Understanding Genomic Evolution of Olfactory Receptors through Fractal and Mathematical Morphology

Sk. Sarif Hassan[1, 2], Pabitra Pal Choudhury[1], B. S. Daya Sagar[2], Shantanav Chakraborty[1], Ranita Guha[1], Arunava Goswami[3]

[1]*Applied Statistics Unit, Indian Statistical Institute, Kolkata, India*

[2]*System Science and Informatics Unit, Indian Statistical Institute, Bangalore, India*

[3]*Biological Sciences Division, Indian Statistical Institute, Kolkata, India*

Emails according to authors ordering above: sarimif@isical.ac.in, pabitra@isical.ac.in , bsdsagar@isibang.ac.in, shantanav89@gmail.com, ranita1990@gmail.com. agoswami@isical.ac.in

Corresponding author email: sarimif@isical.ac.in

Abstract

Fractals and Mathematical Morphology are immensely used to study many problems in different branches of science and technology including the domain of Biology. There are many more unrevealed facts and figures of genes and genome in Computational Biology. In this paper, our objective is to explore how evolutionary network is associated among Human, Chimpanzees and Mouse with regards to their genomic information. We are about to explore their genomic evolution through the quantitative measures of fractals and morphology. We have considered olfactory receptors for our case study. These olfactory receptors do function in different species with the subtle differences in between the structures of DNA sequences. Those subtle differences could be exposed through intricate details of Fractals and Mathematical Morphology.

***Keywords*:** Olfactory Receptor, Succolarity, Fractal dimension, Morphological Skeleton, Bifurcation Dimension. Hurst Exponent.

1. Introduction: Without loss of generality, let us consider the Olfactory Receptors (ORs) OR1D2, CONTIG3463.6-1888, GA_x5J8B7W3YLM-7052533-7051808 of Human, Chimpanzee, and Mouse respectively for our case study. It is to be noted that first we have selected the olfactory receptor OR1D2 from HORDE database and it was blasted in the NCBI database to get highly similar OR sequences in Chimpanzee, and Mouse and we found the CONTIG3463.6-1888, GA_x5J8B7W3YLM-7052533-7051808 of Chimpanzee, and Mouse respectively. In this paper, we have captured the evolution in ORs with the help their textural quantitative views with regards to the fractals and morphological parameters. Also, we have shown that OR1D2 and CONTIG3463.6-1888, GA_x5J8B7W3YLM-7052533-7051808 are associated with two other human ORs namely OR4D2 and OR3A3.

2. *Some Basics and Fundamentals*

In this article we are about to use some standard techniques from Mathematical Morphology and Fractals. So let us warm up about some of the definitions from Mathematical morphology and Fractals.

2.1 *Basics on Mathematical Morphology*

Mathematical Morphology can be used as very fundamental tool for extracting image components that are useful for representation and description. This field was originally developed by *J. Serra* [1]. The Mathematical morphology is based on axiomatic set theory and more relevantly lattice theory. This technique is used for image analysis which provides a quantitative description of geometrical structures. Morphology can provide boundaries of images, their skeletons, convex hulls, watershed for segmentation and many more [2, 3, 4]. The ultimate aim is to extract important features from image data, from which a quantitative significant understanding of the topology of the image can be drawn. In this article, we are emphasizing on morphological skeleton of the images and consequently bifurcation dimension which could be sketched out using Horton law [5, 6]. For morphological transformations we may refer any text book on Mathematical Morphology or some papers as given in the references [7, 8]. Also we have emphasized some other quantitative parameters as reviewed from a well written article [9].

2.2 *Basics on Fractals*

The word *Fractal* is derived from the Latin adjective *fractus*. The corresponding Latin verb frangere means 'to break' to create irregular fragments. In I975, B. Mandelbrot coined the subject Fractals. The precise definition of "Fractal" according to Benoit Mandelbrot is, a set for which the Hausdroff Besicovitch dimension strictly exceeds the topological dimension [10, 11, 12].

2.2.1 *Measuring Fractal dimension*

Mandelbrot founded his insights in the idea of self similarity, requiring that a true fractal "fracture" or break apart into smaller pieces that resemble the whole. This is a special case of the idea that there should be a dynamical system underlying the geometry of the set. This is partly why the idea of fractals have become so popular throughout science; it is a fundamental aim of science to seek to understand the underlying dynamical properties of any natural phenomena. It has now become apparent that relatively simple dynamics, more precisely dynamical system can produce the fantastically intricate shapes and behavior that occur throughout nature. Let us now talk about one fundamental fractal parameter "Fractal Dimension" for a self similar object. There are several methods like box counting method, perimeter area dimension method and so on to compute fractal dimension of an object. Let us focus on the self-similarity dimension as the following.

Given a self-similar structure [10], there is a relation between the reduction factor (scaling factor) 'S' and the number of pieces 'N' into which the structure can be divided; and that relation is…

$$N = 1/S^D, \text{ equivalently, } D = \log(N)/\log(1/S)$$

This 'D' is called the Fractal dimension (Self-similarity dimension)

The fractal dimension alone does not give an idea of what "fractals" are really about. So there was a real need of defining some other fractal parameters. One of the important parameter is ***Succolarity*** which is really meant for the continuous density of the image/ fractal. The primary notion of succolarity was given by Mandelbrot and later R. H. C. de Melo and A. Conci described the method to compute the succolarity of an image/object [13].

It should be noted that another important fractal parameter is known as Hurst Exponent (H) to have fractal dimension for a one dimensional data as explained below.

Hurst Exponent: The concept of Hurst Exponent was introduced by Harold Edwin Hurst and later in Fractal Geometry, B. Mandelbrot had modified it as a parameter of relative tendency of a time series to either strongly regress to the mean or 'cluster' in a direction [14]. In statistical terms, it is sometime referred to long range correlation of a one dimensional time series.

Let us consider a one dimensional finite sequence $x = \{x_i\}_{i=1\ to\ N}$, and then we can define readily the following entities regarding the sequence as follows:

$$m_{x,n} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$X(i,n) = \sum_{j=1}^{i}\{x_j - m_{x,n}\}$$

$$R(n) = \max X(i,n) - \min X(i,n) : 1 \le i \le n$$

$$S(n) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - m_{x,n})^2}$$

Then Hurst Exponent (H) is defined as $\left(\frac{n}{2}\right)^H = \frac{R(n)}{S(n)}$

The relation between Hurst exponent (H) and fractal dimension (D) is $H + D = 2$.

3. *Results and Discussion*

Let us see that how these Human, Mouse and Chimpanzee ORs are evolutionarily connected through Fractals and Mathematical Morphology.

3.1 *Evolutionary Connection of ORs of Mouse and Chimpanzee with Human ORs*

We have considered a DNA as a one dimensional nucleotide sequence and let us define a map $T(A) = 00;\ T(T) = 11, T(C) = 01\ and\ T(G) = 10$. So corresponding to a DNA sequence we now have a binary string. We then calculate the Hurst exponent for the binary string. The result is as shown below:

Olfactory Receptors	*Hurst Exponent (H)*	*Fractal Dimension (D)*
OR1D2	**0.598911**	**1.401089**
GA_x5J8B7W3YLM-7052533-7051808	**0.645594**	**1.354406**
CONTIG3463.6-1888	**0.539152**	**1.460848**

Table-I: Hurst Exponent of ORs

We have classified all the human ORs based on our own classification methodology on the poly-string mean and standard deviation as proposed in [15]. Using the same we have classified GA_x5J8B7W3YLM-7052533-7051808 (Mouse) and CONTIG3463.6-1888(Chimpanzee) and the results are as follows:

Olfactory Receptors	*Class According to Poly-String Mean*	*Class According to Poly-String SD*	*Maps to (With respect to Hurst Exponent)*
OR1D2	**CGTA**	**CGAT**	**OR1D2 (Trivially)**
GA_x5J8B7W3YLM-7052533-7051808	**GCTA**	**GCAT**	**OR4D2**
CONTIG3463.6-1888	**CGAT**	**ACTG**	**OR3A3**

Table-II: Evolutionary Connection of ORs with Human

The Mouse OR (GA_x5J8B7W3YLM-7052533-7051808) maps to a human OR OR4D2 based on classification and closest Hurst exponent. But it is to be noted that GA_x5J8B7W3YLM-7052533-7051808 is more similar to OR1D2. But as far as Hurst exponent is concerned (amount of long range correlation in the sequence) the mouse OR maps to OR4D2. In this connection, it is our strong conviction that, OR4D2 is very much similar to OR1D2 in the sense of structural similarity in sequence, although they belong to different families as per HORDE qualitative classification. Also we could validate that mouse and human ORs are almost similar in structure and consequently in function too.

The Chimpanzee OR (CONTIG3463.6-1888) maps to a human OR OR3A3 according to the classification as shown in Table –I and II. Although OR3A3 and OR1D2 belong to different family but with respect to evolution in connection with Chimpanzee OR CONTIG3463.6-1888, they are structurally almost same as per quantification shown above.

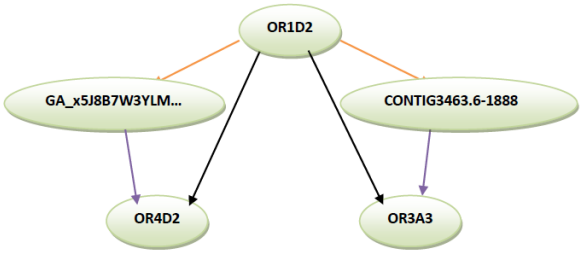


Figure-I: Evolutionary connect ion among Human, Mouse and Chimpanzee ORs

OR1D2, GA_x5J8B7W3YLM-7052533-7051808 and CONTIG3463.6-1888 are also most similar to OR4D2 and OR3A3 as shown above. They are evolutionarily connected and hence through biological evolution CONTIG3463.6-1888 and GA_x5J8B7W3YLM-7052533-7051808 are updated as OR3A3 and OR4D2 respectively.

3.2. Fractal and Morphological Quantification of ORs

Now, we are about to find the same fractal dimension from a different look as stated below.

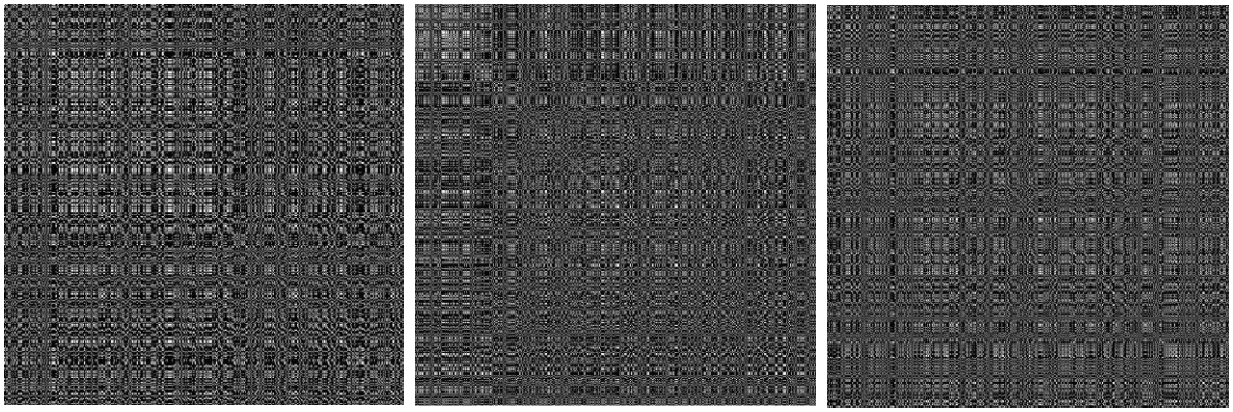
3.2.1 Fractal Dimension of Binary Image Matrix

We consider a DNA nucleotide sequence and plot the sequence in two axes and we define a mapping as follows: $f: \{X, Y\} \rightarrow \{0, 1\}\ as$

$$f(X, Y) = 0\ if\ Y \neq X$$

$$= 1\ otherwise$$

Consequently, we would be able to have a binary square matrix. Now let us consider the DNA sequences of OR1D2 CONTIG3463.6-1888, GA_x5J8B7W3YLM-7052533-7051808 of Human, Chimpanzee, and Mouse respectively. The corresponding matrices of the same are shown in figure-II.



[Matrix for OR1D2] ***[Matrix for* CONTIG3463.6-1888*]*** ***[Matrix for* GA_x5J8B7W3YLM-7052533-7051808*]***

Figure-II: Binary Image matrices

The fractal dimension corresponding to each of the above matrix is given in the table below:

Olfactory Receptors	*Fractal Dimension (D)*
OR1D2	**1.77687**
GA_x5J8B7W3YLM-7052533-7051808	**1.81916**
CONTIG3463.6-1888,	**1.82463**

Table-III: Fractal dimensions of ORs

Here we see that fractal dimensions of ORs of Chimpanzee and Mouse are almost same. Through genomic evolution they got updated into OR1D2 in human and fractal dimension is also reduced by a small amount 0.04 i.e. through genomic evolution amount of complexity or disorderliness got decremented.

3.2.2. *Results on Succolarity*

We have taken three olfactory receptor DNA sequences corresponding to Human, Mouse and Chimpanzee. A DNA sequence can be thought of as a texture of four disjoint template of A, T, C and G. So we have four different templates of each DNA sequence. We found the succolarity for each of those three sequences as shown in the table below:

Olfactory Receptors	*Succolarity Results*	
OR1D2	Template of A	**0.001026**
	Template of T	**0.001690**
	Template of C	**0.001482**
	Template of G	**0.000522**
GA_x5J8B7W3YLM-7052533-7051808	Template of A	**.0008360**
	Template of T	**.0004520**
	Template of C	**.0019240**
	Template of G	**.0003520**
CONTIG3463.6-1888,	Template of A	**.0018160**
	Template of T	**.0020440**
	Template of C	**.0026380**
	Template of G	**.0016020**

Table-IV: Succolarity of ORs

The succolarity of all the textures of A, T, C, and G are almost same for Mouse and Chimpanzee ORs but in case of Human OR OR1D2 are less than same of other two ORs. It is seen that over genomic evolution the succolarity (amount of continuous density) in sequence structure in Human OR gets reduced than other similar sequences in Mouse and Human.

3.2.3 *Results on Bifurcation Dimension of Skeleton*

Let us think of a DNA sequence in terms of a four colored image as (A=Red, T=Blue, G=Green and C=Yellow) as shown below for CONTIG3463.6-1888.

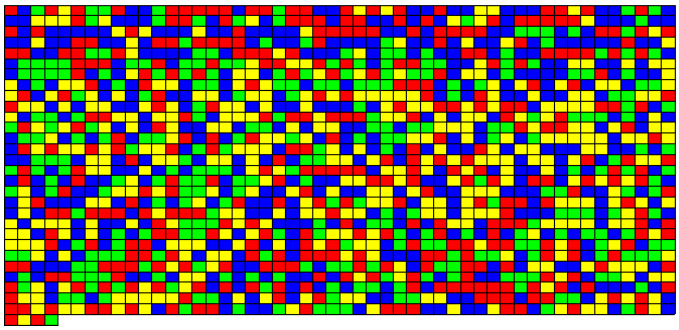


Figure-III: Colored template of CONTIG3463.6-1888

We then decompose the four colored image into four binary images $f(x,y) = z$ using the transformation

$$f^i(x,y) = 1 \;\; z \geq i$$

$$= 0 \; z > i$$

Corresponding to each of the binary image we have obtained their respective skeletons of which the one for CONTIG3463.6-1888 is shown below:

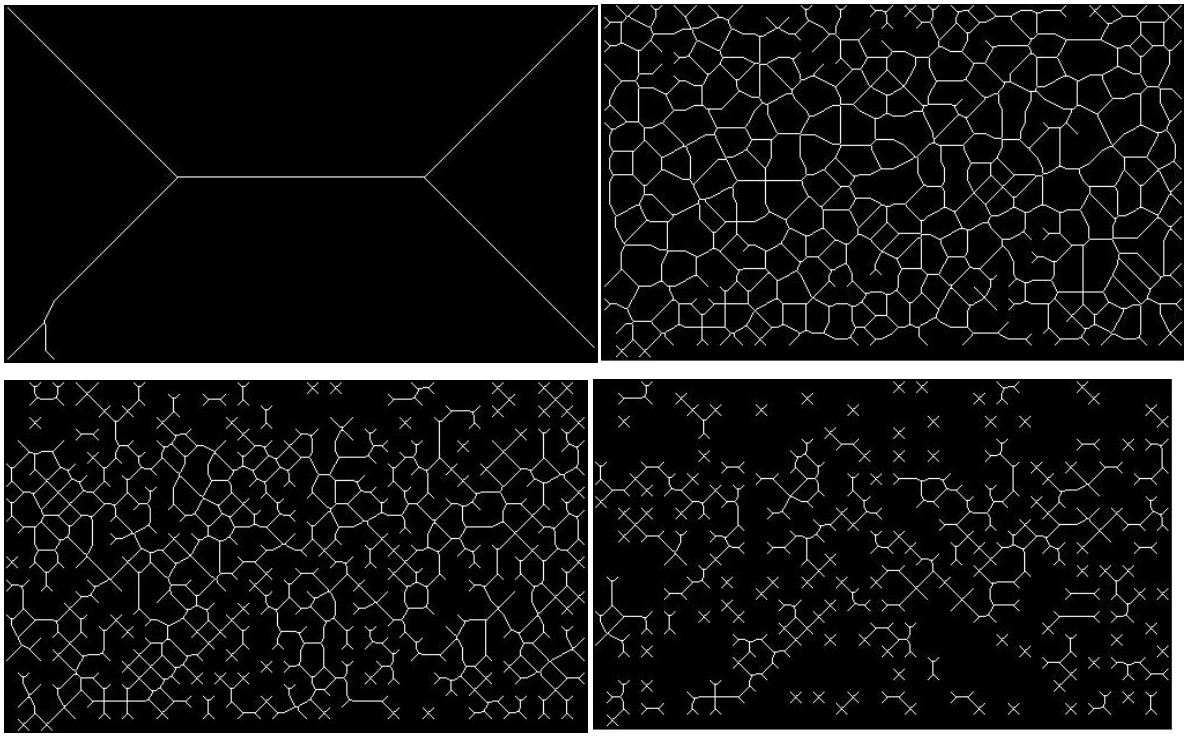


Figure-IV: Skeletons of OR Chimpanzee

Hence, using the technique as explained by Dayasagar et al [3, 4] in for computing ***Bifurcation Dimension*** for the skeletons mentioned above, we have found the same as stated below:

Olfactory Receptors	***Bifurcation Dimension (BD)***
OR1D2	**(3.8699 1.4512 3.2358 3.5306)**
GA_x5J8B7W3YLM-7052533-7051808	**(3.9792 4.8959 2.3225 3.3991)**
CONTIG3463.6-1888	**(1.0337 2.1623 3.7322 2.8787)**

Table-V: Bifurcation Dimension of the skeleton of ORs

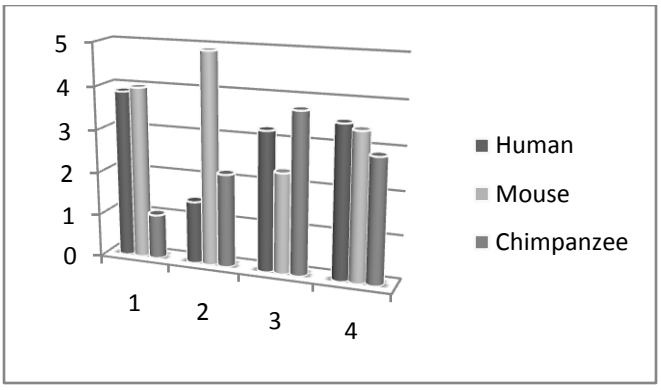


Table-VI: Histogram of Bifurcation Dimension according to Table-V

In table-VI the bifurcation dimensions are shown corresponding to each of OR of three different species. Typically it seems they do not follow a strict order. We believe this parameter would make them distinguished from each other in olfaction functioning.

4. *Conclusion and Future Endeavors:* In this paper we have shown an evolutionary connection among Human, Mouse and Chimpanzee ORs. These sequences have very close sequential similarity but they do function in different species due to their intricate details of the structures in the DNA sequence. Those intricate details are illustrated here. In near future we are about to report a quantitative classification based on Fractals and Mathematical Morphology with some more details about all the ORs of Human, Chimpanzee and Mouse. Also we are about to publish all the data we generated through a Web-Server and a platform as a national facility in Mathematical Genomics.

Acknowledgement: The authors are grateful to **Prof. Jean Serra,** Emeritus Professor, ESIEE-Engineering, University Paris-Est Former Director of Centre for Mathematical Morphology, France, and **Prof. B. S. Dayasagar**, Indian Statistical Institute, Bangalore for their kind advice and suggestions and express their gratitude to **Pratap Vardhan**, MANIT, Bhopal, for his technical help in writing programs.

References:

[1] G Matheron and J Serra, *History of Mathematical Morphology*, **http://cmm.ensmp.fr/~serra/pdf/birth_of_mm.pdf**

[2] B. S. Daya Sagar, 1996, *Fractal relations of a morphological skeleton*, **Chaos, Solitons & Fractals**, v. 7, no. 11, p. 1871-1879.

[3] B. S. Daya Sagar, et al 1998, *Morphometric relations of fractal-skeletal based channel network model*, **Discrete Dynamics in Nature and Society**, v. 2, no. 2, p. 77-92.

[4] **P. Radhakrishnan**, B. S. Daya Sagar, and Teo Lay Lian, (2004) *Estimation of fractal dimension through morphological decomposition*, **Chaos, Solitons & Fractals**, v. 21, no. 3, p. 563-572.

[5] B. S. Daya Sagar and Tay Lea Tien, (2004) *Allometric power-law relationships in a Hortinian fractal digital elevation model*, **Geophysical Research Letters**, (American Geophysical Union), v. 31, no. 6, L06501, doi:10.1029/2003GL019093.

[6] Shih and Mitchell Threshold Decomposition of Gray -Scale Morphology into Binary Morphology, **IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. I I , NO. 1.**

[7] P Radhakrishnan, BS Daya Sagar - *Morphological decomposition of sandstone pore-space: fractal power-laws* **Chaos, Solitons & Fractals**.

[8] J. Serra, *Image Analysis and Mathematical Morphology*, **Academic Press, Inc. Orlando,** FL, USA ©1983 ISBN:0126372403.

[9] Carlo Cattani, *Fractals and Hidden Symmetries in DNA,* **Mathematical Problems in Engineering** Volume 2010, Article ID 507056,.

[10] Choudhury, P.P.; Carry *Value Transformation (CVT): It's Application in Fractal formation*, **Advance Computing Conference**, 2009. IACC 2009. IEEE International.

[11] Choudhury, P.P. et al *Act of CVT and EVT In the Formation of Number-Theoretic Fractals,* **eprint arXiv: 0903.4770.**

[12] B. B. Mandelbrot, **The fractal geometry of nature**. New York, 1982.

[13] R. H. C. de Melo and A. Conci, Succolarity: Defining a Method to Calculate this Fractal Measure, **ieeexplore.ieee.org/iel5/4595514/4604343/04604424.pdf.**

[14] Yu Zu-Guo et al *Fractals in DNA sequence analysis,* **Chinese Physics**, Vol 11, No. 12, 2002, 1313-1318.

[15] Hassan, Sk. Sarif, *Underlying Mathematics in Diversification of Human Olfactory Receptors in Different Loci.* Available from **Nature Precedings <http://hdl.handle.net/10101/npre.2010.5475.1> (2010).**