

Multifractal Analysis of the Pathogenesis of Alzheimer Disease

Tong-han Lan, Zi-yang Lan, Xiao Li, Hao Cai, Fei-fei Yu

Abstract—In this paper the fractal analysis method is introduced to characterize the self-similarity of DNA sequences. Using this method, four whole genomes of the pathogenesis of Alzheimer's Disease provided by NCBI are analyzed. the pathogenesis of Alzheimer's Disease were exhibited multifractal characteristics when calculating the probability measure for four difference the pathogenesis sequences of each base individually. From the results we find that Renyi dimension range between 0.0090 to 3.5043. the multifractal characteristic properties in the whole the pathogenesis of Alzheimer's Disease are verified. The results of this paper suggest that the the multifractal characteristic property is one of the natural properties that DNA sequences possess, and is directly related to the structure and function of the whole DNA molecule.

Index Terms—mutifractal, pathogenesis of Alzheimer's Disease, Renyi dimension

I. INTRODUCTION

Alzheimer's disease (AD) is the most common form of dementia and the most frequent degenerative brain disorder encountered in old age. The risk of developing AD substantially increases after 65 years of age [1]. AD is quickly becoming one of the major universal healthcare problems. While the cause of this disease remains unknown, there is evidence for substantial genetic influence [1]-[5]. With the unclear pathogenesis, several hypotheses about the pathogenesis of AD, such as the Abnormal protein hypothesis, cholinergic hypothesis, oxidative stress theory and the estrogen hypothesis, etc. were proposed. Mutations in three genes, a myloid precursor prote in (APP), presenilin 1 and 2 (PSEN1 , PSEN2), result in an early onset, autosomal dominant form of the disease beginning in the third or fourth decade. The ϵ 4 allele of apolipoprotein E (APOE) increases the risk of both sporadic and familial AD occurring later in life around the sixth decade. Each of these genes is involved in the production or processing of the amyloid β peptide, which is deposited in the brain as dense plaques that are characteristic of the disease [3]. Today, however, there are neither precise diagnostic approaches nor effective therapeutic agents available for Alzheimer disease.

In the past decade or so there has been a ground swell of interest in unraveling the mysteries of DNA. In order to distinguish coding regions from non-coding ones, many approaches have been proposed [6]-[11], at the same times, the nonlinear scaling method, such as complexity [12] and fractal analysis [13]-[17] were used. The word "fractal" was coined by Benoit Mandelbrot in the late 1970s, but objects now defined as fractal in form have been known to artists and mathematicians for centuries. Mandelbrot's definition—"a set whose hausdorff dimension is not integer"—is clear in

mathematical terms. In addition, related concepts are those of self-similarity and sub-divisibility. The length of the coastline of Britain or the length of the perimeter of the Koch curve increases as we measure it at finer spatial resolution. The use of fractal analysis has been applied to many field [17].

The main purpose of this paper is to analyze the pathogenesis of Alzheimer Disease by means of fractal method. DNA research has previously focused on searching for low level patterns directly visible within the sequence, while ignoring high level patterns. The multifractal analysis of the pathogenesis of Alzheimer's Disease is intended to demonstrate that a higher level of pattern information may be available within the the pathogenesis of Alzheimer's Disease [18]. Thus the research has technical significance.

This paper is divided into four parts: the first part is the introduction of the pathogenesis of Alzheimer Disease and DNA series analysis basic state; in the second part we briefly review materials and the fractal method; the third and fourth part are the results and discussion.

II. MATERIALS AND METHODS

A. Data Resources

We will use the tools of the World Wide Web to search the GenBank DNA sequence database

(<http://www.ncbi.nlm.nih.gov>). Homo sapiens amyloid beta (A4) precursor protein (APP), RefSeqGene on chromosome 21 (NCBI Reference Sequence: NG_007376.1, GI:166795291); Homo sapiens apolipoprotein E (APOE), RefSeqGene on chromosome 19 (NCBI Reference Sequence: NG_007084.2, GI:163954918); Mus musculus presenilin-1 gene, alternatively spliced transcripts, complete cds, GenBank: AF007560.1, GI:2463667; Mus musculus presenilin-2 gene, strain 129X1/SvJ chromosome 12 CRA_211000022007779, whole genome shotgun sequence, GenBank: AAHY01101600.1, GI:69874353.

B. Mapping rule

A DNA sequence $\{n_i\}_{i=1,2,\dots,L}$ of length L is comprised a series of 4 types of base as follows: adenine (A); thymine (T); guanine (G); and cytosine (C). In order to apply numerical methods to nucleotide sequence, We can also use $(-2, -1, 1, 2)$ to replace $\{A, G, C, T\}$. We expect it to reveal more information than one dimensional DNA walk [19].

C. Fractal analysis

The use of fractal analysis has been applied to many fields [Kins94], such as ion channel kinetics analysis et al. Fractals are sets which exhibit self-similar properties at different scales and are characterized by their fractal

dimension. The word“fractal” was coined by Benoit Mandelbrot in the late 1970s, but objects now defined as fractal in form have been known to artists and mathematicians for centuries. Mandelbrot’s definition---“a set whose hausdorff dimension is not integer”---is clear in mathematical terms. In addition, related concepts are those of self-similarity and sub-divisibility. The length of the coastline of Britain or the length of the perimeter of the Koch curve increases as we measure it at finer spatial resolution. The use of fractal analysis has been applied to many field [19]-[21].

D. Multifractal Dimension

In practice, it is more useful to consider a signal as a multifractal, namely considering of more than one fractal. Using Renyi dimension can provide comprehension of complexity within the signal. The Renyi dimension is defined as (1) [19]. Where r is the volume element size, N_r is the number of volume elements for a given volume element size, P_j is the probability of occurrence within a given volume element. The q value can be considered as a fractal dimension index.

$$D_q = \lim_{r \rightarrow 0} \frac{1}{q-1} \frac{\log \sum_{j=1}^{N_r} p_j^q}{\log r} \quad (1)$$

III. RESULTS

A useful way of analyzing patchiness arising from the heterogeneous purine-pyrimidine content is DNA walk, defined as above 2.2 mapping rulers. The displacement of

walker after n steps, $y(n)$ is defined as $y(n) = \sum_{i=1}^n u_i$ and will display on a graph of $y(n)$ vs n .

A. The characteristic of DNA walk

We find apparent patchiness in real DNA sequences—both in the noncoding and coding regions. Figure 1 shows a representative DNA walk for four different of the pathogenesis of Alzheimer Disease DNA sequences.

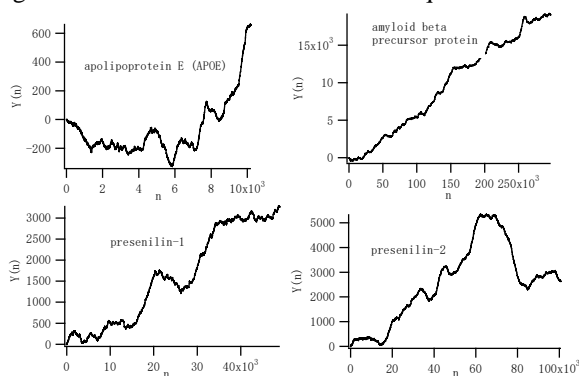


Fig1. shows a representative DNA walk for four different of the pathogenesis of Alzheimer Disease DNA sequences.

B. The characteristic of brown noise and white noise

We use simulation method and produce two series, namely brown noise and white noise. The displacement of walker after n steps $y(n)$ is used, Fig 2. shown brown noise and white noise walker.

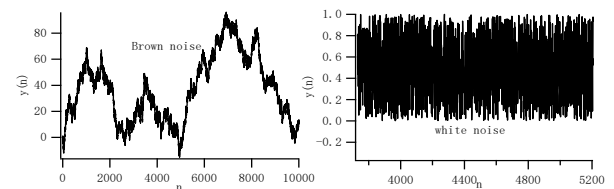


Fig2. shown brown noise and white noise

C. Calculation of probabilities

We can also use $(-2, -1, 1, 2)$ to replace $\{A, G, C, T\}$. The approach implemented is to calculate the probabilities for each individual base of the pathogenesis of Alzheimer Disease separately. Fig 3 shows the pathogenesis of Alzheimer Disease difference base at window 100 condition probability distribution. The results defer gauss distribution.

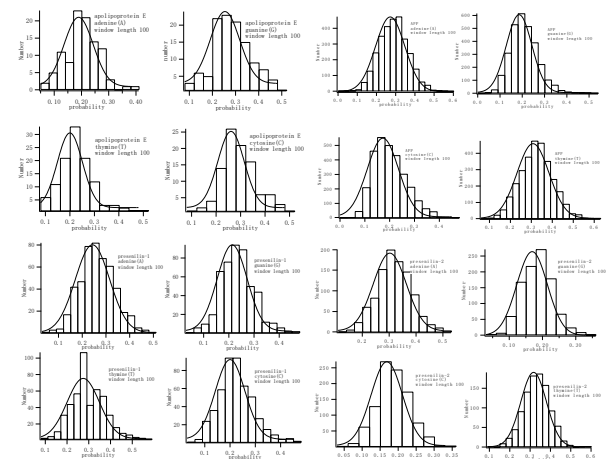


Fig 3 show the pathogenesis of Alzheimer Disease difference base at window 100 condition probability contribution.

D. Estimating the multifractal dimension

We compute Renyi dimension for difference the pathogenesis of Alzheimer Disease of the four bases. The results were shown as Fig 4. Renyi dimension for the amyloid beta precursor protein (APP) between 0.0090 to 3.5043, Renyi dimension for apolipoprotein E between 0.0023 to 1.7731, Renyi dimension for presenilin-1 gene between 0.0028 to 2.6578, Renyi dimension for presenilin-2 gene between 0.0217 to 3.0965.

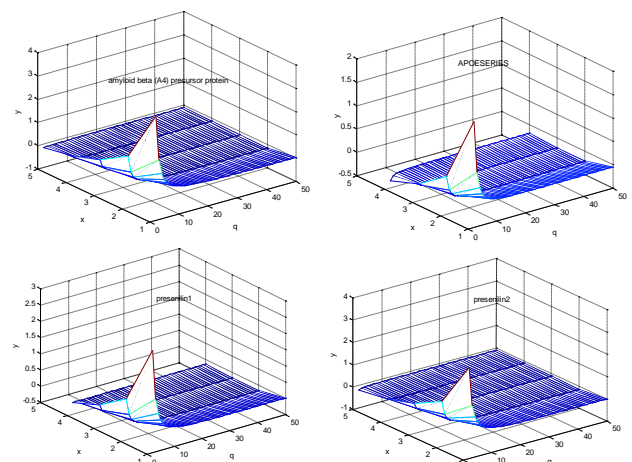


Fig 4. shown Renyi dimension for difference the pathogenesis of Alzheimer Disease

E. After Sliding window compute multifractal dimension

Multifractal characteristic of $\{A, G, C, T\}$ computed over

100 moving windows with a window size of 128 and offset of 1 was shown in Fig.5 for the amyloid beta precursor protein (APP), Multifractal value of A between 0.0357 to 1.0883, Multifractal value of G between 0.0345 to 1.1860, Multifractal value of C between 0.0245 to 0.9557, Multifractal value of T between 0.0220 to 1.1364. The similar apolipoprotein E, presenilin-1 gene and presenilin-2 gene of Multifractal characteristic of {A,G,C,T} can be computed.

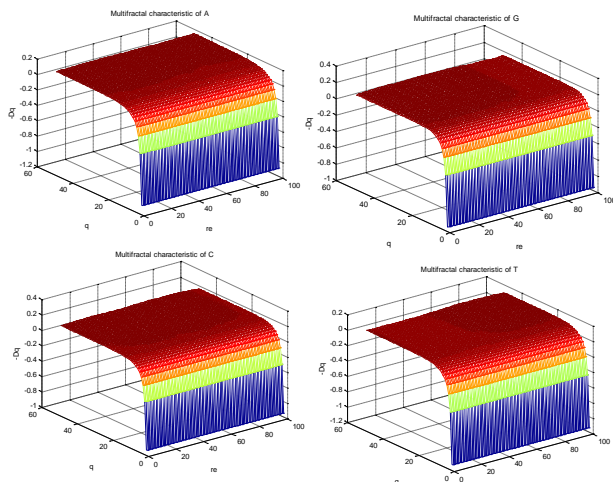


Fig.5 shown Multifractal characteristic of {A,G,C,T} computed over 100 moving windows with a window size of 128 and offset of 1.

IV. DISCUSSION

Multifractal analysis was applied to analyze amyloid beta precursor protein, apolipoprotein E, presenilin-1 and presenilin-2 the sequences of the pathogenesis of Alzheimer's Disease. the pathogenesis of Alzheimer's Disease were exhibited multifractal characteristics, at the same times, it was calculated the probability measure for four difference the pathogenesis sequences of each base individually. From the results we find that Renyi dimension range between 0.0090 to 3.5043. After sliding window compute multifractal dimension, Multifractal characteristic of {A,G,C,T} computed over 100 moving windows with a window size of 128 and offset of 1, for the amyloid beta precursor protein (APP), Multifractal value of {A,G,C,T} can be computed. The results were between 0.0357 to 1.0883, 0.0345 to 1.1860, 0.0245 to 0.9557, 0.0220 to 1.1364 individually. The similar apolipoprotein E, presenilin-1 gene and presenilin-2 gene Multifractal value of {A,G,C,T} can be computed.

In this paper the multifractal characteristic properties in the whole genomes of amyloid beta (A4) precursor protein (APP), apolipoprotein E, presenilin-1 gene, presenilin-2 gene are verified. The results of this paper suggest that the multifractal characteristic property is one of the natural properties that DNA sequences possess, and is directly related to the structure and function of the whole DNA molecule. There are still many discussions concerning the biological meaning and the origin of the multifractal characteristic properties in the DNA sequences. Such as, APP gene encodes a cell surface receptor and transmembrane precursor protein that is cleaved by secretases to form a number of peptides. Some of these peptides are secreted and can bind to the acetyltransferase complex APBB1/TIP60 to

promote transcriptional activation, while others form the protein basis of the amyloid plaques found in the brains of patients with Alzheimer disease. Mutations in this gene have been implicated in autosomal dominant Alzheimer disease and cerebroarterial amyloidosis. Multiple transcript variants encoding several different isoforms have been found for this gene [22]. Chylomicron remnants and very low density lipoprotein (VLDL) remnants are rapidly removed from the circulation by receptor-mediated endocytosis in the liver. Apolipoprotein E, a main apoprotein of the chylomicron, binds to a specific receptor on liver cells and peripheral cells. ApoE is essential for the normal catabolism of triglyceride-rich lipoprotein constituents [23]. act. these properties are considered to be related to the construction of the higher order structure of the DNA molecule.

This paper is preliminary with respect to the multifractal characteristic Exist in potential relevance to the pathogenesis of Alzheimer's Disease. Its aim is to provide the means for the analysis of the multifractal characteristic properties of DNA sequences. although the research results shown the multifractal characteristic Exist the pathogenesis of Alzheimer's Disease, This may led to series of questions, what is the physical phenomena and possible mechanism of the pathogenesis of Alzheimer's Disease? How to categorize different types of the pathogenesis of Alzheimer's Disease, how to understand the function of the pathogenesis of Alzheimer's Disease, how to build the pathogenesis of Alzheimer's Disease kinetics model systematically? Although the data tell us that the multifractal characteristic does exist in the pathogenesis of Alzheimer's Disease, we will still think about the physical properties of the pathogenesis of Alzheimer's Disease. what kind of distribution form do these the pathogenesis of Alzheimer's Disease have? All these studies should be the future direction of research.

ACKNOWLEDGMENT

This project was supported by China National Nature Science Foundation No:30470413, No:31160200. Hubei province Nature Science Foundation No:2004ABA220 and China Postdoctoral Science Foundation. We are also indebted to sir lan zi yang for his valuable support and helpful discussions.

REFERENCES

- [1] Joseph H. Lee & Sandra Baral & Rong Cheng, etc. Age-at-onset et linkage analysis in Caribbean Hispanics with familial late-onset Alzheimer's disease. *Neurogenetics* (2008) 9:51–60
- [2] Nussbaum RL, Ellis CE (2003) Alzheimer's disease and Parkinson's disease. *N Engl J Med* 348:1356–1364
- [3] St George-Hyslop PH, Pettinger A (2005) Molecular biology and genetics of Alzheimer's disease. *Curr Biol* 328:119–130
- [4] Tanzi RE, Bertram EH (2005) Twenty years of the Alzheimer's disease amyloid hypothesis: a genetic perspective. *Cell* 120:545–555
- [5] Pastor P, Goate AM (2004) Molecular genetics of Alzheimer's disease. *Curr Psychiatry Rep* 6:125–133.
- [6] Li W, Kaneko K. Long-range correlation and partial 1/f spectrum in a non-coding DNA sequence. *Europhys Lett*, 1992 17 : 655~660.
- [7] Peng C K, Buldyrev S, Goldberg A L, et al. Long-range correlation in nucleotide sequence. *Nature*, 1992 (356) : 168~170.
- [8] Stanley H E, Buldyrev S V, Goldberg A L, et al. Analysis of DNA sequences Using methods of statistical physics. *J. Physica A*, 1998. (249) : 430~438.
- [9] Herzog H, Grobe I. Correlations in DNA sequences – the role of protein coding segments. *J. Phys Rev E*, 1997(55) : 800~810.
- [10] Lou L F, Tsai Li, Zhou Y M. Informational parameters of nucleic acid and molecular evolution. *J Theor Biol*, 1988(130) : 351~361.

- [11] Luo L F, Tsai Li . Fractal dimension of nucleic acid sequences and its relation to evolutionary level. *ChinPhysLett*, 1988(5) :421~424.
- [12] Jun Xu, Yang Chao, Rensheng Chen. Fractal geometry study of DNA binding proteins. *J theorBiol*, 1994(171) :239~249
- [13] Zheng G J, YU Z G.. Correlation dimension and Kolmogorov entropy of DNA sequences. *Natural Science Journal of Xiangtan University*. 2000 Vol.22No.1.115-119.
- [14] Bai-lin Hao, Hoong-Chien Lee, and Shu-yu Zhang, Fractals related to long DNA sequences and complete genomes, *Chaos, Solitons and Fractals*, 2000.825-836.
- [15] Bassingthwaite, J.B. and Raymond, G.M. Evaluating rescaled range analysis for time series, *Ann. Biomed. Eng.* 1994;22:432-444.
- [16] Churilla, A.M., Gottschalk, W.A., Liebovitch, L.S., Selector, L.Y., Todorov, A., & Yeandle, S. Membrane potential fluctuations of human T-lymphocytes have fractal characteristics of fractional Brownian motion. *Ann. Biomed. Eng.* 1996;24:99-108.
- [17] Caballero R, Jewson S, Brix A .Long memory in surface air temperature: detection, modeling, and application to weather derivative valuation. *Clim Res.* 2002; 21:127-140.
- [18] R.Rifaat, W. Kinsner, and P. McAlpine. Multifractal analysis of DNA.
- [19] William Y.C. Chen and James D. Louck. "Necklaces, MSS Sequences, DNA Sequences". *Adv. in Appl. Math.* 18(1) (1997) 18-32.
- [20] Tong-Han Lan, Bang-Quan Xu, Hui-Jun yuan, Jia-Rui Lin Rescaled range analysis applied to the study delayed rectifier potassium channel kinetics, *Biophysical Chemistry*. 2003a; Vol106/1. 67-74.
- [21] Tong-Han Lan, Xiang-Ming Liu, Hui-Jun yuan, Jia-Rui Lin Gating kinetics of potassium channel in rat dorsal root ganglion neurons analyzed with fractal model. *Biophysical chemistry*. 2003b; Vol.106/3. 203-209.
- [22] Yazawa, T., Tanaka, K., Kato, A., & Katsuyama, T.. The scaling exponent calculated by the detrended fluctuation analysis, distinguishes the injured sick hearts against normal healthy hearts. *Proceeding IAING (WCECS08) International Conference on Computational Biology (ICCB)*, 2008; V(2):7-12. 22-24 October, San Francisco, USA.
- [23] Yazawa, T., Tanaka, K., & Katsuyama, T. Alternans lowers the scaling exponent of heartbeat fluctuation dynamics: A detrended fluctuation analysis in animal models and humans". *Proceeding CSIE2009, World Congress on Computer Science and Information Engineering*. Computer Soc. 2009; April, 221-225, Los Angeles, CA, USA. IEEE DOI 10.1109/CSIE.2009.784.

Tong-han Lan, Gongnongbin Road, Jiangnan District Wuhan 430012, P R China

Zi-yang Lan, Gongnongbin Road, Jiangnan District Wuhan 430012, P R China

Xiao Li, Gongnongbin Road, Jiangnan District Wuhan 430012, P R China

Hao Cai, Gongnongbin Road, Jiangnan District Wuhan 430012, P R China

Fei -fei Yu Gongnongbin Road, Jiangnan District Wuhan 430012, P R China