OPEN ACCESS
Fractals, Vol. 28, No. 7 (2020) 2150031 (11 pages)

DOI: 10.1142/S0218348X21500316

World Scientific
www.worldscientific.com

COMPLEXITY-BASED DETECTION OF SIMILARITY BETWEEN ANIMAL CORONAVIRUSES AND SARS-CoV-2 IN HUMANS

HAMIDREZA NAMAZI,[*,†,¶] ENRIQUE HERRERA-VIEDMA[‡,§] and ONDREJ KREJCAR[†]

[*]*School of Engineering, Monash University Malaysia*
Jalan Lagoon Selatan 47500 Subang Jaya
Selangor Darul Ehsan, Malaysia

[†]*Center for Basic and Applied Research*
Faculty of Informatics and Management, University of Hradec Králové
Rokitanského 62, 500 03 Hradec Králové, Czech Republic

[‡]*Andalusian Research Institute in Data Science and Computational Intelligence*
University of Granada, E-18071 Granada, Andalusia, Spain

[§]*Department of Electrical and Computer Engineering*
Faculty of Engineering King Abdulaziz University
P. O. Box 80204, Jeddah 21589, Saudi Arabia
[¶]*hrn_mechanic@yahoo.com*

Received July 26, 2020
Accepted September 9, 2020
Published October 23, 2020

Abstract

Severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) is the most dangerous type of coronavirus and has infected over 25.3 million people around the world (including causing 848,000 deaths). In this study, we investigated the similarity between the genome walks of coronaviruses in various animals and those of human SARS-CoV-2. Based on the results, although

[¶]Corresponding author.

bats show a similar pattern of coronavirus genome walks to that of SARS-CoV-2 in humans, decoding the complex structure of coronavirus genome walks using sample entropy and fractal theory showed that the complexity of the pangolin coronavirus genome walk has a 94% match with the complexity of the SARS-CoV-2 genome walk in humans. This is the first reported study that found a similarity between the hidden characteristics of pangolin coronavirus and human SARS-CoV-2 using complexity-based analysis. The results of this study have great importance for the analysis of the origin and transfer of the virus.

Keywords: SARS-CoV-2; Coronavirus (CoV); Genome Walk; Fractal Theory; Sample Entropy; Bat; Pangolin; Human.

1. INTRODUCTION

Severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2), as a positive-sense RNA virus, has been the most dangerous virus observed since late 2019. The disease it causes, called COVID-19, has affected the lives of many people in all countries. It is known that there are seven types of coronaviruses (CoVs) that affect mammals and birds.[1] The first coronavirus was discovered in 1930 in domesticated chickens.[2] SARS-CoV-2, as the latest member of this family, has had the worst effect on human life.

The review of the literature shows different categories of studies that have focused on the analysis of the SARS-CoV-2 genome, such as the studies on the phylogenetic analysis of the virus,[3,4] tracking of virus movement,[5,6] characterization of its genome,[7,8] and genomic variance analysis of the virus.[9,10]

An important category of works on the analysis of SARS-CoV-2 is the investigation of its origin. For this purpose, scientists have searched in other species, especially animals, to decode the similarity between their coronavirus genome and the human SARS-CoV-2 genome. Since some types of human coronaviruses that were investigated previously have their origin in bats,[11] most of the works on the analysis of SARS-CoV-2 have focused on bats[12–14] as the origin for SARS-CoV-2. However, some recent studies[15–17] have shown that pangolin CoV also has similar genomic sequences to SARS-CoV-2 in humans. However, all these works looked at the similarities of the genomic sequences of coronaviruses in animals and humans, and none of them tried to analyze the hidden characteristics of its genome. Therefore, to have a more precise tool to compare the genomic structures of coronaviruses of animals and humans, we evaluated the complex structure of coronavirus genomes.

Since stochastic genome walks (from coronavirus RNA) have complex patterns,[18] various complexity methods can be utilized to study the genome changes between animals and humans. Therefore, fractal theory was employed for this approach in this study.

Fractal theory is a well-known technique for the investigation of complex structures of time series and images. Fractal theory is used to quantify the complexity of self-affine and self-similar objects.[19] In general, for a fractal object, the fractal dimension (as a measure of complexity) satisfies the Szpilrajn inequality

$$F \geq D, \tag{1}$$

where F and D represent the fractal dimension and topological dimension (Euclidean dimension) of the object, respectively.

The application of fractal theory is more important in decoding the complex structure of self-affine fractals that cannot be quantified using the Euclidean dimension, due to the variations in their scaling exponents in different directions.[20] Genome walks map genome sequences into random fluctuations and therefore can be quantified using fractal theory.

Fractal theory has been widely applied to analyze various types of time series (e.g. eye movement,[21] EMG,[22] GSR,[23] and voice[24] signals) and images (e.g. MRI,[25] X-ray,[26] and human face[27] images) in biology, medicine, and biomedical engineering. We can also refer to the works on the fractal analysis of genomes, such as the studies that investigated the fractal shape of DNA walks,[28] classified genome sequences by fractal analysis on binary images of DNA,[29] employed fractal geometry and graph theory[30] to analyze lung cancer DNA sequences, classified cancerous cells versus cells from healthy subjects by complexity-based analysis of DNA walks,[31]

and investigated the 3D fractal assembly of DNA.[32] In a recent investigation, we showed the greater complexity of SARS-CoV-2 genome compared to the genomes of HIV and dengue viruses.[33] In other works, we investigated the significant alterations in the complexity of the coronavirus genome among various countries[34] and between different states in the USA.[18]

Sample theory has also been employed for the analysis of genome walk complexity. Sample entropy as a measure of complexity does not depend on the data length. Since coronavirus mutates in different animals and therefore has different nucleotide sequences, the generated genome walks from nucleotide sequences have different lengths. Therefore, to compare different genome walks with different lengths, we calculate their sample entropies that help us verify the results of the fractal analysis, which is dependent on the length of the data. Sample entropy has been applied widely for the analysis of the complexity of various types of data, especially in biology and medicine.[35–38] However, there are only two works that have studied the genome sequences using sample entropy. Singh *et al.*[39] predicted enhancer regions from a DNA random walk using sample entropy. In a previous study,[33] we showed that the SARS-CoV-2 genome walks have greater sample entropy than those of the HIV and dengue viruses.

This study, for the first time, evaluates the similarity of the coronavirus genome between animals and humans using complexity theories. We will present our methodology based on the fractal theory and sample entropy in Sec. 2. Then in Sec. 3, we will discuss the database and the conducted analysis. The obtained results from the analysis will be presented thereafter in Sec. 4. The discussion and conclusion will be highlighted in Sec. 5.

2. METHOD

In this paper, we evaluate the similarity between the genomic structures of coronaviruses in various animals (bat, pangolin, camel, sambar deer, canine, equine, feline, giraffe, and mink) and human SARS-CoV-2. For this purpose, we decoded the genomic structures of different coronaviruses in the form of genome walks. In this research, we used the method developed by Peng *et al.*[40] This method maps nucleotide sequences onto a random walk that has a complex structure, which can therefore be analyzed using different mathematical methods to compute its complexity. Although Peng *et al.*[40] named this random walk a DNA walk, in this study, since the coronavirus is an RNA virus, we named the random walk a "genome walk". In fact, the correlation among nucleotides in the genome distances is decoded on the genome walks.

To explain the process of generating a genome walk, we show a part of the genome sequence for SARS-CoV-2 in Fig. 1. The genome sequence should be read according to the numbers and in the direction of arrows.

As can be seen in Fig. 1, the genomic sequence includes adenine (A), guanine (G), cytosine (C), and thymine (T) bases. Based on the proposed method by Peng *et al.*,[40] we considered pyrimidines (C/T) and purines (A/G) for generating the coronavirus genome walks. For this purpose, we changed each purine to -1 and each pyrimidine to $+1$ wherever we observed them in the genome sequence. Then, "displacement", which is a dimensionless parameter, was defined using the following equation:

$$U(l) = \sum_{i=1}^{N} x(i)\,. \tag{2}$$

As Eq. (2) indicates, $U(l)$ is a combination of up $[x(i) = +1]$ and down $[x(i) = -1]$ fluctuations after N steps in the length (l) of the genome sequence. If we plot Eq. (2), we can obtain the genome walks.

Figure 2 shows a part of the genome walk for SARS-CoV-2. As shown, the genome walk is a random process that has a complex nature, and

```
1  gggtgtctgt cttggtgcag gtcaattcgt gggttgggtc atacccttag cctttatgcc
2  tgtgcaatcc cggaaattta ttgttccatg ggttatgtac ttgcgtaagc gtggcgaaaa
3  gggtgcttac aataaagatc atggatgtgg cggttttgga catgtttatg attttaaagt
4  tgaagatgct tatgaccagg tgcatgatga gcctaagggt aagttttcta agaaggctta
5  tgctttaatt agagggtatc gtggtgttaa accacttctc tatgtagacc agtatggttg
6  tgattatact ggtagtcttg cagatggctt agaggcttat gctgataaga cattgcaaga
7  aatgaaggca ttatttccta cttggagtca ggaactcctt tttgatgtaa ttgtggcatg
8  gcatgttgtg cgtgatccac gttatgttat gagattgcag agtgctgcta ctatacgtag
```

Fig. 1 A portion of the genome sequence of human SARS-CoV-2 extracted in the USA.[41]

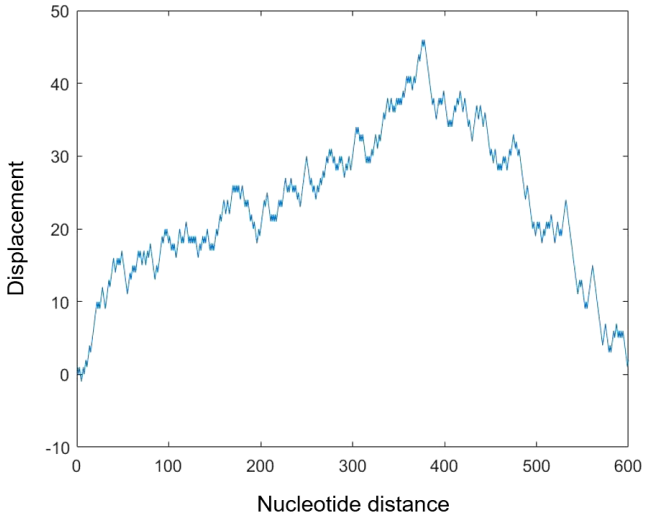

Fig. 2 A segment of the genome walk for SARS-CoV-2.

therefore it can be analyzed using different methods to compute its complexity.

To decode the similarity of the complex structure of genome walks between different animals and humans, we utilized fractal theory and sample entropy. Initially, we calculated the fractal dimension, which shows the complexity of random walks, where there is a direct relationship between its value and the complexity of a random walk.

We chose the box-counting method, which is suitable for the analysis of genome walks.[34] This method works based on segmenting the random walk using boxes that have the same size (ϕ). Then, the number of these boxes (N) is counted. The algorithm changes the size of these boxes in several iterations and finally calculates the fractal dimension using the number and size of boxes in each step:

$$F = \lim_{\phi \to 0} \frac{\log N(\phi)}{\log 1/\phi}. \tag{3}$$

The fractal dimension in the general form is formulated by[42]

$$F_t = \lim_{\phi \to 0} \frac{1}{t-1} \frac{\log \sum_{i=1}^{N} p_i^t}{\log \phi}. \tag{4}$$

In Eq. (4), t is the order of F, and p_i indicates the probability in the ith segment of the random walk, with

$$p_i = \lim_{l \to \infty} \frac{r_i}{l}, \tag{5}$$

where r_i is the number of occurrences in the ith segment. The whole nucleotide distance is denoted by l.

To verify the results of fractal analysis, we calculated the sample entropy of the genome walks. Sample entropy is a modification of approximate entropy that does not depend on the data length. Since the coronavirus genome walks for different animals have different lengths, the calculation of sample entropy helped us to verify the result of the fractal dimension, which is dependent on the length of the data. Similar to the fractal dimension, a larger value of sample entropy indicates a greater complexity of data.

Considering the genome walks in the form of $\{u(1), u(2), u(3), \ldots, u(n)\}$ with a constant interval of ε, we define a template vector of length z (embedding dimension) in the form of $U_z(i) = \{u_i, u_{i+1}, u_{i+2}, \ldots, u_{i+z-1}\}$, and the distance function $d[U_z(i), U_z(j)](i \neq j)$ is the Chebyshev distance. We formulated sample entropy (SamEn) as

$$\text{SamEn} = -\log \frac{A}{B}. \tag{6}$$

Considering e as the tolerance (0.2 $\times$ standard deviation of data), A stands for the number of template vector pairs

$$d[U_{z+1}(i), U_{z+1}(j)] < e. \tag{7}$$

On the other hand, B stands for the number of template vector pairs

$$d[U_z(i), U_z(j)] < e. \tag{8}$$

Therefore, we first analyzed the similarity of the genomes between different samples of coronaviruses obtained from animals and humans by generating the genome walks. Then, to confirm their similarities, we computed their fractal dimension and sample entropy.

3. DATABASE AND ANALYSIS

For our analysis, we used several samples of the coronavirus complete genomes from the open-access nucleotide database.[41] Table 1 lists the genome accession number, name, and isolation location for each sample used in this study. Based on the availability of the genome data, we selected 3–6 samples in the case of different animals and humans. It should be noted that due to the variations in coronavirus genomes between different animals, the genome sequence has different lengths between these species.

We generated the genome walks of different samples of coronavirus genomes using Eq. (2) and then computed their fractal dimension and sample

Table 1 Genome Sequences Used in This Study.

Genome Accession No.	Name	Isolation Location
MT079843	SARS-CoV-2	China
MT079846	SARS-CoV-2	China
MT079847	SARS-CoV-2	China
MT079853	SARS-CoV-2	China
MT428551	SARS-CoV-2	Kazakhstan
MT435080	SARS-CoV-2	India
DQ022305	Bat SARS CoV	China
DQ071615	Bat SARS CoV	China
DQ412042	Bat SARS CoV	China
DQ412043	Bat SARS CoV	China
FJ588686	Bat SARS CoV	China
GQ153539	Bat SARS CoV	Hong Kong
MT121216	Pangolin CoV	China
MT040335	Pangolin CoV	China
MT072864	Pangolin CoV	China
MT040336	Pangolin CoV	China
MT040334	Pangolin CoV	China
MT040333	Pangolin CoV	China
FJ938060	Feline CoV	USA
FJ938062	Feline CoV	The Netherlands
FJ938061	Feline CoV	USA
FJ938059	Feline CoV	The Netherlands
FJ938057	Feline CoV	The Netherlands
FJ938053	Feline CoV	The Netherlands
EF424623	Giraffe CoV	USA
EF424622	Giraffe CoV	USA
EF424624	Giraffe CoV	USA
NC_023760	Mink CoV	USA
HM245925	Mink CoV	USA
HM245926	Mink CoV	USA
MF113046	Mink CoV	China
LC061274	Equine CoV	Japan
LC061273	Equine CoV	Japan
LC061272	Equine CoV	Japan
EF446615	Equine CoV	Unspecified
KY063616	Canine CoV	China
KY063617	Canine CoV	China
GQ477367	Canine CoV	Taiwan
JQ404410	Canine CoV	Unspecified
JQ404409	Canine CoV	Unspecified
KP981644	Canine CoV	Italy
KF906251	Camel CoV	United Arab Emirates
KF906249	Camel CoV	United Arab Emirates
MN514967	Camel CoV	Nigeria
MN514966	Camel CoV	Nigeria
MN514965	Camel CoV	Nigeria
FJ425188	Sambar deer CoV	USA
FJ425190	Sambar deer CoV	USA
FJ425189	Sambar deer CoV	USA

entropy. These analyses were performed in MATLAB (The MathWorks, USA). As mentioned previously, the box-counting algorithm was chosen for the fractal analysis.

To investigate the significance of the difference between different genome walks, we performed statistical analyses. A one-way repeated-measure analysis of variance (ANOVA) test and a post-hoc

Tukey test ($\alpha = 0.01$) were conducted on the calculated fractal dimensions and sample entropies.

4. RESULTS

Figure 3 shows the generated genome walks for the different animals' coronaviruses. Genome walks in different colors belong to different samples of coronaviruses. It should be noted that in the case of some of the plots, some genome walks overlap with each other.

As the plots in Fig. 3 show, most of these animals show similar patterns of genome walks. For instance, camel, sambar deer, equine, and giraffe show similar coronavirus genome walks, which indicates the similarity of the coronavirus genome between them.

As mentioned previously, we also analyzed the genome walks for bats and pangolins, which are shown in Fig. 4. Genome walks in different colors belong to different samples. In addition, Fig. 5 illustrates the samples of the SARS-CoV-2 genome walk for humans. Comparing the genome walks in Fig. 3 with the SARS-CoV-2 genome walks for humans (Fig. 5) indicates that none of these patterns is similar to the SARS-CoV-2 genome walks

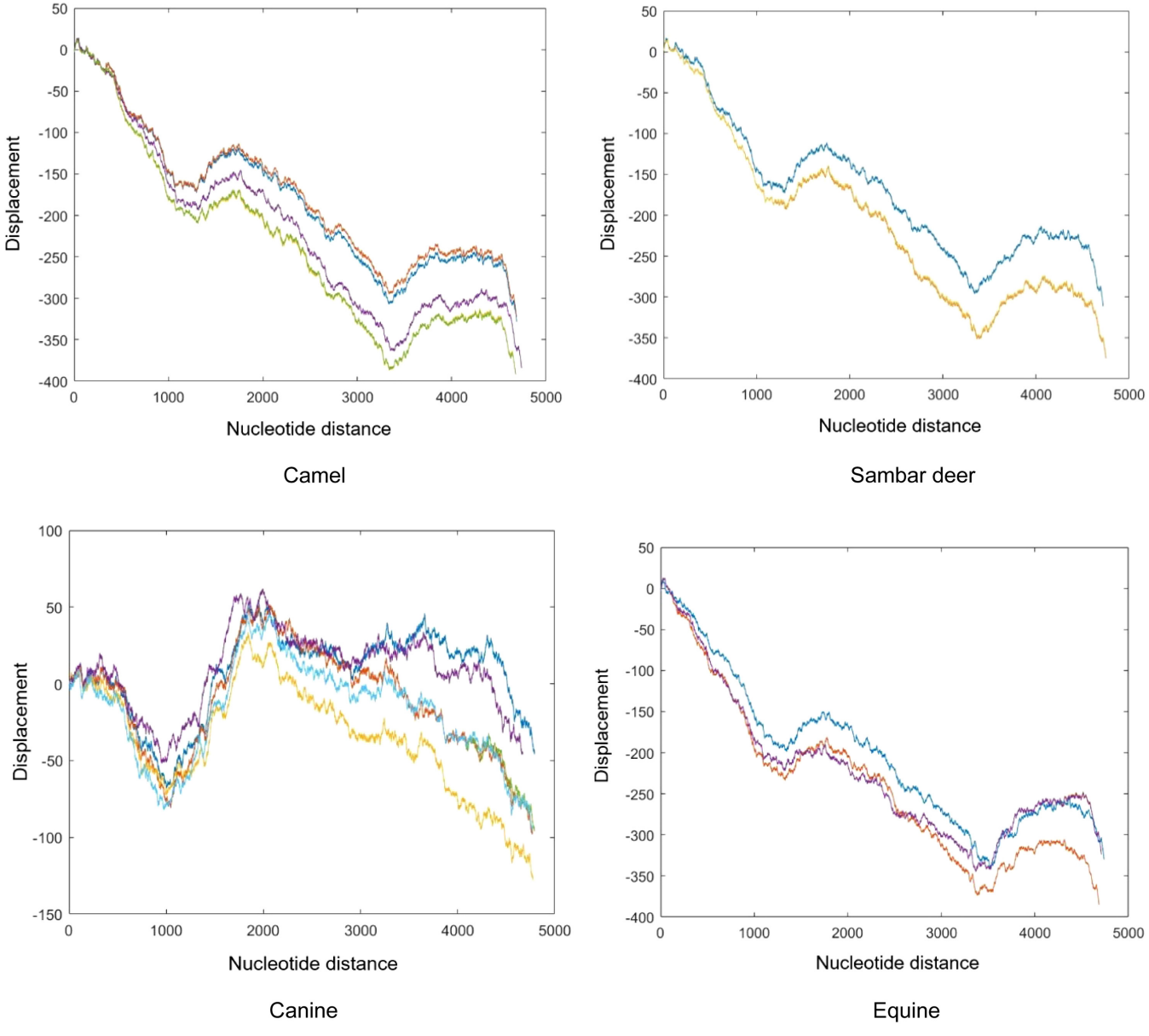

Fig. 3 Samples of the coronavirus genome walks for various types of animals. Each color indicates a sample genome walk.

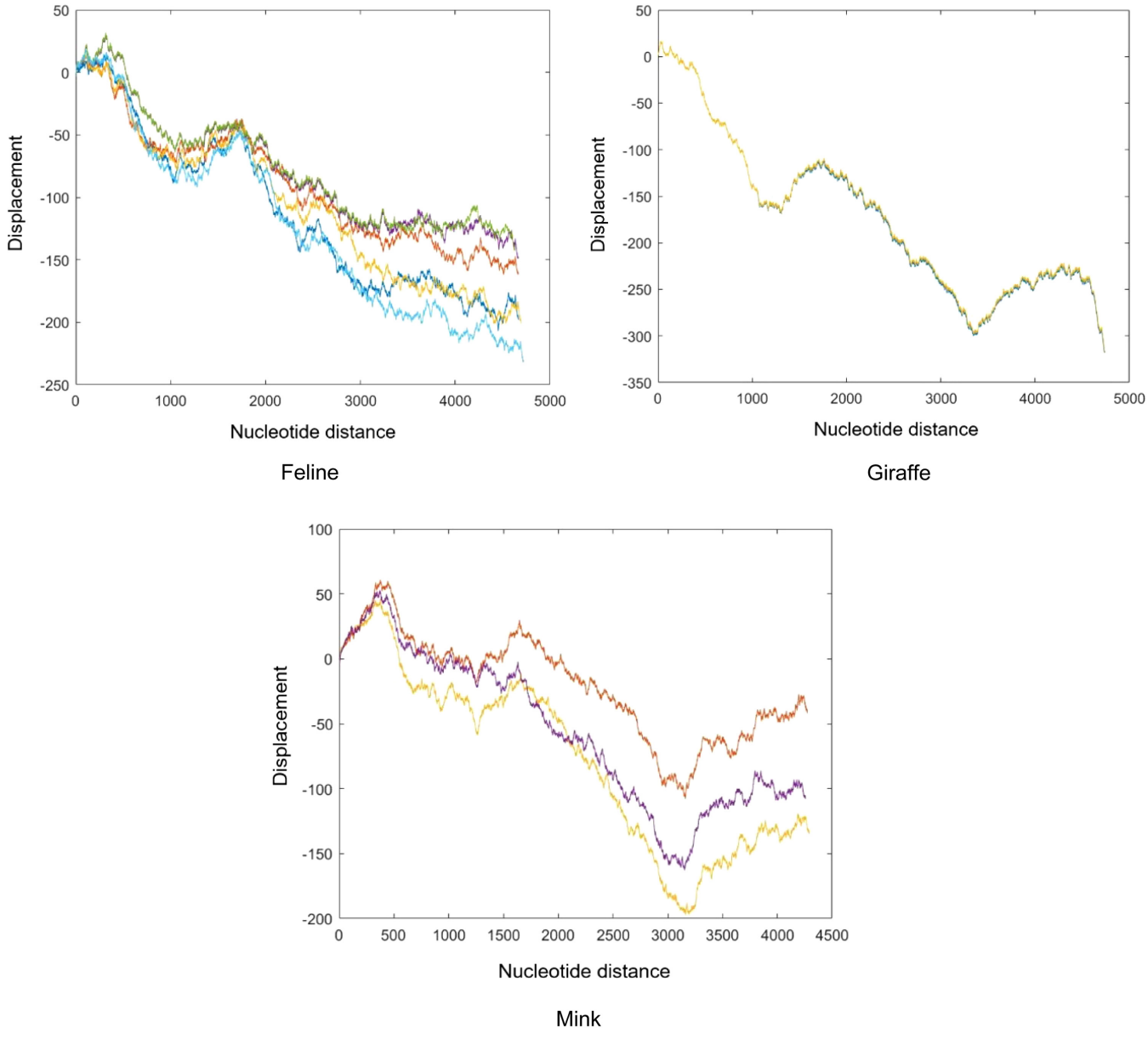

Fig. 3 (*Continued*)

for humans. Therefore, we can state that the coronavirus genomes in the case of these animals are different from the SARS-CoV-2 genome in humans.

However, comparing different plots in Fig. 4 with the SARS-CoV-2 genome walks for humans (Fig. 5) indicates that bat, pangolin, and humans have similar genome walks for their coronaviruses. To mathematically investigate these similarities, we computed the fractal dimension and sample entropy for these genome walks.

Figure 6 illustrates the values of the genome walks' fractal dimension for the human, bat, and pangolin coronaviruses. The reported value on each bar indicates the average of the calculated values between different samples. Standard deviations are shown in the form of error bars.

As shown in Fig. 6, bat CoV genome walks have the lowest fractal dimension. However, the pangolin CoV genome walks and the human SARS-CoV-2 genome walks show similar values of the fractal dimension. In other words, we can indicate that the pangolin CoV genome walks and the human SARS-CoV-2 genome walks are more complex than the bat CoV genome walks.

The ANOVA test results (p-value = 0.0001, F-value = 20.4342) indicate significant variations in the complexity of coronavirus genome walks. To check which pairs of coronaviruses caused this

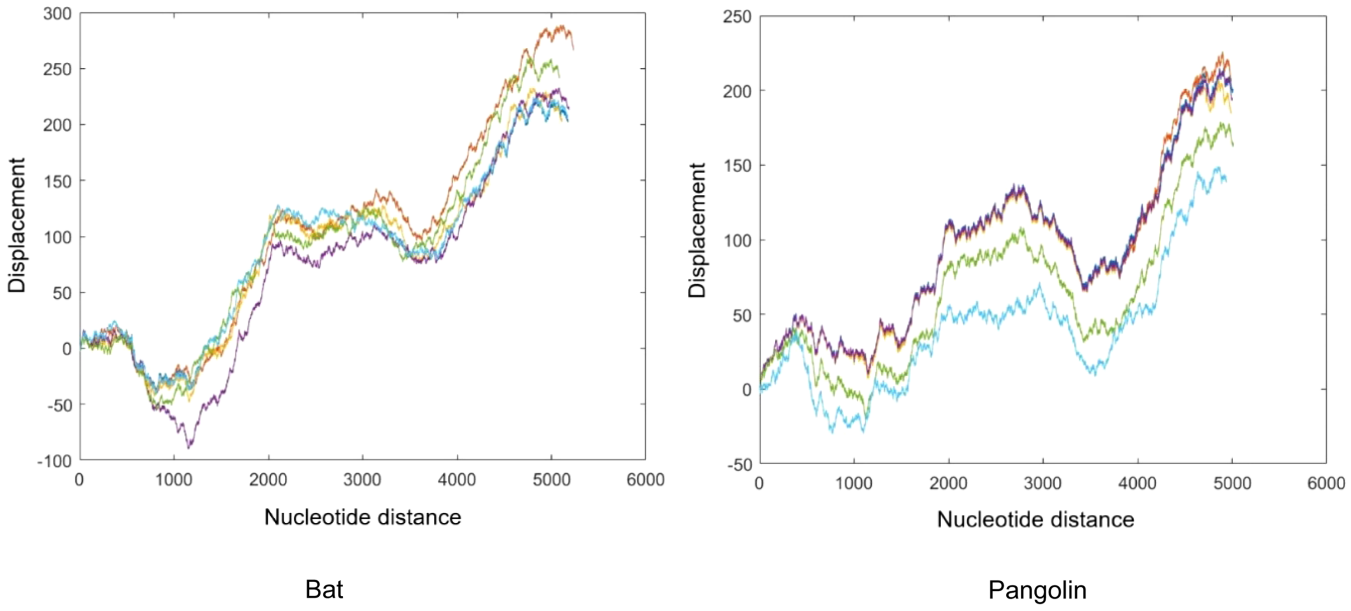

Fig. 4 Different samples of the coronavirus genome walks for bats and pangolins. Each color indicates a sample genome walk.

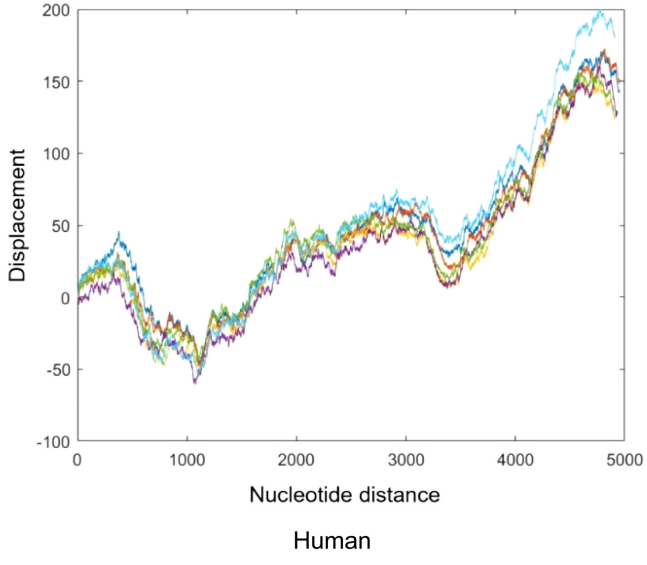

Fig. 5 Different samples of the SARS-CoV-2 genome walk for humans. Each color indicates a sample genome walk.

significant variation, the post-hoc Tukey test results are shown in Table 2. Based on these results, the complexities of bat and pangolin CoV genome walks are significantly different. In addition, the genome walks of bat CoV and human SARS-CoV-2 have significantly different complexities. However, the obtained value (p-value = 0.9979) for the difference between the values of fractal dimension for the pangolin CoV genome walks and the human SARS-CoV-2 genome walks indicates no significant difference. In other words, the complexity of the genome walks for pangolin CoV is 99.79% similar to the complexity of genome walks for human SARS-CoV-2.

As mentioned previously, since the genome walks of bat CoV, pangolin CoV, and human SARS-CoV-2 have different lengths, we also calculated the sample entropies of these genome walks, which are presented in Fig. 7.

Similar to the trend of the fractal dimension for different genome walks, the bat CoV genome walks have the lowest sample entropy. In addition, the

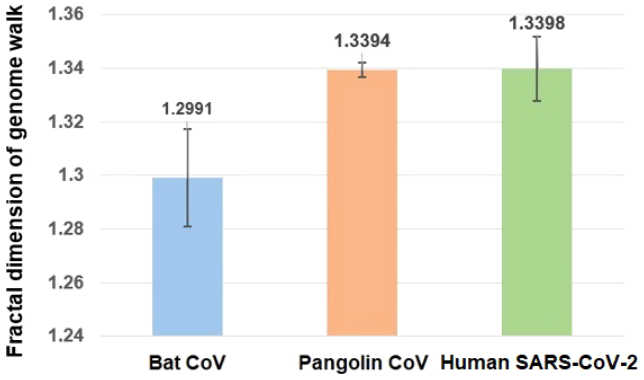

Fig. 6 The fractal dimensions of coronavirus genome walks for bat, pangolin, and humans.

Table 2 Comparison of Fractal dimension of Genome Walks among Various Types of Coronaviruses.

Pair	p-Value
Bat CoV versus pangolin CoV	0.0002
Bat CoV versus SARS-CoV-2	0.0002
Pangolin CoV versus SARS-CoV-2	0.9979

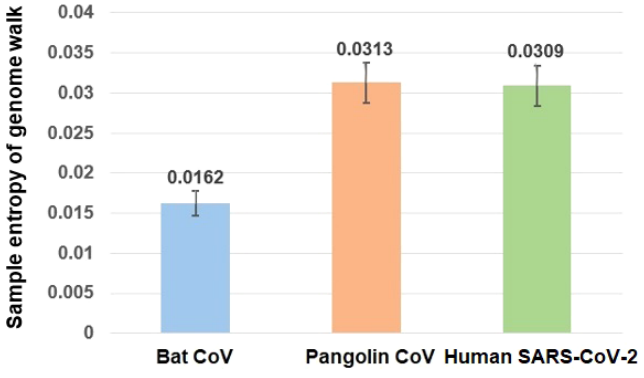

Fig. 7 Sample entropies of the coronavirus genome walks for bat, pangolin, and humans.

pangolin CoV genome walks and the human SARS-CoV-2 genome walks show similar values of sample entropy. In other words, we can indicate that the pangolin CoV genome walks and the human SARS-CoV-2 genome walks are more complex than the bat CoV genome walks.

The ANOVA test results (p-value = 0, F-value = 85.2953) indicate a significant variation in the genome walks' complexities. In addition, the results of the post-hoc Tukey test in Table 3 indicate that the complexities of the bat CoV genome walks and the pangolin CoV genome walks are significantly different. A similar result can be observed in the comparison between the complexities of bat

Table 3 Comparison of Sample Entropy of Genome Walks among Various Types of Coronaviruses.

Pair	p-Value
Bat CoV versus pangolin CoV	0
Bat CoV versus SARS-CoV-2	0
Pangolin CoV versus SARS-CoV-2	0.9490

CoV genome walks and the human SARS-CoV-2 genome walks. However, the obtained value (p-value = 0.9490) for the difference between the sample entropies of the pangolin CoV genome walks and the human SARS-CoV-2 genome walks indicates no significant difference. In other words, the complexity of genome walks for pangolin CoV is 94.90% similar to the complexity of the human SARS-CoV-2 genome walks.

Therefore, we can state based on the presented results that, although bat CoV has a similar genomic structure to human SARS-CoV-2, its characteristics (fractal dimension and sample entropy) are significantly different from human SARS-CoV-2. The results show that the pangolin CoV genome is similar to the human SARS-CoV-2 genome in terms of the structure and characteristics.

5. DISCUSSION AND CONCLUSION

We investigated the similarity of the coronavirus genome between various animals and humans. We mapped the variations in the genome along with the nucleotide distance in the form of genome walks. Therefore, we generated genome walks of coronaviruses for different animals and humans. The findings indicated that the coronavirus genome walks of bat and pangolin are similar to the human SARS-CoV-2 genome walks. To check the inherent characteristics of these random walks, we calculated their fractal dimensions. The results indicated a significant variation among the complex structures of the bat coronavirus genome walks and the pangolin coronavirus genome walks. Similarly, we observed a significant difference among the complex structures of the bat coronavirus genome walks and the human SARS-CoV-2 genome walks. However, no significant difference was observed among the complex structures of the pangolin coronavirus genome walks and the human SARS-CoV-2 genome walks. Based on this result, the complexity of the pangolin CoV genome walks is 99.79% similar to the complexity

of human SARS-CoV-2 genome walks. In addition, since the genome walks of coronaviruses for bats, pangolins, and humans have different lengths, to verify the results of the fractal analysis, we calculated the sample entropies of their genome walks. The results of the analysis of sample entropy were similar to the results of fractal analysis. Based on the results, we could only observe nonsignificant difference among the complexities of the genome walks for the pangolin CoV and the human SARS-CoV-2. Therefore, we can conclude that although the bat CoV genome walk is similar to the human SARS-CoV-2 genome walk, the decoding of the complexity of genome walks indicates that the pangolin CoV genome walk is significantly similar to the human SARS-CoV-2 genome walk. This result is in line with the reported investigations[15–17] that claimed a high similarity between the pangolin coronavirus and the human SARS-CoV-2. In fact, our finding is one step forward compared to those studies, since we also analyzed the hidden characteristics of the coronavirus genome, in addition to merely examining its genome sequences. Although we do not claim that pangolins are hosts for human SARS-CoV-2, our findings challenge the results of studies[12–14] that introduced bats as the hosts for SARS-CoV-2. Finding the host of SARS-CoV-2 outside the human body needs more detailed studies through the analysis of coronavirus genome walks in other types of animals.

Overall, for the first time, we decoded the similarity of the coronavirus genome between various animals and humans by quantifying the complexity of genome walks, beyond merely examining the genome sequences.

ACKNOWLEDGMENTS

This work was supported in part by the project of Kamil Kuca of Grant Agency of Excellence 2020/2205, Faculty of Informatics and Management, University of Hradec Králové, Czech Republic; and by the Spanish Ministry of Science with the FEDER financing of Project PID2019-103880RB-I00.

REFERENCES

1. D. X. Liu, J. Q. Liang and T. S. Fung, Human coronavirus-229E, -OC43, -NL63, and -HKU, in *Reference Module in Life Sciences*, Elsevier Reference Collection in Life Sciences (Elsevier, 2020), pp. 1–11. doi:10.1016/B978-0-12-809633-8.21501-X.
2. T. Estola, Coronaviruses, a new group of animal RNA viruses, *Avian Dis.* **14** (1970) 330–336.
3. P. Stefanelli *et al.*, Whole genome and phylogenetic analysis of two SARS-CoV-2 strains isolated in Italy in January and February 2020: additional clues on multiple introductions and further circulation in Europe, *Euro Surveill.* **25** (2020) 2000305.
4. J. A. Jaimes *et al.*, Phylogenetic analysis and structural modeling of SARS-CoV-2 spike protein reveals an evolutionary distinct and proteolytically sensitive activation loop, *J. Mol. Biol.* **432** (2020) 3309–3325.
5. L. Morawska and J. Cao, Airborne transmission of SARS-CoV-2: The world should face the reality, *Environ. Int.* **139** (2020) 105730.
6. F. Ferretti *et al.*, Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing, *Science* **368** (2020) eabb6936.
7. X. Ou *et al.*, Characterization of spike glycoprotein of SARS-CoV-2 on virus entry and its immune cross-reactivity with SARS-CoV, *Nat. Commun.* **11** (2020) 1620.
8. Y. Wang *et al.*, Characterization of an asymptomatic cohort of SARS-COV-2 infected individuals outside of Wuhan, China, *Clin. Infect. Dis.* (2020). doi:10.1093/cid/ciaa629.
9. O. M. Ugurel, O. Ata and D. Turgut-Balik, An updated analysis of variations in SARS-CoV-2 genome, *Turk. J. Biol.* **44** (2020) 157–167.
10. T. Koyama, D. Platt and L. Parida, Variant analysis of SARS-CoV-2 genomes, *Bull. World Health Organ.* **98** (2020) 495–504.
11. D. Forni, R. Cagliani, M. Clerici and M. Sironi, Molecular evolution of human coronavirus genomes, *Trends Microbiol.* **25** (2017) 35–48.
12. D. Paraskevis *et al.*, Full-genome evolutionary analysis of the novel corona virus (2019-nCoV) rejects the hypothesis of emergence as a result of a recent recombination event, *Infect. Genet. Evol.* **79** (2020) 104212.
13. H. Zhou *et al.*, A novel bat coronavirus closely related to SARS-CoV-2 contains natural insertions at the S1/S2 cleavage site of the spike protein, *Curr. Biol.* **30** (2020) 2196–2203.e3.
14. P. Zhou *et al.*, A pneumonia outbreak associated with a new coronavirus of probable bat origin, *Nature* **579** (2020) 270–273.
15. T. Zhang, Q. Wu and Z. Zhang, Probable pangolin origin of SARS-CoV-2 associated with the COVID-19 outbreak, *Curr. Biol.* **30** (2020) 1346–1351.e2.
16. T. T. Lam *et al.*, Identifying SARS-CoV-2-related coronaviruses in Malayan pangolins, *Nature* **583** (2020) 282–285.
17. K. Xiao *et al.*, Isolation and characterization of 2019-nCoV-like coronavirus from Malayan pangolins, preprint (2020), bioRxiv 2020.02.17.951335. doi:10.1101/2020.02.17.951335.

18. H. Namazi, O. Krejcar and A. Subasi, Complexity and information-based analysis of the variations of the SARS-CoV-2 genome in the United States of America (USA), *Fractals* (2020). doi:10.1142/S0218348X21500237.
19. T. H. Wilson, Some distinctions between self-similar and self-affine estimates of fractal dimension with case history, *Math. Geol.* **32** (2000) 319–335.
20. Y. Jin *et al.*, Definition of fractal topography to essential understanding of scale-invariance, *Sci. Rep.* **7** (2017) 46672.
21. H. Alipour *et al.*, Complexity-based analysis of the relation between fractal visual stimuli and fractal eye movements, *Fluct. Noise Lett.* **18** (2019) 1950012.
22. H. Babini, V. V. Kulish and H. Namazi, Physiological state and learning ability of students in normal and virtual reality conditions: Complexity-based analysis, *J. Med. Internet Res.* **22** (2020) e17945.
23. S. Omam *et al.*, Complexity-based decoding of brain-skin relation in response to olfactory stimuli, *Comput. Methods Prog. Biomed.* **184** (2020) 105293.
24. M. R. A. Ahamed *et al.*, Complexity-based analysis of brains' synchronization in human-human interaction, *Fractals* (2020). doi:10.1142/S0218348X20501029.
25. S. Krohn *et al.*, Evaluation of the 3D fractal dimension as a marker of structural brain complexity in multiple-acquisition MRI, *Hum. Brain Mapp.* **40** (2019) 3299–3320.
26. H. Namazi and V. V. Kulish, Complexity-based classification of the coronavirus disease (COVID-19), *Fractals* (2020). doi:10.1142/S0218348X20501145.
27. H. Namazi *et al.*, The fractal based analysis of human face and DNA variations during aging, *Biosci. Trends* **10** (2017) 477–481.
28. C. Cattani, Fractals and hidden symmetries in DNA, *Math. Probl. Eng.* **2010** (2010) 507056.
29. C. Cattani and G. Pierro, On the fractal geometry of DNA by the binary image analysis, *Bull. Math. Biol.* **75** (2013) 1544–1570.
30. M. Babič, J. Mihelič and M. Calì, Complex network characterization using graph theory and fractal geometry: The case study of lung cancer DNA sequences, *Appl. Sci.* **10** (2020) 3037.
31. H. Namazi, V. V. Kulish, F. Delaviz and A. Delaviz, Diagnosis of skin cancer by correlation and complexity analyses of damaged DNA, *Oncotarget* **6** (2015) 42623–42631.
32. A. Carbone *et al.*, 3D fractal DNA assembly from coding, geometry and protection, *Nat. Comput.* **3** (2004) 235–252.
33. H. Namazi, Complexity-based classification of the coronavirus genome versus genomes of the human immunodeficiency virus (HIV) and dengue virus, *Fractals* (2020). doi:10.1142/S0218348X20501297.
34. H. Namazi and N. B. M. Dawi, Information and complexity-based analysis of the variations of the coronavirus genome between different countries, *Fractals* (2020). doi:10.1142/S0218348X20501340.
35. A. Rizal and S. Hadiyoso, Sample entropy on multidistance signal level difference for epileptic EEG classification, *Sci. World J.* **2018** (2018) 8463256.
36. M. Du *et al.*, Detection of stretch reflex onset based on empirical mode decomposition and modified sample entropy, *BMC Biomed. Eng.* **1** (2019) 23.
37. T. Horie *et al.*, Sample entropy in electrocardiogram during atrial fibrillation, *Yonago Acta Med.* **61** (2018) 49–57.
38. M. Nezafati, H. Temmar and S. D. Keiholz, Functional MRI signal complexity analysis using sample entropy, *Front. Neurosci.* (2020). doi:10.3389/fnins.2020.00700.
39. A. P. Singh, S. Mishra and S. Jabin, Sequence based prediction of enhancer regions from DNA random walk, *Sci. Rep.* **8** (2018) 15912.
40. C.-K. Peng *et al.*, Long-range correlations in nucleotide sequences, *Nature* **356** (1992) 168–170.
41. National Center for Biotechnology Information, Nucleotide database (2020), http://www.ncbi.nlm.nih.gov/nuccore/.
42. H. Namazi and S. Jafari, Decoding of wrist movements direction by fractal analysis of magnetoencephalography (MEG) signal, *Fractals* **27** (2019) 1950001.