. Hence thisgenome has a strong monofractal nature. This result suggests astrong periodicity in the nucleotide sequences of this bacteria.This bacterium does not have many types of sequences repeatsof length 8. We observe also, that the C. elegans genomehas a small fractal dimension compared to bacteria exceptedEubacterium and Lactobacillus, but it is more multifractal thanthe archeobacteria. Human chromosome 22 has the ΔD(q)maximum value. This result suggests a high aperiodicity ofthe nucleotides along the chromosome 22 sequence. Indeed,the human genome is a very complex genome and it is repeatsequences rich.

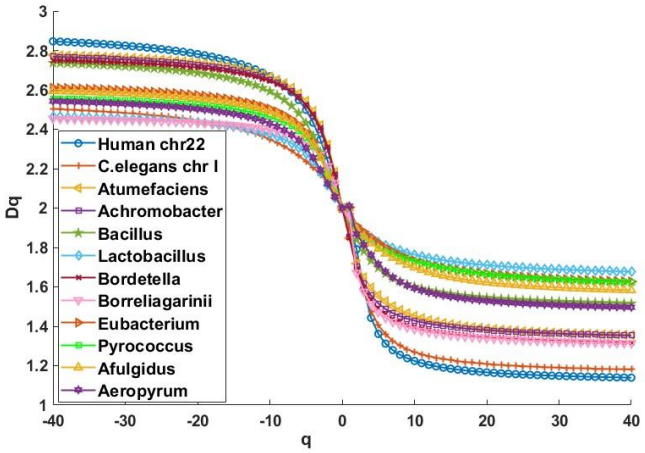


Figure 9: The general dimension spectrum D(q) of each genomes

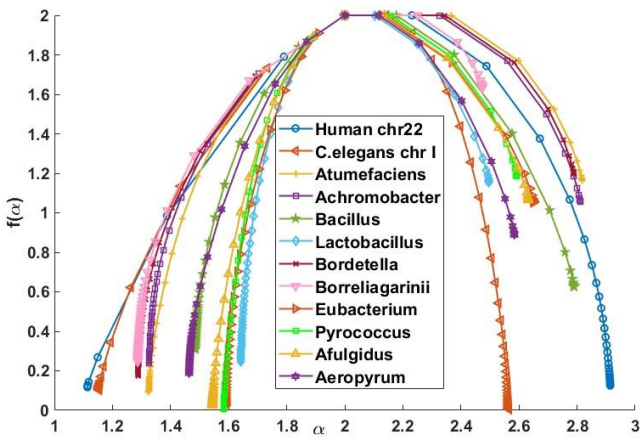


Figure 10: Multifractal spectrum$f(\alpha)$of each genomes

From the figures 9 and 10 we observed that, the D(q) values of all genome, decreasing with increasing q values. Hence the D(q) spectra of all organisms are multifractal-like. Each genome has a very distinct D(q) and $f(\alpha)$ curves. The spectrum $f(\alpha)$ of bacteria except Eubacterium and Lactobacillusare larger than archaebacteria $f(\alpha)$spectrum.

The fractal and multifractal analysis of each genome show that exists a difference between genome. The bacterias havestrong multifractality than the Archaebacteria except for theEubacterium and Lactobacillus, they have a strong monofractality.This shows that the Archaebacteria have a low density ofsequences repeats of different types. We can easily distinguishthe human chromosome 22 from other genomes.

In order to more distinguish the genomes themselves we use two-dimensional points (FD, ΔD(q)) represented in figure 11.

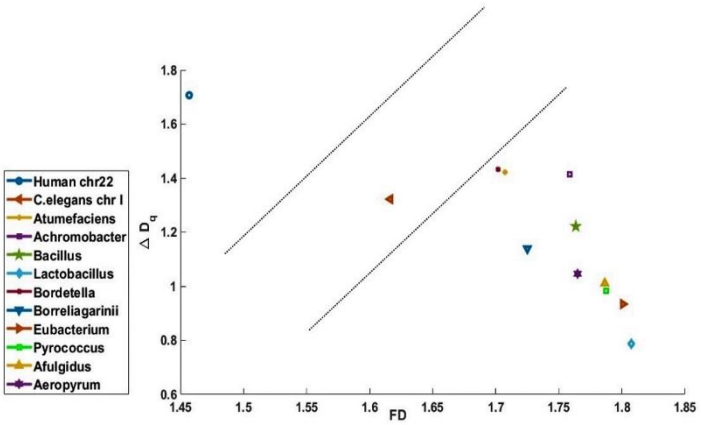


Figure 11: Distributions of two-dimensional points (FD, ΔD(q)) all genomesselected.

From figure 11 it is clear that genomes roughly gather into three classes. The first one is for human chromosome 22, the second is for C. elegans chromosome I and the last class is for bacteria and archaea. Using the distance between the points, one can obtain a classification of genomes. We can see also that all archaebacteria are grouped with each other. The Eubacterium is very closed to archaebacteria. Indeed, its genomic signature obtained by $FCGR_8$ looks like that of the archaea Afulgidus (Figure 2). The Atumefaciens and Bordetellaare very close in the space (FD, ΔD(q)). The same thingis observed for the two Archea Afulgidus and Pyrococcus. Ingeneral, genomes that are close phylogenetically are almostclose in the space (FD, ΔD(q)).

5. CONCLUSION

Frequency Chaos Game Representation provides a powerful tool for visualizing fractal structures that derive from the repetition of some patterns in DNA sequences. In the present paper, we perform a fractal and multifractal analysis of human chromosome 22 and some bacteria based on FCGR images. The fractal dimension, the general dimension spectrumDq and the multifractal spectrum are calculated using the box-counting method. We have used the result to compare fractality and multifractality between the studied genomes and for examining the variation impact of FCGR images scales on the fractal dimension values and the multifractality degree.

The results show the fractality decreases when the FCGR images scale k increases, unlike the multifractality, it increases when the FCGR images scale k increases. This shows that there exist many sequence repeats of length 8 of different types. Indeed, the multifractality degree shows that several types of monofractal coexist in the same set. Also, from FD andΔD(q) values we can conclude that human chromosome 22has a strong multifractality and it is repeat sequences rich.We prove also that the Archea have a strong monofractalitycompared to other bacteria. Moreover, the representation of all studied genomes in the two-dimension space (FD,ΔD(q)),show that the FD and ΔD(q) are useful and powerful toolsfor classification genomes. Indeed, genomes that are close phylogenetically are almost close in the space (FD,ΔD(q)) like archaebacteria. Finally, we will use this analysis in our futures works to classified the bacteria and study other genomes.

REFERENCES

1. A.B. Majumder and S. Gupta.**CBSTD: a cloud based symbol table driven DNA compression algorithm**, Industry Interactive Innovations in Science, Engineering and Technology, Springer pp.467-476, 2018
2. CK. Peng, SV. Buldyrev, AL. Goldberger, S. Havlin, F. Sciortino, M.Simons and HE.Stanley, **Fractal landscape analysis of DNA walks**, Physica A: Statistical Mechanics and its Applications, Elsevier,Vol. 191, pp. 25-29, 1992
3. B. Mandelbrot, **The fractal geometry of nature**, New York: WH freeman, 1983, Vol. 173
4. K. Falconer, **Fractal geometry: mathematical foundations and applications**, John Wiley & Sons, 2004
5. Du. Ristanovic and NT. Milosevic, **Fractal analysis: methodologies forbiomedical researchers**, Theoretical biology forum, 2012, Vol.105, pp.99-118

6. M. Hamidi, and H. Ghassemian and M.Imani.**Classification of heart sound signal using curve fitting and fractal dimension**, Biomedical Signal Processing and Control, Elsevier 2018, Vol.39, pp.351-359
7. SA. David, JAT. Machado, CMC. Inàcio Jr and CA. Valentim Jr, **A combined measure to differentiate EEG signals using fractal dimension and MFDFA-Hurst**, Communications in Nonlinear Science and Numerical Simulation, Elsevier 2020, pp. 105170
8. R. Silva, JRP. Silva, D. Anselmo, JS. Alcaniz, WJC. da Silva, MO.Costa, **An alternative description of power law correlations in DNA sequences**, Physica A: Statistical Mechanics and its Applications, Elsevier 2020, Vol.545, pp. 123735
9. LGA. Alves, PB. Winter, LN. Ferreira, RM. Brielmann, RI. Morimoto and L.AN. Amaral,**Long-range correlations and fractal dynamics in C.elegans: Changes with aging and stress**, Physical Review E, APS 2017, Vol. 96 ,pp. 022417
10. RF. Voss, **Evolution of long-range fractal correlations and 1/f noise in DNA base sequences**,Physical review letters, APS 1992, Vol. 68, pp. 3805
11. A. Karmi, A. Najafi, P. Gifani and S. Khakabimamaghani, **Fractal Analysis of DNA by Nonlinear Genome Signal Processing for Exon and Intron Separation**, Annual Research & Review in Biology, SCIENCEDOMAIN International 2014, Vol. 4, pp. 699
12. H. Salat, R. Murcio, E. Arcaute, **Multifractal methodology**, Physica A: Statistical Mechanics and its Applications, Elsevier 2017, Vol. 473, pp. 467-487
13. J. Luo, Z. Zheng, T. Li and S. He, **Spatial heterogeneity of microtopography and its influence on the flow convergence of slopes under different rainfall patterns**, Journal of Hydrology, Elsevier 2017, Vol. 545, pp. 88-99
14. C.L. Berthelsen, J.A. Glazier and S. Raghavachari, **Effective multifractal spectrum of a random walk**,Physical Review E, APS 1994, Vol. 49, pp. 1860
15. Z.G. Yu, V. Anh and K.S. Lau, **Measure representation and multifractal analysis of complete genomes**, Physical Review E, APS 2001, Vol. 64, pp.031903
16. J.A. Glazier, S. Raghavachari, C.L. Berthelsen, and M.H. Skolnick,**Reconstructing phylogeny from the multifractal spectrum of mitochondrial DNA**,Physical review E, APS 1995, Vol. 51, pp. 2665-2668
17. E. Maiorino, L. Livi, A. Giuliani, A. Sadeghian and A. Rizzi, **Multifractal characterization of protein contact networks**, Physica A: Statistical Mechanics and its Applications, Elsevier 2015, Vol. 428, pp. 302-313
18. M. Pal, B. Satish, K. Srinivas, P. M. Rao and P. Manimaran, **Multifractal detrended cross-correlation analysis of coding and non-coding DNA sequences through chaos-game representation**, Physica A: Statistical Mechanics and its Applications, Vol. 436,Elsevier 2015, pp. 596-603
19. P.A. Moreno, P.E. Vz, E.Martinez, L.E. Garreta, N. Diaz, S. Amador, I.Tischer, J. M. Gutiez, A. K. Naik, F. Tobar and F. Garcia, **The human genome: a multifractalanalysis**, BMC Genomics, BioMed Central 2011, Vol. 12, pp. 506
20. P. E. Vz, L. E. Garreta, E. Mart´ınez, N. Diaz, S. Amador, I. Tischer, J.M. Gutiez and P.A. Moreno, **The Caenorhabditis elegans genome: a multifractal analysis**,Genetics and Molecular Research, pp. 949965, 2010
21. R. Touati, I. Messaoudi, A.E. Oueslati, Z. Lachiri, M. Kharrat, **New Intraclass Helitrons Classification Using DNA-Image Sequences and Machine Learning Approaches**, IRBM 2020
22. R. Touati, S. Haddad-Boubaker, I. Ferchichi, I. Messaoudi, A. E. Ouesleti, H. Triki, Z. Lachiri and M. Kharrat, **Comparative genomic signature representations of the emerging COVID-19 coronavirus and other coronaviruses: High identity and possible recombination between Bat and Pangolin coronaviruses,** Genomics, Elsevier 2020, Vol. 112, pp. 4189-4202
23. ZC. Babachia, AE.Oueslati, **Studying Multifractality in C. elegans sequences based on multi order FCGR Images**, In: 2019 International Conference on Control, Automation and Diagnosis(ICCAD). IEEE,2019 pp. 1-7.
24. D.Holste, I. Grosse and H.Herzel, **Statistical analysis of the DNA sequence of human chromosome 22**,Physical Review E, APS 2001, Vol. 64, pp.041917
25. J. Li, **Comparative analysis of human chromosome 22 NF2 region and gene expression profiling of human orthologs in zebrafish early developmental stages via whole mount in situ hybridization**, The University of Oklahoma, 2008
26. B.B.Mandelbrot, **Self-affine fractals and fractal dimension**, Physica scripta, IOP Publishing 1985, Vol. 32, pp. 257
27. J. Theiler, **Estimating fractal dimension**, JOSA A, Optical Society of America 1990, Vol. 7, pp. 1055-1073
28. L.Renaud and N. Betrouni. **Fractal and multifractal analysis: a review**, Medical image analysis, 2009, vol.13, pp.634-649
29. R.B. Moruzzi, A.L. de Oliveira, F.T. da Conceiçào, J. Gregory and L.C. Campos**, Fractal dimension of large aggregates under different flocculationconditions**, Science of the Total Environment, Elsevier 2017, Vol. 609, pp.807-814
30. LS. Liebovitch, T. Toth, **A fast algorithm to determine fractal dimensions by box counting**, physics Letters A, Elsevier1989, Vol.141, pp. 386-390
31. FE. Lennon, GC. Cianci, NA. Cipriani, TA. Hensing, JH. Zhang, CT. Chen , SD. Murgu, EE. Vokes,MW.Vannier and R. Salgia, "**Lung cancer a fractal viewpoint**, Nature Reviews Clinical Oncology, Publishing Group 2015, Vol.12, pp. 664-675, Nature
32. N. Sarkar, B.B Chaudhuri, **An efficient differential box-counting approach to compute fractal dimension of image**, IEEE Transactions on systems, man, and cybernetics, Vol. 24, pp. 115-120, 1994
33. K. Deng, D. Pan, X.Li, F. Yin, **Spark testing to measure carbon content in carbon steels based on fractal box counting**, Measurement, Elsevier 2019, Vol. 133, pp. 77-80

34. B.G. Li, Z.G. Yu, Y. Zhou**, Fractal and multifractal properties of a family of fractal networks**, Journal of statistical mechanics: Theory and experiment, IOP Publishing 2014, Vol. 2014, pp. 2020
35. Z.Q. Jiang, W.J. Xie, W. X. Zhou and D. Sornette, **Multifractal analysis of financial markets: a review**,IOP Publishing 2019, Vol. 82, pp. 125901
36. YU. Zu-Guo, Vo.ANH and LAU. Ka-Sing, **Multifractal characterisation of length sequences of coding and noncoding segments in a complete genome**, Physica A,Statistical Mechanics and its Applications, 2001, vol. 301, pp. 351-361
37. I. G. TORRE, J. C. LOSADA, R. J. HECK, AM. Tarquis, **Multifractal analysis of 3D images of tillage soil**, Geoderma 2018, vol. 311, pp. 167-174,
38. NCBI website. [Online]Available:http://www.ncbi.nlm.nih.gov