



IMAGE: A MAP OF THE STARS OF THE ORION CONSTELLATION

Print ISSN: 2514-863X Online ISSN: 2514-8648

# JournalPreview

London Journal of Research in Computer Science and Technology  
Volume 21 | Issue 1 | Compilation 1.0



# JournalPreview

LONDON JOURNAL OF RESEARCH IN COMPUTER SCIENCE AND TECHNOLOGY

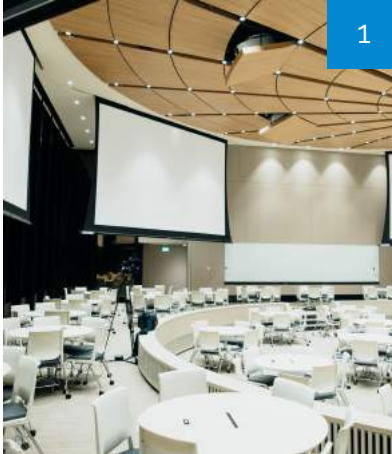
This document is a pre-published view of London Journal of Research in Computer Science and Technology Volume 21, Issue 1 and Compilation 1.0. For any minor changes and updations kindly follow your paper's live editing URL given in sent email or get in touch with our support team at [support@journalspress.com](mailto:support@journalspress.com) or visit our website to use live chat support. This is a beta document thus order, content or existence of papers may alter in the published eJournal. You are requested to kindly acknowledge and approve your research paper in this JournalPreview within three days.

# Journal Content

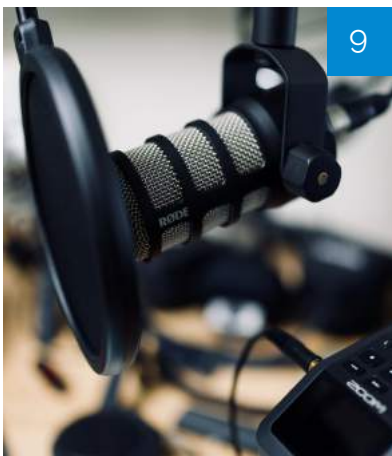
In this Issue



London  
Journals Press



- i. Journal introduction and copyrights
  - ii. Featured blogs and online content
  - iii. Journal content
  - iv. Curated Editorial Board Members
- 



1. Supreme Theory of everything: A New...  
**pg. 1-8**
  2. Novel Remote Evidence Collaborative...  
**pg. 9-18**
  3. Effect of three Modes of teaching on Students...  
**pg. 19-28**
  4. Overview of Interactive Chatbot for Modelling, Predicting...  
**pg. 29-38**
  5. Study of Certain Corona Family Related Viruses Based...  
**pg. 39-111**
- 



- V. London Journals Press Memberships



Scan to know paper details and  
author's profile

# Supreme Theory of everything: A New Possibility to Open the Hysteresis

*Ulaanbaatar Tarzad, Jargalan Narmandakh, Batgerel Baltin & Sangaa Deleg*

*Mongolian University of Science and Technology*

## ABSTRACT

The problem of mathematical description for the hysteresis loop of ferromagnetic materials has been studied for a long time but unsolved until the present. Some models as particularly the Preisach model of general characters of hysteresis and the Jiles – Atherton model of ferromagnetic phenomenology, used in a wide range. But these can describe only some part of hysteresis. The eddy current generates heat in nanomagnetic materials because of hysteresis loss. Nowadays, the necessity to study hysteresis increases by nanomagnetic properties, but it is described only by magnetic domains and their changes. In this paper, the opening of the ferromagnetic hysteresis loop is shown firstly based on the trigonometric model. Finally, its confirmation is indicated by the experimental results of the  $\text{CuFe}_2\text{O}_4$  compound.

**Keywords:** open hysteresis, projection of hysteresis, magnetization, magnetic domain, nonmagnetic material, trigonometric expression of hysteresis.

**Classification:** B.3.2

**Language:** English



London  
Journals Press

LJP Copyright ID: 975821  
Print ISSN: 2514-863X  
Online ISSN: 2514-8648

London Journal of Research in Computer Science and Technology

Volume 21 | Issue 1 | Compilation 1.0



© 2021. Ulaanbaatar Tarzad, Jargalan Narmandakh, Batgerel Baltin & Sangaa Dele. This is a research/review paper, distributed under the terms of the Creative Commons Attribution-Noncommercial 4.0 Unported License <http://creativecommons.org/licenses/by-nc/4.0/>, permitting all noncommercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

# Supreme Theory of everything: A New Possibility to Open the Hysteresis

Ulaanbaatar Tarzad<sup>α</sup>, Jargalan Narmandakh<sup>σ</sup>, Batgerel Baltin<sup>ρ</sup> & Sangaa Deleg<sup>ω</sup>

## ABSTRACT

*The problem of mathematical description for the hysteresis loop of ferromagnetic materials has been studied for a long time but unsolved until the present. Some models as particularly the Preisach model of general characters of hysteresis and the Jiles – Atherton model of ferromagnetic phenomenology, used in a wide range. But these can describe only some part of hysteresis. The eddy current generates heat in nanomagnetic materials because of hysteresis loss. Nowadays, the necessity to study hysteresis increases by nanomagnetic properties, but it is described only by magnetic domains and their changes. In this paper, the opening of the ferromagnetic hysteresis loop is shown firstly based on the trigonometric model. Finally, its confirmation is indicated by the experimental results of the  $CuFe_2O_4$  compound.*

**Keywords:** open hysteresis, projection of hysteresis, magnetization, magnetic domain, nanomagnetic material, trigonometric expression of hysteresis.

**Author α:** School of Applied Sciences, Mongolian University of Science and Technology.

**σ ω:** Institute of Physics and Technology, Mongolian Academy of Science.

**ρ:** Institute of Mathematics and Digital Technology, Mongolian Academy of Science.

## I. INTRODUCTION

Hysteresis can be found everywhere as in physics, astronomy, chemistry, engineering, biology, and economics.

Sir James Alfred Ewing studied firstly the hysteresis of magnetic materials around 1890. M. Krasnosel'skii turned his attention increasingly to

discontinuous processes and operators, in connection firstly with nonlinear control systems and then with a mathematically rigorous formulation of hysteresis which encompasses most classical models of hysteresis and is now standard [1].

One of the most common magnetic phenomena is magnetic hysteresis, which shows the relationship between the magnetization of the ferromagnetic material and the magnetic field [2-3]. Many years of theoretical research have provided detailed theories, mathematical expressions, and models to explain the physical mechanism of the phenomenon of hysteresis [4-6].

Of these, the Preisach model [7], which determines the general characteristics of hysteresis, and the Giles-Atherton model [8], a theory of ferromagnetic phenomenology, are widely used. Theoretical physicists and mathematicians have been working for many years to understand the physical mechanism of hysteresis and to create sufficient mathematical expressions to describe it. Many experimental studies of hysteresis carried out in modern techniques and technologies. For example, only the types of Hall effects include the Quantum Hall effect [9], the quantum anomaly Hall effect [10], and the spin Hall effect [11-12]. For each topic involving hysteresis, there are appropriate models, indicating the need for a unified model.

All modern magnetic materials, from the soft Fe-Ni alloy to the hard magnetic Nd-Fe-B permanent magnet used in electronic circuits, are all associated with hysteresis. It is necessary to model the hysteresis phenomenon and create the conditions with the highest efficiency by selecting the optimal parameters.

Saturation of the magnetic hysteresis loop, coercive force, and residual magnetism are the main parameters that determine the properties of ferromagnetism (Figure 1).

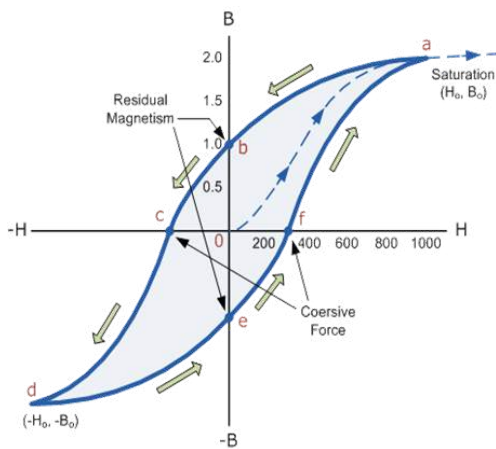


Figure 1: Ferromagnetic hysteresis loop

The Phenomenon of Electromagnetic energy into thermal energy by the action of some magnetic nanoparticles is widely studied. The amount of heat emitted depends on two types: the loss of ferromagnetic hysteresis and magnetic susceptibility of superparamagnetic nanoparticles [13].

This paper focused on opening the hysteresis of magnetic materials and on explaining the characteristics of the hysteresis loop using a mathematical model (The Supreme Theory of Everything) based on the transformation of trigonometric formulas [14-16].

## II. MATHEMATICAL MODEL OF HYSTERESIS

Hysteresis is the most complicated of all the nonlinearities presented because  $y$  is not simply a function of  $x$ , as it has been in the other cases. Rather,  $y$  is also a function of  $x'$ . Unlike the other nonlinearities, there is no single mathematical expression for hysteresis. It is typically expressed graphically. The first thing to note about the relationship between  $x$  and  $y$  is that for a given  $x$ , there may be two possible values for  $y$ . The way to interpret the relationship is as follows. Consider a specific value for  $x$  shown in Figure 2 as  $x_0$ . The question is whether  $y$  value will take on value  $y_1$  or  $y_2$ . Just knowing  $x_0$  is not enough information.

The value of  $y$  depends on where  $x$  came from. If  $x$  is increasing, then  $y = y_2$ . If  $x$  is decreasing, then  $y = y_1$ . The curve is followed in the direction given by the arrows and indicates whether to use the upper or lower part.

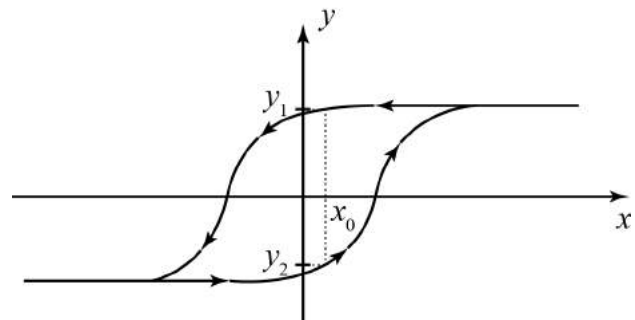
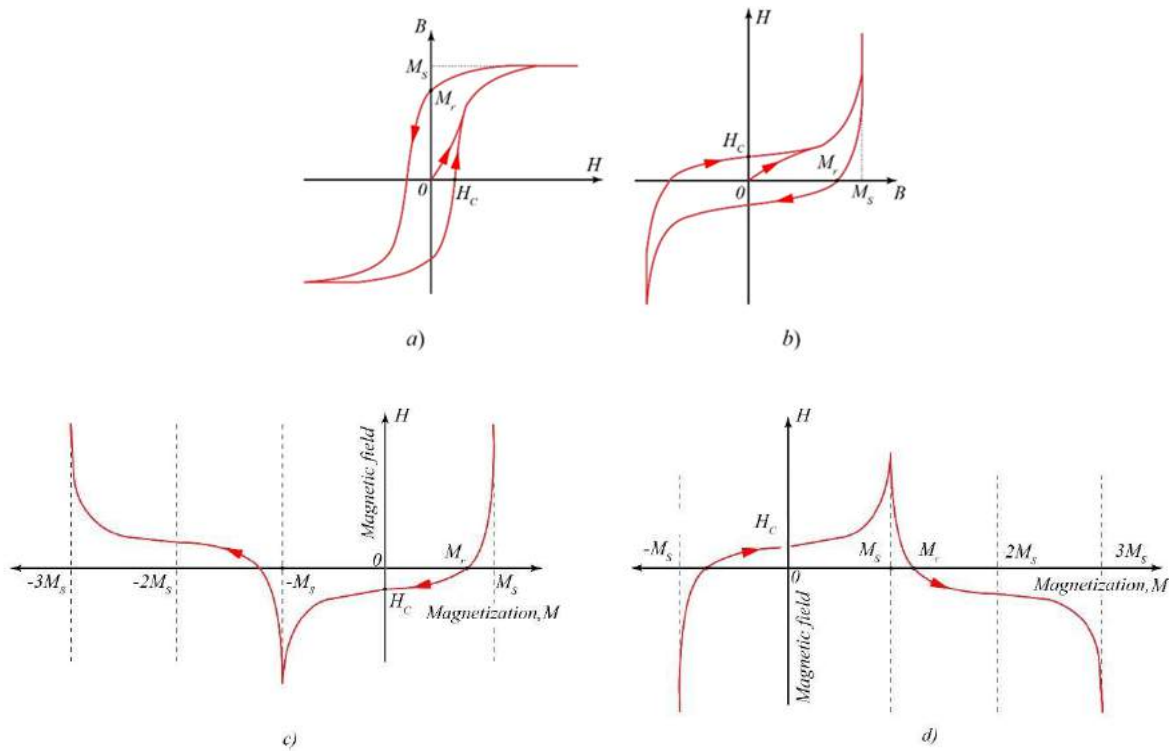


Figure 2: Demonstrating how to interpret a hysteresis curve

A system with hysteresis is one that has memory. Its output depends on where it came from". For this reason, leading scientists worked hard to express the mathematical formulations of the magnetic hysteresis, but generally experimental. [17].

The hysteresis loop has two directions: upward and downward. Since the question is whether  $y$  value will take on value  $y_1$  or  $y_2$  [17]. To avoid this problem and open the hysteresis  $y$  only must take one  $y$  value. To do this, we need to transform the function  $B(H)$  into  $H(B)$  (Figure 3b), and then it is possible to convert the left or right direction of the function by the mirror method at the saturation point ( $M_s$ ). Depending on which saturation point, it may be called left or right open hysteresis (Figure 3).



**Figure 3:** Algorithm for the opening of ferromagnetic hysteresis a) Traditional hysteresis, b) Reverse hysteresis, c) Left open hysteresis d) Right open hysteresis

For magnetic hysteresis, the magnetization  $B$  is a function of the magnetic field,  $B = B(H)$ . It is not possible to open hysteresis because the quantity  $H$ , which directly controls the process, varies in all numerical regions. However, since the dependent variable  $B$  is changed in the interval  $[-M_s, M_s]$ , it is possible to open the inverse hysteresis  $H = H(B)$ . To construct a mathematical model, let us move on to the dimensions that do not depend on (1).

$$x = \frac{\pi B}{2 M_s}, \quad x \in \left[-\frac{\pi}{2}, \frac{3\pi}{2}\right] \tag{1}$$

$$y = \frac{H}{H_c}, \quad y \in [-\infty, \infty]$$

The mathematical expression for open hysteresis is described by Formula (2) [14-16] shown in Figure 3 c.

$$y = \frac{a \sin(x-\theta)}{|\cos(x)|} \tag{2}$$

Where  $x$  is the value of the angle along the circle,  $\theta$  is the parameter that determines the phase shift or residual magnetization value, and  $a$  is the parameter that determines the amplitude or coercive force.

Thus, by replacing the two axes of the hysteresis graph, the design becomes clear and easier to use. It indicates that the x-axis seems to take on the plus and minus infinity and trigonometric transformation from a time-domain function to a frequency domain. We understand the partial frequency interval as time. The vertical axis represents the magnetic field (electric field), and the horizontal x-axis represents the magnetization (polarization) (Figure 3b). Formula (2) looks like the ratio of the upper and lower axes of an ellipse. If the amplitude of the sine function is greater than the cosine ( $a > 1$ ), the ellipse is vertically oval, and if the amplitudes are equal ( $a = 1$ ), it is circular. At  $a < 1$  it is a horizontal oval ellipse. Since Formula (2) has the form cosine function  $1/0$ , two large singulars

appear to be monotonically increasing and decreasing functions. The positive singular is  $\pi/2$ , and the negative singular is  $-\pi/2$ . One complete turn of a circle consists of two parts: the left-wing is in the interval  $[-3\pi/2, -\pi/2]$  and the right-wing is between  $[-\pi/2, \pi/2]$ .

While the magnetic field ( $H$ ) can increase infinitely, the magnetization ( $M$ ) is finite or

saturated at a certain  $M_s$ -point. When  $H$  equals 0, the residual magnetization ( $M_r$ ) is determined. The magnetic field then becomes negative and appears as infinite at  $270^\circ$ , and the magnetization reaches saturation point again at  $M_s$ .

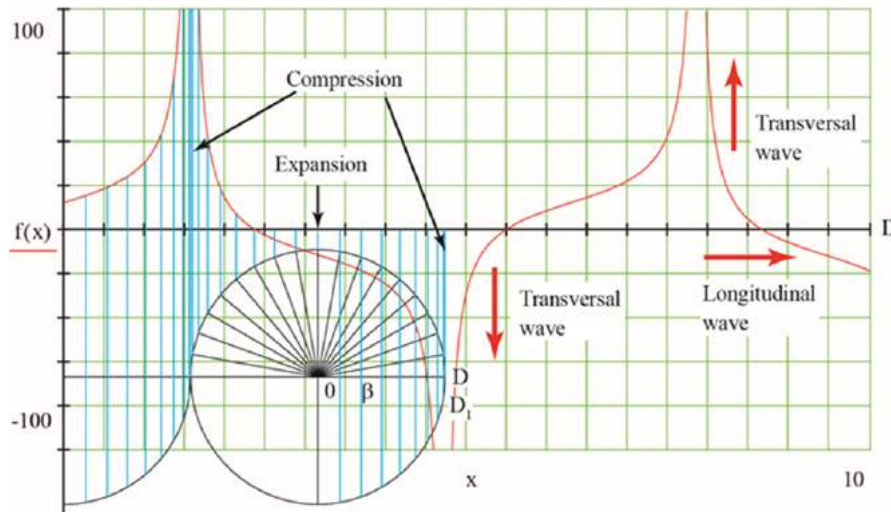


Figure 4: Open hysteresis defined by Formula (2) (For simplicity,  $a = 0.5$  and  $\theta = -\pi/4$  are calculated)

Figure 4 shows the projection on the x-axis of the y function [14] and the magnetic state of open hysteresis.

The function  $y$  (optically, the distance of light transition in the transmitting medium) is directly proportional to the amplitude (thickness of the light-transmitting medium) and is inverse

proportional to the cosine of the difference of incident and refraction angles [14]. Thus, the Supreme Theory of Everything shows no space or time, but that it is by expressions such as amplitude, frequency, and phase shift.

$$c(x) = \frac{10 \sin(x - \frac{\pi}{4})}{|\cos(x)|}; \quad c1(x) = \frac{10 \sin(x + \frac{\pi}{4})}{|\cos(x)|}; \quad c2(x) = \frac{15 \sin(x)}{|\cos(x)|} \quad (3)$$

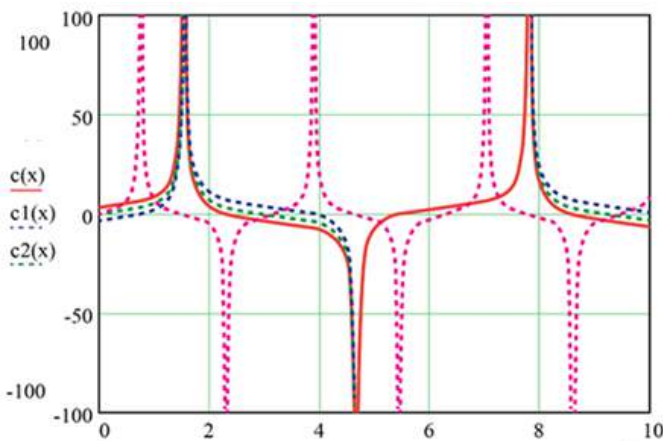


Figure 5: Behaviors of open hysteresis

Let us consider some features of open hysteresis based on equations (3):

- When the phase shift is negative ( $c(x)$ ), the right-wing is higher than the left-wing. If the phase shift is positive ( $c1(x)$ ), the left-wing of the hysteresis is higher than the right-wing.
- If the formula has no phase shift ( $c2(x)$ ), the two wings are symmetrical.

The characteristics of open hysteresis compared with the experimental results.

### III. HYSTERESIS OF MAGNETIC $\text{CuFe}_2\text{O}_4$ COMPOUNDS

The measured values of the magnetic properties of the  $\text{CuFe}_2\text{O}_4$  magnetic material cited from the work [17] (Figures 6, 7, 8).

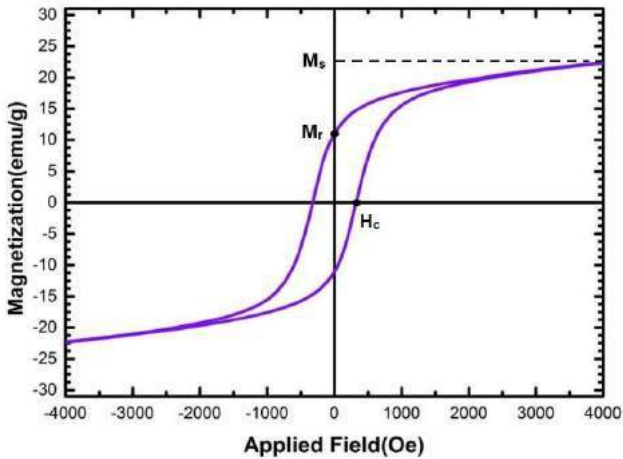


Figure 6: The hysteresis loop of  $\text{CuFe}_2\text{O}_4$  nanoparticles fired at  $850^\circ\text{C}$

The results of the hysteresis measurements are analyzed using the methods described above. To approximate, transfer the physical quantity to the normally non-dimensional dependent quantity according to Formula (2).  $M_s = 22 \text{ emu/g}$ , and  $H_c = 325 \text{ Oe}$  were obtained from the measurement results [18]. To do this, use the values of  $M_s = 22 \text{ emu/g}$  and  $H_c = 325 \text{ Oe}$  obtained from the measurements, normalize them according to Equation (2), and convert them to non-dimensional quantities.

Since the opening point is  $x_{open} = -\frac{\pi}{2}$ ,  $x_{open} = -\frac{\pi}{2}$ , the lower curve of the hysteresis loop conversion leads to an open hysteresis curve (Figure 7 and Figure 8).

$$\bar{x}_i = -\pi - x_i \quad \bar{x}_i = -\pi - x_i, \quad \bar{y}_i = y_i \quad \bar{y}_i = y_i$$

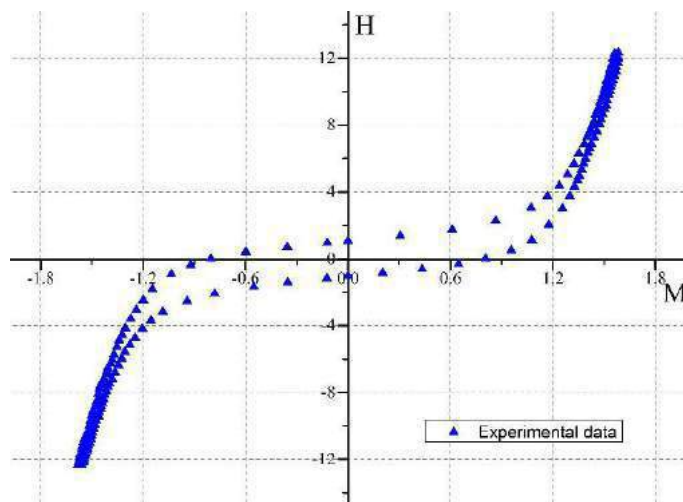


Figure 7: Reverse hysteresis of CuFe<sub>2</sub>O<sub>4</sub> nanomagnetic compounds

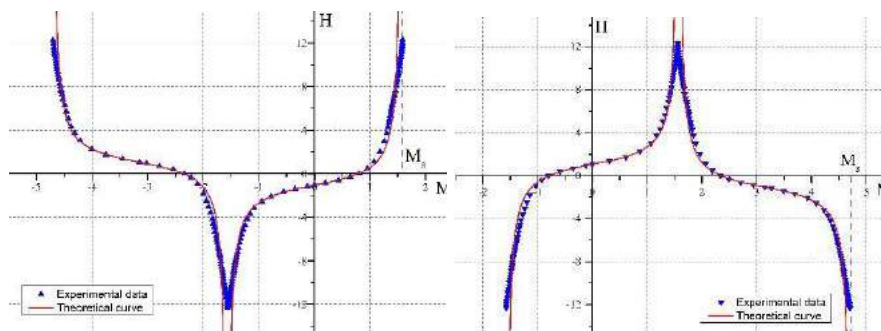


Figure 8: Fitting of the theoretical and experimental curves of the open hysteresis  
a) left open hysteresis, b) right open hysteresis

The results show that the mathematical model of The Supreme Theory of Everything can approach the magnetic hysteresis curve with high accuracy and to explain the physical process.

There are the values of the approximate parameters  $a$  and  $\theta$  and the standard errors shown in Table 1.

Table 1: Standard deviations of the theory and measurement curves

	$a$	Standard error	$\theta$	Standard error
Left open hysteresis	1.48395	0.02339	0.79116	0.01557
Right open hysteresis	1.49068	0.02091	-0.78845	0.01359

The standard approach error was found to be 1.57% for the parameter  $a$  and 1.97% for phase transition  $\theta$ .

The results show that the mathematical model of The Supreme Theory of Everything can approach the magnetic hysteresis curve with high accuracy and explain the physical process.

### III. CONCLUSION

1. The experimental values are converted to non-standardized units by Formula (2).

Subsequently, the hysteresis opens successfully after the transformation.

2. It is possible to explain the hysteresis of the magnetic material determined experimentally with the help of Formula (2). It is possible to extend the formula by adding physical parameters in the future.
3. With the help of the open hysteresis formula of the magnetic materials, it is possible to express many properties of the magnetic materials, such as ferromagnetic, Paramagnetic, ferrimagnetic, antiferromagnetic.

## REFERENCES

1. Krasnosel'skii, Mark A., Pokrovskii, Aleksei V., *Systems with Hysteresis*, 1989, Springer, Berlin.
2. G. Bertotti, *Hysteresis in Magnetism*, Academic Press, San Diego (1998).
3. F. Fiorillo, C. Appino, and M. Pasquale, *Hysteresis in Magnetic Materials, The Science of Hysteresis, Volume III, Chapter 1*, pp. 1-190, (2006).
4. Mayergoys, Isaak D. *Mathematical Models of Hysteresis and their Applications:(2003)*, Academic Press. ISBN 9780124808737.
5. A.Tena, D.Fotiadis, Ch. Massalas, *Hysteresis modeling and applications, Advances in scattering and biomedical engineering*, pp. 313-322, 2004.
6. N. Pop, *A model for magnetic hysteresis, The European Physical Journal Plus volume 134*, 2019.
7. F. Preisach, *Uber Die magnetische nachwirkung, Z. Phys. 94*, 277-302, (1935).
8. D. C. Jiles and A. L. Atherton, *Theory of ferromagnetic hysteresis, J. Magn. Magn. Mater. 61*, 48-60 (1986).
9. Yong-Chang Lau, Davide Betto, Karsten Rode, JMD Coey, Plamen Stamenov, *Spin-orbit torque switching without an external field with a ferromagnetic exchange-biased coupling layer*, arXiv: 1511.05773 [condmat. mes-hall], 2015.
10. Cui-Zu Chang, Jinsong Zhang, Xiao Feng, Jie Shen, Zuocheng Zhang, Minghua Guo, Kang Li, Yunbo Ou, Pang Wei, Li-Li Wang, Zhong-Qing Ji, Yang Feng, Shuaihua Ji, Xi Chen, Jinfeng Jia, Xi Dai, Zhong Fang, Shou-Cheng Zhang, Ke He, Yayu Wang, Li Lu, Xu-Cun Ma, Qi-Kun Xue, *Experimental Observation of the Quantum Anomalous Hall Effect in a Magnetic Topological Insulator, Science, Vol 340*, 2013.
11. Kane C. L. and E. J. Mele, (2005), *Quantum Spin Hall Effect in Graphene, Phys. Rev. Lett. 95 (22):226081*. arXiv:cond-mat/0411737. doi:10.1103/PhysRevLett.95.226801.
12. Mark Wardle (2004). "Star Formation and the Hall Effect". *Astrophysics and Space Science*. 292 (1).
13. Mahnaz Amiri, Masoud Salavati-Niasari, Ahmad Akbari, *Advances in Colloid and Interface Science*, vol. 265, pp. 29-44, 2019.
14. Ulaanbaatar Tarzad., *Formula Extraction in Supreme Theory of Everything, Advances in Theoretical and Computational Physics, Vol 2, Issue 4, 2 of 3*, 2019a, doi.org/10.33140/ATCP.02.04.01.
15. Ulaanbaatar Tarzad, *Supreme Theory of Everything, Advances in Theoretical and Computational Physics, Vol 2, Issue 2, 2-6*, 2019b, doi.org/10.33140/ATCP.02.02.05
16. Ulaanbaatar Tarzad, *Supreme Theory of Everything: Whole Universe in a Simple Formula, London Journal of Research in Science: Natural and Formal, Volume 20 | Issue 5 | Compilation 1.0*, p. 73-90.
17. Mellodge, P., *Characteristics of Nonlinear Systems, in A Practical Approach to Dynamical Systems for Engineers*, 2016, <https://www.Sciencedirect.com/topics/engineering/hysteresis>.
18. Reza Peymanfar, Farzaneh Azadi, Yousef Yassi, "Preparation and Characterization of CuFe<sub>2</sub>O<sub>4</sub> Nanoparticles by the Sol-Gel Method and Investigation of its Microwave Absorption Properties at Ku-band Frequency using Silicone Rubber", *Proceedings of The 3rd International Electronic Conference on Materials Sciences*, 2018, doi.org/10.3390/ecms2018-05218.

*This page is intentionally left blank*



Scan to know paper details and  
author's profile

# Novel Remote Evidence Collaborative Framework to Enable Efficient Making Digital Artifacts for RDF Formats on Web Resources

*Naseema Shaik & Dr. E. Sreenivasa Reddy*

*Acharya Nagarjuna*

## ABSTRACT

In real time information monitoring systems, applications are used in services to make parameters described in their file configuration, there are different codes were used to denote these parameters at various applications. When any user change their file configurations present in service are modified, it is very necessary to update them in configuration information processing systems. Because of immutability of text it is a complex task to update the modified content in web related applications. Making digital artifact is the one of the major issue to verifiable and reliable in update content. De-centralized Hash URI approach (DHURI) is one of the approach to handle efficient making of digital artifacts on web resources in nanopublications The main drawback of above approach is that the digital artifacts supported in the project context happens to use RDF formats (with html links) only and no other formats of data such as CSV, PNG are allowed since they are quite hard to verify in the current architecture. So to support multi labeled data in making digital artifacts, we propose and introduce Novel Evidence based Acquisition Collaborative Framework (NEACF), this framework support other types of artifacts in web resources. Sharing of encoded information between publisher and verifier aids in establishing remote evidence process establish-ing with respect file type. Experimental evaluation of proposed framework to minimizes the page loads by reducing start up delays using the above buffer heuristics mentioned and also supports multiple file types as publishing's besides RDF Html files.

*Index Terms:* file synchronization systems, remote evidence, digital artifacts, and centralized systems.

*Classification:* D.3.3

*Language:* English



LJP Copyright ID: 975822  
Print ISSN: 2514-863X  
Online ISSN: 2514-8648

London Journal of Research in Computer Science and Technology

Volume 21 | Issue 1 | Compilation 1.0





# Novel Remote Evidence Collaborative Framework to Enable Efficient Making Digital Artifacts for RDF Formats on Web Resources

Naseema Shaik<sup>α</sup> & Dr. E. Sreenivasa Reddy<sup>σ</sup>

## ABSTRACT

*In real time information monitoring systems, applications are used in services to make parameters described in their file configuration, there are different codes were used to denote these parameters at various applications. When any user change their file configurations present in service are modified, it is very necessary to update them in configuration information processing systems. Because of immutability of text it is a complex task to update the modified content in web related applications. Making digital artifact is the one of the major issue to verifiable and reliable in update content. De-centralized Hash URI approach (DHURI) is one of the approach to handle efficient making of digital artifacts on web resources in nanopublications The main drawback of above approach is that the digital artifacts supported in the project context happens to use RDF formats (with html links) only and no other formats of data such as CSV, PNG are allowed since they are quite hard to verify in the current architecture. So to support multi labeled data in making digital artifacts, we propose and introduce Novel Evidence based Acquisition Collaborative Framework (NEACF), this framework support other types of artifacts in web resources. Sharing of encoded information between publisher and verifier aids in establishing remote evidence process establishing with respect file type. Experimental evaluation of proposed framework to minimizes the page loads by reducing start up delays using the above buffer heuristics mentioned and also supports multiple file types as publishing's besides RDF Html files.*

*Index Terms:* File synchronization systems, remote evidence, digital artifacts, and centralized systems.

*Author α:* Research Scholar, Acharya Nagarjuna University, Guntur, AP, India.

*σ:* PhD Dean, Faculty of Engineering, Acharya Nagarjuna University, Guntur, AP, India.

## I. INTRODUCTION

Principle vision behind semantic web with substance of machine competent with various documentations, allowing, notwithstanding different things, for electronic assortment and present day request frameworks over a great deal of associated data. As even human customers are from time to time easy to trap by spam and beguiling substance that can be found on the web, we should be extensively logically stressed because of motorized counts that self-lessly explore semantic web content. Without reasonable counter-measures, vindictive on-screen characters can harm or control such estimations by including just a few carefully controlled things to colossal plans of information data. To address a part of these issues, different intelligent data vaults have appeared, for instance, Figshare and Dryad (<http://figshare.com>, [http:// datadryad.org](http://datadryad.org)). Plus, Digital Object Identifiers (DOI) has been upheld to be used for articles just as for intelligent data (Paskin, 2005). While these approaches verifiably improve the condition of intelligent data, explicitly when gotten together with Semantic Web strategies, they have by the by different hindrances: they have united models, they give us no likelihood to check whether the data have been (deliberately or coincidentally) changed, and they don't reinforce

find a workable pace from informational indexes, (for instance, particular data sections). We battle that the joined thought of existing data vaults is clashing with de-concentrated conventional with various logical relations, and that it has authentic outcomes concerning enduring quality and trust. The affiliations running these stages may in the long run fizzle be picked up by money related authorities who don't feel concentrated on the principles of science, or for various reasons become unfit to keep their destinations good to go. Regardless of the way that displayed data accomplishes, ensure relations on informational collections with available ways to deal with test that hash esteem is trustworthy or not. Particularly in cloud based web related services, file synchronization become most powerful in recent years to maintain backup of data with automatic updates across multiple devices. for automatic changes applied in updated content in different web services in today's world. So need efficient solution for digital remote evidence problem by the prevalence of web related services. While some work has been led on the recuperation of proof from document synchronization administrations, it principally centers on the social affair of records and logs put away on the neighborhood machine's hard drive. Recuperating proof from cloud-based arrangements is commonly led through program interface or customer application synchronization, requiring the validation subtleties for the administration. At the hour of composing, the creators couldn't recognize any productions concentrated on a "sometime later" recuperation of privately undermined or unrecoverable proof from a decentralized document synchronization administration.

So that in this paper, we present Novel Evidence based Acquisition Collaborative Framework (NEACF), this framework support other types of artifacts in web resources. Sharing of encoded information between publisher and verifier aids in establishing remote evidence process establishing with respect file type. We also present a methodology for remote recovery with variable and reliable of digital evidence on decentralized file synchronization related

services. This is very useful to forensic related investigations to overcome counter tracks of cybercriminals, to getting basic knowledge regarding data recovery communication over digital evidence and digital investigation for the verification of verifiable and reliable data evaluation in web related resource systems.

## II. REVIEW OF RELATED WORK

This section describes the description of file synchronization in cloud related web resource with respect to content update and modify based on evidence activities. Also describes different relations present in acquisition of remote evidence with file synchronizations on web related applications.

### 2.1 Recovery of Remote Evidence

A client server based framework for remotely assembling forensically stable circle pictures over the Internet is plot in an article by Scanlon and Kechadi [2010]. This framework depended on a live legal sciences situation whereby the presume machine is booted utilizing a Linux based live CD or USB key so as to take an evident, remote clone of any capacity gadget on the machine. The proof social occasion and check process used regular SHA512 hashing of "pieces" of the remote hard drive in blend with a hash of the remote volume completely. The creators found that the division of enormous proof models, its transmission over a scrambled Internet association with a server in a measurable research center and the consequent recombination process didn't meddle with the resultant hash esteems when thought about against those of the first volume. The errand of performing proof obtaining from cloud-based document synchronization frameworks is the most practically equivalent to the work exhibited as part of this paper, in spite of utilizing a concentrated server. Chung et al. [2012] proposed a novel procedure model for the examination of distributed storage administrations illustrating best practices for criminological agents. Research directed into the proof recuperation from cloud-based record synchronization apparatuses parts into two proof

social occasion strategies: nearby cloud proof obtaining and remote cloud proof procurement.

## 2.2 File Synchronization in Web based Services

Neighborhood cloud proof recuperation centers around recouping distributed storage leftovers through hard drive investigation. In a volume of work led by Quick [2012], the neighborhood remainders of erased documents is broke down across Dropbox, Google Drive and SkyDrive (presently OneDrive). The metadata that remained was adequate to demonstrate that a cloud based record was available on the neighborhood drive after erasure. The creators likewise demonstrated that the demonstration of downloading the information from the remote area utilizing a program or synchronizing utilizing the customer application doesn't change the hash of the document or any related cloud metadata [Quick also, Choo, 2013]. The main metadata that was diverse on the neighborhood machine when contrasted with its cloud put away partner was the document creation/change dates.

Getting to remote advanced proof put away on the servers of these cloud document synchronization specialist organizations is a

laborious errand for computerized scientific agents. In 2014, Federici portrayed an apparatus worked for the assortment of proof from the cloud called Cloud Data Imager (CDI) worked as an augmentation to the work directed on nearby cloud related remainders by Federici [2014]. CDI encourages the read-just access to remote proof put away on Dropbox, Google Drive and SkyDrive. This apparatus depends on the recuperation of the cloud administration's username and secret key or an entrance token string from the neighborhood machine for remote verification.

### III. NOVEL EVIDENCE BASED ACQUISITION COLLABORATIVE FRAMEWORK (NEACF)

This section describe the basic steps involved in collaborative framework to maintain automatic updates for synchronized data for both centralized (cloud related web services) and decentralized services (service less/cloudless). Basic steps are describes as follows to enable services related centralized synchronization with respect to processing configuration of files described in figure 1.

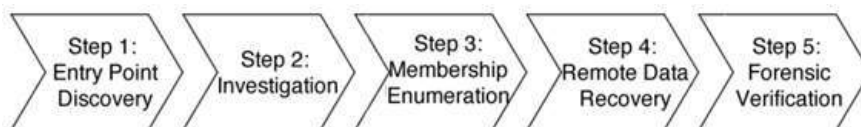


Figure 1: Basic steps to process synchronized configuration files

In identification of synchronized configuration files at web related services described in figure 1, first, identify basic entry points to enable accurate profile for data and verify its being replicated to or suspect the device from server. Identifier (Configuration Manager) must investigate which data is suspect and replicated from other system to verify either it is verifiable or reliable accessing system. Verify user enumerate with same system or other system configurations in remote. If data is updated then recover remote evidence with forensic verifiable. General structure and basic operations performed in proposed framework shown in figure 2,

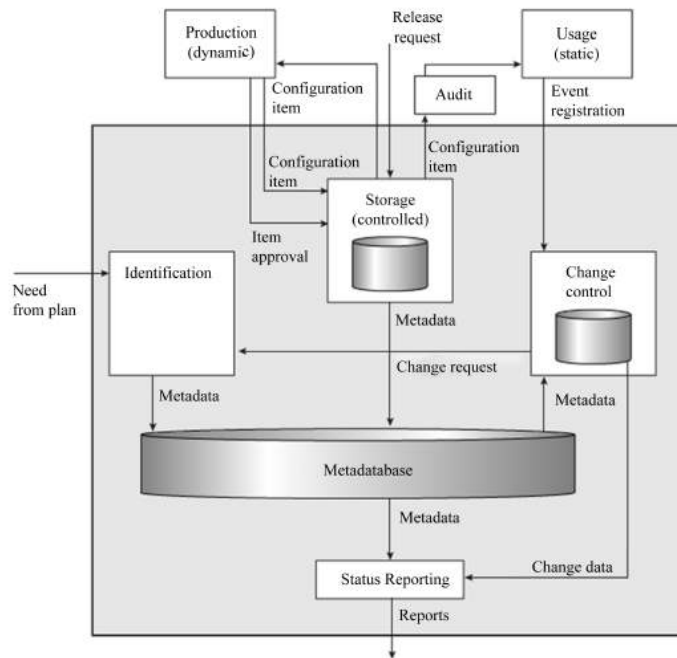


Figure 2: Operations of file synchronization procedure in proposed approach

In different types of multidisciplinary science related applications consists collaborative server (Configuration Manager) to do technical operations to its components as well as evaluate the system procedure. In computer science related applications monitoring different versions of files in web resource system, if any user can change the document or use that document at another implementation then track all the changes appeared in documents. Proposed configuration framework follows following modules in identification of synchronized files.

- a. Identification of Unique Content: This module describes the identification of configuration of metadata, check whether it is unique and specify the relations of content occurrence in throughout the world and other configuration related outside elements.
- b. Storage of Updated Data: This module describes configured data appeared or not and also describe either it is damaged then explore state of the storage of particular data.
- c. Control of Change Content: This module is used to explore all the implemented changes in computer science web related applications.
- d. Reporting Status of Data: It describe effective management of legible and useful information

in maintenance of web related data.

Analysis of web related network relational approach for the entity relationships use implemented solution as follows:

- Application (code, name, portrayal)
- Layout (code, name, filepath8, portrayal)
- Application\_template(application\_code,template\_code)
- Parameters\_group (code, group\_name, portrayal)
- Parameter (code, identifier9, parametersgroup\_code, parameter\_type, begin\_date, end\_date)
- Layout parameter (template\_code, parameter\_code)
- Esteem (code, esteem, creationhour)
- Parameters\_group\_value (parametersgroup\_code, value\_code)

At the point when a parameter p is refreshed, the accompanying activities are performed:

- a. Another worth is made for the parameters bunch containing p, and spared. The accompanying activities are reshaped for all design parameters q in a similar gathering with p.
- b. A duplicate of q's layout content is made.

- c. In that duplicate, all parameters identifiers are supplanted by their last spared qualities in the framework.
- d. The outcome at that point overwrites the layout's objective document, which is the first

arrangement record. The methodology for refreshing a setup parameter, algorithmic procedure to enable web related services as per the following:

```

updateParameter (Parameter: p, String: newValue)
/*p: parameter that needs to be updated; newValue: new value of the parameter */
1: p.getGroup().setLastValue(Value (newValue));
/* creating a new value and adding it in the list of parameter's group values */

/* Updating all parameters in the same group with p; p.getGroupedParameters() returning at least {p} */
2: for all par ∈ p.getGroupedParameters()do

    3: Template template = par.getTemplate();
/* reading the template file containing the parameter */

    4: String file = template.filepath; /*reading file effectively containing the parameter */

    5: String templateCopy = readFile (template.name); /* reading template file content */

    6: for all param ∈ template.getParameters() do

        7: templateCopy = templateCopy.replace(param.identifier, param.getGroup().getLastValue());
/* setting values of parameters in the template*/

    8: end for

    9: writeFile (file, templateCopy); /*updating the target file of the template */

10: end for

```

*Algorithm 1:* Procedure to check updated content in processing of synchronized files

Basic description of proposed approach for synchronized file configuration with respect to updated content described in Alg 1. This procedure assigns parameter to explore parameter with new updated values and also update all the parameters with configured files in centralized web service related application.

also update all the parameter present in same group with configuration respective files.

Let us consider P1, P2 are two basic applications with respective configuration files ab11.txt, ab12 with different properties ab2.xml for P2 as shown in figure 3.

## IV. IMPLEMENTATION & PERFORMANCE EVALUATION

This section describes the implementation procedure with performance evaluation for synchronized configuration systems. Let us discuss about implementation of basic example with reliable and verifiable checking of updated configuration file systems.

### 4.1 Implementation

Before going to discuss about procedure of configured synchronized system with their respective updated template files, after performing this different common related to be modified or synchronized with appropriate group data, each parameter is updated, this updated operation not directly done in configuration file,

```

CF11.txt
1 appli_name=A1
2

CF12.properties
1 application_name=A1
2

CF2.xml
1 <?xml version="1.0" encoding="UTF-8"?>
2 <persistence version="1.0" xmlns="http://java.sun.com/xml/ns/persistence" xmlns:xsi="http://www.w3.org/2003/10/20/xmlschema-1" xsi:schemaLocation="http://java.sun.com/xml/ns/persistence http://java.sun.com/xml/ns/persistence/persistence_1_0.xsd">
3 <persistence-unit name="Application.Persistant.Unit" transaction-type="RESOURCE_LOCAL">
4 <provider>org.hibernate.ejb.HibernatePersistence</provider>
5 <properties>
6 <property name="hibernate.cache.provider_class" value="org.hibernate.cache.NoCacheProvider"/>
7 </properties>
8 </persistence-unit>
9 </persistence>
10
11

```

Figure 3: Sample file systems with .txt and .xml file systems

Modified RDF documents with different notation described in figure 4.

```

CF11_template.txt
1 appli_name=A1_CF11_appli_name
2

CF12_template.properties
1 application_name=A1_CF12_application_name
2

CF2_template.xml
1 <?xml version="1.0" encoding="UTF-8"?>
2 <persistence version="1.0" xmlns="http://java.sun.com/xml/ns/persistence" xmlns:xsi="http://www.w3.org/2003/10/20/xmlschema-1" xsi:schemaLocation="http://java.sun.com/xml/ns/persistence http://java.sun.com/xml/ns/persistence/persistence_1_0.xsd">
3 <persistence-unit name="Application.Persistant.Unit" transaction-type="RESOURCE_LOCAL">
4 <provider>org.hibernate.ejb.HibernatePersistence</provider>
5 <properties>
6 <property name="hibernate.cache.provider_class" value="A2_CF2_provider_class"/>
7 </properties>
8 </persistence-unit>
9 </persistence>
10
11

```

Figure 4: Modified documents with template content with different notations

Basic implementation described as follows:  
parameters group (code, group name, portrayal)(1, P1\_application\_name, "") (2, P2\_ab2\_provider\_class, "")

- parameter (code, identifier , parametersgroup\_code, parametertype, begin\_date, end\_date) (1, P1\_ab11\_appli\_name, 1, standard, 24/02/2020, invalid) (2, P1\_ab12\_application\_name, 1, normal, 24/02/2020, invalid) (3, P2\_ab2\_provider\_class, 2, normal, 24/02/2020, invalid)
- template\_parameter (template\_code, parameter\_code) (1, 1) (2, 2) (3, 3)
- esteem (code, esteem, creationhour) (1, P1, 10/02/2015;07:30) (2, org.hibernate.cache.NoCacheProvider,24/02/2020;07:30) parameter\_sgroup\_value(parametersgroup\_code, value\_code) (1, 1) (2, 2)

#### 4.2 Simulated Results

To develop efficient architecture to explore services of file synchronization with respect to configuration between files. For that, use JAVA and NETBEANS latest versions with 4-8GB RAM and 500-1TB capacity of hard disk, based on this criteria construct dynamic user interface to process different RDF documents described in figure 5.

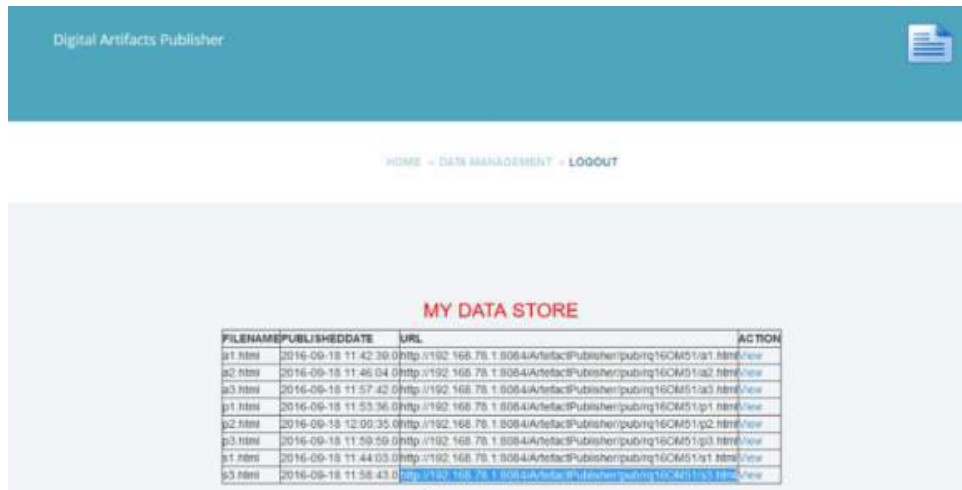


Figure 5: Description of different URLs to process different formats relates to RDF documents.

As shown in figure 5, it describes the total hash URLs for different RDF documents formats, there are different types of parameters that appear to explore web related RDF documents with updated temple results. We compare the performance evaluation of proposed frameworks with traditionally used approaches like De-centralized

Hash URI approach (DHURI) [1] with respect to processing time evaluation and resource utilization for processing synchronization file configuration in web related services with making of digital artifacts. Execution time for processing efficient web related data shown in figure 6 and table 1.

Table 1: Execution time to process web related services

No. of documents	DHURI	NEACF
10	3.42	2.91
20	5.26	3.24
30	6.35	5.92
40	6.94	6.78
50	7.64	6.91
60	8.42	7.35

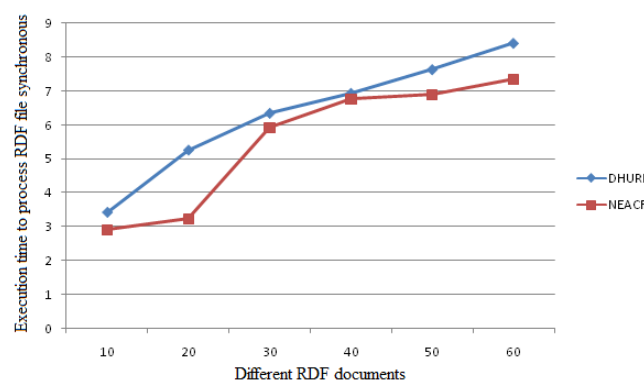


Figure 6: Performance evaluation of time with respect to RDF document processing

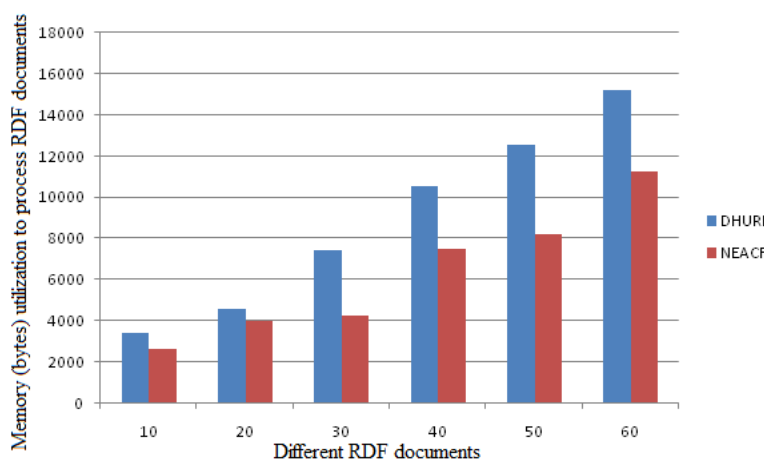
Execution time to process RDF documents with updated content at web server in making digital artifacts with respect to synchronized file con-

figurations. Resource utilization for synchronized web content with making of efficient digital artifacts in evaluation of web resource documents.

Memory utilization to process RDF documents described in figure 7.

*Table 2:* Performance evaluation of memory utilization in processing RDF documents

No. of Documents	DHURI	NEACF
10	3457	2634
20	4587	3985
30	7458	4251
40	10524	7524
50	12563	8245
60	15246	11254



*Figure 7:* Performance evaluation of memory utilization in processing documents.

As shown in above figures proposed approach give better and efficient results with respect to processing of synchronized files with making of digital artifacts in RDF documents with reliable and verifiable data recovery from updated data in web resources.

## V. CONCLUSION

This paper presents a novel collaborative framework to provide verifiable secure remote recovery of making digital artifacts from centralized web resources network. Verification of secure evidence by hash key verification with respect to protocol hierarchy describes frequent hashing to integrity of synchronized RDF document evaluation with different formats. Frequent hashing key evaluation for different updated RDF documents with different notations

described in web resource development. Experimental evaluation of proposed approach with respect to updating each document with different memory utilization and evaluation time of processing RDF documents in reliable data processing in web resource implementation. Further improvement of this research is to evaluate the extensive performance of proposed approach with different machine learning approaches in reliable web environment.

## REFERENCES

1. Naseema Shaik, E. Sreenivasa Reddy, "De-Centralized Hash URI Procedure for Making Digital Artifacts as Immutable, Reliable, Verifiable on Web", International Journal of Recent Technology and Engineering (IJRTE)

- ISSN: 2277-3878, Volume-8 Issue-3, September 2019.
2. Mark Scanlon Jason Farina Nhien An Le Khac Tahar Kechadi, "Leveraging de-centralization to extend the digital evidence acquisition window: case study on bittorrent sync", arXiv: 1409.8486v1 [cs.CR] 30 Sep 2014.
  3. Emmanuel, M.M. and Ibrahim, M.N. (2015) Automatic Synchronization of Common Parameters in Configuration Files. *Journal of Software Engineering and Applications*, 8, 192-200. <http://dx.doi.org/10.4236/jsea.2015.84020>.
  4. Ramsey, N. and Csirmaz, E. (2001) An Algebraic Approach to File Synchronization. *Proceedings of the 9th ACM International Symposium on Foundations of Software Engineering*, 175-185. <http://dx.doi.org/10.1145/503229.503233>.
  5. N. Bitouz'e, F. Sala, S. M. S. T. Yazdi, and L. Dolecek, "A practical framework for efficient file synchronization," in *Proc. 51st Annual Allerton Conf. on Comm., Control, and Computing*, pp. 1213–1220, 2013.
  6. A. Kontorovich and A. Trachtenberg, "String reconciliation with unknown edit distance," in *Proc. IEEE Int. Symp. on Inf. Theory*, pp. 2751–2755, 2012.
  7. S. M. S. Tabatabaei Yazdi and L. Dolecek, "Synchronization from deletions through interactive communication," in *IEEE Trans. Inf. Theory*, vol. 60, pp. 397–409, Jan. 2014.
  8. N. Bitouz'e and L. Dolecek, "Synchronization from insertions and deletions under a non-binary, nonuniform source," in *Proc. IEEE Int. Symp. Inf. Theory*, 2013.
  9. N. Ma, K. Ramchandran, and D. Tse, "A compression algorithm using mis-aligned side-information," in *Proc. IEEE Int. Symp. Inf. Theory*, 2012.
  10. Mark Scanlon, Jason Farina, and M-Tahar Kechadi. Bittorrent sync: Network investigation methodology. In *Proceedings of Ninth International Conference on Availability, Reliability and Security (ARES 2014)*, Fribourg, Switzerland, September 2014. IEEE.
  11. Darren Quick. Forensic Analysis of Cloud Storage Client Data. Master's thesis, University of South Australia, Adelaide, Australia, 2012.
  12. Darren Quick and Kim-Kwang Raymond Choo. Forensic collection of cloud storage data: Does the act of collection result in changes to the data or its metadata? *Digital Investigation*, 10(3) :266 – 277, 2013. ISSN 1742-2876. doi:<http://dx.doi.org/10.1016/j.dii.2013.07.001>. <http://www.sciencedirect.com/science/article/pii/S1742287613000741>.
  13. Luciana Duranti, Weimei Pan, Joy Rowe, and Georgia Barlaoura. *Records in the Cloud (RiC)*. University of British Columbia, 2013.
  14. Jason Farina, Mark Scanlon, and M-Tahar Kechadi. "BitTorrent Sync: First Impressions and Digital Forensic Implications". *Digital Investigation*, 11(S1):S77 – S86, 2014. ISSN 1742-2876. doi: <http://dx.doi.org/10.1016/j.dii.2014.03.010>. <http://www.sciencedirect.com/science/article/pii/S1742287614000152>.
  15. Corrado Federici. Cloud data imager: A unified answer to remote acquisition of cloud storage areas. *Digital Investigation*, 11(1):30 – 42, 2014. ISSN 1742-2876. doi:<http://dx.doi.org/10.1016/j.dii.2014.02.002>. URL. <http://www.sciencedirect.com/science/article/pii/S174228761400005X>.
  16. George Grispos, William Bradley Glisson, and Tim Storer. Using Smartphones as a Proxy for Forensic Evidence Contained in Cloud Storage Services. 2013 46th Hawaii International Conference on System Sciences, 0: 4910–4919, 2013. ISSN 1530-1605. doi: <http://doi.ieee-computersociety.org/10.1109/HICSS.2013.592>.

*This page is intentionally left blank*



Scan to know paper details and  
author's profile

# Effect of three Modes of Teaching on Students' Learning Retention in Electronics Works in Technical Colleges in Rivers State

*Reagan Nnabio Robinson*

*Ignatius Ajuru University of Education*

## ABSTRACT

The understanding and learning retention of electronics works being one of the major subjects offered in technical colleges is a key confirmation on the general performance of electronic subjects in technical colleges. This study was set out to determine the effect of three modes of teaching on students' learning retention in electronics works in technical colleges in Rivers State. The study also examined the influence of location on students' learning retention in electronics works. Two research questions guided the study, while two hypotheses were tested at 0.05 level of significance. The design of the study was pre-test, post-test non-equivalent control group quasi- experimental study. A sample size of 174 students drawn from a population of 267 students in Rivers State technical colleges was used. To obtain the sample, two technical colleges were purposively selected. Three intact classes were randomly selected and assigned to experimental group 1, experiment group 2 and control group.

*Keywords:* modes, teaching, students, learning, retention, electronics, works, technical and colleges.

*Classification:* J.2

*Language:* English



LJP Copyright ID: 975823  
Print ISSN: 2514-863X  
Online ISSN: 2514-8648

London Journal of Research in Computer Science and Technology

Volume 21 | Issue 1 | Compilation 1.0



© 2021 Reagan Nnabio Robinson. This is a research/review paper, distributed under the terms of the Creative Commons Attribution-Noncommercial 4.0 Unported License (<http://creativecommons.org/licenses/by-nc/4.0/>), permitting all noncommercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

# Effect of three Modes of Teaching on Students' Learning Retention in Electronics Works in Technical Colleges in Rivers State

Reagan Nnabio Robinson

## ABSTRACT

*The understanding and learning retention of electronics works being one of the major subjects offered in technical colleges is a key confirmation on the general performance of electronic subjects in technical colleges. This study was set out to determine the effect of three modes of teaching on students' learning retention in electronics works in technical colleges in Rivers State. The study also examined the influence of location on students' learning retention in electronics works. Two research questions guided the study, while two hypotheses were tested at 0.05 level of significance. The design of the study was pre-test, post-test non-equivalent control group quasi-experimental study. A sample size of 174 students drawn from a population of 267 students in Rivers State technical colleges was used. To obtain the sample, two technical colleges were selected using purposive sampling technique. Three intact classes were randomly selected and assigned to experimental group 1, experiment group 2 and control group. The Electronics Works Retention Test (EWRT) of 30 question items was the instrument used for the study. The instrument was validated by three experts in technology and vocational education. A reliability coefficient of 0.67 was obtained using Kuder-Richardson 20 and Pearson's Product Moment Correlation methods. Data collected were analysed using mean with standard deviation to answer the research questions while analysis of covariance (ANCOVA) was used to test the hypothesis at 0.05 level of significance. It was found that students taught with improvised instructional resources mode of teaching retained better in electronics works.*

**Keywords:** modes, teaching, students, learning, retention, electronics, works, technical and colleges.

**Author:** Department of Technical Education Ignatius Ajuru University of Education Rumuolumeni, Rivers State.

## I. INTRODUCTION

The academic and vocational preparation of students for job placement that involves practical and applied skills in the areas of science and technology is one of the key objectives of technical education programme. It emphasizes the understanding and practical application of basic principles of science and mathematics, rather than the attainment of proficiency in manual skills that is properly the concern of vocational education. Technical education programmes are provided for learners to acquire knowledge and practical skills for effective nation building. Technical education according to Nwachukwu (2006) is an education that is specifically designed to provide an individual with the acquisition of practical and applied skills as well as basic scientific knowledge in order to get the individual adequately equipped for self-reliance or employment in the industries. Technical education programmes are offered in technical colleges and other related vocational institutions at post primary level.

Technical colleges according to Okoro (2006) are vocational training institutions in Nigeria that admit junior secondary school leavers to senior secondary schools and provide them with full vocational courses for three years duration. Technical colleges are regarded as the principal vocational institutions in Nigeria that provide

craftsmanship training. They give full vocational training to prepare students for entry into specific engineering trades by offering several subjects that will enable the students' to gain employment after graduation, and one of such subjects is electronics works.

Electronics works in technical colleges involves the repairs, maintenance and construction of basic electronics systems. It enables students in technical colleges to learn basic electronics theory that are needed to understand circuit designs in order to install, operate, maintain and repair electronic systems (Medugu, 2011). According to Robinson (2012), electronics works as a subject also deals with the study of the properties and behaviour of electrons under different conditions, especially with reference to technical and industrial applications. This is because devices which are used in electronics systems, control the flow of electrons which came from atoms. It has electronics devices and circuits as one of its branch subjects that involves the design and interconnection of electronic components with conductive wires or traces through which electrons can flow.

Electronics works is a skill oriented subject that requires adequate presence and use of instructional resources during teaching and learning processes (Robinson, 2012). Aina (2000) opined that the adequate use of instructional resources in a technical college is the heart of craftsmanship training. This invariably means that instructional resources are very important in technical college training. The lack of adequate use will hinder the acquisition of sufficient practical skills which are the core objectives of the programme. Instructional resources are usually arranged to give occupational direction so that acceptable work habits and procedures are successfully executed.

Moreover, Okujagu (1992) defined instructional resources as those educational apparatus, that are curriculum oriented, audio-visuals, teaching and learning materials and basic tools the teacher uses to assist learners in their learning process. It ensures that the learners see, hear, feel, recognize and appreciate learning, utilizing the five senses

modalities at the same time. At all levels of the nation's educational system and for all known and existing school types, instructional resources are indispensable factors in the attainment of set goals and technical colleges is not an exception. It is known that no technical college programme can be functional without instructional resources being utilised during the classroom experience in order to inculcate technology skills for maximum nation's development. According to Akaniwor (2005), instructional resources can be real or improvised.

Real instructional resources are those resources that are ready made; produced by manufacturers for a specific task. In electronics works, real instructional resources are purchased from sellers or gotten from technicians' dumps. However, due to consistent inadequacies in the provision of real instructional resources in the teaching of electronics works in technical colleges; it becomes imperative for technical teachers to think of how best to make use of their creative skills to improvise in order to achieve their lesson objectives.

Improvisation refers to the technique of using materials obtainable from the local environment to produce a product that will serve the purpose of the real one in its absence (Olagunju, 1998). Thus in electronics works; improvisation refers to the use of materials available in the local environment to produce basic electronics systems in the absence of the real ones. This invariably means that when the real instructional resources are not available, improvisation takes their place by producing similar instructional resources that will serve the same purpose. This is to enhance the teaching and learning processes in technical colleges by ensuring that the persistent problem of inadequacy of instructional resources will not seriously affect students' learning retention.

In the same vein, the real and the improvised are all instructional resources and any can be used in the class for teaching learning process but depending on the subject matter, topic and choice of selection by the teacher (Umunadi, 2009). Umunadi however suggests that using the real instructional resources is more profitable in the

class because it has the advantage of being manufactured properly and correctly to suit a teaching and learning process. Udosen and Ekukinam (2013) are of the opinion that improvised instructional resources can be more profitable in the class since it uses things that are familiar to the students to produce materials that can enhance teaching and learning processes. They emphasized that if the students are involved in the improvisation and/or the improvisation of an instructional resource will be done before the students during teaching periods, better impartation can be obtained with respect to students' learning retention. In electronics works, the appropriate selection of instructional resources lies on the shoulder of the teacher (Umah and Maaji, 2010). It is the teacher's responsibility to choose either real or improvised instructional resources that will best suit a teaching learning process.

In another development, Towe (2000) and Umar (2002) in their separate studies pointed out that, despite the inevitable need for the use of instructional resources for teaching in technical colleges, the teachers whose responsibility are seriously lacking in the use of neither the real nor the improvised instructional resources for teaching in technical colleges could lead to very poor students' learning retention. These according to Umar and Maaji (2010), have become a persistent problem for some years in the technical colleges. Apart from the problem of students' learning retention in relation to real and improvised resources, the issue of location of a school could have a significant effect on students' learning retention.

A school could be located in the urban or rural area. Odo (1999) stated that the location of a school (urban or rural) affects a child's ability to study and perform at the level expected of him. Jahun and Momor (2001) noted that different aspects of the school environment influence students' learning retention. They stress that the individual students' learning retention behaviour is influenced not only by the motivating forces of his home, scholastic ability and values but also by the social pressure applied by the participant in the school setting. In Nigeria, rural based schools

commonly lack good infrastructures which could serve as an inhibiting factor of a good learning retention. However, Umunadi (2009) stated that school location among other factors is a necessary prediction in electronics works learning retention. Similarly, Odo (1999) found school location to be a significant factor on student's difficulties in mathematics. He emphasises that students in rural areas have less difficulties in mathematics than their urban counterparts due to less distraction. However, Jahun (1999) reported that urban students performed better than their rural counterparts in Ahmadu Bello University Mathematics Test (ABUMT). Therefore the question is, whether students in urban and rural locations have different learning retention?

In education, learning retention can be referred to as having information stored in the memory in such a way that it can readily be retrieved in response to standard prompts (Karpicke and Roediger, 2007). It involves the ability to retain facts and figures in memory after teaching. Students must retain information taught in classes in order to show the benefit of learning. The teachers' job is not over according to Lieb (1991) until he has assisted the learner in retaining the information. For students to retain information taught, they should be able to interpret and apply the taught information to new situations (Abbamondi, 2014). Therefore Abbamondi concluded that students' scores on delayed tests was a good measurement of learning retention.

This problem of poor students' learning retention is seen in their performances in various examinations, particularly the National Business and Technical Education Board (NABTEB) examination. Federal Ministry of Education (FME, 2010) reported that there was a decline in students' performance in electronics works. It stresses that students' performance in electronics works in technical colleges have been dwindling in recent time and the situation calls for immediate attention in the institutions. FME (2010) stated that technical colleges are expected to produce craftsmen who are highly needed in the industries. According to Aworanti (2011), for some past years till date, electronics works had

recorded a high failure rate of over 50 per cent in National Business and Technical Examination Board (NABTEB). The situation calls for urgent attention in order to avoid a total depletion of students' enrolment in electronics craft trade in technical colleges.

According to Akaninwor (2005), the use of inappropriate modes of teaching is one of the reasons for students' poor learning retention in technical colleges. Mode of teaching in this context refers to the manner or form of imparting knowledge or skill to learners which includes; conventional instructional resource mode of teaching, improvised instructional resource mode of teaching and real instructional resource mode of teaching. However, Akinfolarin, Ajayi and Oloruntegbe (2012) observed that technical teachers' give more preference to conventional instructional resource modes of teaching over other modes. Hence, it is the view of the researchers that technical teachers' preference for conventional instructional resource mode of teaching over result oriented ones like real instructional resources mode and improvised instructional resources mode in teaching might be a major reason for students' poor learning retention in electronics works.

Conventional instructional resource mode in this context refers to the use of chalk boards for teaching learning processes. It is without the use of real instructional resources or improvised instructional resources (Akaninwor, 2005). It is a talk and chalkboard method of teaching. On the other hand, real instructional resource mode refers to the use of real instructional resources in a teaching and learning process while improvised instructional resource mode refers to the use of improvised instructional resources in a teaching and learning process.

The conventional instructional resource mode of teaching is employed by many electronics works teachers because it enables the teachers to cover a larger content area at a time and the students are given the same content at a time. Besides, this mode of teaching fails to encourage manipulative skill and creative thinking in the learner, hence leading to poor learning retention of the students.

It is regrettable that many electronics works teachers prefer the use of the conventional instructional resource mode which is devoid of the use of neither real instructional resources nor improvised instructional resources. This invariably means, electronics teachers are no longer recognising the potency of using real instructional resources nor improvised instructional resources in the teaching of electronics works in technical colleges. Hence, the need arises to determine the effect of these three modes of teaching on students' learning retention in electronics works in technical colleges in Rivers State.

## II. PURPOSE OF THE STUDY

The purpose of the study was to determine the effect of teaching electronics works with real instructional resources, improvised instructional resources and conventional instructional resources on students' learning retention in technical colleges in Rivers State. Specifically the study sought to;

1. Determine the effect of using real instructional resources, improvised instructional resources and conventional instructional resources modes of teaching on students' learning retention when taught devices and circuits in electronics works in technical colleges in Rivers State.
2. Determine the effect of using real instructional resources, improvised instructional resources and conventional instructional resources modes of teaching on urban and rural students' learning retention when taught devices and circuits in electronics works in technical colleges in Rivers State.

## III. RESEARCH QUESTIONS

The following research questions guided the study:

1. What are the mean learning retention scores of electronics works students when taught devices and circuits using real instructional resources, improvised instructional resources as compared with those taught on the same

- topic using conventional instructional resources in technical colleges in Rivers State?
2. What are the mean learning retention scores of urban and rural electronics works students when taught devices and circuits using real instructional resources, improvised instructional resources as compared with those taught on the same topic using conventional instructional resources in technical colleges in Rivers State?

#### IV. HYPOTHESES

The following null hypotheses (Ho) were tested at, 0.05 level of significance;

1. There is no significant difference in the mean learning retention scores of electronics works students when taught devices and circuits using real instructional resources, improvised instructional resources as compared with those taught on the same topic using conventional instructional resources in technical colleges in Rivers State.
2. There is no significant difference in the mean learning retention scores of urban and rural electronics works students when taught devices and circuits using real instructional resources, improvised instructional resources as compared with those taught on the same topic using conventional instructional resources in technical colleges in Rivers State.

#### V. METHODOLOGY

This study adopted a quasi-experimental research design. Specifically, the design is a pretest posttest non-equivalent control group quasi experimental group design. The population for the study consisted of 267 students. There are three technical colleges offering Electronics Works in Rivers State which are; Government Technical College Port Harcourt, Federal Science Technical College Ahoada and Government Technical College Tombia. The population for each college was as follows: Government Technical College Port Harcourt, 99 students; Federal Science Technical College Ahoada, 93 students and Government Technical College Tombia, 75 students. The sample for the study consisted of 174 students of Electronics Works, which represents 65% of the population. The sample was

obtained using purposive sampling technique to select two technical colleges for the study considering location as a factor. There are three intact classes; the researcher randomly assigned these intact classes to experimental group 1 (E1), experimental group 2 (E2) and control group (C) using balloting method. The experimental group 1 had a sample size of 62 students, experimental group 2 had 52 students and the control group had 60 students.

The instrument used for data collection was the Electronics Works Retention Test (EWRT) which was constructed and developed by the researcher. It is a well-structured test from the content of Electronics Works in the NABTEB syllabus with five items for each selected subject topic. The instrument contained a total of 30 items with four-point multiple choice responses. The development of the instrument was based on a Table of Specification.

The instrument was subjected to face and content validation by three experts from technology and vocational education. Their observations were used to improve the instrument in content, grammar, spellings and language. The reliability indices of the instrument was determined by subjecting EWRT to internal consistency and stability using Kuder-Richardson 20 formula and Pearson's Product Moment correlation methods respectively. The average reliability coefficient for both was 0.67, which was considered adequate to be used for the study. The experiment took duration of six weeks

A pretest was given to the students in each group a day before the teaching began using the Electronics Works Retention Test (EWRT). A posttest was also administered in the sixth week during the revision of all taught topics using the same Electronics Works Retention Test (EWRT). The Research questions were answered using mean with standard deviation, while the Hypotheses were tested at 0.05 level of significance using one-way and two-way Analysis of Covariance (ANCOVA).

## VI. PRESENTATION OF RESULTS

The analysis of data in relation to each of the research questions and hypotheses are presented as follows;

### Research Question 1

What are the mean learning retention scores of electronics works students when taught devices

and circuits using real instructional resources, improvised instructional resources as compared with those taught on the same topic using conventional instructional resources in technical colleges in Rivers State?

*Table 1:* Mean learning retention scores with standard deviations due to modes of teaching

Modes	Pretest		Posttest		N
	Mean	SD	Mean	SD	
Real	2.53	0.98	19.95	2.98	62
Improvised	1.92	0.69	23.63	3.06	52
Conventional	2.40	0.81	13.94	2.76	60

Table 1 shows the mean learning retention scores of the modes of teaching. From table 1, the mean retention score of real instructional resources (experimental group 1) was 19.95 with a standard deviation of 2.98. The mean retention score of improvised instructional resources (experimental group 2) was 23.63 with a standard deviation of 3.06. While the mean retention score of conventional instructional resources (control group) was 13.94 with a standard deviation of 2.76. The mean learning retention score of improvised instructional resources was higher, followed by real instructional resources and then

conventional instructional resources followed. Students taught with real instructional resources. Students taught with conventional instructional resources retained least.

### Research Question 2

What are the mean learning retention scores of urban and rural electronics works students when taught devices and circuits using real instructional resources, improvised instructional resources as compared with those taught on the same topic using conventional instructional resources in technical colleges in Rivers State?

*Table 2:* Mean learning retention scores with standard deviations due to modes of teaching on location

Modes	Urban		Rural	
	Mean	SD	Mean	SD
Real	19.68	4.11	19.86	3.89
Improvised	23.59	3.86	23.32	4.21
Conventional	14.12	5.06	13.89	4.95

Table 2 shows the mean learning retention scores of the modes of teaching on location. From table 2, the mean retention score of real instructional resources (experimental group 1) for urban was 19.68 with a standard deviation of 4.11 while the rural mean retention score of real instructional resources (experimental group 1) was 19.86 with a standard deviation of 3.89. The mean retention score of improvised instructional resources (experimental group 2) for urban was 23.59 with

a standard deviation of 3.86 while the rural mean retention score of improvised instructional resources (experimental group 2) was 23.32 with a standard deviation of 4.21. The mean retention score of conventional instructional resources (control group) for urban was 14.12 with a standard deviation of 5.06 while the rural mean retention score of conventional instructional resources (control group) was 13.89 with a standard deviation of 4.95.

The mean learning retention score of urban students was highest when taught with improvised instructional resources, followed by real instructional resources and then conventional instructional resources. Similarly, the mean retention score of rural students was highest when taught with improvised instructional resources, followed by real instructional resources and then conventional instructional resources. This implies that the urban students retained best when taught with improvised instructional resources followed by those taught with real instructional resources. The rural students retained best when taught with

improvised instructional resources followed by those taught with real instructional resources. The urban and rural students retained least in conventional instructional resources.

*Hypothesis I*

There is no significant difference in the mean learning retention scores of electronics works students when taught devices and circuits using real instructional resources, improvised instructional resources as compared with those taught on the same topic using conventional instructional resources in technical colleges in Rivers State.

*Table 3:* One-way ANCOVA on students' learning retention due to modes of teaching

Source	Type III sum of squares	df	Mean Square	F	Sig.	Dec.
Corrected Model	22.488	6	3.748	6.578	0.000	
Intercept	2852.386	1	2852.386	5006.000	0.000	
Modes	4.800	2	2.400	4.211	0.041	S
Location	0.164	1	0.164	0.288	0.661	NS
Modes and Location	0.564	2	0.282	0.495	0.591	NS
Error	96.330	172	0.570			
Total	4222.000	173				
Corrected Total	220.952					

Table 3 shows students learning retention with respect to modes of teaching. For modes of teaching, the SPSS computed F – value of 4.211 was found significant at 0.041 level of significance, which is less than the 0.05 level of significance set for the research. Therefore, mode of teaching is significant. The null hypothesis 1 is rejected. This means that there is a significant difference in the mean learning retention scores

of electronics works students when taught devices and circuits using real instructional resources, improvised instructional resources and those taught on the same topic using conventional instructional resources in technical colleges in Rivers State?

*Table 4:* Two-way ANCOVA on students' learning retention due to modes of teaching on location

Source	Type III sum of squares	df	Mean Square	F	Sig.	Dec.
Corrected Model	22.488	6	3.748	6.578	0.000	
Intercept	2852.386	1	2852.386	5006.000	0.000	
Modes	4.800	2	2.400	4.211	0.041	S
Location	0.164	1	0.164	0.288	0.661	NS
Modes and Location	0.564	2	0.282	0.495	0.591	NS
Error	96.330	172	0.570			
Total	4222.000	173				
Corrected Total	220.952					

Table 4 shows students learning retention with respect to location. For location, the SPSS computed F – value of 0.288 was found significant at 0.661 level of significance, which is higher than the 0.05 level of significance set for the research. Hence, location is not significant. The null hypothesis 2 is not rejected. This means that there is no significant difference in the mean retention scores of urban and rural electronics works students. On interaction for hypothesis two (modes and location) , the SPSS computed F – value of 0.495 was found significant at 0.591 level of significance which is higher than 0.05 level of significance set for this research. Hence, the null hypothesis 2 is not rejected. This means, there is no significant difference in the mean retention scores of urban and rural electronics works students when taught devices and circuits using real instructional resources, improvised instructional resources and those taught on the same topic using conventional instructional resources in technical colleges in Rivers State. It implies that location was not found significant with respect to retention.

## VII. DISCUSSION OF FINDINGS

The study revealed in Research Question 1 that students taught with improvised instructional resources (experimental group 2) had a higher mean retention score in electronics works compared to those taught with real instructional resources (experimental group 1) and conventional instructional resources (control group). Hypothesis 1 further confirmed a high retention rate by indicating that mode of teaching was a significant factor in the retention of students in the electronics works content. This means that the students who were taught using improvised instructional resources have the better ability to retain after a long while. The reason for the better retention with improvised instructional resources may have been as a result of students' ability to link new concepts to improvised instructional resources that are related to their everyday life. Thus, the result of the study revealed that the adoption of improvised instructional resources that are related to students' everyday life enhances learning

retention in electronics works. This study is supported by Karpicke and Roediger (2007) view that students develop greater ability of learning retention of technical education subjects when they are involved with the teacher in the preparation and use of improvised instructional resources. This is at variance with Abbamondi (2014) view, who believed that in the teaching of technology education, improvised instructional resources cannot be of much help at the initial stage, rather enough use of drawings will bring about better retention.

The study also revealed in Research Question 2 that urban and rural students taught with improvised instructional resources (experimental group 2) had a higher mean retention score than their counterparts taught with other modes of teaching. Moreover, Hypothesis 2 reveals that location has no significant effect on students' learning retention in electronics works. This agrees with Jahun and Momor (2001) findings that there is no significant difference in the mean retention score of urban and rural in the technology. The results in this study however disagree with Umunadi (2009) that urban students show significantly greater retention in mathematics than the rural students. On the average, it implies that retention does not depend on location rather it depends on the mode of teaching in electronics works.

## VIII. CONCLUSION

The results of the study provided the empirical evidence that improvised instructional resources mode of teaching enhanced students' learning retention better than real instructional resources mode of teaching and control resources mode of teaching. This implies that, for effective teaching that will bring about a better learning retention of students in electronics works, improvised instructional resources mode of teaching should be used. This finding is irrespective of the location of the technical college; rather it is depended on the mode of teaching.

## IX. RECOMMENDATIONS

The following recommendations were made based on the findings of the study:

1. Since the use of improvised instructional resources mode of teaching has been found to be effective in better learning retention of students in electronics works content; teachers in electronics works should accept the intricacies and develop a better attitude of using it often.
2. Workshops/seminars should be regular for electronics works teachers in order to enable them to use improvised instructional resources mode of teaching excellently in the technical college.
3. Since improvisation enhances learning retention in electronics works, State and Federal governments should legally enforce its mode of teaching in technical colleges.

## REFERENCES

1. Abbamondi, D.(2014). Elements of Learning. Retrieved March 27, 2008. <http://www.etc.net/tech/adult learning/elements.htm>.
2. Aina, O. (2000). Technical and Vocational Education in Nigeria: Vision and action; Blueprint and master plan – Federal Ministry of Education (2001 – 2010).
3. Akaniwor, G. I. K. (2005). Industrial Education and Technology in Nigeria. Development and Current Trends (3<sup>rd</sup> ed.). Port Harcourt: Wilson Publishing Company Ltd.
4. Akinfolarin, C. A., Ajayi, I. A., & Oloruntegbe, K. O. (2012). An appraisal of resource utilization in vocational and technical education in selected colleges of education in southwest Nigeria. *Journal of Education*, 2(1), 41-45.
5. Aworanti, O. A. (2011). Why Candidates Fail in Public Examinations. Being a paper presented at the federal ministry of education national stakeholders' consultative meeting on improving performance in public examinations, at the National Universities Commission (NUC), Abuja.
6. Federal Ministry of Education (2010). The national master plan for Technical and Vocational Education (TVE) development in Nigerian, being a 21<sup>st</sup> century presentation at FME Abuja.
7. Jahun, J. U. (1999). Farlex depart collection, London: Harrap Press.
8. Jahun, J.U & Momoh, J.S (2001). The effects of environment and sex on the mathematics achievement of junior secondary school. Three students in Kwara State. *Journal of the Mathematical Association of Nigeria*,8 (1) 53-58.
9. Karpicke, C. & Roediger, E. (2007). Repeated retrieval during learning is the key to long-term learning. *Journal of Memory and Language*, 57, 151–162.
10. Lieb, S. (1991). Principles of adult Learning. Retrieved on march 27, 2008 from <http://hone:///u.hawal.edu.Initranet/Committe-/fac.Dev.Com/guideuk/taec-htipladunits2.thm>.
11. Medugu, J. D. (2011). Effectiveness of a Digital Oscilloscope for the Teaching of Some Radio, Television and Electronics Work Concept at Technical Colleges Level. Unpublished PhD Thesis, Abubakar Tafawa Belewa University, Bauchi, Nigeria.
12. Nwachukwu, C. E. (2006). Designing appropriate methodology in vocational and technical education for Nigeria (1<sup>st</sup> ed.). Nsukka: University Trust Publishers.
13. Odo, T.O.(1999). Gender and school location as factors for students' difficulty in secondary school geometry. Unpublished M.Ed Thesis. Department of Vocational Teacher Education, University of Nigeria, Nsukka.
14. Okoro, O. M. (2006). Principles and methods in vocational and technical education (ed.). Nsukka: University Trust Publishers.
15. Okujagu, T. (1992). Context and Content of Teacher Education in the Twenty-First Century Nigeria. Perspectives on Teacher Education in Nigeria. Association for Promotion of Quality Education in Nigeria
16. Olagunju, Y. A. (1998). Entrepreneurship: Small scale business enterprises development in Nigeria (2<sup>nd</sup> ed.). Ibadan: Ibadan University Press Plc.
17. Robinson, R. N. (2012). Technology education: The means to the realization of Nigeria's vision 2020. *Global Journal of Educational Research*, 11(1), 7-13.

18. Towe, P. E. O. (2000). An In-depth review and assessment of the present state and focus of Technical and Vocational Education in Nigeria. A paper presented at National Seminar on TVE in Nigeria; Vision and Action 31<sup>st</sup> October – 2<sup>nd</sup> November at National Center for Women Development, Abuja.
19. Udosen, I. N., & Ekukinam, T. U. (2013). Improvisation of Technological Instructional Media and Students' Performance In Primary Science In Nigerian Schools. Retrieved in [www.schoolar.lib.vt.edu/ejournal](http://www.schoolar.lib.vt.edu/ejournal).
20. Umah, I. Y., & Maaji, A. S. (2010). Repositioning the facilities in technical college workshops for efficiency: A case study of north central Nigeria. *Journal of sTEem Teacher Education*, 47(3), 40-52.
21. Umar, Z. (2002). The use of resource materials in primary schools: Reflection in action (2<sup>nd</sup> ed.). Yola: Education and Management Services.
22. Umunadi, K. E. (2009). Teacher Utilization of Instructional Equipment and Materials in Teaching Basic Electricity in Urban and Rural Technical Colleges. *International Journal of Scientific Research in Education*, 2(2), 88- 95.



Scan to know paper details and  
author's profile

# Overview of Interactive Chatbot for Modelling, Predicting and Reporting Covid-19 in Mano River Union

*Olutola Olaide Fagbolu, Mohamed Sellu & Tolulope Adeola Fagbolu*

*University of Sierra Leone*

## ABSTRACT

COVID-19 is a phenomenal pandemic that wreaked havoc and still affecting all facets of the human race globally, a novel virus since December 2019 with a record of millions of confirmed cases and associated mortality of hundredth of thousands in well over 230 countries of the world and these cases rise daily. With the incessant increase in the number of cases in Mano River Union (Sierra Leone, Guinea and Liberia), second wave and new variants. The Basic Reproduction Number ( $R_0$ ) over a time frame using recorded incident cases of the World Health Organization (WHO) by governments of Sierra Leone, Guinea and Liberia.

*Keywords:* chatbot, covid-19, NLP parser, RNN health.

*Classification:* K.8.1

*Language:* English



LJP Copyright ID: 975824  
Print ISSN: 2514-863X  
Online ISSN: 2514-8648

London Journal of Research in Computer Science and Technology

Volume 21 | Issue 1 | Compilation 1.0

© 2021. Olutola Olaide Fagbolu, Mohamed Sellu & Tolulope Adeola Fagbolu. This is a research/review paper, distributed under the terms of the Creative Commons Attribution-Noncommercial 4.0 Unported License <http://creativecommons.org/licenses/by-nc/4.0/>, permitting all noncommercial use, distribution, and reproduction in any medium, provided the original work is properly cited.



# Overview of Interactive Chatbot for Modelling, Predicting and Reporting Covid-19 in Mano River Union

Olutola Olaide Fagbolu<sup>α</sup>, Mohamed Sellu<sup>σ</sup> & Tolulope Adeola Fagbolu<sup>ρ</sup>

## ABSTRACT

*COVID-19 is a phenomenal pandemic that wreaked havoc and still affecting all facets of the human race globally, a novel virus since December 2019 with a record of millions of confirmed cases and associated mortality of hundredth of thousands in well over 230 countries of the world and these cases rise daily. With the incessant increase in the number of cases in Mano River Union (Sierra Leone, Guinea and Liberia), second wave and new variants. The Basic Reproduction Number ( $R_0$ ) over a time frame using recorded incident cases of the World Health Organization (WHO) by governments of Sierra Leone, Guinea and Liberia. The exponential growth method estimates the growth rate of COVID-19 and  $R_0$  using R Survival Analysis Packages and functions to report infection rate, mortality rate and offers live information for planning and preventive measures. With WHO speculations that the virus has come to stay with the human populace, there is an urgent need to explore how computing statistics with Natural Language Processing (NLP) will salvage the infection rates, mortality rates, FAQs. NLP parser is used to extract related information from Emergency Department reports that serve as dataset coupled with the death toll and patient counts as of July 24, 2020, to develop an interactive chatbot that gives preventive measures, symptoms, predict  $R_0$ , report routine statistical data, FAQs about COVID-19, emergency contacts for all the provinces in Sierra Leone, Guinea and Liberia and general toll-number for Ministries of Information and Communication, Health, etc. This research work is done through intensive and extensive assessments, observations, and information on the case by case of patients to develop the chatbot Covid19Mano. Dialog flow*

*open-source environment is used with PHP for documenting the content of the database for reprogrammed questions, phrases, or words about COVID-19, and NLP parser was integrated with Facebook Messenger and Whatsapp to test the efficiency and accuracy of the chatbot. It offers encoder decoder models for sequence-to-sequence prediction problems in question answering, text and speech translation, and many more magical and exciting features when trained with Recurrent Neural Networks(RNN).*

*The chatbot will enable users to get up-to-date information about the Coronavirus pandemic, its spread in Mano River Union, who to contact, what to do, and many more related challenges such as predictive analytics of infection, transmission, death, and recovery records and consequently model these mathematically. All the information, analytics, graphics would be embedded into government websites of Mano River Union to combat the ravaging fake news, myth, and stigmatization about the COVID-19 and offer the general overview of modelling, reporting, and predicting the pandemic effect on these nations of Mano River Union.*

**Keywords:** chatbot, covid-19, NLP parser, RNN health.

**Author  $\alpha$   $\sigma$ :** Department of Mathematics and Statistics Fourah Bay College, University of Sierra Leone, Freetown Sierra Leone. e-mail: fagbolu.olutola@usl.edu.sl

**$\rho$ :** Department of Health Policy and Management, Faculty of Public Health, University College Hospital, Ibadan, Nigeria.

## I. INTRODUCTION

A novel coronavirus has been a global threat since December 30 2019 with a record of many millions of confirmed cases associated with many thousand

ds of deaths in over 204 countries of the world, however, these cases rise daily (WHO, 2020a). Coronaviruses are family viruses that range from the common cold to MERS (Middle East Respiratory Syndrome) coronavirus discovered in 2012 and SARS (Severe Acute Respiratory Syndrome) coronavirus discovered in 2003. They are circulating in animals and some of these coronaviruses can spill over from animals to humans. SARS-CoV-1 was first identified in China in 2003, caused more than 8000 cases in 33 countries over eight months, it was discovered to be transferred to humans from civet cats. In 2012, WHO was notified of 2,494 confirmed cases of MERS-CoV with 858 associated deaths, first identified in Saudi Arabia and it was transferred to humans from camels. COVID-19 is a zoonotic virus from phylogenetic analyses undertaken with available full genome sequences, bats appear to be the reservoir of the COVID-19 virus, but the intermediary host(s) has not yet been identified (WHO, 2020c). The current outbreak dynamics is human-to-human transmission. Current estimates of the incubation period of the virus range from 1-14 days and these estimates will be refined as more data become available. Understanding the time when infected patients may transmit the virus to others is critical for control efforts. Detailed epidemiological information from more people infected is needed to determine the infectious period of 2019-nCoV, in particular, whether transmission can occur from asymptomatic individuals or during the incubation period (WHO, 2020b).

The World Health Organization strategies through a combination of public health actions, such as rapid identification, diagnosis and management of the cases, identification and follow up of the contacts, infection prevention and control in healthcare settings, implementation of health measures for travellers, raising awareness among the populace, communication of risks, development of chatbots but without any ability to predict the transmission rate, mortality rate and model them but WHO chatbot only gives overview report of cases without capability to model, and predict appropriately (WHO, 2020b).

The novel COVID-19 exhibits non-specific symptoms and the disease's appearance can range from

no symptoms (asymptomatic) to severe pneumonia and death. Most asymptomatic infections have relatively rare cases that are asymptomatic on the date of identification that went on to develop the disease. The proportion of truly asymptomatic infections is unclear but appears to be relatively rare and does not appear to be a major driver of transmission. Most people infected with the COVID-19 virus have mild disease and recover. For most people, COVID-19 infection will cause mild illness however, it can make some people very sick and it can be fatal in other people. Older people aged over 60 years, those with pre-existing medical conditions such as cardiovascular disease, chronic respiratory disease, hypertension, or diabetes are at high risk (WHO, 2020a). Approximately 80% of laboratory-confirmed patients have had mild to moderate disease, which includes non-pneumonia and pneumonia cases, 13.8% have severe disease (dyspnea, respiratory frequency  $\geq 30$ /minute, blood oxygen saturation  $\leq 93\%$ , lung infiltrates  $>50\%$  of the lung field within two to four days) and 6.1% are critical (respiratory failure, septic shock, and/or multiple organ dysfunction). Disease in children appears to be relatively rare and mild with approximately 2.4% of the total reported cases reported amongst individuals aged less than 19 years. A very small proportion of those aged under 19 years have developed severe (2.5%) or critical disease (0.2%) (WHO, 2020c). Anosmia (inability to smell) and in some people inability to taste have been reported as a symptom of COVID -19 infection (ECDC, 2020).

The importance of global surveillance is to monitor trends in the disease where human-to-human transmission occurs, hurriedly discover new cases in countries where the virus is not circulating, provide epidemiological information to conduct risk assessments at the national, regional, and global levels, provide epidemiological information to guide preparedness and response measures.

A case can be identified when a patient with acute respiratory illness (that is, fever and at least one sign or symptom of respiratory disease, for example, cough or shortness of breath) with no other etiology that fully explains the clinical presentation, a history of travel to or residence in a cou-

ntry, area or territory that has reported local transmission of COVID-19 disease during the 14 days before symptom onset. Specimens from the lower respiratory tract, such as sputum, endotracheal aspirate, or bronchoalveolar lavage, can be collected for investigation. If patients do not have signs or symptoms of lower respiratory tract disease or if specimen collection for lower respiratory tract disease is clinically indicated but the collection is not possible, upper respiratory tract specimens, such as a nasopharyngeal aspirate or combined nasopharyngeal and oropharyngeal swabs should be collected. If initial testing is negative in a patient who is strongly suspected to have COVID-19 infection, specimens should be collected again from multiple respiratory tract sites (such as the nose) and should also include sputum and endotracheal aspirate. Additional specimens may be collected, such as blood, urine, and stool, to monitor the presence of virus and shedding of virus from different body compartments. A confirmed case is a person with laboratory confirmation of infection with the COVID-19 virus irrespective of clinical signs and symptoms (WHO, 2020d, 2020b).

It is known that the virus gains access to the body through the eyes, nose, and mouth, so it is advised to avoid touching one's face with an unwashed hand. Washing of hands with soap and running water for at least 30 seconds, or cleaning hands thoroughly with alcohol-based solutions, tissues, or gels. It is also recommended that social distancing (to stay one meter or more away) from people infected with COVID-19, who are showing symptoms, to reduce the risk of infections through respiratory droplets. Practicing voluntary self-isolation by infected people, basic hand hygiene, and respiratory hygiene (sneezing or coughing into a tissue which is immediately disposed of properly or covering one's cough with the elbow) has been recommended to reduce the effect on humans, avoid unprotected contact with live animals. There is no specific treatment for coronaviruses but symptoms can be treated.

In Africa, there are cases of COVID-19 in well over 50 countries with almost six thousand deaths and Mano River Union has thousands of confirmed cases with many casualties hence there is a need

to give preventive measures, symptoms, FAQs (Frequently Asked Questions) on COVID-19, Emergency contacts for all Districts in Mano River Union and the general toll-free number of Ministries of Health in Mano River Union, live info on the cases of the virus, quiz game and graphical infection rate, predictive analytics for deaths and recovery on chatbot called Covid19 Mano to reduce mortality or avoid more causality. A chatbot is Artificial Intelligence oriented natural language processor that communicates indistinguishably from human or another participant of communication, the essence of a chatbot is to make users feel that they are communicating with any living human and it becomes imperative in the management of this global outbreak of Coronavirus because of its ability to guide during and after this pandemic with natural language using voice or text. Although several efforts have been put in place to promote health literacy, Most outpatient information for COVID-19 in Mano River Union is provided on mass media, print media, public places, hospitals and other fora despite present publicities, health campaigns, efforts, and methods that are being used to promote health literacy on COVID-19 to the public, the rate of deaths as a result of COVID-19 continue to increase at an alarming rate. The majority of the populace gets to know information about COVID-19 when it reaches their nation, district, city, or town and at that time it might already be too late for them. The versatility of chatbot has found its usage in relatively all domains including customer care, marketing, teaching and learning, training. It handles the day-to-day tasks of office assistants, booking hotel rooms, making reservations, etc. A chatbot is a smart assistant that can be a software tool, algorithm or artificial intelligence that communicate with a human or another participant on a particular topic(s), area of interest, customer's query or satisfactions, or any other specific domain, it can be text-based which follow rule-based approach to respond to queries or voice-based which interact through voice and respond to written commands at times through voice communications or combined both text and voice features (Smutny & Schreiberova, 2020). The technology avails end- users the platform to

interact with digital entities on a user-friendly and renowned social media - Facebook and Whatsapp which are the preferred option in Sierra Leone, Guinea and Liberia because of their wider acceptability among the populace and ease of use. The chatbot on Facebook Messenger will be able to provide accurate and consistent information about COVID-19 in Mano River Union in both English and French Languages, answer any bothering questions, discuss related topics, inform the society about this pandemic and any other related task on COVID-19 such as predictive information about transmission, deaths and recovery (Sanni S.A, Rajabu N and Fagbolu O.O, 2018).

This research work cut across several interdisciplinary domains of public health informatics, mathematical modelling and natural language processing that design and implement an interactive chatbot to offer adequate, timely information about the COVID-19 outbreak in Sierra Leone, Liberia and Guinea or any other future health and wellness challenge(s) to increase awareness of health literacy and policies, prevention and wellness of human. Emphatic mentions are previously developed chatbots such as ELIZA in 1956, ALICE, Claude, Hex, TAY, Xiaoice, etc. ALICE is based on NLP and pattern matching while relating responses from knowledge records while Claude picks inputs, responds based on its database and finally gives answers. Hex possesses an exemplary feature of being able to introduce new topics after answering any posed question. Xiaoice (literally Microsoft little ice) acts as a seventeen-year-old girl on Weibo and conversation can be established with her (Smutny & Schreiberova, 2020) (ZEMČÍK, 2019). The chatbot will not only comprise content managed technology that runs only on Facebook messenger and Whatsapp but can be deployed on email, SMS and any other chatbot compliant social media with the ability to estimate using gathered data and statistics (Sanni S.A, Rajabu N and Fagbolu O.O). The source of motivation for this research work is the need to provide reliable, accurate and straight-forward information about the growth and spread of this pandemic to curb the ravaging fake news, myth and stigmatization about the virus.

## II. STATEMENT OF PROBLEM

Several chatbots for COVID-19 have been developed with various levels of success reported in the literature. However, no available chatbot is supported with adequate estimation, predictive model of infection, transmission and recovery based on the data and statistics rather they are only reporting daily situations. The research work will address the problem by developing an interactive chatbot that will estimate, model and report Basic Reproduction Number ( $R_0$ ) for COVID-19 over time using incident number cases that are recorded in Mano River Union. It will rapidly proffer a digital time capsule for future researchers of COVID-19 with innovative responses to FAQs, information, and other sundry matters on this pandemic while setting standard measures for any other related challenges.

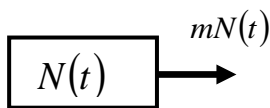
Developing a light-proof bilingual chatbot to support empirical information of the World Health Organization (WHO), governments in Mano River Union and report, predict and model the information on Facebook messenger and Whatsapp platform for the populace. This will serve as an improvement in previously created WHO chatbot and boost modelling techniques of the pandemic, estimation, and provide an accurate projection of infection, recovery and transmission.

The long-term impact of the project will be on health policy and management, planning, economic and technological application with its benefits. The fundamental benefits being in the health sector, this is expected to permeate other aspects of the society as Mano River Union can plan, project and estimate accurately the rate of recovery, mortality and transmission using mathematical models. Community engagement would further be strengthened by the partnership between the governments of Mano River Union and academia. Aspects of this work will foster study exchanges in both lingua franca, support ongoing undergraduate and postgraduate students' training for long term manpower development.

### III. RESEARCH OBJECTIVES

The aim is to model predictive analytics of gathered information about COVID-19 in Mano River Union periodically, estimate the rate of growth of the COVID-19 in Sierra Leone ( $k_s$ ), Guinea ( $k_g$ ) and Liberia ( $k_l$ ) and offer situation reports. The study will evaluate numerous factors affecting the detection and prevention of the spread of COVID-19 in Mano River Union, extract relevant information about infected patients from clinical reports to travelling history, contacts and design an interactive chatbot that will combat the consequent effect of inadequate information about this challenging pandemic, predict and model COVID-19 based on the size of the population of individuals that get infected at a particular rate, and increase awareness of health literacy. The specific objectives are-

- 1) To gather cumulative cases of COVID-19 in Mano River Union with age and sex.
- 2) To formulate a reality-based model for the chatbot using Recurrent Neural Networks (RNN).



Here,  $N(t)$  is the number of individuals at the time  $t$ ,  $m$  is the rate of growth of the pandemic and so  $mN(t)$  is the number of individuals who get infected per unit time. The differential equation for the model is as follows:

$$\frac{dN(t)}{dt} = -mN(t) \dots\dots\dots(1)$$

The assumption that once infected, individuals become infectious at a constant rate, or that once infectious individuals recover and become immune at a constant rate is often incorporated into infectious disease models. This assumption is convenient since the average pre-infectious and infectious periods are unknown.

The result of equation (1) solves to give

$$N(t) = N(0)e^{-mt} \dots\dots\dots(2)$$

as illustrated below:

- 3) To develop interactive chatbot by including accuracy of Natural Language Processing and present its general overview.
- 4) To test and prototype the chatbot on Facebook messenger and Whatsapp platforms.

### IV. METHODOLOGY

The source of motivation for this research work is the need to provide reliable, accurate, and straightforward information about the growth and spread of this pandemic to combat the ravaging fake news, myth and stigmatization about the virus. The exponential decline or growth method will be employed to estimate the size of the population of Mano River Union that is Liberia, Sierra Leone and Guinea population assuming that individual gets infected at a constant rate (Vynnycky, E and White, 2010). The model has the following structure:

By differentiating the expression  $N(t) = N(0)e^{-mt}$  using the rules of differentiation, we obtain the following expression:  $\frac{dN(t)}{dt} = -mN(0)e^{-mt}$ , since  $N(t) = N(0)e^{-mt}$  we obtain our desired result:  $\frac{dN(t)}{dt} = -mN(t)$

Attack and Case-Fatality Rates -in general, the Attack Rates for an infectious disease is measured as:

$$AR = \frac{\text{Number of who get sick}}{\text{Number of people at risk}} \text{ in a timeframe: duration of the pandemic.}$$

For the Case-Fatality Rates, where the cases are the people with disease and fatality is how many deaths is measure as:

$$CFR = \frac{\text{Number of deaths}}{\text{Number of cases}} \text{ in a timeframe: duration of the pandemic}$$

To achieve all these objectives, an appropriate flowchart that represents all the requirements were done as depicted in figure 1. The statistical data from different sources were obtained and these sources include WHO, Ministries and Parastatals and District Health Information Desks which are pre-processed by cleaning (removing noise, stop words, and converting to either uppercase or lowercase) furthermore, the research approach would proceed to tokenization if cleaning is accomplished and stemming and lemmatization are employed to split sentences and texts into constituent words afterward vectorization was done through TF-IDF (Term Frequency-Inverse Document Frequency) instead of the commonly used bag of words. The test data created Long Short Term Memory (LSTM), train and save model before allowing any input from the users and finally, responses were generated.

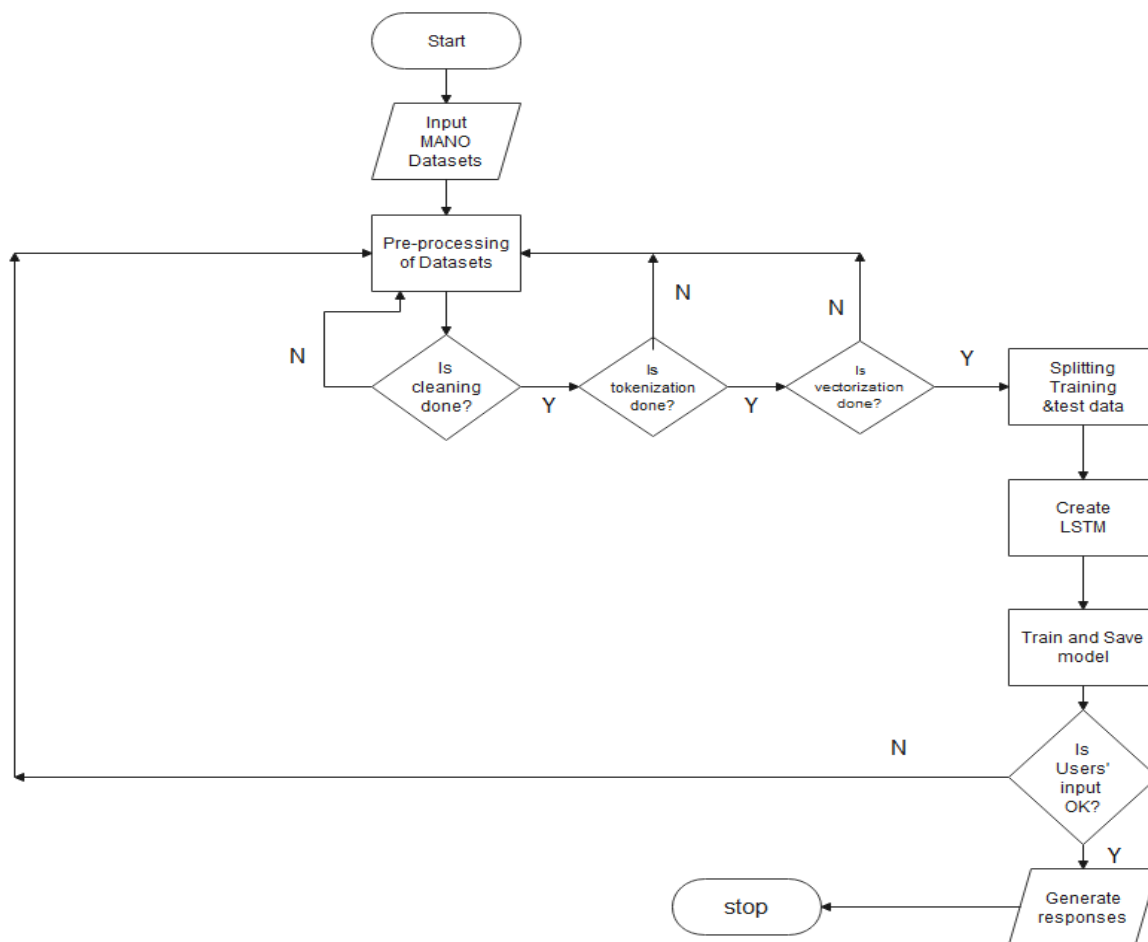


Figure 1: Flowchart for the Design methods

As seen in figure 2, the rectangular shape and flowlines depict vector and functions respectively as in matrix multiplication, input vectors are in red, output vectors are blue and RNN's states are

sequence output, sequence input, both sequence output and input and synced sequence input and output (Karpathy, 2015)

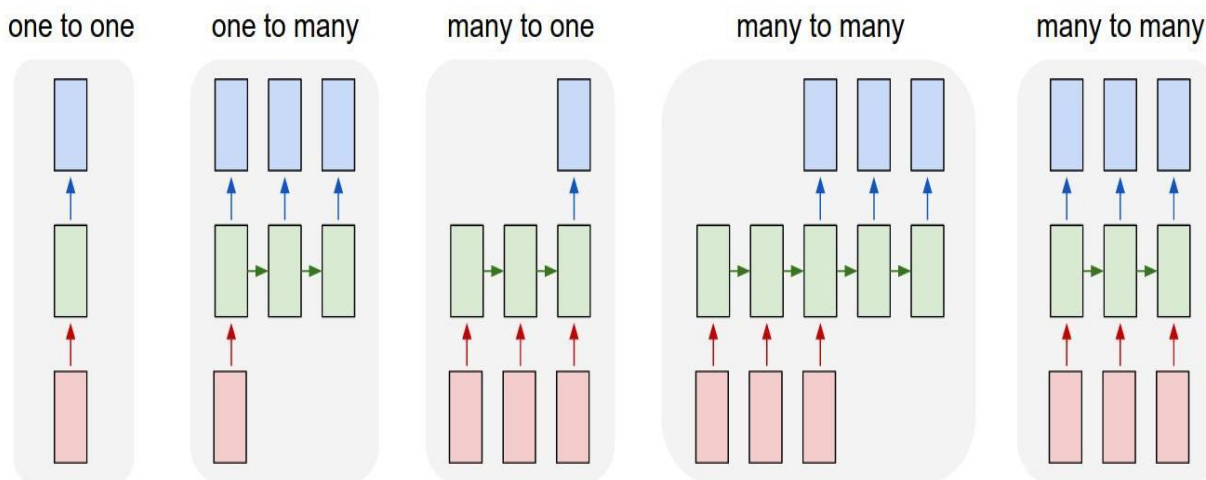


Figure 2: Sequences of Recurrent Neural Networks (RNN)(Karpathy,2015)

The reality-based model will rely on rules, statistical methods, or templates of encoder- decoder models to situate interactive chatbot as an example of RNN for sequence-to-sequence prediction problems in question answering, text and speech translation, video captioning, and many more.

The programming tools needed are Facebook, Twitter, WhatsApp, Python, R, AWS Aurora, Node.js, Express JS, AWS AI/ML. The collected data and derived mathematical models will be used as a data set to train chatbot (Covid19Mano) over a specific period.

#### 4.1 Design activities and output

Activity	Expected Outcome
Collection of data and frequency of infections in population as described in terms of incidence and prevalence	Data and statistics based on projection for the susceptible population for a particular time during the pandemic, mathematical equations, and subsequent models
System Specification	Study of requirement and technical documentation of interactive chatbot.
System Design	System architecture, user-computer interface, web-interface, algorithms for the chatbot processes.
System Implementation	Formatting data and statistics results and creation of program data files, programming using Python with other open-source tools and applications as well as Java Android platforms.
System evaluation	Report on analysis of gathered cumulative cases in Mano River Union questionnaire administered to patients, health workers, and cross-sectional and sample population.

#### 4.2 Activity indicators

- 1) Demonstrated design of Chatbot.
- 2) Documentation of the mathematical model and system specification.
- 3) Presentation of system architecture, user-computer interface, web-interface, algorithms for the chatbot processes.
- 4) System testing (running chatbot on different platforms).
- 5) Deployed chatbot for use by selected stakeholders within Mano River Union to determine the real-life and practical impacts and the chatbot (Covid19Mano) hosted on the web.

- 2) Training and retraining are imminent once new data are added and by considering the robust nature of COVID-19.

### CONCLUSION

The data collection procedure involved primary and secondary data. The primary data were collected using both qualitative and quantitative techniques. The quantitative technique involved the administration of structured questionnaires and observation while the qualitative technique involved conducting in-depth interviews. Pre-tests on chatbot were administered by the researchers using tests containing questions selected from FAQs. The Natural Language Understanding (NLU), semantic frame, and Dialogue Management (DM) were used to train the chatbot. R Survival Analysis Packages and functions reported infection rate, mortality rate and offer live information for planning and prevention measures. This chatbot with RNN architecture was developed, evaluated with datasets from Mano River Union and remember the contexts, queries and keep them to generate responses, the model is

### V. LIMITATIONS

Despite the inherent advantages and contribution to social, scientific, and health-related knowledge, the following are some of its limitations that have made the chatbot to be at its overview stage.

- 1) The training cost is exorbitant and without training the chatbot would only be generative and whenever is in a dilemma for a certain query, it will generate a response by recalling fixed answers without evidence as prevalent in the retrieval-based approach.

basic without any complexity and offers high accuracy for large data such as Covid19 information in Sierra Leone, Guinea and Liberia. Sequence input and sequence output provided a bilingual feature, for example, text or datasets can be in the English language and the responses are displayed in French and vice versa, most especially for Guinea as a francophone nation. In the future, Bidirectional Recurrent Neural Networks (BRNN) and attention mechanism can be deployed in the development of interactive chatbot with recommender systems on what to do, food to eat, and many more precautionary approaches would be part of the chatbot.

### REFERENCES

1. ECDC, E. C. for D. P. and C. (2020). Question and Answers on Covid-19. European Centre for Disease Prevention and Control.
2. Karpathy, A. (2015). The Unreasonable Effectiveness of Recurrent Neural Networks. Web Page, 1–28. <http://karpathy.github.io/2015/05/21/rnn-effectiveness/>.
3. Sanni S.A, Rajabu N and Fagbolu O.O Interactive Chatbot to support Interns during Internship Program, 8th East African Education Quality Assurance Forum (EAQAN), May 7-10, 2018. Rwanda.
4. Smutny, P., & Schreiberova, P. (2020). Computers & Education Chatbots for learning: A review of educational chatbots for the Facebook Messenger. *Computers & Education*, 151(June2019), 103862. <https://doi.org/10.1016/j.compedu.2020.103862>.
5. Vynnycky, E and White, R.. (2010). *An Introduction to Infectious Disease Modelling*. Oxford University Press.
6. WHO, W. H. O. (2020a). Coronavirus disease 2019 (COVID-19) Situation Report – 70 (Vol. 70, Issue March).
7. WHO, W. H. O. (2020b). Novel Coronavirus (2019-nCoV) Situation Report - 7 (Issue 7).
8. WHO, W. H. O. (2020c). Report of the WHO-China Joint Mission on Coronavirus Disease 2019 ( COVID-19 ) (Issue February).
9. WHO, W. H. O. (2020d). The First Few X cases and contacts (FFX) investigation protocol for coronavirus disease 2019 (COVID-19) (Issue February).
10. ZEMČÍK, M. T. (2019). A Brief History of Chatbots. *DEStech Transactions on Computer Science and Engineering*, aicae. <https://doi.org/10.12783/dtcse/aicae2019/31439>.

*This page is intentionally left blank*



Scan to know paper details and  
author's profile

# Study of Certain Corona Family Related Viruses based on Percentage Nucleotide Concentration and Golden Ratios and a Novel Sonic Attack Technique to Deactivate all Mutating Viruses

*Ethirajan Govindarajan*

## ABSTRACT

A set of 130 complete genome data of Corona type viruses are analyzed and classified based on their Percentage Concentrations of Nucleotides and Golden Ratios. The purpose of this study is to extract finer and hidden details about the evolutionary nature of Corona type viruses. It was found that all Corona type viruses evolve by random mutation but with their basic structural and functional genetic property maintained in the system biological sense. This has already been verified by extracting common digital spatial and spectral patterns during a study on “Pairwise Spatial Correlation of SARS-Corona Viruses” and on “Pairwise Spectral Correlation of SARS-Corona Viruses”. The scope of this paper is limited to the study of a set of 130 virus metadata in terms of certain quantificational measures such as (i) Golden Ratios (GR) and (ii) Percentage Concentrations of Nucleotides (pA, pT, pG and pC), and classification of those viruses based on these measures.

**Keywords:** corona type viruses, golden ratios, percentage concentrations of nucleotides, artificial intelligence, sonic bursts.

**Classification:** H.5.5

**Language:** English



LJP Copyright ID: 975825  
Print ISSN: 2514-863X  
Online ISSN: 2514-8648

London Journal of Research in Computer Science and Technology

Volume 21 | Issue 1 | Compilation 1.0

© 2021. Ethirajan Govindarajan. This is a research/review paper, distributed under the terms of the Creative Commons Attribution-Noncommercial 4.0 Unported License (<http://creativecommons.org/licenses/by-nc/4.0/>), permitting all noncommercial use, distribution, and reproduction in any medium, provided the original work is properly cited.







# Study of Certain Corona Family Related Viruses based on Percentage Nucleotide Concentration and Golden Ratios and a Novel Sonic Attack Technique to Deactivate all Mutating Viruses

Ethirajan Govindarajan

## CAUTION

*This is a theoretical investigation carried out on certain genome data of Corona type viruses archived in NCBI and results reported in this document are outcomes based on certain subjective definitions, notions and genome data analysis. Results presented here are to be deliberated and interpreted by Genetic Scientists, Agriculture Scientists, Biologists, Biotechnologists, Biochemists, Biophysicists and Microbiologists. This research report is meant for public use.*

## Chief Scientist and Team Leader

**Ethirajan Govindarajan**

Chairman, Pentagram Group of Companies, India, USA and Canada  
Advisor, Galileo Innovations Pvt. Ltd., Bangalore, India; Advisor, Remedium Therapeutics Pvt. Ltd., Chennai, India  
Franklin Fellow, London Journal Press, UK; Fellow, African Center for Cyberlaw and Cybercrime Prevention, Paris  
E-Mail: dr.rajane@gmail.com

## Senior Research Scientists

**Dr. Andras Pellionisz**, Chief Executive Officer, HolGenTech, California, USA  
**Dr. Jean Claude Perez**, IBM European Research Center on Artificial Intelligence, France  
**Dr. Michael Patrick Coyle**, Chief Executive Officer, Avatar MedVision US LLC, NC, USA  
**Dr. Manish Prateek**, Vice Chancellor, Dev Bhoomi Uttarakhand University, Dehradun, India  
**Dr. Amit Agarwal**, Director, APJ Abdul Kalam, Institute of Technology, Dehradun, India  
**Dr. Madhavi Pingili**, HoD, Department of IT, CMRN, Hyderabad, Telangana State, India  
**Dr. K. Janarthanan**, School of Biomedical Engineering, Jimma University, Ethiopia

## Research Team

**Dr. S. Nageswara Rao**, Alchemist and Siddha Physician, Chennai, India  
**Sathya Govindarajan**, Director, PRC Global Tech., Inc., Ontario, Canada  
**Prashanthi Govindarajan**, Director, PRC Global Tech., Inc., Ontario, Canada  
**Preeti Kashyap**, Data Scientist, Pentagram Research Centre Pvt. Ltd., Hyderabad, India  
**Dr. K. Thiagarajan**, Professor, Dept. of Math., K Ramakrishnan Coll. of Tech., Trichy, India  
**Dr. Mohd. Sherfuddin Khan**, Scientist, Pentagram Research Centre Pvt. Ltd., Hyderabad, India  
**L. Lakshmi**, AP, Dept. of Arch., Periyar Maniammai, Inst. of Sci. and Tech., Tanjavur, India  
**Bhavya Sharma**, Data Scientist, Pentagram Research Centre Pvt. Ltd., Hyderabad, India  
**Pat Krishnan**, CEO, Remedium Therapeutics Private Limited, Chennai, India  
**Dr. A. Madhumathi**, Professor, School of Architecture, VIT, Vellore, India  
**Yashaswi Vemuganti**, Consultant, Avatar MedVision US LLC, NC, USA  
**N. Suriya Prakash**, Aptean India Private Limited, Bangalore, India

## Advisory Committee Members

**Pritam Kumar Sinha**, Founder Director, AntiZyme OPC Pvt. Ltd., Bangalore, India  
**Dr. Ashok Alur**, Director, Centre of Excellence for FPO, Government of Karnataka, India  
**Dr. Dharm Singh**, Director General, Indian Medicinal Plants Marketing Federation (Retd.)  
**B. R. Badrinath**, Head, Special Projects, Galileo Innovations Pvt. Ltd., Bangalore, India  
**Dr. M. Shankar Lingam**, National Institute of Rural Development, Hyderabad, India  
**H. B. Nayak**, CEO, Pentagram Research Centre Pvt. Ltd., Hyderabad, India

## ABSTRACT

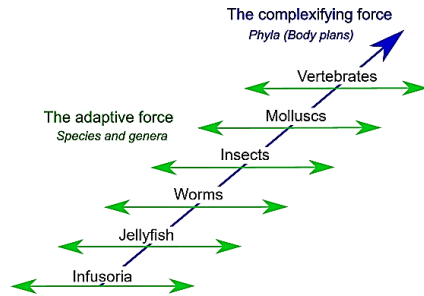
A set of 130 complete genome data of Corona type viruses are analyzed and classified based on their Percentage Concentrations of Nucleotides and Golden Ratios. The purpose of this study is to extract finer and hidden details about the evolutionary nature of Corona type viruses. It was found that all Corona type viruses evolve by random mutation but with their basic structural and functional genetic property maintained in the system biological sense. This has already been verified by extracting common digital spatial and spectral patterns during a study on “Pairwise Spatial Correlation of SARS-Corona Viruses” and on “Pairwise Spectral Correlation of SARS-Corona Viruses”. The scope of this paper is limited to the study of a set of 130 virus metadata in terms of certain quantificational measures such as (i) Golden Ratios (GR) and (ii) Percentage Concentrations of Nucleotides (pA, pT, pG and pC), and classification of those viruses based on these measures. Despite many drugs, vaccines and immunity boosters that move around the world with the idea of bringing down the pandemic state of affairs, the unpredictable ramification due to structural and genetic functional variations of Corona type viruses causes a serious concern, more specifically, a doubt/fear whether God has created these viruses or Man has created them. However, one should continue doing research in developing various diagnostic tools and therapeutic measures. Dismembering virus structures with ‘sonic’ frequencies that resonate with the natural frequencies of RNA forms and dismember them is one way of looking into solving this problem. This paper advocates a novel technique of generating pure tones as bursts and explores the possibilities of breaking the virus RNA sequences that are present inside the body or in the environment.

## Curtain Raiser

सूक्ष्मयोनीनि भूतानि तर्कं गम्यानि कानि चित  
पक्ष्मणो ऽपि निपातेन येषां सयात सकन्धपर्ययः

*There are many creatures that are so minute that their existence can only be inferred.  
With the falling of the eyelids alone, they are destroyed. (Mahabharata)*

RNAs and DNAs are the fundamental units used in this process of consumables and consumers interchanging their roles. Life Forms are constructs made up of these units. RNAs could be viewed as ‘Viruses’. They are constructive as well as destructive. Viruses play a constructive role in the forward process of evolution when the hosts can consume them, whereas, the same viruses play a destructive role in devolution, the reverse process of evolution when the hosts are unable to consume them.



“Lamarck's theory of evolution involved a complexifying force that progressively drives animal body plans towards higher levels, creating a ladder of phyla, as well as an adaptive force that causes animals with a given body plan to adapt to circumstances. The idea of progress in such theories permits the opposite idea of decay, seen in devolution”.

[https://en.wikipedia.org/wiki/Devolution\\_\(biology\)](https://en.wikipedia.org/wiki/Devolution_(biology))

At present, SARS-Corona type viruses act as destructive viruses and they do their job. It is up to humans to understand them, their role and try to delay the process of becoming consumables for them in the reverse process.

उद्यन्नादित्यः क्रिमीन्हन्तु निम्नोचन्हन्तु रश्मिभिः ।  
ये अन्तः क्रिमयो गवि ॥ १ ॥  
विश्वरूपं चतुरक्षं क्रिमिं सारङ्गमर्जुनम् ।  
शृणाम्यस्य पृष्टीरपि वृक्षामि यच्छिरः ॥ २ ॥  
अत्रिवद्धः क्रिमयो हन्मि कण्ववज्जमदग्निवत् ।  
अगस्त्यस्य ब्रह्मणा सं पिनष्यहं क्रिमीन् ॥ ३ ॥  
हतो राजा क्रिमीणामुतैषां स्थपतिर्हितः ।  
हतो हतमाता क्रिमिर्हतभ्राता हतस्वसा ॥ ४ ॥  
हतासो अस्य वेशसो हतासः परिवेशसः ।  
अथो ये क्षुल्लका इव सर्वे ते क्रिमयो हताः ॥ ५ ॥  
प्र ते शृणामि शृङ्गे याभ्यां वितुदायसि ।  
भिनसि ते कुषुम्भं यस्ते विषधानः ॥ ६ ॥

May the sun as he rises up kill the virus with his rays and also while setting let him kill the virus. Let these rays destroy the virus that lives within the beings. **I destroy the structure of the virus and demolish the virus in all of its forms.** Oh virus ! I will crush you in the same way as the Rishi Atri, Rishi Kanva and Rishi Jamadagni - the lords of cleanliness - crush you with a great grinder and with blazing fire. I crush you thoroughly with the knowledge of cleanliness. The king of the virus has been killed and their progenitor chieftain also is killed. The virus is killed along with his mother, his brothers and sisters. Neighbors of the virus are killed along with the virus' acquaintances. The subsidiary viruses and smaller viruses have been killed as well. I crush your antennas with which you torment others and I pierce your vicious pouch which is the store of your poison.

<https://bhawnayagya.org/atharva-veda-hymn-for-eradication-of-virus/>

This report is a result of intensive research carried out by devoted scientists who work with passion. The team acknowledges **Professor Jean-Claude Perez** for his suggestion to study partially deleted genomes like strains evolved in WA, Seattle, USA referenced in his research paper titled “COVID-19, SARS and Bats Coronaviruses Genomes Unexpected Exogeneous RNA Sequences”, (<https://osf.io/d9e5g/>). His timely advice and technical references in connection with Corona viruses is highly appreciated. The team also acknowledges “Edouard Aime, Alessandro Masciarelli, Alexandra HC, Alexandra Simplissima, Béatrice LOPEZ, birgitta.neher, Cerdan, Christian PELLEGRINI, Christian Perronne, Daniel Favre, Diego Lucio Rapoport,

DR Gérard GUILLAUME, Eric Caumes, Ester Sabino, fdo c, Francisco de Asis de Ribera Martin François GODEAU, Friedman Robert, Gilles ROQUETTE, Guy BERNARD, Hélène Banoun, Igor Gauquelin Jacques Demongéot, Jean François Toussaint, Joseph G Tritto, Lawrence Sellin, Leonidas Kleanthous, Mackenzie Sam, Marc F. Paya, Marc Niaufre, Marie-Paule KIENY, Martine Wonner, Matthias Lacoste, MEGARBANE Bruno nfaria, Patrice Serres, peng.zhou, Philip Risby, Philippe Charignon, philippe JUVIN, Ray Cavanaugh, Richard Castanet, rossana.segreto, Sebastien Mercier, SEYRAT Emilie, Simon A. Troost, Trinquetaille Serge, Valere Valerian, Volkmar-Weisser, Walt Dod, William Quenoy, Xavier Azalbert”, to name a few, for their untiring work and concern shown on human health, especially during the pandemic situation. Their views and suggestions have triggered a great motivation behind the research carried out by our team.

**Keywords:** corona type viruses, golden ratios, percentage concentrations of nucleotides, artificial intelligence, sonic bursts.

## I. DATA USED FOR ANALYSIS

More than 70,000 virus genome data have been listed in NCBI website. However, one hundred and thirty complete genomes of various categories of viruses are drawn from NCBI web site <https://www.ncbi.nlm.nih.gov/genbank/sars-cov-2-seqs/> for data analysis. The data given by NCBI is expected to be genuine.

### Arteriviridae

- S1. >NC\_001639.1 Lactate dehydrogenase-elevating virus, complete genome
- S2. >NC\_001961.1 Porcine respiratory and reproductive syndrome virus, complete genome
- S3. >NC\_002532.2 Equine arteritis virus, complete genome
- S4. >NC\_003092.2 Simian hemorrhagic fever virus, complete genome
- S5. >NC\_025112.1 Mikumi yellow baboon virus 1 isolate MYBV\_M58, complete genome
- S6. >NC\_025113.1 Southwest baboon virus 1 isolate SWBV\_16986\_11/4/2013, complete genome
- S7. >NC\_026439.1 Forest pouched giant rat arterivirus isolate PREDICT-06509, complete genome
- S8. >NC\_026509.1 DeBrazzas monkey arterivirus isolate PREDICT-06530, complete genome
- S9. >NC\_027124.1 Pebjah virus isolate I621, complete genome
- S10. >NC\_029053.1 Kafue Kinda chacma baboon virus isolate KKCBV-1, complete genome
- S11. >NC\_029992.1 UNVERIFIED: Free State vervet virus isolate VSAT1003, complete genome
- S12. >NC\_035127.1 Olivier's shrew virus 1 isolate Gkd-1, complete genome
- S13. >NC\_038291.1 Porcine reproductive and respiratory syndrome virus 2, complete genome
- S14. >NC\_038293.1 Simian hemorrhagic encephalitis virus isolate Sukhumi, complete genome
- S15. >NC\_043487.1 Lelystad virus, complete genome
- S16. >NC\_048209.1 Zambian malbrouck virus 1 isolate SHFVagmMal\_seqID\_01, partial genome

### Duplornaviricota

- S17. >NC\_003555.1 Giardia lamblia virus, complete genome
- S18. >NC\_005883.1 Chalaro elegans RNA Virus 1, complete genome
- S19. >NC\_007523.1 Coniothyrium minitans RNA virus, complete genome
- S20. >NC\_009224.1 Botryotinia fuckeliana totivirus 1, complete genome
- S21. >NC\_009890.1 Black raspberry virus F, complete genome
- S22. >NC\_014609.1 Armigeres subalbatus virus SaX06-AK20, complete genome
- S23. >NC\_024151.1 Beauveria bassiana victorivirus NZL/1980 isolate 6887, complete genome
- S24. >NC\_025214.1 Botryosphaeria dothidea victorivirus 1 isolate GY25, complete genome
- S25. >NC\_027212.1 Camponotus yamaokai virus genomic RNA, complete genome
- S26. >NC\_030295.1 Golden shiner totivirus isolate GSTV/US/MN/2014, partial genome
- S27. >NC\_030867.1 Fusarium poae victorivirus 1 genomic RNA, complete genome
- S28. >NC\_035674.1 Australian Anopheles totivirus isolate AATV 150840, complete genome
- S29. >NC\_038928.1 Aspergillus foetidus slow virus 1 CP gene and RdRp gene, genomic RNA
- S30. >NC\_038929.1 Beauveria bassiana victorivirus 1, complete genome
- S31. >NC\_040431.1 Diatom colony associated dsRNA virus 13 genomic RNA, complete genome
- S32. >NC\_040632.1 Gigaspora margarita giardia-like virus 1 isolate GmGIV1-BEG34, complete genome
- S33. >NC\_040653.1 Fusarium asiaticum victorivirus 1 isolate F16176, complete genome
- S34. >NC\_040659.1 Diatom colony associated dsRNA virus 10 genomic RNA, complete genome
- S35. >NC\_040660.1 Diatom colony associated dsRNA virus 11 genomic RNA, complete genome
- S36. >NC\_040775.1 Diatom colony associated dsRNA virus 12 genomic RNA, complete genome
- S37. >NC\_040793.1 Alternaria arborescens victorivirus 1 genomic RNA, complete genome

### Gresnaviridae

- S38. >NC\_046959.1 Guangdong greater green snake arterivirus strain LPSG2430 1ab protein, putative glycoprotein, and hypothetical protein genes, complete cds

### Kitrinoviricota

- S39. >NC\_000939.2 Pothos latent virus genes for replicase, capsid protein and movement protein
- S40. >NC\_001461.1 Bovine viral diarrhea virus 1, complete genome
- S41. >NC\_001512.1 O'nyong-nyong virus, complete genome
- S42. >NC\_001513.1 Ononis yellow mosaic virus, complete genome
- S43. >NC\_001564.2 Cell fusing agent virus strain Galveston, complete genome

- S44. >NC\_001642.1 Bamboo mosaic virus, complete genome  
 S45. >NC\_001728.1 Odontoglossum ringspot virus, complete genome  
 S46. >NC\_001948.1 Rupestris stem pitting associated virus-1, complete genome  
 S47. >NC\_002604.1 Botrytis virus F, complete genome  
 S48. >NC\_002729.1 Banana mild mosaic virus, complete genome  
 S49. >NC\_002795.1 Aconitum latent virus, complete genome  
 S50. >NC\_003557.1 Garlic latent virus, complete genome  
 S51. >NC\_003603.1 Groundnut rosette virus complete genome, strain MC1  
 S52. >NC\_003608.1 Hibiscus chlorotic ringspot virus, complete genome  
 S53. >NC\_003634.1 Physalis mottle virus, complete genome  
 S54. >NC\_003679.1 Border disease virus X818, complete genome  
 S55. >NC\_003852.1 Obuda pepper virus, complete genome  
 S56. >NC\_003900.1 Aura virus, complete genome  
 S57. >NC\_004724.1 Grapevine rootstock stem lesion associated virus, complete genome  
 S58. >NC\_005062.1 Omsk hemorrhagic fever virus, complete genome  
 S59. >NC\_005132.1 Botrytis virus X, complete genome  
 S60. >NC\_006939.1 Olive mild mosaic virus, complete genome  
 S61. >NC\_007733.2 Angelonia flower break virus, complete genome  
 S62. >NC\_009028.2 Ilheus virus, complete genome  
 S63. >NC\_009892.1 Peach chlorotic mottle virus, complete genome  
 S64. >NC\_011535.1 Grapevine Algerian latent virus, complete genome  
 S65. >NC\_011538.1 Nemesia ring necrosis virus, complete genome  
 S66. >NC\_011552.1 Peach mosaic virus, complete genome  
 S67. >NC\_011559.1 Anagyris vein yellowing virus, complete genome  
 S68. >NC\_012533.2 Kedougou virus strain DakAar D1470, complete genome  
 S69. >NC\_012534.1 Bagaza virus strain DakAr B209, complete genome  
 S70. >NC\_012812.1 Bovine viral diarrhea virus 3 Th/04\_KhonKaen, complete genome  
 S71. >NC\_013006.1 Kalanchoe latent virus, complete genome  
 S72. >NC\_015782.2 Grapevine Pinot gris virus complete genome, genomic RNA  
 S73. >NC\_016038.2 Brassica yellows virus isolate BrYV-ABJ, complete genome  
 S74. >NC\_016404.1 Actinidia virus B, complete genome  
 S75. >NC\_016440.1 Garlic common latent virus, complete genome  
 S76. >NC\_016959.1 Ndumu virus, complete genome  
 S77. >NC\_018713.1 Pestivirus strain Aydin/04-TR, complete genome  
 S78. >NC\_020470.1 Andean potato latent virus, complete genome  
 S79. >NC\_020471.1 Andean potato mild mosaic virus isolate Hu, complete genome  
 S80. >NC\_023439.1 Kama virus strain LEIV-20776Tat polyprotein gene, complete cds  
 S81. >NC\_023892.1 Gaillardia latent virus isolate 5/18-05-2010, complete genome  
 S82. >NC\_024458.1 Pitaya virus X isolate P37, complete genome  
 S83. >NC\_024887.1 Middelburg virus isolate ArB-8422, complete genome  
 S84. >NC\_026620.1 Jutiapa virus strain JG-128, complete genome  
 S85. >NC\_028793.2 Phasey bean mild yellows virus isolate NSWCP15, complete genome  
 S86. >NC\_030693.1 Grapevine Red Globe virus isolate Graciano-T101, complete genome  
 S87. >NC\_031327.1 Anopheles flavivirus variant 1, complete genome  
 S88. >NC\_040842.1 Potexvirus sp., complete genome  
 S89. >NC\_040837.1 Grapevine associated tympo-like virus, complete genome  
 S90. >NC\_040800.1 Actinidia seed-borne latent virus isolate 01227, complete genome  
 S91. >NC\_040788.1 Kampung Karu virus isolate SWK\_P44, complete genome  
 S92. >NC\_040776.1 Rocio virus strain SPH 34675, complete genome  
 S93. >NC\_039237.1 Bovine viral diarrhea virus 2 polyprotein gene, complete cds  
 S94. >NC\_039218.1 Kyasanur forest disease virus polyprotein gene, complete cds  
 S95. >NC\_039217.1 Phascolus vulgaris endornavirus 1 isolate PvEV-1\_Brazil polyprotein gene, complete cds  
 S96. >NC\_038966.1 Atractylodes mottle virus isolate SK, complete genome  
 S97. >NC\_036587.1 Babaco mosaic virus isolate Tandapi, complete genome  
 S98. >NC\_035462.1 Ocimum basilicum RNA virus 1 isolate DV1 RNA-dependent RNA polymerase and movement protein genes  
 S99. >NC\_035453.1 Actinidia virus 1 isolate K75, complete genome  
 S100. >NC\_035116.1 Lake Sinai Virus TO ORF1, RNA-dependent RNA polymerase, ORF3, and ORF4 genes, complete cds  
 S101. >NC\_035071.1 Apis flavivirus isolate RI-A polyprotein gene, complete cds  
 S102. >NC\_034833.1 Agave tequilana leaf virus isolate agave azul-Mex1, complete genome  
 S103. >NC\_034242.1 Ochlerotatus caspius flavivirus isolate 1608 polyprotein gene, complete cds  
 S104. >NC\_034216.1 Lagenaria siceraria endornavirus-Hubei isolate JZ, complete genome  
 S105. >NC\_034207.1 African eggplant yellowing virus isolate eMA4, complete genome  
 S106. >NC\_034205.1 Grapevine rupestris vein feathering virus isolate Mauzac, complete genome  
 S107. >NC\_033828.1 Peach virus D isolate SK, complete genome  
 S108. >NC\_033725.1 Bamaga virus isolate CY4270 polyprotein gene, complete cds  
 S109. >NC\_033724.1 Kadam virus from Uganda polyprotein gene, complete cds  
 S110. >NC\_033723.1 Gadgets Gully virus from Australia polyprotein gene, complete cds  
 S111. >NC\_033699.1 Jugra virus strain P-9-314 polyprotein gene, complete cds  
 S112. >NC\_033693.1 Bouboui virus strain DAK AR B490 polyprotein gene, complete cds  
 S113. >NC\_032088.1 New Mapoon virus, complete cds  
 S114. >NC\_031752.1 Botrytis cinerea endornavirus 1 strain HBtom-372, complete genome  
 S115. >NC\_031463.1 Ceratobasidium endornavirus B isolate Murdoch-2 polyprotein and ORF2 genes, complete cds  
 S116. >NC\_031462.1 Ceratobasidium endornavirus A isolate Murdoch-1 polyprotein gene, complete cds  
 S117. >NC\_043110.1 Banzi virus strain SAH 336 polyprotein gene, complete cds

GTGCCACATC CTCTACCCGC CTGGTTACAC CGTATGATGT GCACCTGGTG ACCTCCACAC CCGCGGGCGG GAGCTCTTTG GCTTCGGCTC CAGAAGGGAC CTACTTGGCC  
CGAGTCGCGC GATCGGGGTT GACTGGCCGG TCGTGCATGT TTGTCCCAC TTGTCCTCTG AACTTTGGCG AGTGCTCTTG AAGGGAGCCT GCGTGTGCTA AGAATGGTGT  
TAGTGTCTTC GGGTCTGCAT CCGGTCTCGG TGGGTTTTC ACTATTCATG GCAACCCCTG GGTAGTGACT GCTACGCACC TGCTGTGAGA TGGAAGGCT AGGGTTTCTC  
GTGTCCGCTT CTCTCAGTGG TTGACCTTGA AGAGCGTTGG TGATTACGCA TTCCGCCAGT TGGGOTGAGT GAAAGGGTAC GCGCCCAAGG TCGAGCTGTG AGATAGGAGA  
GGGGAGCCTT ATTTGCTACC CGAAGTGGAG TGGAGCTGGG TTTTGTGTGG TCCAAACACT TCGCAATGCT TACTAAGTG TGGCCACTCT TGGTCCCAG TAGTTGATGA  
GGATGGCACAC CTTATCCGGT TCCACACTGG TTCAAACAAA CCGCGCTCAG GAATGATCAC AACCACAATT GGAAAAACC TGCGCATGAG TAATGTGAAA TTGAGTGAAA  
TGTGCCAGCA CTATGGCCGC TCTGGCGTGC CTGTTTCAAC TGTCGGATTG CCAAGCATC TGATCGTCCA CGTGGAGGCG GTGGCTAGTG ACCTGGTAGC TGTAGTAGAA  
TCCTTTGCCA CACCTGAAGG CGCATTGTCC TCGGTTCACT TGTTGTGTGT CTOTTTCTTT TTGTTGTGCT CTGTGGCGCC TCATTCATGT TCCCTTTGTC CCAGTGAATT GAGTTGCCAT  
TTCTTCTCTG AATGAGATCT TGCCGTGGT GTTGGCCAGA TTAATGTTTT CTTTTGGCCT ATCTTGTGTT TCGAGTCTCA CAGGCTTTTC GGTGCAAGTG TTGTGTGTC  
GCCTAGTAAT AGCAGCTTTG AATCGGTTCT CTGTGTCAAT CCGTTCITTT TTGTTGGGCG AGTTGTTTCA TGTGTTCTGT ATGCCGTCTC ACTAGGTCTCT  
GTACCAGGTT ATTTTTTACCC GAGCACGACG GAAGTCGCCT CAAAAGAAAT ATTTGTTACC TTGCTTGCCA TCCATGTTTT GGCTTTGCTC CTAAGCTTGT TTAAGCGCCC  
CATGCTTGTG GATGTGTTAG TGGGAATGG CAGCTTTGAT GCCCGATTTT TCCTGAAGTA TCTTGCTGAA GGGAAATTTG GTGACGGAGT TTCTGATTC TGCAACATGA  
CACCTGAGGG GTTGACCGCC GCCCTGGCCA TCACCTTTGC TGATGATGAC TTGGAATTC TACAACGGCA TTCAGAGTTC AAGTGTCTG CACGCTTCTG AAACATGAGG  
AATGGAGCTA AGRAATTTAT TGATCCCGC TATGCTCGTG CTCTTAGAGC TCAGCTGGCT GCCACAGACA AAATTAAGGC CTCRAAAGC ATCCTGGCCA AGCTAGAGAG  
CTTTGCTGGA GGTGTTGTCA CCAAAGTTGA ACCAGGTGAT GTTGTGGTGG TACTCGGAAA AAGATTTGTT GCGCACCTCG TAGAGATCAC CATCAACGAT GTCACACATG  
TGATCCGAGT AATTGGAGCA CGAGTAATGG CCGGCACCTA GTTCTCTGTT GGTACCATCT GCGGAGACTG TGTGAAGCTG TGTGAAGCTG TGTGAAGCTG CCAAGTGAAT  
TCARAAAGAA AACCTCGCCG CCAAAAGCGC ACCCGACTGG GCAGTAAAGT GGTGCGCAAG GTCCGAATTG TACAACGGCA TTCAGAGTTC AAGTGTCTG CACGCTTCTG  
TGATGAGACA TATGAAGCTT TCTGGTGTG TGAAAACCTG AGACTCGCCA CCGTTGGCAC TTGTTGGGCG GTCCACAATT GAGCAAAAT TTGTTGCTG TCCATGTTTT  
TAAAGACCCT AGTGGAGAAA TACCCTGTGG GAAAAATAA GCATATTGAG TTTGCAGTGA CCACCTACAA CCTGGATGGT GAAGAGTTTG ATGTCCCCGA CCATGAGCCT  
CTGGAGTGA CCAATTACCAT CGGAGATTTCT GATCTGGGAG CAGAAAAGCT CACAGTGGAC CACAGCTGGG TCGCAAAATG CCCTTGGACT GCACACTGAT GTCACACATG  
GGAAAAGCTG GCTCGCATCA TTGAAAGGCT CAATGGTTTG CAGCAGTATG GTGCAATTC GTCGTAATTC CACTTGGCCA TGAACGCTG TGAACGCTG CAGCTGACTG  
GTCCAAGGAC GCGGTCAAAA TTGTGAGTA TCACAGCCGG ACTTTCFCCA TTGTTGATGT CAACTTGAAG GTCATGTCC TGTATGAGTA CCGTCGTACG ATGGAAAAC  
CCGGCATTTT GCTTGTGGCC AAACGTACTG ATGGTGTGGT TGTCATGCCA AAGCATGAAC CCGTCTGGT AGACGTACTT TTAGTGGTG TTAGTGGTG AGGATGCCGA GCTTTCCCC  
CGTACCCACG GCCCGGGCAA CACCGGTTTT TGTTGGGACT TGAATCCCCG CCTGTGACT TAGAACTTGA GTTTCAGAAA GTTTCAGAAA GATTCATGTA GATGACTGCT  
CATCGCCAGG GGTGATGCC CAGCCTTGA TCTTCTTACC AAATTAOACT CTGAAGAGG TGATCCGCTG CACATCGCGG TGTATGTTG TGTATGTTG GATTCATGTA  
TAACTACCTT AATACCTGAA AAGACAAGAG AGCCGCTGCA TCGCGCCGCT GTTCAACAA GAGGGTTCC TGTATGTTG TGTAACTT TGTAACTT TGTAACTT  
CACGGTTTTG AGTTGATATG TCCGACATTG CCGCTAGTG TTTTAGAGTA CTGGACAGT CTGGACAGT GCGCCGATA CTCAAGAAT GCTAACAAAA CATGGTTGCG CTAGCGCAGC  
CGAAAAGGAC TTACAATAAT TTGACTTGTG CAGGACGGGG TTCTGTTTGC TGTGTTGCT GATGTTGCTG TGTGTTGCTG TGTGTTGCTG TGTGTTGCTG TGTGTTGCTG  
TGTTCAATGCC TCCGCATATC CCGCTAAGA ACTCGATGGC CCGCATAGG CAGCATAGC GTGCGGACTG TATTTCCGCTA TATTTCCGCTA TATTTCCGCTA TATTTCCGCTA  
CGAGCCTGTG AGGAGCACTG GCAAAACGCT ACCCCCTGCA CCGTGAAGAA CAGCAGTATG TGTGTTGCTG TGTGTTGCTG TGTGTTGCTG TGTGTTGCTG  
AGGCCTAAGG TCCGCTCTCA CGGTGTGAC CCAGGGCTTT ATGCGAAAAG GCATTTGAA GCGATTTGAA GCGATTTGAA GCGATTTGAA GCGATTTGAA  
CGCGCCGTTG TCTGGAGGCT CATGTGACCG CATGTGACCG GCGATTTGAA GCGATTTGAA GCGATTTGAA GCGATTTGAA GCGATTTGAA GCGATTTGAA  
ATTCCTTCCT ACGTTTGAAG TTGTTGCTG GACGTGTGCA GCACCATGAG TGGTGTCTT GATCAAAACC CCAATTTGCTG TGTGTTGCTG TGTGTTGCTG TGTGTTGCTG  
CACAGTTTTAT AGTTAATAA TATATGCCA ACACATGGTG TTATCAGCAT TCAAGTGTG TCGTACATG TCAAGTGTG TCGTACATG TCAAGTGTG TCGTACATG  
TCTTTGAGTT ACAACCGTTG TTGTTGTTT CCGACGAGT GGTTCCTAC AACGAGAGTG ATGAGTTGCC AACTACCAT TTCTTTGTTG ACCACCTGGA TTTAATGTTG  
GGTTTTAAAA CAGACAGGCT CAAAACCTGT ATCACTCAG AACCACAAT TCTGTTGCTG TGTGTTGCTG TGTGTTGCTG TGTGTTGCTG TGTGTTGCTG  
GGCCCTAGCG TATCAAAATG AAGCCACTGA TATTTCCGCTA TATTTCCGCTA TATTTCCGCTA TATTTCCGCTA TATTTCCGCTA TATTTCCGCTA TATTTCCGCTA  
TGGACCTTGT CTGTTGGTATT GGTGAGTGG CCGGATCCCC CTGTTGAGA TTTCCCAGC CTAGCTTTT CTGTTGAGA TTTCCCAGC CTAGCTTTT  
AAATGTAGAA CTGCGCACAA TTGCGGGGCC CCAGCCACCC TTGTCAGCAG TTGTTGGCTA AACTTTGTTG ACTACCATG GCATGGACT CCACATTTGCC TGTGTACT  
CCCATGTGGG TCTGCTGTAG GATCGGCGCT CTGTTGACAG GTTGAACAG AGGTGAGCAG CTTACCCGCT GGGGCAATAC AGCCAGGCT TATCTTTGAT TGCGTCCCG  
CGAAAGTTGA TTGCTGTGCT GTTGAACAGAA AGGTGAGCAG CTTACCCGCT GGGGCAATAC AGCCAGGCT TATCTTTGAT TGCGTCCCG  
GTTGATCTC CGATGGTGA TTATCAAGTA ATGAAGTTG CCCCACAATG TCGTACATG TCAAGTGTG TCGTACATG TCAAGTGTG TCGTACATG  
CGCCCCGGG ACTGGAAAA CAACCTTACT GTTGAAGTGT GTCCGGGATG ATGATGTAAT CTATACCCCG ACCCAGAGA CCATGCTTGA TGTGTTCAA GCCTTAAAG  
TGTGAGATT TGATCCACCA AAGACACCCC CGCTGGAATT CCCCCTCTT CCGCGGACTG GCGCGGACTG TCGTTGATT TGTGTTGATT TGTGTTGATT  
TATCTGATG AAGCAGCTTA TTGCAACCCA TTGACTGTGC TTGACTGTGC TTGACTGTGC TTGACTGTGC TTGACTGTGC TTGACTGTGC TTGACTGTGC  
CGGCGCTTGC TTTGCAATTT CCGTGTAGCC CCGGAGACAG CTGATTGAGG TTTCCGTTT CCGACTGTCT GTGGTTAAT CCATAGTAA ATTTTATAA GAAAGTTGG  
TCCCTCGCGG CCCCGACACT TGGTCAAGT TTTTGAACA GTTACCAAG GTTACCAAG GTTACCAAG GTTACCAAG GTTACCAAG GTTACCAAG GTTACCAAG  
TCCAGCCAA GCTGACCCCTA TGATGTGGTC ACAGTGTATC TACCCACCC TAAATCCTG TAAATCCTG TAAATCCTG TAAATCCTG TAAATCCTG TAAATCCTG  
CATCTATGAC CCAATATGAT AGTGTGACGA ATATTACAA GTTTTGAAC ATGAGCCAGC TGATGGTGG TGAATGGTGG TGAATGGTGG TGAATGGTGG TGAATGGTGG  
GTGTTGAAA CAGCCTGGCA AGCAGCTCCA GGACACTGTA TTTAAAATTG TTTAAAATTG TTTAAAATTG TTTAAAATTG TTTAAAATTG TTTAAAATTG  
GTTTTCTATT ATTCCTCTGA TTTGATTGAG TTTGCAAAA TACCCCAGA ATTTGTGAAA CATTGGCCCC TCGTAAACCG CCAGAATAGA ACTGAGTGGC CCGCAGGCT  
GGTGTGGCGG ATGAAATAAA TGAATAAAAA TTACCGCGCT GCTTTTGGG CCGGATATTA CCGGATATTA CCGGATATTA CCGGATATTA CCGGATATTA  
TGACCAAAAT TTTAAAAGGG GAGAGTGTAC CCGCTGCTGA TCAATCATG TCAACTGGGA GAGTGTGAT TGGTGTGAT TGGTGTGAT TGGTGTGAT TGGTGTGAT  
CGAAGAAAAT GCCCTCAGCC CTTCATTTGCC GAAGTCAAGG GAAGCAATG TGGTGTGAT TGGTGTGAT TGGTGTGAT TGGTGTGAT TGGTGTGAT  
TGTAAAAGGTT GCGCTCAGTT CCGTCAAGAA GCGCGCAAAA GCGCTGTGTA CTGTGACTGA CTGTGACTGA CTGTGACTGA CTGTGACTGA CTGTGACTGA  
TGGATTACAA GCTACTGGTT GATTTTCAGC CCGTCAAGCT TATGGTGGG AAGATGCTA CTGTGACTGA CTGTGACTGA CTGTGACTGA CTGTGACTGA  
TTAAAGTTG CTGAAGGGGA GCGCGTCTC TTTGATTGG ATGAAATTTG CACCATTTG TGAAGTGTG TGAAGTGTG TGAAGTGTG TGAAGTGTG  
GTTCTGTTCA GTCCGTGCAA AAGTGAAGCT CACCTATGAG AGCTAAGCT ATGAAATTTG CACCATTTG TGAAGTGTG TGAAGTGTG TGAAGTGTG  
CATACATTTA CAACGTTGG CCATCCACT ATACCACAAG TAATTTGCC CCGCTTGA AAGTGTG TGAAGTGTG TGAAGTGTG TGAAGTGTG  
AAGTAATGG TAATGTGCA ACTCTTATA ACAAGGCTAT TGAAGTGTG TGAAGTGTG TGAAGTGTG TGAAGTGTG TGAAGTGTG TGAAGTGTG  
TCTACCATG CCGCGTGGCT GGTTTTGGG TGTATCGAG TGTATTTGG ATGAAATTTG CACCATTTG TGAAGTGTG TGAAGTGTG TGAAGTGTG TGAAGTGTG  
ATGCCATCAA GTTATGTCAA CCAGATGCCA TCCATTATGG CAAAGGTTTT GGGTGTGTT GGGTGTGTT GGGTGTGTT GGGTGTGTT GGGTGTGTT  
ATGTTCTATA TGTCTCACAA AGGTCGCGAG ACCTGGGGTA CCGTCCGCCA TTCAGGCGG TCTGTTGCA TCGCATATG TCGCATATG TCGCATATG  
AGGCTTGGAA AGGGAGGCGT GCACATATCT AATGTCTCGG CTTCCCGCCA TTCAGGCGG TCTGTTGCA TCGCATATG TCGCATATG TCGCATATG  
TCAGCTTTGG ACCATCAAC CACATCCCTCA CAGTAGCTGA GCACATATG TCGCATATG TCGCATATG TCGCATATG TCGCATATG TCGCATATG  
GTAATGTTAC TGGGCTTACC CAGCTCTGCT TTTGTTTCC TATCCATAGG ATCCGATAGG ACCGGTAA GA CCGCATATG TCGCATATG TCGCATATG  
GTTGAGAATT CTTATTRACT CCACCCITTT TGTGTGACTG ATGATACTG TGAGGAGTCA TCCCTCAGT ACTTGGATT AATGTATGAC GAGTGTGTT  
CTCAATATAT GGCACAATCG CCATCAATCG COTGGTGTG CCGTACTTAC AGTGTATG TCGCATATG TCGCATATG TCGCATATG TCGCATATG  
ATAACGCAC CCGAGTGGAG CACAATCAGA ATGTAAACC ACCTTGGCC ATCTTGTAGT ATCTTGTAGT GATGTTGATG GATGTTGATG GATGTTGATG  
TCTGGTTGG TCTTGAATGT CTCTACTTTT CTCAGCGCT CACCTGGCG CCGTGTCAA CAGGAGTGA CCGTGTCAA CAGGAGTGA CCGTGTCAA  
CTGCTGATCG ATCCCAAGG GTGTCAATAT ACACATAAAA AAGAGGTTGA CCGTGTCAA CAGGAGTGA CCGTGTCAA CAGGAGTGA CCGTGTCAA  
TGCTGGACAT GAACATGTTG CCAACTTAGT GACCTTGGCT CCGTGTCAA CAGGAGTGA CCGTGTCAA CAGGAGTGA CCGTGTCAA CAGGAGTGA  
ATTTGGCCGT TGTATAAAT TTACTGAGT ATGTGAGTCA TATTCAAAT AAAAACTTGA AAAAACTTGA AAAAACTTGA AAAAACTTGA AAAAACTTGA  
AAATGGGGAA CGGTTATAAC CTGGTTTTG GCCATATTAT AGGCAATATG AAAAACTTGA AAAAACTTGA AAAAACTTGA AAAAACTTGA AAAAACTTGA  
ATTTTGTATT TTCTCTCTAC GAAAATGCCC TGCGCCGAG GCAACAGCT CCAAAAGT TTGATTTA CTTAATTTT CTTAATTTT CTTAATTTT CTTAATTTT  
ACATTTCCGT TACGCGGTAG AGACATTTGT AATTTTTCCA CTTTACTG ACTTAATTTT CTTAATTTT CTTAATTTT CTTAATTTT CTTAATTTT  
TGTAGCTGG AGGAGGCTAT AGTATGTCA TGTGCAATG TGTGCAATG TGTGCAATG TGTGCAATG TGTGCAATG TGTGCAATG TGTGCAATG  
ATGAGCTGGA GATATAAGT CACAGCAATTT ACAAAATTTG TCCCTGATC CAAAGAAAAA GCTTACCGGA GCTTACCGGA GCTTACCGGA GCTTACCGGA  
GCTACAAGG CACCCAATGG AGTTAAGC CTTTGTGCTG GATGCTGTGA AAGCGCTGG GCTTACCGGA GCTTACCGGA GCTTACCGGA GCTTACCGGA  
AACTTCCCTG TATCAGATAC TGATTGCCCT TAGTTTGACT TATACCCCAA TTGCCATCTA GAGCTTAA GAGCTTAA GAGCTTAA GAGCTTAA GAGCTTAA  
TTTTTATAA TTGTTGTGTT AGTTTGTGTT ATTTGTGTTA CCGATCTTTT GTTATGCTG GTTATGCTG GTTATGCTG GTTATGCTG GTTATGCTG  
ACTCTGTGCA AGATAGTAAA CTGGATGGTA CTCAGATGCC GTTGTGCTT CTTAGTGTG CTTAGTGTG CTTAGTGTG CTTAGTGTG CTTAGTGTG  
GAGCCTAACC ACATCGTCAA CAACCGCTTT TGTGTTAGG AAGCCAGTA GTACCCCTGT GAACGAGTA GAACGAGTA GAACGAGTA GAACGAGTA  
AGGCCAGTAT AAAAGGAGCG GTTAACTT TTGAAGTATG CTCAAAAATG AAGAAAGT AGCCCAAGCT GACTTCCCT GACTTCCCT GACTTCCCT  
GCTTCCGCAAT GCCCGACAGA ATAAAAGAAA GGGCGACAAA AAGAAGTAA AGCCCAAGCT GACTTCCCT GACTTCCCT GACTTCCCT  
ATGAGGTACA AATGTGTAGA TCCCTATTGG TAACTTGTG TAACTTGTG TAACTTGTG TAACTTGTG TAACTTGTG TAACTTGTG TAACTTGTG  
CCAACGATG CCACGTGTG CACTGTTGG CTTAATTAAT GCTTCTGCTA ATTTCTTCTG TTAAGAGTTA CAATGTAAGT CAATGTAAGT CAATGTAAGT  
TTGCGATTTG GCTGGGCCG AATT

Now, the numerical equivalent of the above genome S1: >NC\_001639.1 is given below.

1224431111	3122323231	1444323433	2114432421	2224113313	4414233314	1323344434	1142231331	2132134431	1424421421	2144244424
2421141222	4433323234	1244212443	4412441442	1211441234	1324333122	4314433234	2232321444	4411234213	1323222233	3134123334
1132421223	4141433234	4213423414	3344442242	4444422244	4134131144	4243322333	4344223333	2232221212	3344234414	1441424143

Study of Certain Corona Family Related Viruses Based on Percentage Nucleotide Concentration and Golden Ratios and a Novel Sonic Attack Technique to Deactivate all Mutating Viruses

2333133434	4212214121	4444144113	3112324314	4332333142	1234233223	2433341222	1444112234	4434123143	2443331124	1411422241
3342431424	1142223243	4424132232	3214433314	3331114231	4313211141	4622313311	3112223133	2121424143	4332344433	2332124421
2414233444	3324441332	1223414232	1234111234	1323412323	1323123114	4122444233	2234144414	3232231341	1442444322	3442413413
4444244341	1133222323	2443224243	3123424323	4331323233	3434241411	1341141411	2222232424	4311413331	2122232331	4114133224
1112312224	3224443114	4123333131	3341433133	2232343131	1312412434	1241424442	3443144144	2232221112	2313222443	1142144432
2231221224	2222142332	3323143323	4133221432	4424224441	2341131111	1231414411	1224241113	2322333111	1132322231	3314414134
1322331133	2334423231	3313342314	4213423144	3412333421	4431144444	3413313233	4233444421	2223343434	4344234112	1443433322
3131342323	3221114423	1141111231	2212322323	2333444233	3232444331	3242214124	1334141222	4443134333	4313423231	3131122222
2232343232	2332313224	2441243222	4111423232	3344144444	2124411132	2123322441	1141133113	2122143342	1424244444	1333314332
3442323314	2341223212	2242342124	1241123141	2422232321	2342422232	1441134231	4211423413	4133314134	3242323434	2324231441
1312123212	4112233241	2314334413	1424444344	1444223344	1423244324	2342144214	1112233122	3232111241	1414233142	3231411111
3311133131	2431344424	3233334242	4244231123	2321432332	3232323343	3231123411	1224321443	1442144414	3231112111	2223311422
3413411111	3244412413	2344223344	2241344223	3232331344	2212232314	2342333142	2214411424	1323421323	1143121344	2434232431
1142342131	3423323133	4321413113	2342123333	1244441133	1231423332	1422223341	3114242131	4231111223	4423144433	4313132334
2331411322	3413432331	4241314244	1234314324	4113442334	2434423342	4132313244	1342134134	1223142241	3234234133	242342423
1323341413	1124231411	3313342324	3233433231	1411431423	4214141324	4223332223	1211311424	2342144111	323141113	1342234214
1334223224	2423134243	2331344412	3313142241	1432421134	1231124213	1311124423	3113423422	4143234222	3113132341	1134421134
3411343411	2113141113	1131441342	4444114144	2144443442	3111444411	4443331332	4442334343	1223113223	313233231	313311344
3333442323	1324334214	2124321444	3121133444	1134443244	4443443431	4411114324	1334442211	1113341313	4133322241	2142111331
2124124233	1124411423	1124233332	2111234223	1441243223	1331232343	3341222223	1111341211	4244444233	2333344423	3341424323
421312414	3342414122	4123331432	2441432312	4432313241	1414423112	2332242334	4412342311	4124133441	1311242341	4323234323
4231323111	1134111142	2124422331	1223313133	1414442411	1113423323	1342444134	3424422144	4133134122	332223343	442322313
4314413413	1131443443	1413423233	4313131311	2223233113	2244342111	3134331344	3134441414	4324411343	3312421331	3221344223
3431441342	2241411314	3333434331	332313224	1244142342	2323131241	1331313332	1441442121	1113122443	3324333222	3321411113
2221233422	3322314441	4413322222	4242232442	4441242442	4344123233	2423243234	2433124243	4411133422	1122333423	1221232221
4141232322	2342212322	3232323242	1234222414	2244413441	2234122412	4444422344	2233232222	2233232222	4424212423	4344212423
222224232	2233422232	3223413233	2332322441	3313324421	4423114423	3324232423	2244132342	2433423132	3234334134	2422313433
2122443321	1233232241	4431442322	1144121223	3232233242	3232334144	3233414312	22322324243	4323233334	3344414334	1122123322
2214224422	4344221233	2332222331	4223331424	3232224223	4432334242	3432331133	3222411132	3222423411	3232323131	3233421444
1231332141	2221131322	2244444231	4111332334	2431441141	4213133413	2223231412	3212132341	4444343223	1244412442	2323342144
3341241113	342323341	3332113323	1344424144	1332214231	4113441321	1342141321	1442131131	1111111122	2441141131	4132323444
4441442142	3144441434	1134332311	3234221113	324224132	343233124	1122413312	3242333244	4431132211	113322241	1133224412
1411334444	122424441	1143242411	2231244231	3234214124	3232223142	4232214421	4244342344	1234333323	3221233432	3243414423
1223233341	443231222	2342311322	1123332231	3312232333	4331333412	3223413122	2232223442	3331223212	321234221	1222242412
1242342233	4232421324	4232313223	2331431331	1441143144	4423323411	3112441224	2422124423	2332233441	3331322123	2321142444
1442423412	4242313312	3334221412	4444412332	1223241214	1344231223	1223334242	2123122344	3231223411	4132334412	4222122341
1133222442	2322322231	4342121232	3232244222	3442242212	444424224	1444442422	4232112234	4412341224	441223442	4123211332
3414423414	4443241424	2332223332	3411222224	2234233413	2311222223	3342323232	3413223433	2312442312	2241242233	2224233434
3234414124	2424144334	4233221414	4321231232	3414423323	1442441414	443433434	3134242223	3422433424	4131133314	4212232344
3413244344	3124344322	3142344433	2342341232	2232444414	1114222334	1232444223	1133313442	1133314213	3342323421	1311232332
2132342424	3332423412	4433324223	233322224	1421224123	3411444232	3321323142	3421434144	2342324131	2332133422	1333222442
3232433422	4242412324	2231442241	1313432233	13142214341	2243444313	2331333412	3111332314	3444441133	244242324	3112133131
3334313422	1223424144	3411323313	2331344233	2222322333	2441114142	3432222342	2214211323	2334314242	3322444413	2132231231
3312344114	4221243323	2441414233	2241114111	4343342413	3112312414	1144441412	3311111444	2333412313	2112323111	2231323111
2323441341	4223434324	2423344234	4232224114	4232231223	4441134124	2312432413	3233413343	3232323134	1442323134	2321321311
2422234421	4142311133	4341223244	2433224132	2322323232	4224224422	4232334344	2412241232	2444222324	4413211223	4132234122
2224224423	1123131242	2344232332	3223344131	2211232222	4222234422	1422232222	2413242241	4133422224	3323411323	2242322234
3442132112	1341342223	1124332423	4232324122	4332242222	2232233334	1322322241	2232232423	1234432424	1222331314	1241332442
3214413322	2222241444	3134143143	3111324342	4111131112	1222322144	2234223441	2441232222	3342223424	4211342232	2211343444
4123422342	1323232133	2333311233	4134223122	3443412222	2442311321	4232423111	3331122234	3231433132	2242312244	2341141231
4144231333	3223144344	3444233431	2414422324	2312312314	2233112244	2141143341	2241313224	1132342243	2413434224	114123133
1123313421	1331122212	3313244344	2123424323	4242213134	2413423342	1441413141	1112221334	4241111134	1244233441	1342131313
4222342331	3232323241	4411132231	1441332312	3223233233	2142433111	3111312332	3343144243	2312312414	4124114312	1141214123
2312443132	1123213141	4313211233	4433414241	3224242322	3321441242	3433131442	2313114344	2323113124	4413233142	4324114134
2411131114	1143243443	4411113234	1443314233	3414231132	3324341243	3133111223	1433132113	2231411123	342334121	1334213244
2312323141	2123113342	2242332324	2311114243	1314244341	4442233414	3241341122	3344332224	1111122412	4131111412	3341341133
2111314424	1323313111	2144442323	3111111211	3412122313	2223413231	4414421411	4423312332	3113132223	1232444431	4412313442
4233132331	4412214412	4331312242	3124233113	4131113342	4131113342	4133432314	3141412333	2412314244	2234231424	4211331311
3311113423	3424341241	2231113442	4112332223	4134132412	4232222111	4232214441	4424233442	2311434234	2421313323	3232414232
3244113314	3433241111	2232311321	2414134433	1422242441	2232321232	4114223113	3241232442	2231231321	4432432143	1233311114
4433341222	2443423344	1114231423	1233232332	2324123431	1134123114	4424442332	1314324124	2231423323	1442334331	2442244444
4324144413	3444433411	4144333124	4144332223	2323331422	2311244443	443223142	2131142231	3223241311	4131221241	4134123424
4133434133	3323123444	4134422331	2422442214	1112214124	4232113133	4314124212	4314124343	3232122322	411214434	2223313411
2114121422	1121442311	1131411113	1344342341	2343343342	2322141141	1133332244	2322132312	2423111422	2332334214	3142421444
4143322223	1232321232	2443141223	44344421323	2222131321	4223341432	4344443121	4244113112	3421141111	4123322343	4213434344
431113314	2214111112	2241423234	4133413333	2242323234	4232323242	4214123323	431332144	2313234342	3124343424	1314324134
2322412344	3243141214	4443421131	1424312334	1334124114	3323341312	2444142414	3233223411	2434124443	1412231234	4423234111
4313442323	1331341423	3411144323	1444442341	4422311111	3413214232	2411131341	1314413314	1121423333	1441121142	2322324142
1334421133	2443424241	3433232314	4433232314	1234311133	4411231133	1444124232	2213311131	1211122414	4441223443	3441331223
3433443223	2423313342	1344213442	4123231443	3241144443	3431221241	332322214	1141112423	1421223112	234233234	2311311233
1224422442	1432222311	2232234412	3143233411	3414412313	2334234222	3121134333	313222324	1213131312	4242322144	3412413411
4141322212	1322211211	2121234441	1414123323	2212413412	2413323233	2414111122	3313342232	2222341131	4244421311	1233134114
2422231322	1411443223	2233232122	4431431432	3322224214	1143131323	1331323344	3114214412	2242223223	1441442331	2221123223

2211113224 4231133331 3334324224 2223122233 1231122232 4144112342 1133232441 1122144223 4113214241 3231323424 1241322242
3142311324 3244232411 2314344214 1132311342 4414423142 1231142342 3222341131 3442122343 2344333342 2312324334 1434211123
4121412221 4111432333 4412441442 1214414113 2112222344 4344223211 1131214234 1321444322 3342321113 3413233321 4123222444
1132112333 2112322341 1424221211 1411334212 2311312341 2221224134 2412232224 4124422312 1224233222 2422333412 2422222333
4221441243 4433432233 3423322233 2321241314 2332222434 4343321234 3111314231 4413242312 4113242312 1214131133 2142213133
1234412411 3221232411 4413123441 2441221233 4113332222 3331234222 3341412311 1122342144 1234223123 1232112141 4132432421
1232424121 2324241441 1332434313 1442333321 3332322412 4413133421 2233131141 2433143114 2431132232 2414124244 1141312134
1334223311 1333133432 3414121242 1123242433 4224443441 2241334434 2242341323 2121123211 4412411322 2134112311 1441213441
2413422233 1441241114 1412444241 4132134231 3414421233 4233232411 2341124224 2324222241 2423223413 4334122234 2223233122
3211232214 2343242144 4133242324 2223322244 2124412313 1433321131 4341442131 3223344421 1122241442 2123211312 3232334323
3223131122 4221221424 4414442222 2323231423 1231214223 2313313241 2442414132 1422331222 2314224444 3123123244 1322322324
4241121212 3341411243 4412411234 4423323224 3441214124 1132444124 3123222334 1123212413 4132422422 2132414112 2422323222
1211413414 4441323313 4141112413 1232114414 3244223344 1242223123 3231432331 1111233212 3232213132 1412243144 1424221241
2442332233 2422311232 4244214222 4241334342 4414234433 4432342141 4311334232 3132322222 1124121341 1412444122 1331112332
4234231234 1214441133 3232412122 1414121121 1133132231 4412322124 1412331243 3441133231 1333322331 1221122241 1421144433
2342331412 3114123223 4411422132 3144223342 3432342211 2141443432 2313421113 3111423312 2241244141 1222222112 3232423344
1222334432 2232121112 2221423132 1232313241 2122411112 1111442334 1231134222 3113422431 4433444142 2242224444 2132321323
1112333311 4332212114 4223322223 3441212212 2334112123 1112322231 1311122333 4132332233 1214412414 3322122344 1222232222
1222232122 2242424214 4311112344 2343443413 3411413424 1144113112 2231242141 1422114422 3232313423 1123231443 332241141
1412224332 2143433213 1314122232 1122222441 3422231424 1422112224 4223112222 4221444343 4441422322 131222422 2412233341
2232134233 1331334212 2334121114 1321232412 2134241121 2123411322 3234442242 4344224122 2242242342 3213134132 2133112232
1231342331 3121211323 4144131222 1411122223 2442231214 4111331111 3242144331 1443413244 3322423323 3134134123 3443132412
3421411334 4144411223 1332211314 4322323424 3123323231 1134432343 3342111144 3244423443 1111123331 3344213112 2223231241
1142244233 2124131214 2312234422 2132223142 2121444411 2234412421 4134422113 3222224323 3414222334 2332122321 1121222221
2222212111 2232232322 1322223232 1222312321 4412412242 3221441124 4132234342 3242424331 3423221223 4442432323 3334122214
1424223241 1312132111 4233123321 4241312344 332323422 4421332131 1322141244 2134444144 4134414323 3141442423 1433143241
3134421144 1412432411 4114434422 2323322133 1134413321 3214442232 3114333413 2233244433 1422241113 1424323422 333341131
1334432124 1111331343 3221144222 2311321232 4241111211 3113111134 3344131121 1133334111 4413411422 1124132231 2211234122
3422434112 3443341413 1211133111 333413111 1131131134 1131131134 1344411342 3414224442 1233433314 4124231242 44344412323 1231444421
1231332141 1123232131 2442412233 2114422322 4112411332 3323314132 3214422132 2314244331 3331241122 2414232413 2224123422
4411434123 4414232243 3421122112 3422423421 1224224234 2211313221 4112321132 4123241324 1312341343 1424134422 2232112211
2234312223 3423334433 1122

Important Note:

The concepts and tools presented here in this paper have been tried on all 130 virus genomes. However, basic results pertaining to S1: >NC\_001639.1 Lactate dehydrogenase-elevating virus, complete genome only are provided here due to space constraints. So, let us consider this genome for further analysis. The major aim of this paper is to provide certain feasible computational techniques for classifying any given large set of genome sequences. The first technique is "Classification of 130 virus genomes based on Four level Auto Correlation Measures (ACMs)". The second technique is "Classification of 130 virus genomes based on Percentage Concentration of Nucleotides and Pairwise Correlation Coefficients". The third technique is "Classification of 130 virus genomes based on JCP Golden Ratio (GR) values of the sequences.

II. CLASSIFICATION OF 130 VIRUS GENOMES BASED ON AUTO CORRELATION MEASURES

Consider the virus genome sequence S1: >NC\_001639.1. The graphical inverse S1^-1 of S1 is given below.

TAAAGCCGCG GTCCGTTTAC CGTTAATATA TGTTTTCCGA CTCAGCCGAC TAGACTGAT GACTGTAATG TAACATTGAG AATTTCGCTT TTAATCGTC TTCGTAATA
ATCCGGCTGT CACCGTACCC AACCCTCGTA CTTTGACTGT CACTTAAACT AGGAGGCCCT CAGTGTGATC CATGTGGGAA ACCTGGTTC CAATGGTTC
TCCTAGATGT GTAACATATG AGTAATCCCC AGTAGTGTAC GCCTCTAGT CTACACGGGC GGTATCCCTT CACGTCGAA CCGACGAAGA AGAAAAGAC GGGGAARAGGA
AATAAGACAC GCCGTAAACC TTCGTTACGT AATAGTTTGA CTAATTAAC GACCAAACCG GGAAATAAGA CCGGGCAAAA GAAGAATAAA ACTCTGTATG AAGTTTCCCA
ATTGGCCAGG AAAACTATGC CGGAAGAACC GGGGTCTGTG CTCAGAACCT TTGGTAGC GGACGAGCCG CAAGTGTTC CATGATGGAC GAAGGATGT GTGTTTCCCG
CAACAACCTG TACACCAATG CGAGACTGCA GGCAGTCTCC ACAGGTGCAC CGACGACACC CAGTCTACA TTGAAGATGG ATCCCTCGTG TTGGCCGTAG ACTCATGGTA
GGTCAATGA TAGCAATGTT TCACATTATA CCGGGTGTCC TCCGTTATT GTCAGAGCTC TCCTGTGCTT TGACACAAC ATTTCTITAC TTTATGTAG TTTATGTAGT
TTGATTGTGT TGTAAATAT TTTTATTTTT ATAAATGTTA TGGTCGGTTC CACGGTGTCT TTTGGAATTC CGACATCTAC GTTAAACCCC ATATTCAGTT TGAATTCGTT
TAGTCATAGA CTATGGTCTA TCAAACTAGT GTTTTAAGAT CGGAGGGTCA AAAAGCCGTC CCGGCCGTGC CGAAAGTGTG GTAGCTCGTT TGGCCAGAT
ATGGCTTAA CCAACGGAAC ATCGTACTGA CCGGTGTCGA CGAGGTGGTA TTGGCTGAC GCTGAAAGCA TCTGAAAGTC AACCATATG AACATTTAGA
CCACGTGAAT ATAGAGTCG ACTATGTTAA GGATGACGA GATGTCTGCT TCTTTACTT CGCTCTTCC CCGTTGAAC GTATATACT CGATTACTGT ATGACAATA
CGGTTATTCG AGGAGTCTGA TGTACGGGTT TACTTTCTTT TAGATTTGTC ACCCCGCCAC ATTCTTTTAA GTTCTTTTAA TTCCTACTGT TTCGACCTTT TTAATGTTTA
CAGAGATGCC CAATTGGCTT TACAACAAC TTTGGCCGAT GTAAGTCGAT TGTGTTCCTA TTTGGTTCCTA ACTGTTCCAA AGTTTAAAGA CCAACTCGAC AACCGGACCG CCCTCCGTAA
AAGCCATCTC TCTTTATGTT TTTATTTGTT TTTACCCTTA TTGGCAGTAC CATAGGTTGG TGACGGGTTA AAGAAGTTTG TAAAGTATAA CGGTTATTTG ACCGGTTTGG
GTTCCAATAT TGGCAAGGGG TAAAGTGATG TGATCCCCTT TCCTTACACC GGCCAGCTCT GAAGTTTCCA AGTACGGTCC AAAATAAGCA TTAATGATG TGTATGAGTC
ATTTTAAATA TGTTTGCCGG TTTACGGTTC TGTGTAATTT TTTAACACCT ACTTTAGTTC AAAGGAAATC GAGTTGGCCG ACATAATTCG TGCCTCGGTT CCAGTGATTC
AACCCTGTGA CAAGTACAGG CCGTGGCCAC ATCAAACTTTA ATTAAGTTGG GGGAAAGTGA ACCCGCTAGG TACACATTTG ACCCAGTGGG GGAAAAATA CACATATACA
TGTGGGAAAC CATACGTAGT TGCTGTGGTA GGGATATCCC TACAACGATA CTAATTTTTG TGAAGTCCGG AAGCACATCG TGCCGGGCTG CACTCCGGCG ACTTCTCAT
CCTCTGTAAG TTCTGGTTGG TCTTACTTTT CTCACCAGCT TACATGATG TGTGTATGTT AAAAAGGTGC AGTGGTAGTT TCTACCAGTT CCTGCCCAA TGTAAAGTAA
ACACGAGGTG ACCCCACGAC AATATTGTGT TTCTTAAACAC TGATTTCTTT TGACGACTAT GTAACGGTTT GTAGTATACC TGAACACTAC ACCGCTTGTG GTCCCGTAA
TACCGTAAAC ACGGTATATA ACTCCCTGTT TTGACCTGTA GTAGCCCCTT CAGTTTTAGG TFCATGATAC CTTCTACTGAG TAGTGTACAT AGTATGACT GTGTTTTCC
CACCTCATTT ATPTCTAAGA GTTTGGTCGG TGTTGAGAAT GTATTCACAT TTAATCCCCT GTTGTAGATCC ATCGAGATCC ACGCAGAAAT GGCAGATATC CTATCCCTTG GTTCTGCTG
GGACCATCTG GCGFCATCTG AATGTTAGTT GTTCTGGTTA CCGGAGCTTT GTCGACTTTT CTGTACTTCTA AACTACTGT GGTCCGATAT CACCGACTCC TGAACATCC
TACAACAAC ACCAGGTTTC GACTACCGAT ACCAAGTAA CGATTTGAAC TACCAATGTA ATATGTGAGC TCTTCCGGG ACTTACCGCC CTTCGGCTCT GTAATCTATA
CACGTCGCGA GGGAAAGGTT CCGAGATAG CAACCTCTTA CACTGTGTTG AACTCAGA GAATCACAAGA GGTTACCGAA GGGTACCGAT ACCTACTGTT TGGGAGGCG
TGGAAACACT CTGTATATCT CTTGATCTGC TGACACATAA CTGTAGTAGT CAGTACAGTC GCTAACATAG TTAAGTATC ACGGTTCCGT AGGCTTTTGG GAAACGTTAT ACCACTCGTA
GACCAACTGT ATTTGAATCAT CGTAGAGATT TCAAGGAAGA CATAAGTCT GACCACTGAC AAGCTGAGC GTGCCTATG GCGCCGTATT TGGTCAGACT ATGTTGTTGG
GTCGTCGGTC GGCCGTACC ATTCGGGTTT TTTCTACGGG TTTCTTTGTT CTTATAGTCT CTACTTTTGT TACTCGACT ATTTACGTT AAGTTATCCG AACATAATTT
CTCAACGTTG TAATGGGATA TGAACCCCTT GTACATGGGT GACGGAAATG TCGGTGTGCC ATGACGTCAT AGAAAATGTT ACCCGCGTTT TAAITAACCAT CACATCCACC
TACCGGCTGC AAAACTTTAC ATACGTAAT CCGCAGCTGC TAGTTCGGGG CCGTCCGTTA TCCGAGAACG TTTGTCTGCA AGTATCTAGC CACCTCGAAG TGAACATCCG
CAGTAACTGT CCTGCCTGAG TCGATCTTTG ACTACTCGTG AGTGAATCCA TTAACCTGTT GGAATCTGTA AACCTGTAA AGTAGGTTA GTTCTCTGTA GTCCTCTG
CGGAGGGGA AGTCTTGA AATCTCTGCT TCTGACCGA AGTACCOCG CATACGGGAG CACTTTTATG CACTATGTA GAAAGGTGTG GTATTCGAAC TGCOCGACTT
TTAGTTGGTC ATCCAACAT AGGTACGAGA AACCTTCCCC CACTTCCATTT CTTAGGTTAA GACCATCCAT CTGTAGTAC TGTCATGTGT CGGGAAAAAG CCGGAAARAGG
CCTGTTGAAT CGGTTGGA AATGTTACTT GTGCTGTTT GTCACCCC ATCCCTGGAT ATCCCTGACT CTTACTGTAG TGGTTGTAA GAAAGCAACT GAACGGGTTA
CTTCCCGACT CCGGTAAGA AACGTTTAAG TTAAGTAAG AGTAGGTTA TAAAGCCTGT TAACCTGGCT TACTGGTTA ATGATTTAG ATGACCCCGG ATGATTTAG
AGAGGGGAAA ATTTAATAAC CAGTTTATA TTCTATGGTG GGGTCCCTGA GGTTCTTTAT ATGAGCCCGG ATGATTTAG GGGCCCTGTT TGTCTGCGG CTCTAATAAT
AGGTAATAA AGTACGGCGT GTGGTCGGTA AGCCCGGTGA GTCGAAGATA GACCCGCGAA TGCTGCCCG TTACAATGT GTTAAAGCC CCATAAAAC GTTTGACTTA
GTTTAGTCCC CTATATATCT TTGGGTTTAA TACTGGTTGA ACCCCGTTTC CTCTTCTGTA CCGGAGGTGC GGGTATGCTA CGACGTTAAA ATTTAGTCAA CAGGACGCTA
CAGCACGGTC GAAAAAGTGT TGTGGTGGTT GTTTAAAAA ACAAATACG GTTGTGGTT

TTGACTAGTA TACCCAGTAT CTACTTTTGT ATTACGGACC GAGCCCAAGT ACGTTGTTCCT CGAGAACGCT TTAAGTTCCT AAATCCGCAC CCATCTATGT GACACTGGTG  
TAGTATCCAC GTCCGAACCG ACCCTAGTTA CCACATCCGA GGTAGTGTGG CTAGCGGTAC TATTTCCGAG TCCGTGGACAG GTATGCCAAC CATGACAAGG TTTTGAAC  
GGGGTCCACAG CCCCGGCGCT CCGTGGTTGA GAAGAAATAT TTTAAAGAAA TACCTTAATT GGTGTCTGCC AGGCTTTGCC TTTTGGAGTT AGTCGACAGA GGGCCGCTAG  
TTCCCTTTTAC GTTTTCGTTCC GGGCAACTTC GGGTGTCCCC CCTCGACTAA GTTTAGTGGT TCGCTGTGGT TACCCCAAAA ACTGTCTTGA AACTCGTGTG GTTACCACC  
CGTCAATCCGA CGAAGTCCAG GGCAGTAGTT TGGTTGGGTT GGGTTGAGTA GCGGTTGCCAG CCGGTTGCCAG CCCTGTCCAG CCCTTAAGG TCGCCCAACA  
GAAAACCACC TAGTITAGAC GTGTGAAAAT TCCGAAACTG GTGTAGTTCG TACCAAGCAA TGCTGTGGT AGCTCTACAG TCGTGTACAA ACCCGTGGG AGTAATAAGC  
CAACAAAAGG GTCAGGGGCC CCGCGGTCAA TGTTTGACG TGGACTCCTA CAACGACAAA TGCTGTGGT AGCTCTACAG TCGTGTACAA ACCCGTGGG AGTAATAAGC  
TATTAGTGGT AGGCCCTCTA GTTGGTGAAA AGTTCCTAC AGAGCAGAGT GCCTGTGGT AACTACTGTT GCGGTTGAGC CGAGCTGGT ATTCGACAG TGAACAGCA  
AGTGGCTGTC GTTAGTTGA AAGCCCCCTA CCATGCCCTG CGTGTAGTTA TCATACAGTT CAAGACAAA ATTTAAGTAC CGACTCTAC TCGTGCAGG TGCTCCGGA  
CTAGGATGTC GTACGGGTGT ACCCTCATGT TGTCGCCGTA TGTCCAGG CACCTACAGG TCAAAATCCG TTTTGGACAG TGTTCACACC GACCCCGGG  
CGTTACACCG GTCCAAGATG TAAAAAGAA AAGAAGCGG CTTTCCGCCA AGGTGTACAG TTCCCTTTTC GATCCAGGCC CTTTAGACTT GGTCCCCTA GCGCCGCTGA  
TTCGTTATGG TGTCTGTTCC AGTTTTCATGG TTGAGAGTAG CACTAGTGTT GTGGCGGTC CGAGTAGTT CTATCGGGG CGTGTACGCT TTATAAGCGG CTGGCTCGAC  
CGGAAGTAAA CTATGCCGAT CGCGCGCTGC TATGCTAGGG CAACGCCGTC GTTCCTGGAT GCGCGTCTAT ACGCGTGGG CCCTTCAAAC CCAAGACTCC ACTACTGTCA  
AAACCTGGAC AGACAAAAT GTGGGTTGTA ATTTAGTGCC ACCAGTGTGT TCTTTACCAT CAAGCCGTTG AGTAGTGAGA GCAACATCTT TTTGTTGCAG TTTTCCAGC  
GGTTGTTGCC AACATTTAGT TTCTCAACGA GGTAAGATC CCTCAGAAGC TTTTGTGTCG GAGGTTAAA CACTGGTAAA GACTTACGAC TATTGTGTA CACAACCGT  
ATATAATAAT TTGATAATTG ACACAACGCAT TACCACCATT GTCTTAGAGG AACTACTGTT GCGGTTGAGC ATAGTTTCGT TAGTGAGTAC CGAGCTGGT GCGATACCGT  
TGTTAAGTTT TGCACTCCTC CTTAGGTAGG AAGTCGTGGT CCGTTAAAGT TATCGTCAA GCAACATTTG GTGGAATATT AGCCGCCCA ACTGGCCACT GTTCCAGCA  
CCAGTGGGAT GTCGTGTGGC GGCGTGTAGG ACCCCACTTA AATAAGAAA GATGAGGCTA CCCCCAAGT TGGTGTGCTA CCCCAGAGG AACTTTGCAG TACTTGCAG  
GGCGACTCTC GCCTGGAATC GCGATCACGT TGTTCATAA ACCAGGGTC ATAACAGAGC TTTTGTGTCG GAGGTTAAA CACTGGTAAA GACTTACGAC TATTGTGTA CACAACCGT  
AACGCTCACG AGGAGTGTCC GAGCAAAAGT GTCCCGTAGT TACAGCCCTA CGCTAACGTT GGTGCACCTA CCGTTAGACG GTGGCAACTA CGGACGATG CTCAGAAATC  
GCCCTACACA GTTCGCCGATG TTGTCCGCTG CAGACTCGGG TTAGTCCTAGG TTAGTCCATG GAGTCCATG CATCTCGTGT CATCTCGTGT GTCGCGGAG CGACCTGGT  
AGTTTAAAAA CATTCAGGAA AAGCCGACGC GATCGCTGG GTACAAAAA ATCCGTAAGAA CTTCAAGAAC GCTTGTAGTA TTCTTATGTA TTTGTTGTCG CCGCTGACG  
GCCTGTATG TTGAGTTTGG CCACCCCTAC TCAGACTCGG TGGTTTCAA GCTTGTAGTA TTCTTGTGTA TTCTTGTGTA TTCTTGTGTA TTCTTGTGTA TTCTTGTGTA  
AGAAAAGTCC ATAATTGATA GTTTGCGCAG AATAACAGAG TTTCGCCCA TAACTTGTTA TGTGGGCTA CAGCTATTCC TAGTGGAGAA TGTCCATCAT TAACACTTCC TCCGAGAAAC  
GACCCCGTGA GTTGGGAGCC GTACCTCGTA CGACACTATT AGACAAGACT GTTGAAGTTA GTTAGTCTTA AGATCAGTT CAGATTAGTT GCGCGCCCT AAGTTTCCAG GTGTGTTGG  
CCACAACGCG CCGGCAACC CTGGCCCTTG CTTGAGCGGT AGGATGGTCC ACCTGACTGAC TAGCTACTGCT AGGTAAGTGA CTTCCCAAGT CTCCCAAGT CTGTTTGGT  
CAAACCGGTC TTCGTTTACG GCCCCAAAAG GGTAGCATGC TGCCATGAGT AGTTCCTCT AACTGGAAGT ACTGGAAGT CAACCTGAGT GGTTACCTCT TTCAGGCGGA CACTATGA  
TGTTAAAAAT GGGCGAGGAA CCTGTGTCAG TGTGTTGGAG ATCTCGTCCG AAGTTCCGCTA CTTCAACAGT GTCCACAGT TGTCAAATTT CBTCTACTGA CGAGCTTGG  
AGTTACTACG TCCCGTCGAA AAGGAAGAGG AATCCCTCAGT CGTTCCTCAG CGTTCGCTAC ACAGAGTCCC GAGCAGAGTT ACACGGAAGT GACACTGGT TAAGCTCGGA  
AGGCTACCAT TACCAGTGA GTCTCCGAG TACCAGCCCC TGTAGTTTGA GAAGTGTAG GTCCAAACAT CCAOAGTGC TTTGAGTTA TACGAATAAA AAGGTTGTC  
CCATAAAGAG GTACTCCAG AANTGAAAC CCGTACAAA AGACTACTTA AGAATTTTGG CCGGTTAAGC ACTGCAGCT TCCCGCCCT CAGAGCTAA AAGCTGTGG  
TCTTTCCGAA GTATACAGTG TAGTGTGTCAT CCGAATAACG TCTCGAACA CAATGAATGA GACCTGATA CATTGAATGA GCTGGTAAA CACCGTTAAG CACCGCGAA  
AACCGCGCT GCAACAAGA AACTCCAGAA CTGTCCAGT GACCCTAGAA GTGTCGCGAA GAGTTCAGA GCGCTTACG ATGTGTTGT CTGTACTAC GCGCCGTAAT  
GAGCACAGAG TTAATGAGCC TTAGTAGCAA ACTGTAGCAA TACCACATG AGATGCCTA CAGGTCFCA CGGTTGTGTTA CAAAAGAGG CTATGTTGTT CCGCAAGTGA  
AACCCTGTT GTGGAGTCCG TTTCCAGAGA TCGAACCGGT CCTACGAAA ACTCCGAAT TAAAACAGAC ACCGTCGCT GACTCGAGT TCTCGTCTC GTATCCGCT  
GAGTTATTTA AGGAATCGAG GTAAGGAGTA CAACTCCG GACTGTCTG TGAACTTGAG ACTTACGGCA ACATCCTTAA GGTTCCAGT TAGTCTGTCT CACTAGCGGT  
CCCCCGGACA GTTGGGAGT CCAACAGTACA ACGTCTTAG TCTTTGAGC AGTCCGTTA AGTCCGTTA ACAGGAACTG GGCGAATCG TTTTCATGAAG TCCCTTTTAC  
GGTGTATGTT GTAGTCTGTC GATCCCGCG AATTTGTTGG AATCCCTGTT TCGGTTTGT AACTACCGTT CGTTCATATG TTTTATAAG AAACCTCCGT GAAGCGACA  
CGAGCCCAT TTTATTTGGC CATGTCTG ATCACAGAG TTACTCTG GTGACTGTG TGACTTTTGT GACTTTTGT TCTATCCCGT TTTCTTTTGT AATTAGACCG GTTGTGGT  
CTAAGTTTCG ACGATAATGA TCCGCGTGT CGTTGTGAAC GTGGCTTTCT GGACACTTCT GACTTTTGT TCTATCCCGT TTTCTTTTGT AATTAGACCG GTTGTGGT  
CGTTTCTAGA GTAAGTCCIT CTTTTTACGT TGACGTTAA CTCCCTGTTT CCGTGTACT TACTCCGCG TGCTTTTCT TTTCTGTTGT TTTCTGTTGT GGCTCCGTT  
ACGGGAAAGT CCAACTCCGT TTCTAAGATG ATGTCGATGG TCCAGTATC GAGGCGGAG GTGCAAGTCT GCTCATACGA ACCCGTACG TAGTCTACGA CTGTGAGCT  
GTCTCGCGG TATCACGACC GTGTAAGATG AGTAAAGTG TAATGATC GGGTCCAAA AAGTATAC CCAACACTAG TAAGGACTCG GCGCAAACAA ACTTGTCTAC  
ACCTGTGCTT ATTCCAACG TAGGAGTAGT TGATGACCCC TTGGTCTCAG CGGTGTAA TAACTGTTT TCGCTCACA ACCTGGGTTT TTTTGGTTG GAGGAGGTG  
AACCGGACAG GTTATTCGAG CCGGAGAGTA TAGACTGTG ACGTGGAAAC CCGCAGTGG GCGCAAGT GCGTACGACT GCGTACGACT TAGTGTGTTG GAGACTTCC  
AGTGTCTGAC TCTCTCGCG GTGTCCTTT TGGATAGTG GGGATCGGAA TGGTAGACT TCGTCCAGC ATGCTCAGT ATGTTGTCOC AACCGTACT ATCACTTTG GGGTGGCTG  
GGCTACGTC TGGCTCTG TGATTGTGT GAAATCTGT TCGGATACA GGAATCCGAGT GGAATCCGAGT GAAAGTCTT CAAACACTTT CAAACACTTT CAAACACTTT  
TCAGTTGCGG CTAGCCGCT GACCGGGTT CATCCAGGA AGACCTCGC TTCCGTTTCT CGAGCGCGG CCCCCACC TCCAGTGGT CACGTTAGT ATGCCAATT  
GGTCCGGACA TCTCTACAC CGTGCGGGT CTTTGGTCT ACTTGTACT TAGTGGGTT GACTTTGTT GTGGTTTTT AAGTAGCGT CGTTCTTTT AAGTGGGTT  
TGGTCTCACT GCGCCAGCT CAGTGGAA GTGTTCGTT TACCCTIAG TTAGCTATA TTAGCTTCT TTAGCTTCT TTAGCTTCT TTAGCTTCT TTAGCTTCT  
CTGAGTTGG TTGTTCTTTT GAAACCTTA TTTTACCGG TGACAAGCTT AGTCCCTGTA GATTTCTCGG GTTAGTTAGT GGTAGTCTCG GGTAGTCTCG GGTAGTCTCG  
CGGTAGGAG TCTCATGCT CCAACCTCA ATGTTGATTG AGGGACCGT TGGTGTCTA TTCTCTTACC TAAGAAGTGT GTCCCCAGT ACCAAGGAG AGGTTGTGAG  
TGTCTGATC TGTCCGTTG TCTATACTCT TTTAATTCTG AATGTATGTT AGGGTCCGTT GTTTTTAGG GTTTTTAGG GTTTTTAGG GTTTTTAGG GTTTTTAGG  
TCGTTTTAGT GGCACGGGT GTTAGTCGCT GCTGTGGGTT ATTTGGTGG GTACACCTG CCTCATCCAT TGTCCAGTT GGTGCTACT CCGTGCATC AGTTAAGT  
GCAACCCCTT TTTACCCG GAAACTCTT GAAACCTTTG GAAACTTTG AAGCCCTGG GAGTGTAGG CCGTGTAGG TTAACTAGT TAACTAGG TGGTGCAT TGGGAATC GTGAAGTT  
CGAACGCAC CAGTCTAC ACCCCCTGG TGACAGAACA ACTTTAAA AAAAAGAAGA FCCAATGACA TCGAATGACC GAACAGTCA TGGACACTC CCGAGTGGAA  
TGGGACGGT GTCCGAACT ACCGCCATCG GTGTTCTAC CTTAGTGGG CCCCACGTA TATGTACAG GTTTGACGGA GATCAACA AGCTCGGTG AAAACGTC  
CCTTTTGAAGA ATTTACATGG AGTACCCGAG GGTGAGAGT TGTGAGAAA TTTGTGAAA ATGTGAGAA CCGTCCGAGG TGCGGTCTCG GTGGCTCTG TGTGCTCAG GTTCAGTTG  
TGGTGGTATT CCGTCTCTT CATTTTGGT ATTACGGCAC CGCGGGGTC GTGCGTCTG TGTGTTGAT TGTGTTGAT TGTGTTGAT TGTGTTGAT TGTGTTGAT  
AGCCACTGTT GGTAAAGGTC TTAGGCGGAG TTCTCGACGG CGTGTGAGT GGTTCTGTG GTTTCTTGTG TTTGGTCTC TTTGGTCTC TTTGGTCTC TTTGGTCTC  
GTTGTGTTCT TTTTGTCTT ACACATTGT ATTAGTGGGG TTAATTGAAA AACCGCTCA ACCGCTGCT CCGTGTGAG GGTGCTGAG GGTGCTGAG GGTGCTGAG  
GACCACCAAG TTGTTCTGGT ATTTGAAAAC TGGTGGCCTA GGGTGGCCTA GAAACTATCC ACCATAGG AGGAACTAGA AGGAACTAGA AGGAACTAGA AGGAACTAGA  
GGGCAGAAC ACTTTGACC ACGGTTCCC ATTAGGATC TAGGGGAC CTGCCACC CGAGCCGAG CGAGAAATCG CTTGAAAGT GTTTAAGAGA GAGCGGTCTC  
GACAGCCGCG AGAAGAGCAG CAGCGAGTTG TCTCCGGTT TGGGTTACGA GGACCCACTC CTCGCGACC TCGAGTGGT GAAACACTG CACAGGAGAG TTAAGGTTT  
CTAATCAAAA CGAAAAGT GACTGTCTGTT GAGCACGTA CACAAGCTA CTCACCGTCT TGTTAAGTC CACAACCTG TCGCTAGTGC ACCATGACCA GCTTATGACA  
GGTGTACAC TCGCAGTAA GATCGTGCT ACGGTCGCG GGTGGTCC CCTCAATACC AAGAGTTTT ACGGGCTGT AGGAGTTGCT ACCRAGTGT AATTTGGGTT  
CTAAGTCAAC TAAAGTCTA CTATAGGAAA TCATACTGG GGACGAGAC GAAAATTCG CGACTGCAA AACAGCGCC CCCCCTCCG CCAACCCGCA ATATCCCATG  
CTATCTCGG GGGGCCGAGG GAGTAGTGGT GAGTAGTGA GATTAGCGGG GTTAGCGGG GATTAGCGGG TCCCTGAGG AAAATCCGCC CCAATCCAAA CCCCCTGACC  
AGAAGAAAAA GAATAACGCG AACCGAATC CGAAGCTGA GAAAGTTCGTT GAAGTTCTGT GCACTCTGC ACAGAGTCTA GAGTACCCG GAGGATCCCG  
AGGTCTCGA GTCTCTGTT CGGACTCGT TCGAGAACAA GTCTAAACCA TCGTCTCAG AATAGTTGG AATAGTTGG ATAGTTGG ATAGTTGG ATAGTTGG  
GAGGAACAGT CTAGACAGG GTAGTCCCT GTCTGGAGT CTGACTTCA CTTGACTTCA CAGTACTTCA CAAGCGGTT TTTAGTACTA CAGTACTTCA CAGTACTTCA  
ACTCAGTGC GACGTTGAA AGGTCTGGA GAGCGGCCA GTCCGTAAA AGTCACTAC CAAGACTCA CAATTTAAG CTTTCAAGG ATTTCAAGG ATTTCAAGG  
ATCGCGAGTG TCCGAGATCG TCAAAGCTG CGTCCGATA CCAAGTGTG GCAAGTGTG TCCAGCTCA CCTTCAAGG AACTTTAAGG AACTTTAAGG AACTTTAAGG  
CGGTGACTAC CCGTAAACAG GACGTTCAAG GTTTAAATAA AGTCAACCCA TCCAGCATG CCAACTAGT GGTGTAGT TAAACTAGG ATTTAAGG AACTTTAAGG  
GTGGCTCCG AGCTAGAGGA AAGGAAAAC AGTGTGAGT CACAATCAA TGTGTTAGT TAACACTTAA TGTGTTAGT TAACACTTAA TGTGTTAGT TAACACTTAA  
CAGTACTGGT TAACATGTA TAGAACAGT CTGTCGGTGA GTGACAGCAG GGACGAGC GAAACTTCG CGACTGCAA AACAGCGCC CCCCCTCCG CCAACCCGCA ATATCCCATG  
TCTTTATGTT ACGTCAAGTT TCCGTGGCAG GGGACCCCT CTACTCGGA TTATGAAGGA ACAAACCTAG GATTTGAAC CTATGAAAC CTATGAAAC CTATGAAAC  
ACCTTTGAG TGGTGTGGG TGGTTTTTTA AGAGATTCC GAGCGGCCA GCGCTTTACA CCGCTTTACA CCGCTTTACA CCGCTTTACA CCGCTTTACA CCGCTTTACA  
ACAAGTCCAA ATTTGTGTG AGATTTGGG GGCATAAGT GGCATAAGT GGCATAAGT GGCATAAGT GGCATAAGT GGCATAAGT GGCATAAGT GGCATAAGT  
GAGGAGTGC CCGTGGAGG TTGACGACG CAGGAGTTG TGA AAAAAGG TGTGAAATG TGTGAAATG TGTGAAATG TGTGAAATG TGTGAAATG TGTGAAATG  
GTTGGTGTCA TTTTCTATT AGTTTTGCCA TCAAGCTTTT GAGTTAAAAA TGTTCCACCA CCGCTTCCTC TTTTAAACA ATAGCTCTGT TTTTAAACA ATAGCTCTGT  
GTACCAAGG CTTGCTTGT TAACTTGA CAACAGGTT TTAAGGACT TTAAGGACT TTAAGGACT TTAAGGACT TTAAGGACT TTAAGGACT TTAAGGACT  
GCCGTGTTT GGAAAACGGT CCCCAGCAC TCCGTGCC CCAACAGG TCCGTTGCC TCCGTTGCC TCCGTTGCC TCCGTTGCC TCCGTTGCC TCCGTTGCC  
GTTATGACG CTGGCCCGT CACTATCCTA TGGTGGCCGT TGGCCACTC GCTCTGACT GCTCTGACT GCTCTGACT GCTCTGACT GCTCTGACT GCTCTGACT  
CTCCGCTGTT TCAACTCAGC TGGGACTTTC AACACTAAG AACTAAGG AACTAAGG AACTAAGG AACTAAGG AACTAAGG AACTAAGG AACTAAGG AACTAAGG  
CCCCATACAT TATCCCGGGA GCGGCACTC ACCACACCGT CCGGATATT TGTGGGTT TGTGGGTT TGTGGGTT TGTGGGTT TGTGGGTT TGTGGGTT  
GATCCGTGTC ACACGTTTCA TGAACGGGA CAGGGGTTT GTGAGATCT AACCCTCAT TCAACCTCAT TCAACCTCAT TCAACCTCAT TCAACCTCAT TCAACCTCAT  
CTCACTGTC GCGCTTTACA ATCTCTCCCT CCATATCAT CTCGAAGCGA TGATAGGAGT TCAACGCCCC GTGACAGGGT CACCAGGAA CTTTATCTGC CAATGGCCTG  
CCCAAGTGTG TTAGAAAAGC CTTA

## 2.1 Dyadic intersection of a sequence and its inverse

The following proposition holds for any type of alphanumerical sequence.

### Proposition 2.1.1

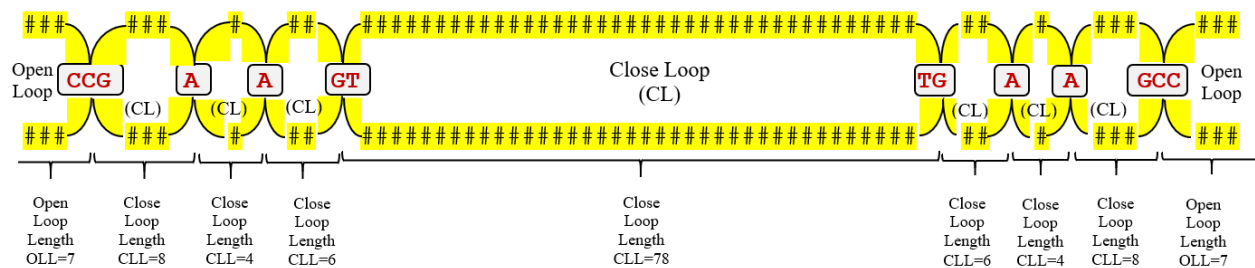
“Given any arbitrary numerical or alphanumerical sequence  $x_1$  of length  $N$ , one may obtain its inverse as  $x_1^{-1}$  of length  $N$ . The sequence  $x_1$  and its inverse  $x_1^{-1}$  are pairwise intersected and all matching symbols are identified. Two matched symbols are fused as one. The sequence of all fused matched symbols forms a palindrome sequence”.

#### Example 2.1.1

Comparison of the first 70 nucleotides  $s_1$  of the virus genome S1 and its inverse  $s_1^{-1}$

$s_1$	ATT <b>CCG</b> AAAA <b>GATTGT</b> GTGA ACCCGT <b>GCGG</b> TAACCGTCTA TTTCAAGGAG CCA <b>CTGGGAC</b> <b>AGTGCCCGC</b>
$s_1^{-1}$	CG <b>CCGGTGA</b> <b>CAGGGT</b> CACC GAGGAACTTT ATCTGCCAAT GGCGTGCCCA AGT <b>TGTTAG</b> <b>AAAAGCC</b> TTA

The pairwise intersection of  $s_1$  and  $s_1^{-1}$  denoted as  $s_1 \cap s_1^{-1}$  is **CCGAAGTTGAAGCC**. This sequence **CCGAAGTTGAAGCC** is a palindrome sequence. The unmatched symbols are denoted by the don't care symbol #. Now, the resulting pair of sequences is represented as:



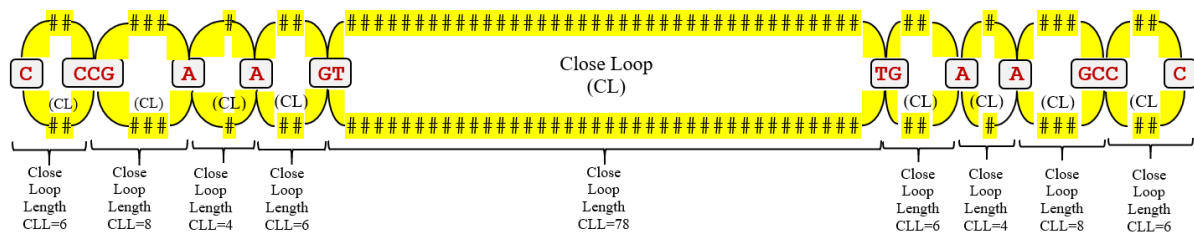
where # refers to don't care symbol, OLL refers to 'Open Loop Length', and CLL refers to 'Closed Loop Length'. The value of an OLL is calculated as  $2n+1$ , where 'n' is the number of don't care symbols present in the pair of sequences. The value of a CLL is calculated as  $2n+2$ , where 'n' is the number of don't care symbols present in a single strand. The OLL and CLL values calculated for the above example pair  $s_1 \cap s_1^{-1}$  is given as  $\{OLL=7, CLL=8, CLL=4, CLL=6, CLL=78, CLL=6, CLL=4, CLL=8, OLL=7\}$  or simply written as  $\{7, 8, 4, 6, 78, 6, 4, 8, 7\}$ . Note that the value of an OLL is always an odd number and the value of a CLL is always an even number. **Note that the sequence pair  $s_1 \cap s_1^{-1}$  is an open-ended pair.** Now the graphical inverse of the sequence  $s_1 \cap s_1^{-1}$  denoted as  $(s_1 \cap s_1^{-1})^{-1}$  is **CCGAAGTTGAAGCC**. It is observed that  $s_1 \cap s_1^{-1} \stackrel{\oplus}{=} (s_1 \cap s_1^{-1})^{-1}$ , where the symbol  $\oplus$  denotes the binary relation 'graphical equivalence', meaning the sequence is a palindrome. Moreover, the sequence of loop values also exhibits palindrome property. In this case, the maximum CLL value is 78 and the minimum CLL value is 4.

#### Example 2.1.2

Consider the sequence  $s_2$  given below and its inverse  $s_2^{-1}$

$s_2$	CTT <b>CCG</b> AAAA <b>GATTGT</b> GTGA ACCCGT <b>GCGG</b> TAACCGTCTA TTTCAAGGAG CCA <b>CTGGGAC</b> <b>AGTGCCCGC</b>
$s_2^{-1}$	CG <b>CCGGTGA</b> <b>CAGGGT</b> CACC GAGGAACTTT ATCTGCCAAT GGCGTGCCCA AGT <b>TGTTAG</b> <b>AAAAGCC</b> TT <b>C</b>

The pairwise intersection of  $s_2$  and  $s_2^{-1}$  denoted as  $s_2 \cap s_2^{-1}$  is **CCCGAAGTTGAAGCCC**. Now the graphical inverse of the sequence  $s_2 \cap s_2^{-1}$  denoted as  $(s_2 \cap s_2^{-1})^{-1}$  is **CCCGAAGTTGAAGCCC**. It is observed that  $s_2 \cap s_2^{-1} \stackrel{\oplus}{=} (s_2 \cap s_2^{-1})^{-1}$ , where the symbol  $\oplus$  denotes the binary relation 'graphical equivalence'. The sequence **CCCGAAGTTGAAGCCC** is a palindrome sequence. If the unmatched symbols are replaced by the symbol #, then then the resulting pair of sequences is represented as:



where # denotes don't care symbol and CLL refers to 'Closed Loop Length'. The value of a CLL is calculated as  $2n+2$ , where 'n' is the number of don't care symbols present in the pair of sequences. The CLL values calculated for the above example pair  $s_2 \cap s_2^{-1}$  is given as  $\{CLL=6, CLL=8, CLL=4, CLL=6, CLL=78, CLL=6, CLL=4, CLL=8, CLL=6\}$  or simply written as  $\{6, 8, 4, 6, 78, 6, 4, 8, 6\}$ . **Note that the sequence pair  $s_2 \cap s_2^{-1}$  is a close-ended pair.** In this case also, the maximum CLL value is 78 and the minimum CLL value is 4.

**On the notion of ‘Self Similarity Index (SSI)’**

Let us consider the sequence  $x_1 = 1, 2, 3, 4, 5, 6, 7, 8, 9, 10$ . Now  $x_1^{-1} = 10, 9, 8, 7, 6, 5, 4, 3, 2, 1$ .

Both the sequences are compared. It is observed that not even a single match is found. The number of elements in the sequence is 10. One finds the ratio of ‘0’ to ‘10’ to be ‘0’. For convenience, one would call this ratio as ‘Self Similarity Index (SSI)’.

$$\begin{matrix} x_1 = & 1, & 2, & 3, & 4, & 5, & 6, & 7, & 8, & 9, & 10 \\ x_1^{-1} = & 10, & 9, & 8, & 7, & 6, & 5, & 4, & 3, & 2, & 1 \end{matrix}$$

So, SSI of  $x_1 = 0$ . Let us consider another sequence  $x_2 = 10, 10, 10, 10, 10, 10, 10, 10, 10, 10$ . Now, graphical inverse of  $x_2$  is  $x_2^{-1} = 10, 10, 10, 10, 10, 10, 10, 10, 10, 10$ . So,  $x_2 = x_2^{-1}$ . All elements of  $x_2$  match with all elements of  $x_2^{-1}$ . In this case, SSI of  $x_2$  is 1. This means that the SSI values range from 0 to 1 for all sequences. Therefore, one could attribute a real number SSI value to any sequence, from the cantor set [0.1]. Now the parameter SSI is formally expressed as  $SSI = (n/N)$ , where n is the number of matches found and N is the number of elements present in the sequence.

*Proposition 2.1.2*

**One can interpret a discrete Hilbert space  $l_N: (N \rightarrow \infty)$ , which is a sequence space, as a disjoint union of potentially denumerable subspaces, each having a unique SSI value.**

This proposition is a self evidential truth.

*Proposition 2.1.3*

**Every SSI based subspace of the discrete Hilbert space  $l_N: (N \rightarrow \infty)$  is potentially denumerable.**

This proposition is a self evidential truth.

*2.2 Dyadic pairing of nucleotides in a sequence and its inverse*

The following proposition holds for any type of virus genome sequence.

*Proposition 2.2.1*

Given any arbitrary nucleotides sequence  $x_1$  of length N, one may obtain its inverse as  $x_1^{-1}$  of length N. The sequence  $x_1$  and its inverse  $x_1^{-1}$  are compared and all of the natural pairing of nucleotides are identified. The sequence of all such natural pairs of nucleotides  $\langle A, T \rangle, \langle T, A \rangle, \langle G, C \rangle$  and  $\langle C, G \rangle$  forms a palindrome sequence.

**Example 2.2.1**

Natural Pairing of  $s_1$  and  $s_1^{-1}$

$s_1$	ATTCCGAAA GATTGTGTGA ACCCGTGC GG TAACCGTCTA TTTCAAGGAG CCACTGGGAC AGTGGCCCGC
$s_1^{-1}$	CGCCCGGTGA CAGGGTCACC GAGGAACTTT ATCTGCCAAT GGC GTGCCA AGTGTGTTAG AAAAGCCTTA

The sequence of natural pairs of nucleotides along with don't care symbols is given below.

##### $\langle A, T \rangle$ ##### $\langle G, C \rangle$  $\langle T, A \rangle$  $\langle G, C \rangle$ ### $\langle C, G \rangle$  $\langle C, G \rangle$  $\langle T, A \rangle$  $\langle G, C \rangle$ ### $\langle T, A \rangle$  $\langle A, T \rangle$ ### $\langle C, G \rangle$  $\langle G, C \rangle$ ### $\langle T, A \rangle$  $\langle A, T \rangle$ ### $\langle C, G \rangle$  $\langle A, T \rangle$  $\langle G, C \rangle$ ### $\langle G, C \rangle$  $\langle G, C \rangle$ ### $\langle C, G \rangle$  $\langle A, T \rangle$  $\langle C, G \rangle$ ##### $\langle T, A \rangle$ #####

The sequence of natural pairs of nucleotides along with loop lengths is given below, and this is a palindrome sequence.

15 $\langle A, T \rangle$ 18 $\langle G, C \rangle$ 18 $\langle T, A \rangle$ 8 $\langle C, G \rangle$ 8 $\langle C, G \rangle$ 4 $\langle T, A \rangle$ 4 $\langle G, C \rangle$ 8 $\langle T, A \rangle$ 6 $\langle C, G \rangle$ 6 $\langle C, G \rangle$ 6 $\langle T, A \rangle$ 8 $\langle G, C \rangle$ 4 $\langle G, C \rangle$ 8 $\langle C, G \rangle$ 8 $\langle C, G \rangle$ 4 $\langle A, T \rangle$ 18 $\langle T, A \rangle$ 15

*Proposition 2.2.2*

Given any arbitrary nucleotides sequence  $x_1$  of length N, one may obtain its inverse as  $x_1^{-1}$  of length N. The sequence  $x_1$  and its inverse  $x_1^{-1}$  are compared and all of the unnatural pairing of nucleotides are identified. The sequence of all such unnatural pairs of nucleotides  $\langle A, C \rangle, \langle C, A \rangle, \langle T, G \rangle$  and  $\langle G, T \rangle$  forms a palindrome sequence.

**Example 2.2.2**

Unnatural Pairing of  $s_1$  and  $s_1^{-1}$

$s_1$	ATTCCGAAA GATTGTGTGA ACCCGTGC GG TAACCGTCTA TTTCAAGGAG CCACTGGGAC AGTGGCCCGC
$s_1^{-1}$	CGCCCGGTGA CAGGGTCACC GAGGAACTTT ATCTGCCAAT GGC GTGCCA AGTGTGTTAG AAAAGCCTTA

The sequence of unnatural pairs of nucleotides along with don't care symbols is given below.

$\langle A, C \rangle$  $\langle T, G \rangle$ ##### $\langle T, G \rangle$  $\langle T, G \rangle$ ##### $\langle A, C \rangle$ ### $\langle C, A \rangle$ ##### $\langle G, T \rangle$  $\langle G, T \rangle$ ### $\langle A, C \rangle$ ##### $\langle C, A \rangle$ ### $\langle T, G \rangle$  $\langle T, G \rangle$ ### $\langle A, C \rangle$ ### $\langle C, A \rangle$ ##### $\langle G, T \rangle$  $\langle G, T \rangle$ ##### $\langle G, T \rangle$  $\langle C, A \rangle$

The sequence of unnatural pairs of nucleotides along with loop lengths is given below, and this is a palindrome sequence.

$\langle A, C \rangle$  $\langle T, G \rangle$ 22 $\langle T, G \rangle$ 12 $\langle T, G \rangle$ 12 $\langle A, C \rangle$ 4 $\langle C, A \rangle$ 14 $\langle G, T \rangle$ 6 $\langle G, T \rangle$ 6 $\langle A, C \rangle$ 10 $\langle C, A \rangle$ 6 $\langle T, G \rangle$ 14 $\langle A, C \rangle$ 4 $\langle C, A \rangle$ 12 $\langle G, T \rangle$ 4 $\langle G, T \rangle$ 22 $\langle G, T \rangle$ 4 $\langle C, A \rangle$

This computational tool of dyadic intersection of a sequence and its inverse is applied to the actual virus genome sequence S1: >NC\_001639.1. The ‘single-strand sequence’ obtained after point-wise intersection of S1 and its graphical inverse  $S1^{-1}$  is given below. The don't care symbol # represents those characters which are not matched.

#T##G###	G###GT#T#	#####	T#####CA	#T###G##	#####G##	#####	#A#####	#A#T#G#C##	#T#T#T#T#	T#C#T##T#	T#C#T##T#
#####C#T#T	C###GT##GC	A#CCT#	C#T#####	#####A##	AG##GG##T	C###G#T#C	#####AC##	#####TG#A#	#####TT#T#	#A#####C	#A#####C
#####AT#T	#####TG#	#####T#C#C	#G#####T#	C#CC#T###	C#C#####	#####G#T#C	#####C#C#T#G##	#####G#A#A	#####G#A#A	#####G#A#A	#####G#A#A
#####AG#C#	#####TA#	#####C#	#####T#	C#T#####	#####C#G#	#####A#	#####A#	#####A#	#####A#	#####A#	#####A#
#####AA#T#T#	C###AG###	G#####G#	#####A#	#####C#	#####T#G#	#####A#T#	#####A#T#	#####A#T#	#####A#T#	#####A#T#	#####A#T#
#####AC###	#####CA#	#####G#	#####T#C	A#G#####	#G#####C	C#TCC###	TTG#A#####	#####T#G#	#####CC###	C#G#####	C#G#####
#####CA#A	AG##TG#T	#####T#	#####G#T#	#####G#T#	#####G#T#	#####G#T#	#####G#T#	#####G#T#	#####G#T#	#####G#T#	#####G#T#
#####T#T#	#####T#	#####T#	#####A#G#	T#GT#G##	#####C#	#####T#	#####A#C#C	#####T#A#	#####A#T#	#####A#T#	#####A#T#
T#G#####	#####T#A#T	#####AC###	#####T#A#T	CC#####A	#####A#	#####C#	#####C#	#####G#	#####G#	#####G#	#####G#
#G#G#A#	#####CC#G##	#####G#	#####G#T#	#####A#G#G#T#	#####G#C#C#C	GC###G##	#####G#	#####G#	#####G#	#####G#	#####G#
#####G#	#####T#A#	A#A#####	#####A#	#####G#C#T#	#####T#	#####T#	#####G#	#####A#	#####A#	#####A#	#####A#
#####GT#G#	#####G#G#	T#####CG#T	A#####T#T#T	#####C#	#####C#A#	#####T#	#####A#	#####T#	#####T#	#####T#	#####T#
#####G#G#C	#####T#G#T	T#C#C#A#	#####A#	#####T#	#####T#	#####T#	#####T#	#####T#	#####T#	#####T#	#####T#
A#####	#####TT#T#	T#####	#####T#CC###	#####C	#####T#	#####T#	#####T#	#####T#	#####T#	#####T#	#####T#
G#####	T#G#	#####G#	#####G#T#	T#T#####	#####G#	#####G#	#####G#	#####G#	#####G#	#####G#	#####G#
#####AA#A#	#####C#G	T#####GC#C	T#####T#T#	T#####A#C#	#####T#G#C#	#####G#AA#	#####C#C	A#####G	#####G#	#####G#	#####G#
AAC#C#T#A	#####G#	T#####CG#T	A#####T#T#	#####G#	#####G#	#####G#	#####G#	#####G#	#####G#	#####G#	#####G#
TG#####A#	#####CGT#G#	C#####T#	#####G#	A#####C#	#####T#	#####T#	#####T#	#####T#	#####T#	#####T#	#####T#
#####T#	#####G#	#####C#T#T#	#####T#C#	#####A#G#A#	#####G#T#	#####G#A#T#	#####C#	#####C#	#####C#	#####C#	#####C#
A#C#C#G#T	C#####C#	#####A#T#	#####T#T#C	#####C#A#	G#AA#####	G#AGT#C#	#####T#	#####T#	#####T#	#####T#	#####T#
#####GC#A#	#####G#A#A#	A#####	#####C#C#T	#####A#	#####G#A#	#####C#T#G#G	#####G#	#####G#	#####G#	#####G#	#####G#
#####C#T#	A#T#C#####	#####T#G#C#	#####T#G#	#####C#C#	#####A#C#C	A#GC#####	#####G#A#	#####G#A#	#####G#A#	#####G#A#	#####G#A#
#####C#	#####CA#	AAT#T#GGT	#####G#T#	#####T#	#####T#	#####T#T#	#####C#	#####C#	#####C#	#####C#	#####C#
#####A#	#####C#	G#A#C#T	#####CA#	C#T#G##	#####C#	#####T#	#####C#T#	#####C#T#	#####C#T#	#####C#T#	#####C#T#
C#####A	#####G#AA#	A#####	#####C#	#####A#	#####C#G#	#####G#T#	#####C#T#	#####C#T#	#####C#T#	#####C#T#	#####C#T#
G#####	#####G#C#	#####T#TG#	G#A#A#A#A	T#T#####	#####C#	#####C#	#####G#C#	#####G#C#	#####G#C#	#####G#C#	#####G#C#
G#####T#	#####A#AC	G#####	#####TG#	#####G#	#####G#	A#####G#G	#####C#T#	#####C#T#	#####C#T#	#####C#T#	#####C#T#
T#T#T#G#T	GG#G#C#	#####G#T#T#T	T#C#T#T#T	#####C#T#T#T	C#####G#T#	#####G#	#####G#	#####G#	#####G#	#####G#	#####G#
C#CA#G#T#T	T#####	#####T#	T#C#T#T#T	#####C#A#	T#GG#T#C#	#####C#T#	#####G#	#####G#	#####G#	#####G#	#####G#
T#####TG#	#####T#T#	#####T#TG#	G#####C#	#####C#	CC#####C#	#####G#C#	#####G#C#	#####G#C#	#####G#C#	#####G#C#	#####G#C#
#A#T#####	#####G#	#####C#T#T#	A#####T#TG#	#####T#	G#####C#	#####T#	#####T#	#####T#	#####T#	#####T#	#####T#
#####G#C#T#	#####T#T#G#	#####T#G#	#####T#C#	#####G#T#C#G	C#####G#	#####T#A#G	#####G#A#	#####G#A#	#####G#A#	#####G#A#	#####G#A#
#####T#	T#A#A#T#T	A#####G#A	#####C#C#A#	#####C#	#####G#A#	#####C#A#T	#####C#A#T	#####C#A#T	#####C#A#T	#####C#A#T	#####C#A#T
#####T#A#	GC#GT#G#A	#####T#	G#CCT###	#####C#	#####G#	#####C#A#GTA	#####G#A#G	#####G#A#	#####G#A#	#####G#A#	#####G#A#
C#C#CC#ACT	C#C#A#AG#	A#####T#A#	T#####AAG	#####G#	#####G#	#####G#G#C	#####G#G#C	#####G#G#C	#####G#G#C	#####G#G#C	#####G#G#C
A#A#GG###	ATT#T#	A#####T#A	TT#####A	G#G#C###	G#G#T###	T#T###CC#	#####G#	#####G#	#####G#	#####G#	#####G#
A#####A	#####G#T#	#####G#	A#####G#TGA	G#####G#A#	C#####G#A#	#####GG#A#	#####T#	#####T#	#####T#	#####T#	#####T#
T#T#T#G#T	CT#A#	T#G#A#T#	T#####G#G#A	ACC#####CC	C#####G#A#	#####G#A#	#####T#	#####T#	#####T#	#####T#	#####T#
#####G#	C#C#AG#T	G#####T#T#	#####T#	A#####C#	G#T#G#T#T#	#####G#C#	TA#####C#	#####A#	#####A#	#####A#	#####A#
#G#G#T#T#	T#####T#	C#C#T#TG#T	T#####	G#####C#	#####C#T#	#####T#	T#####T#	A#T#C#C#C	CCAT#T#T#T	#####G#	#####G#
#####T#CAC	#####CG#AC#	#####T#G#	#####C#A#	#####G#	#####T#	#####T#	#####T#	C#####G#	#####C#G#	TTT#G#	#####G#
G#G#T#T#C	C#C#CC#G#	C#G#####	#####A#T#T	#####G#	#####G#	#####G#C#	#####G#C#	#####G#C#	#####G#C#	#####G#C#	#####G#C#
#####T#T#	G#T#G#	G#####C#	G#T#G#	#####C#A#	#####T#G#	#####G#G#C	A#####G	#####G#C#	#####G#C#	#####G#C#	#####G#C#
G#G#T#T#T	G#G#T#T#T	C#####G#G	A#####G#T	G#####C#	#####G#T#T	#####G#T#T	#####G#T#T	#####G#T#T	#####G#T#T	#####G#T#T	#####G#T#T
G#####	#####T#	C#C#	#####T#C#	T#C#	#####C#	#####C#	#####C#	#####C#	#####C#	#####C#	#####C#
#####CA#A#	#####C#	C#G#A#	TG#####G	T#####T#	#####A#A#	#####T#G#T	#####A#G	T#G#G#	#####G#	#####G#	#####G#
#####T#G#	#####C#T#	#####T#	#####C#A#	G#G#####	G#T#G#	#####G#A#A#C	#####G#G#G#	#####G#A#	#####G#A#	#####G#A#	#####G#A#
#####GT#C#G#	T#####G#G#	#####C#	C#A#T#C#	#####G#T#T#	#####C#A#T#	#####T#A#G#	#####T#A#G#	#####T#A#G#	#####T#A#G#	#####T#A#G#	#####T#A#G#
#####T#T#	#####C#G#G#	#####C#T#T#	T#####	#####G#	#####G#	#####T#G#	#####T#G#	#####T#G#	#####T#G#	#####T#G#	#####T#G#
#T#T#C#	#####A#	T#####G#	#####G#	T#####C#	#####T#C#	#####T#C#	#####T#C#	#####T#C#	#####T#C#	#####T#C#	#####T#C#
G#C#####	G#####T#T#	A#T#####G	TG#G#####	C#T#T#T#T#T	TG#G#G#C	#####GT#T#	T#T#G#C#G	#####G#	#####G#	#####G#	#####G#
#####A#	#####T#T#C	G#G#####	A#####G#	CA#A#####	#####G#T#	#####T#	#####T#	#####T#	#####T#	#####T#	#####T#
A#A#CT#G#	#####G#	TGGG#####	#####T#	#####C#	#####T#T#	#####C#	#####C#	#####C#	#####C#	#####C#	#####C#
#####G#	#####G#	#####CGA	#####T#C	#####T#	#####T#	#####T#	#####T#	#####T#	#####T#	#####T#	#####T#
A#####A#	#####GA#T#	#####C#G#	TA#G#C#T#	#####T#C#T#	#####C#T#G#A#	#####G#A#A#	A#A#T#G#	#####A#C#	#####A#C#	#####A#C#	#####A#C#
#####T#T#	#####G#T#T#	C#A#G#T#	A#####G#T#G#T	#####T#T#	#####T#T#	#####T#T#	A#A#T#T#T#	G#G#C#C#	#####G#C#	#####G#C#	#####G#C#
#####A#C#G#	#####T#G#C#C#	#####G#G#G#G	C#G#C#C#	#####C#	#####C#	#####C#	#####C#	#####C#	#####C#	#####C#	#####C#
#####A#C#	#####C#	C#A#A#CG#	#####C#	#####C#	#####C#	#####C#	#####C#	#####C#	#####C#	#####C#	#####C#
#####GT#AC	#####G#C#	#####G#C#	#####T#	#####C#	#####C#	#####C#	#####C#	#####C#	#####C#	#####C#	#####C#
#####A#	#####T#G#A#	T#####G#T#G	A#####G#	#####A#C#G#	#####G#G#	#####G#G#	#####G#G#	#####G#G#	#####G#G#	#####G#G#	#####G#G#
#####A#A#	C#A#A#A#	#####G#A#A#	GATC#G#G#	#####A#C#	#####C#G#A#	#####C#G#A#	#####C#G#A#	#####C#G#A#	#####C#G#A#	#####C#G#A#	#####C#G#A#
G#####TG	#####G#T#	#####T#	#####T#G#	#####T#A#A#	T#T#T#T#	#####T#G#	#####T#G#	#####T#G#	#####T#G#	#####T#G#	#####T#G#
#####AAG#C	#####T#A#A#	#####T#A#G#	#####GCC#	#####T#A#A#	T#G#####	#####C#A#T#	#####G#	#####G#	#####G#	#####G#	#####G#
#####C#T#T#	G#G#G#T#	#####C#	#####T#T#T#	#####G#C#A#G#	#####A#	#####C#G#T#	#####G#C#G#T#	#####G#C#G#T#	#####G#C#G#T#	#####G#C#G#T#	#####G#C#G#T#
C#####C#	#####C#C#	#####C#	C#####T#	#####G#G#T#T#	#####G#G#T#T#	#####G#G#T#T#	#####G#G#T#T#	#####G#G#T#T#	#####G#G#T#T#	#####G#G#T#T#	#####G#G#T#T#
CA#C#C#G#	#####G#T#C#	#####C#C#	#####T#C#T#C	#####A#	#####T#G#	#####T#G#	#####T#G#	#####T#G#	#####T#G#	#####T#G#	#####T#G#
T#####A#A#T	#####C#G#A#	#####A#	G#G#G#G#	#####T#	#####T#	#####T#	#####T#	#####T#	#####T#	#####T#	#####T#
#####T#T#G	#####T#G#	#####G#A#G#G	#####CCT#	#####T#	#####G#	#####G#	#####G#	#####G#	#####G#	#####G#	#####G#
G#A#A#A#	T#CA#	#####C#	A#####T#G#	#####T#G#	#####G#T#G#	#####T#C#A#	#####G#T#C#	#####G#T#C#	#####G#T#C#	#####G#T#C#	#####G#T#C#
#####T#A#G#	GT#####A#	#####A#A#	AC#####A#	AG#C#C#	#####G#	#####G#	#####G#	#####G#	#####G#	#####G#	#####G#
#####C#G#	#####C#A#T#G	#####	#####	#####C#T#GAA#A#A	#####A#T#G#	#####A#A#	#####A#A#	#####A#A#	#####A#A#	#####A#A#	#####A#A#
A#C#####	C#C#C#A#	#####G#A#	C#####T#	#####C#A#	G#T#G#A#	#####C#G#	#####C#G#	#####C#G#	#####C#G#	#####C#G#	#####C#G#
G#C#C#G#G	T#####GC#	A#####C#	#####T#G#T#C#	T#T#C#C#G	#####C#A#	G#A#T#T#	#####G#A#A#A#	#####G#A#A#A#	#####G#A#A#A#	#####G#A#A#A#	#####G#A#A#A#
A#C#T#T#T	G#G#####	T#####	#####C#C#A#	#####C#C#A#	#####C#G#T#	A#A#A#G#	#####G#T#C#	#####G#T#C#	#####G#T#C#	#####G#T#C#	#####G#T#C#
A#####T#T#	AG#####A#	#####A#G#A#A#	#####T#	#####T#A#T#G#	C#A#A#	#####A#	#####A#	#####A#	#####A#	#####A#	#####A#
C#####G#	#####G#	#####G#T#A#	#####CG#C#	#####T#T#T#C	A#G#G#T#	#####C#	#####C#	#####C#	#####C#	#####C#	#####C#
GGGT#####	#####G#T#C#	A#####T#G#	AA#CC#####T	TC#####G#	#####C#	#####C#	#####C#	#####C#	#####C#	#####C#	#####C#
G#G#C#####	T#T#####	#####G#	T#T#T#T#T#	#####G#C#	#####G#C#	T#T#T#T#T#	#####C#G#	#####C#G#	#####C#G#	#####C#G#	#####C#G#
T#A#T#T#	#####G#	C#G#G#T#	C#####T#C#	TGG#T#	#####C#C#	#####T#T#T#	#####C#C#	#####C#C#	#####C#C#	#####C#C#	#####C#C#
#####T#A#A#	#####C#	T#####	#####A#	#####C#G#	#####C#G#	#####A#T#G#G#G	#####A#T#G#G#G	#####A#T#G#G#G	#####A#T#G#G#G	#####A#T#G#G#G	#####A#T#G#G#G
C#C#G#G#C	C#####T#	#####A#G#	T#C#	#####G#A#	#####G#A#	A#G#	#####G#A#	#####G#A#	#####G#A#	#####G#A#	#####G#A#
#####G#G#	#####T#G#C#	#####G#C#	#####G#	#####G#G#C#A#A#	#####G#G#C#A#A#	#####G#G#C#A#A#	#####G#G#C#A#A#	#####G#G#C#A#A#	#####G#G#C#A#A#	#####G#G#C#A#A#	#####G#G#C#A#A#
#####T#C#	#####A#	T#####G#	#####G#G#G#G#	#####G#G#G#G#	#####G#G#G#G#	#####G#G#G#G#	#####G#G#G#G#	#####G#G#G#G#	#####G#G#G#G#	#####G#G#G#G#	#####G#G#G#G#
#####C#T#	T#####T#	#####G#C#	G#####T#	#####T#G#T#	#####T#G#T#	#####T#G#T#	#####T#G#T#	#####T#G#T#	#####T#G#T#	#####T#G#T#	#####T#G#T#
#####G#G#G#	T#G#C#T#T#	#####G#	C#####	#####T#G#	#####T#G#	#####T#G#	#####T#G#	#####T#G#	#####T#G#	#####T#G#	#####T#G#
G#C#G#C#	C#C#G#C#	#####G#G#G#G#T	#####T#T#G#C#	#####T#T#G#C#	#####T#T#G#C#	#####T#T#G#C#	#####T#T#G#C#	#####T#T#G#C#	#####T#T#G#C#	#####T#T#G#C#	#####T#T#G#C#
#####CA#A#	GC#C#A#C#T#	#####G#	TACCC#####	T#A#A#T#T#	#####T#T#C#	#####T#T#C#	#####T#T#C#	#####T#T#C#	#####T#T#C#	#####T#T#C#	#####T#T#C#
C#C#C#T#T#	T#####T#	G#####C#A#	T#T#T#C#	#####C#	A#####C#	#####A#A#G#G	#####A#A#G#G	#####A#A#G#G	#####A#A#G#G	#####A#A#G#G	#####A#A#G#G
#####G#T#GA#	C#####G#	#####C#CA	#####T#	#####A#A#	G#G#C#A#	T#####T#	#####T#	#####T#	#####T#	#####T#	#####T#
GT#T#AT#	#####T#T#	T#T#####	TT#####	A#####	A#GT#####	#####T#T#	#####T#T#	#####T#T#	#####T#T#	#####T#T#	#####T#T#
#####T#G#	#####A#	#####A#A#	#####GCC#T#	T#####							



6C18G8T10C	4TG4T18C2	A4T8C10CC	0G14G24G4	8TCC10AG6	22A4T4C4C	40C8A6A4A	G10A6CG10	4A6CC12GT	G8T24C4CT	A4G4T10AA
4A4T14C8T	4D2T812A1	C10C6T46C	6A10G8G6A	CA4G7C2C1	0P44A44G8	6P4816G22	4F4C18G6T	4C20012G2	4T20C4T6T	4G4G8T3C8
8A30T20C8	4C14G10T1	4C4G4TGC1	10GTC4T14	12A20A12G	12A20CA4G	8AG6G16TA	2A4C8A18G	A8G8A6T4G	CAA10GT6C	G8GCG6G18
4C20AA4G10	4C4G6T10C	A10T6T1T4	4C4G12TG4	16A4AA14T	4A8T4T4A4	4A6T4CA16	4C12T4AA4	4T10G20G1T	T4T4T28G2	G4T20G22T
0T10A4T16	T6G8AA5T1	G16T6C28C	C4T24T6T6	T2AA20C4G	8G6AC6C12	4A22T6AAC1	4T8GT26A12	T8A4C6C4T	6G4T10C4G	G36CC12AT
T4A18A6T8	T8A8C8C20	4T4T8A4T6	C10T12G4G	6G10TG8T10	AG20AGC12	2G64C10T22	T12T4C4G20	T6A4A78A6	G10A6T14G	4C4A16C8G3
0G14G10CC	8C6G8T4G4	4A14T4G20	14G4CC12A	G6G8A4ACC1	6T20G8G12	12AA8C22A8	A16CCT6A4T	18CA10T6T6	T4T18G4TTG	CC6T18T4A
4A4T8C6C4	26T8C22G6	G4T8G6A22	8A12T6T4T1	8T6G28A2T	C18T6C4C3	G4T4T4G18	4G12T10C8T	GT8GA4A12C	14G10C12C	14G6T6T14A
4GTT10CC7	CC16G16G6	C4T28G14C	C10C12CA10	CCT4TG6G8	8GC10T6A1	4T4AA8G4G	TT8C26A4A1	2G14G10GA	C6T4T8A2A	AC4T710A4T
A4G8C8A62	6C10A14G12	C12A6T8C14	T4A18C14T	14C4GA30CA	14G4A4A10	12G4T4C4C2	2C12C6C8T6	CC04G6T12	C4C4T14G8	T12A76T8C1
6A8T4TT12	6G14CA14C	4T4G10CT6	G8GA8A14T1	2T6C12TCC	ACG6T8C8T	T4C14T8T4C	4T4T6T6T10	C4G4T4A8G1	2A30G14G6G	10T4A4C14T
26T4TG8G10	G8T44T8G10	G8GT4T26T1	4C4A4T10G6	G14G30A12G	8A4T4G4C10	T6T6T4T4C4	T8T14C4T4T	C8G76GCA4C	CT12C6T12T	14A8AG8GG6
4A4C14TG6A1	2TT4T8A16C	8T6AT12A8T	G14T4C4C4G	12T6C4CC6T	8C6C12C22C	4C4T4G12G1	0A4A4G14C	30AG4C14TA	14C18A4T14	
C8T6A12C12	G14A10C26A	6A8C8G4AT4	A10TT4CA22	AA8T4T6C8A	G10G14G12A	4A26C8TTGA	4G8AA4T14A	T6T10C8G8	G6G74TCC10	AC12C10CCA
14G28T4CA6	G16G16CC6T	CC10TTG4A1	4T6T6G14CC	12C10G14C1	2A4A8GT8T8	C10T12G4T1	8G4T4T4G30	C4C6T18C22	TA28G6T18T	4T6T12A8T2
2A6G8T4GT6	G22C8T26A4	C6C8T4A14	4T18T6CCGT	T4G18T4T6T	6T10AC18T4	A6TCC16A8T	22C8AA12C1	2G8G20T16C	CA4A8GG6G6	12CC4G14G2
0G4T14A4G4	G4T8G6C8CG	C10G14G30G	8C16A4C24G	14T6A10GA6	A8T4A6A20T	G4C4T12T22	T10C4G26A1	2CG2A0GA10	T8G710G6GG	4G12T10CC6
4A48T4T4T2	0C8C8A8T8T	T6A18A4T18	T12CC3G6G6	4T10T48T12	4C6C4AG8T12	A26T78T14C	AA6T22A4C1	2C6A8G68G4	C20AA26T6T	76T24T4CC2
8C6T16TG12	T6AA8G6T16	AT4A10T10T	22G20T4G24	G28T4T4T16	G20G10T4C4	AA4T12C4A1	6AC4T6A4T4	A4T4T8A4T1	4AAA4A16C4	GT12GG4C1T
4TT6T10A4C	10TT6G4C10	G4AA20C4CA	18G6CG8G8	6T7G10AAC4G	4T6A88G8AG	18A8C4A12A	T16G6GA8G	4AC20A12TG	12A20A12C1	4T4CTG10A1
6CGT4AG4C14	T10G14C4A8	C20T30A8A3	8CT8GG4CT6	T4C10T12AG	12G20CT4T6	T6C18A4GA2	6T16G4T6A8	2G4T4T10C	30AG4C14TA	68G10A66C6
G4T6C10C14	A12T28T4T8	C14CTA4G4A	A10T4G4AGC	T4C24T8G6T	G12CC6A4C1	0GC6A10G4A	4A6A8C40CC	4C4T4A22C6	GA10CCT8G4	G24G14G20C
C10C8T4A4T	C18T4GT4C1	078G18C6C1	8C14A10CA4	GC22G8A4C1	2TT12T20C8	T14CA10AAT	4T6GGT12G4	TT20T16T18	TT4T4A14T1	2C20GG8C8G2
8A14C12T4G	8A4C6T4CCA	10A4C6T6G1	4C20T16C4T	18T4A820C	8A4T8T4C18	A6G4AA8T10	C12A40AA6C	6G8GC4C4A	G4C4T18A10	4T4C6G6T8A
4G4C6G14G2	2G10C14G7	G6GA4AA4A4	AT4T26C28C	6G10C8G16G	G4A26C4A7	G18T4T8A6A	C4G24T62G8	T10AG10A12	G4G12C4AT	28G4C4G6T6
G124T4T6T4	G4T4G6G8T4	12G6TTT4TC	TT16C4T4T6	TC12T20TG	4T28A4G8A4	4T8C4GA6A6	T4T4C4C6A6	TT4T6T32T4	T4C4T14C6A	10T4G4T4C
14C4T10G14	T76G10T6T1	2A8C8T12T1	4T6T4TTT8T	20G12C26CC	14T14G8GTT	14C30C8A6T	10A4C4GA44	T20G16C8CT	4T4A10T6TG	20TG8G8C6T
16T10T6T6T	22G20T7T6T6	T2G6C4AT16	T4T4G8T6G	C10G76C6CG	10G16T8A4G	4G6A12G20G	T4T6GA14C6	C8T20T6A8T	T126GAA4A1	6C12C4A6A2
8G6A18CA7C	T4TAGT14C6	G11A6G10G	TTGGA4G6T1	076ACCT22	12G20CT4T6	8C2G10C4A6	GTA12G4AG1	8A18A8A8G	10C8CC4ACT	6C6A4G46C7
10T4A6T12A	AG28G4G6T6	A10GG4T8G4	G6G6C20A14T	4C2A4AA4G	10ATT76T12A	6T10ATT18G	4G8C8G4G4T	14TG8C8C12	0T12GCG4G4	T8T4T4CCG1
4A10A18A12	G8T16G6A10	G4TGAG10G6	A14GGCA24T	T16T6GA18A	18TTT16T4T	6T8CT8TA10	TG4G6T6T10	G6GAAC10T2	CC12G6A16T	12TA6C8G16
1A1716AC6C	16G12C6AG6	T4G8T8T22A	10C12G6TG4	T110GG6G4C	10A14C10A2	6C6T74A8C8	G6T10T14T6	C6C4T4G7A4	T18G10C8T1	6T6C12C4T6
T10T8A6T4C	8C8CAAT4T8T	16G16T4CAC	6CG6AC12T	4G10CAAT3T	G6A10T12T1	6T12C4G6AT	4C6A14C8GT	T4TG10G4G	6C8C8CG6C	4G28AT4T14
G12C8A14G6	G6G6CCT4T1	2T14G6T30C	6G16T10CG4	T6G10T12C1	0G4T6G24C4	1C7CG6G4G4T	CL12A24G4C	4GA6CT10TA	4G18T12G8T	8T218G44T
16T4T8T4GC	16TGG20C10	CA18C32G12	C4G26T12C4	TG22G12G8T	10C4T4C38A	30A4GG4C6G	26A12C4A12	144CC4C6G8	A4TG14G4T1	0T20ACA14T
4TG4T14A6G	T4G4G58T4G	20C6T18T18	C4A6G4G14G	6T7G34G4A6	6C4GG8G814	8A4T74GA12	4AGT6C4C18	T8G6G14C10	G5C6T10G1T	0T76C6A8T8
A10C4A6T6A	G30C4TG6A	T0C710G4A18	T10C4G6G8C	4T6T6T36G2	6T6G12P3G2	2G4T14G4C2	4A22T10C14	A12T14G22G	62T4C8C8CT	4TT4TT20C4
C16T4G6T4C	8T10C8T7G4	4C18G10TT6	A4T14G4TG4	G12C6T6T4T	T4T8G8G610	G12T4T4G4	C4G24T4T22	G34T4T12CG	4G18A4G14C	AA18T26T1
2C8GT12C2T	10CC4AA6GT	12A6G4C6A6	CT4G26TGGG	G34CC8T4TT	TC4T4A10C1	0C12G4A4T6	G4G6A6C10T	T16C4GC8AT	G1830C30C	GA18T2C22T
10TT6C61A6	6C4C8T6GA4T	14T16T6A6A	8G4A10A14G	A6T110A8CG	4CTA4AG6C4T	6TC4TAGAG6	C6C7AG6G4T	6GAA44T8G	4C10A8C6C4	0T116G6T6T
6C6A4G8T8A1	0GTC4T10T1	2T14A4AA4A	8T4T4G6GAC	18G6C10T8A	6T7G10C12A	6CG14T8C6C	G4G10G4C4A	G6C14C10T8	T6A6A8G10C4	C16A4G6T6G
8A6C8T18C4	AG16A10C6C	16C6A8C26	C8G38A12A1	0GT4A12AA4	AAG4T8C44GT	AC16G4C54	C6C24CCT14	C6C20C10G	TA68A10A8C	6A8A16A14T
G6G4A4T8C4	G4TG4A4A18	A10G14C6TG	10C4T8AC4T	12GTG10GTT	12T18A8G26	AC6T4A6A8G	16GATC6G6G	10AA6C8C10	GA4C38G22T	T14TCC8G10
AG16T10G6	T30A8T6G6G	12A6T6T10A	4TGT30C4T2	2C10G6G6GG	A18AA6G6C10	26A4A8T6AG	16GCC14T10	AT6G14CA4C	T22G10T12A	16C4T4C6T1
6C10C6T6A	8C0C6G10G	1A6T6T4G4C	A4G16A22C4	G4T4G4G4G4	12G6T8G6G1	1G2A4T10C	12C10C4C10	12T6C68T6G	12C4G4G4T	
4G4C22A16G	4C4G4T6T1	6T10CA6C10	G8G4T6C10C	C16T6C4T4C	16A12T10G2	2T7C4A14G6	TA10T14CCG	16G6A8BA4	6T10C6GAA1	8A6G6G6G10
C22T4C30TG	T4A10T6T6A	12G6G6T8A3	0T6G10GT16	GA10G8C10T1	4TT22G38C4	AG10C8C6A	10G6G6CTAG	16G8A6A4T6	C2A6G8A18T	12T4T10GTG
12T4C8B4C	10GT6G14G1	0A18A4A4GT	4G4C8T4A4G	6GT14A16A	6AC8A10AG4	AC610G20C2	4G14TCA6A6	CC5C4G616C	A4T64G4CTAG	AA4AA12A4T
G10A12A38G	8C26G8BA6C	16C6C10A16	GA4C18T8C6	A8G6T6G4A1	6C4G10G8A6	A78T10C14C	6G4C4G6G10	4G6C8G714G	G6A12C10TG	T6C8T10CC6
18CAG6G4T4	T8AA4A14A	T12T10T4GT	G10A8T4C4G	C6T6T6G16T	60C6G8A10	C4G8T4A4A	6G12T6T6CT	6C6GAG4A4A	T6T4C6G4AT	G4C8A10T4
6AG14A10A4	G8A6A6T16T	14T4A6T8G4	C6A16A6CCT	10TT22CT18	AGC10C30G1	8GTA8CG4C1	6TTT10CA6G	4G6T4A4G12	C10C10A4T4	CTTT4T8CC3
4G6GGT26G4	TC6A6G4G6A	12T6GAA4CC	10TT22T6T18	AC12T6T18T	AA314G4A18	G4G12T4T3	4G22T4T24G	4C4G4T4T12	TG10C6G8GT	4T4T6T6C1
264GT4G14T	4A6T10G18	4G4G4T8C10	C1C220TT4T	T4C8T8BT24	T4C8T8BT24	G6G22G14T1	2A14C10T22	42C4G14T4	G12A2C12G6	2TT63T6T6T
T4C8G6G4C1	0T18A4G10T	G76A6GT4C3	0GA6T6A4C1	0A8T8A6C6T	T10TG10C6T	6C10C14G6G	8T8G4C4G6T	4A12AG4T4A	8G14G8GGG4	C6A6A4G34G
T6G14G4G6A	4C18T18T6C	20G4T58G4G	4TG6A14T4G	T4T14ACA20	T10T4G14GT	4A8G6C4CC1	4T12A4C12A	26G6C4GG4A	30A38C4T4C	10T8G12G22
GT4C12T2F6G	4C12G32C18	AC10C20G6T	16CG4T8T4T	16G6G4G18C	T8T8G1T2T8	4A4T0T3C6A	G4T04T6C1A	2CTG4G6G6T	16A4C24G74	G10C12G10G
6T4GC10T16	G6C30TG64T	T12T4TCG6C	6G6G14A8C1	2G14T4T2A8	6G4C68C4C	16G4G10G7A	TTT8G14A6	6T4A6G4C12	C16T12T10A	6G30T4AC10
G4TT12C2A6G	CC6C4T116G	16T8T4TACC	C8C4T6A8T1	0T6T4T12C6	T16T8C10G1	84T6T4T4C	6C6T14T10T	G6C8A8A4T6T	6C2A610C14	4T4C6GG10T
TT4GT6G12C	10A22T8T8G	4T6GA6C12G	16C6CA16T1	4A16G8C6AT	12T16A6G12	CCC12CCAAG	6G10T6T6G4	GT10AT8T8C	T6T4T16TTT	18A18A6GT1
6TT24ACGG1	4A6G10GAGT	4G10A6G16T	8G12A18A10	A14GCG4T4T	8T4G4GCG12	T10T8C8GT	14T4G4G8C8	G4G18T7A10	T6A12T6TT4	10G6GAA4A24
C4T14A20C6	G0A4G2AGG1	10A6T6G4G2	GA12T6A4T	10AA4G4A6	6CTCA4CC8C	10G8A8AA18	A18GA4G12A	T6G4A410G2	2C22T2C4G6	T10T6GA6A8
TTG4CG6A6T	22G10C6A16	TG6C14TGAT	4T4TAC18A6	G8TCC6A4C12	C16A4AG6C1	2TT8A6T20T	8C6C14AG6T	4T2G2012A6	G4G4A8T16G	10C6G6G6G10
0C8GT8G4TT	4T16T4C4G2	6T6T6TT20G	22C6T6T10T	16T6C8G8CT	20GT6T10A4	4T4T8C16G2	0T4A4G4C4A	10T6A8C30C	14TTG8G14T	14CC26C12G
20T8T7T14G	T14T12T8C8	A12T6T10G6	TT14G10T4C	14C4T4G6G4	10A6C14T4C	4T4T32T6T4	TTG6AC4C4T	4T6A6AG4C8	4A48A8A428	74GT20T7G12
CT6T4T4C16	TTCT4T7T6G	12C8G6G4G2	4G4T6T4G12	TG6T6G4C12	2BTA4T4C12G	4G12A10GA1	0TC28GT24G	4CA6A8T4T1	8G4T4CC26A	4G61G8C10
6G6C2C26T4	TAA4A4A4AG	G22T14G6C4	G4G4T6G6G6	4C10A18T4G	4C10A18T4G	4G4C4C4TCG8	6G6A4A0A1	2C10T8A4A4	6A18C4T8T4	8C20C8A4T
18T4C16T20	C14G6T6C4A	10ACC4T6C4	A8G4T12C14	A28G8C8G20	C12T14A4T4	TT18T16T20	TT4G12TGG6	T4TAA10AC1	4T8C20T12T	T12C4A8G22
CG4AC10A14	C18C6C18G8	T10C4G4T11	8CT4A4T8C1	0CC20G14G2	4G4G8TCC10	AG6C22A4T4	C4CC40C8A6	A4A4G10A6C	G10C4A6C31	2GT6G8T24C
4TCGA4G4T1	0AAAG4ATC1	4C8T4T28T1	2A14C10C6T	4G6G6A10G8	G6A4CAATC	12C10T4T4A	4G8A6T4G16	G22AG4A18C	6T64T4CC20G	12GA12T10C
4T6TC4G8G7	C38A8A30T2	0C8A4C14G1	0T14C4G4GT	C16A10GT4C	T14C12A20A	12G12A20C	A4GG8AG6G1	6TA12A4C8A	18G8A8G8A6	T4G4CAA10G
T6C8G8G6CG	18AC4C20AA	4G10C4G6T7	10CA410T6T	T14TC4GG12	TG4C16A4AA	A14T4A8T4T	4A4T4A6T4C	A16A4C12T4	AA4C4T10G2	0G16T4T4T2
8G24G4T20G	22T10T10A4	TA16T6G8AA	6T12GT16T6	C28CC4T24T	6T6TT26AA2	4C4G8G6AC6	C12C4A22T6	ARCI4T8G2T	6A12T8A4C6	C4T76G4T10
4G6G36C6C1	2T18T4A18A									

TACATAGCTG ATGTGATCGG CTTTGTGTCA GTGCATCATG GTTCTCTTTT CCTGGTCTCT GGCCTTATGT GGCTTTTGG CGTTTGCCTT GTTCGGAGCA ATTTCTGCTT

CCAAGTAGCA CFTGGGGGCC TTTTCTACCG ATGGACTTTC GCATGGCCGA TCTTTTCAAC GTATTTAAGA AGATTACGCT AGCTCTAGA CCCTGGGAA ATGCACCTTG

TTCAGTAGTG TTTAAAAAAT GGACGCTCTG TCACGTGCGG GGGCCCTTAA GGCAGTGA CT CAGACCCAG CGAAGTAAAA AGTCGTACGC CCACCTGCGC CGAACAAAA

GGATCGTGAA AGCTGCTACT GTGGTTTAGA CTAAGGATCG GAACCCAGCT TCCGAGTGG TATGGATTAT GTCTCGGGAA AGCTAATAGG CCTATGCACT GTACTCTCCC

TGGCACTTTG CAGACGTTGG CBTGGGGTCC CCCCCTGGGT GCGCGTGCAG ACGTFTTACAG GTCCCTCTCA TGTACAGTAT CCGGATAATC GAAAGGGCTC GTTATTAGT

ATGTGTAGCC TTTGCAGCCA AGCTTAGGAA TCAGATTGGT GTCTATCGTC GAAAGTGCCTA GGTAAAAACA CCGCGGTCCA CCGCATGCTT GAAAAATGAA GCGCACCCAG

ACTCAGTGA C GGAATCCGC GGGGCGTGCA CTGTCTCGCA GGTAAAAAT TGTGTATGAC TTTGTCCAG TAAAGGTGTC CGAGATCTTC GATCGCATT GAAAGATTTA

TGCACTACTT CTAGCCGGTA CGCTTTCAGG TAGCCATCTT TTCCGGGGTC TCACAGTAA CCTTCTGCTT TAACGAGGCT TTTGCGTTT GCGGTTTTTC GGTGTATTG

GGTCTCTGGT CTTTTTCTCT TGGTACTACG TGACTGTGTT TCGGCTAGTG TAGTGCATAC ATACTTTGCT CCGGTGCGGT AAGTAGGGGG CAAGTGGGGA CTTCGTGGT

ATGTTACATT GGTAGCCCTA CAGCGGAAC CTGGCTCTGC GCACCGCTGG TTTGGGCTTG TGATTACGTC GACTGGGGTA CGTGGCGTGC TGCTTTTTTC GCGCACCTTA

GCGCCCGGT TTTGCACTAG CCTTAGTACG TTCAGCCACT GTTTACCCTT ATTTCCCTTG TTTGTCTT GCATTACAT CCGGTTGTGC ATTGTGACGC CATAGCATT

GCCCCCAAG TTTGTTATT TTTTAAAGTT TACGGAGGAG TGAGTAAAG CTTTGGCGT TCCGTTGGCG GTTATATTTA GGAACACTCG GTGGATGGGA ATATAAGAAC

TCACCCGAAA AGAGATAGC CTCCGTTGAG GTTGCAGTGC ATGCTGATTT ATGCTGATTT ACAGTACCCA AGCTTATTC AGTTGGAGGA TGGCGTCCGT GTTTTCCG

TGTTATTCCG TAGCATACCT CCAGCCGCTT TGTTTTCATT GTTTCCTGGT TACTCTTTT TGACTTAAG CTAAAGTGT GTTTTTCTT TTTGCGGTT TTTGTGCTT

ATCGGATGC GTGCATTTGT TAAAAAGCT TCGGCGGCC GCATCGGCTT TCGGAAACT ATTCCTGCTA ACCTCAGTCA CCGCTATTTT

TTTGTGTTA AACTCTTTCA GCGACACCCG TCTGTCTATC CCGGGGTCCA GCATCCCCAA AGACGCACCG TGTCTCGAGT AAGATTCTT ACCTCGAGGA CAGTCTTAG

ATGGAGACT TCGGATCTC GGTCAATCAC GTCCGTGCA GTCAAGTACA GAGAGTAA CA GAGATGCAA GTCCGCGACC AAGCCTTCA TTTGCGTGA AATATTATA

TCAACTAAT GGTTFGGTGT TATATGAATG TTCCCTTTT AACGGACCCA TAACGTGATA CCTATTTAC CTTATTTGCTG GTTTATGCTG GCGGTTAGG

ATAAGATGC ACGGGCCCGG TGGATGGGCC AGGAACCTGG CAACAACTA TCATTTTGTT GCCTTAATCC ATCGTGTGAT ATTTGTATCT CCGTTFGTCT TGTGAACGCC

CGTTAGTTCC TCCGACTGA CCCCACCTTG GGGCTTATA GAGTTCAAAG GACTTAAACT TATAGCAACA GCATCTACAT GCACAGAAG TCCCCCTCC TGCTGTATA

TCATTAGTCT ACTGCTGGGA ATCTCCAGG TGCTTCTTCT TTTCTCAGG GGTACTTTGG TATGGGTTTC ATGGGTTTC CTTTTCCCT CTTAAGGCT CTTAAGGCT

GTCACTGAT TACTATATCT CGCTCCCTCC CCGTGAAGA CAGCTACACT TAGCACAACG ATATCAAAT TCAGGGAAT TGAGAAAT CC GGGGTCCCA CCCCAGTCC

GCCTCTTGA TFGCCCGCAA TGCCCTGTGT TTTATAGTGT GCTACCTAAT TCCGTTGTTT TACTATCCAA CAACGGTCCA AGCACGGCTG AGCCTCCGG

GGCAGTAGA ATAATGCTTT CGACGAGATG TGGGGTCGTA TTTCCATTTA ATTCGCGCTG TTCCATATGT CAATACCAG GCAATTTTT CTTGTAAGT ATATTGTTG

TGGTCAACT AACTATATTA TAAAACGTGG CTTTTACTTG CGAACCCGCG GCTGAAGTGA GAGGACAATG GAGGACATGA ACTCTGACT GTCTGCACTAA TCGGCTCTA

GGCTTCAAG GGTATGATTC TGACAGGACG TCCATTTCC TGAATGACTT CAGTGCACCG CAGAAACCCC TACGCATCTT GCGCCACTA CTTGCGACTA AGCCTTTTC

TCAAATGTT GTTTTTTTAT CCGGACTGAC TCCAACCTGCT CTTACCATTC AGAATCAAC GGCTCAGGTA CTGGTGGGCG GCGTTGGA AAATTCCCGG GACCTGTATA

CGTGTAGAG GCTATGCGTG TGTGTGGGCT GTTTTCTTCT TCCGTTGAG AATCGAATC CAGTTTTTTC TCATGGTCC TTTGTTACTT TGTTTTTGCC TGGTTCCTA

CGATGCCCTA TGTGCGCTT TCGTTTTGCT TTTGTGCGCT GCGTAGGAGG TTGACCTTAT TCGAACCACT TGCAAGCATG GACATTTAGT CTTGAGGTTG CTTCCGAGTA

GAGAAAGCC CACTCAAGAA TATAAGGGTA GGTGGGCATC AAGGATTTAT ATTTGCGGTT GCCTTCCGCT TTTGCGAAAGT GAGTGGAGG GATTTGAAT TTTTCTTAT

GGTTGGAACC CCGGATTAG ATACCGCAGT GTTACGTGTT GGGCTACACT TACGTTTCCCT TCCCTTCCCT TTTATCCCCA TTTGTCACCG ACTTGCACTA TTTCCGATCC

GTTTTGGGCC CCGGATTGCA GGGCTTTTT TCGTCTGGC GTGCATGGG TCAGTGA CT TAGTGTCCG GTTTGCTGCC CACGCGTCTG GGTCTCAAG CGACATCCC

ATGTTACAT TGTAGTGGTG CTTCAAGGTG GAACGGGGGA TGAAGTGCCT GGCTCTGTT CATACATAGC TGATGTGATC GGTCTTGTGT CAGTGCATCA TGGTCTCTT

TTCCGTGCT TCGGGCTTAT GTGGCTTTTT GCGTTTTGCG TGTFTGCTC CAAFTTGTCT TTCCAAGTAG CACTGTGGGG CTTTTTCTAC CCGATGACTT TCGCATGGC

GATCTTTTCA AGTATTTTAA GAAGATTACG CTAGTCTTCA GAGCCTGTGG AAATGACCTT TGTTCAAGTAG TGTAAAAA TTTGACGCTC TGTACCTGC GGGGCGCTT

AAGGCAGTGA CTCAGACCCA CCGCAAGTAA AAGTCTGAT CCCCCCTG CCGCAACAAA ATGGATCGTG AAAGTCTGTA CTGTGGTTTT GACTAAGGAT CCGAACCCG

GTTTCCGAGT GGTATGGATT ATGTCTCGGG AAAGCTAATA GGCTATGCA GTTACTCTCT CCGCTGACTT GCAGACCTG CCGCTGGCTT CCCCCCTCG GTCCGCGTGC

AGACGTTTCA CCGTCCCTCT CATGTCACGT ATCCGGATA TCGAAGGGG TCTGTATTAG GTATGGTGAG CTTTGGACG CAAGGCTAGG AATCAGATT GGTGTCTATG

TCGAAGTGC TAGGTAARAC AAGCCGCGTC CACCCGATG CTGAARAAAT AGCCGCACCC AGCCGACTGT CCGAAATTC CCGGGGCTG CACTGTCTCG CAGTATAA

ATTTGTGATG ACTTGTCCA CGTAAAGGTG TCCGAGATCT TCGATCGCAT TAGAAGAATT TATGCAACT TTCTAGCCG TACGCTTTCA GGTAGCCATC TTTTCCGGG

TGTCACGATG AACCTTCTCG TTTAACGAGG CTTGTTGCGT TTGCGGTTTT TCGGTTTACT GCGGTTCTCT GCCTTTTCT CTTGTTACTA CTTGACTGTG TTTCCGTTG

TGTATCGAT ACATACTTTG TCOCAGTGGG TGAAGTAGG TGCAAGTGGT GACTCTGTTG TGATGTGATA TTGGTAGCCC TACAGCCGGA CTTGCGTCT TCGCACCTG

CGTTTGGGCT TGTGATTCAG TCGACTGGGG TAGTGTGGGT GCTGTGTTT TCGGCGACT TACGCCCGG GTTTGCACT AGCCTTAGTA CTTTCAAGCA CTTTACC

CTATTTCTT CTTTGTCTT TGCATTAC ATCCGGTTGT GCATTGTGAC GCCATAGCAT TACCCCCA GGTGGTATT CTTTTTAA TTTACGGAGG AGTGTGTA

AGCCTTGGCT GTTCGTTTG TTTGATATT TAGGAACCTAC GGGTGGATGG GAATATAAGA ACTACCCGA AAAGAGATGA CGCTGGCTG AGTTTGGAT CATGCTGAT

TTACAGTCA CAACTTATT CCAAGTGGAG GATGCGCTGC GTGTTTTCTG TTTGTTTTTC GGTGTTATTC CGTATGATAC CTTGTTCCG TTTTGTATA TTTGTGCTT

GGTACTCTT TTTGACTTA ACCTAAGATG TTTGTTTTCTT CTTTTGCGGG TGTFTGTTG TGTGCTGAGA TGTGCAACT GTCCAGGGG CTTTAAAAA GTTCGGCGG

GTTGTCATG ACTCGGCAA CTAAGACTTA CCATTCTCGT CAACCTCAGT CAGCGCTATT TTTTGTGGT TAAACTCTT CAGCGACCC CGTCTGCTA TCCCGGGTC

CAGCATCCCC AAAGACCCAC CBTGTTCTGA GTAAGATCCT TTACCTCGAG GACAGTCTT AGATGGAGAC TTTTCGGATCT TCGGTTCACT ACGTCTGCA GTCTCAAGTA

CAGGAGTAA CAGGAGATG AAGTCGGCA CCAAGCGTTC ATTTCCGGTG CAAAATATA TATCAACTAA CTGGTTTGGT GTTATATGAA TGTCTTCTT TTAACGGAC

CATAACTGA TACCTTCTG CCGTTAATTT ACCTTTATGC TGGGTTGAG AGCAGTCTG GTAATAAGAT GCACGGGCG CTTGGATGG CAGGAACTT GGCAACAA

TATCATTTTG TTGCCTAAT CCATCGTGTG ATATTTTATG CTCCGTTGTC TCTGTGTATC CCGCTTAGTT CCTCCGACT GACCCTACT TGGGCTTAT AAGATCTCA

GGGACTTAAA CTTATAGCAA CAGCATCTAC ATCGCAGAA TCCGCCCTC CCGTCCGTA TATCATTTAG TCACTGTGCG GAATTTCTCA CBTGCTCTT

AGGGACTT GGGT

### 2.3 Classification of 130 virus genomes based on loop lengths due to autocorrelation

Out of 130 virus genome sequences, one hundred and two have been identified as open-ended sequences such as S1, S2, S3, S4, S5, S7, S8, S9, S11, S13, S14, S19, S20, S21, S22, S23, S24, S25, S26, S28, S30, S31, S32, S33, S34, S35, S36, S37, S38, S39, S41, S42, S43, S44, S45, S46, S47, S49, S50, S51, S52, S53, S54, S55, S56, S57, S58, S59, S60, S61, S62, S63, S64, S65, S66, S67, S68, S69, S70, S71, S72, S73, S74, S77, S78, S79, S81, S83, S84, S85, S86, S88, S89, S90, S91, S92, S93, S95, S97, S98, S100, S101, S102, S103, S104, S106, S107, S108, S109, S110, S111, S112, S114, S115, S116, S118, S122, S123, S125, S126, S127, S130 and twenty eight have been identified as close-ended sequences such as S6, S10, S12, S15, S16, S17, S18, S27, S29, S40, S48, S75, S76, S80, S82, S87, S94, S96, S99, S105, S113, S117, S119, S120, S121, S124, S128, S129. The MLL values of 102 open-ended sequences is viewed as a sequence of MLL values and presented as a graph. The maximum and the minimum MLL values in the graph are identified and the difference between the maximum and minimum values is treated as a band. This band is divided into four equal level sub bands. Now all the 102 virus genomes are classified under the four level sub bands. The same procedure is followed to classify 28 close-ended sequences. Table 1 presents values of maximum and minimum loop lengths for all open and close-ended sequences and the respective four sub bands.

Table 1: Maximum and minimum values of loop lengths of open ended and close ended virus genomes

Sl. No.	Open Ended Sequences	Maximum Values of Loop Lengths	Minimum Values of Loop Lengths	Sl. No.	Close Ended Sequences	Maximum Values of Loop Lengths	Minimum Values of Loop Lengths
1	S1	58	3	1	S6	88	4
2	S2	54	4	2	S10	56	4
3	S3	46	3	3	S12	50	4
4	S4	50	3	4	S15	58	4
5	S5	70	3	5	S16	66	4
6	S7	66	4	6	S17	52	4
7	S8	52	4	7	S18	58	4

8	S9	48	4
9	S11	66	4
10	S13	54	3
11	S14	58	3
12	S19	44	4
13	S20	50	4
14	S21	48	4
15	S22	64	4
16	S23	44	4
17	S24	44	4
18	S25	32	4
19	S26	46	4
20	S28	40	4
21	S30	52	4
22	S31	40	4
23	S32	42	4
24	S33	62	4
25	S34	52	4
26	S35	46	4
27	S36	54	4
28	S37	52	4
29	S38	82	4
30	S39	32	4
31	S41	48	3
32	S42	82	4
33	S43	50	4
34	S44	44	4
35	S45	54	4
36	S46	50	4
37	S47	46	4
38	S49	56	4
39	S50	50	3
40	S51	38	4
41	S52	46	4
42	S53	44	4
43	S54	54	4
44	S55	46	4
45	S56	80	3
46	S57	52	4
47	S58	62	4
48	S59	44	3
49	S60	68	4
50	S61	40	4
51	S62	56	4
52	S63	58	4
53	S64	48	4
54	S65	54	4
55	S66	48	4
56	S67	44	4
57	S68	50	3
58	S69	48	4
59	S70	40	4
60	S71	50	4
61	S72	56	4
62	S73	48	4
63	S74	58	3
64	S77	46	4
65	S78	56	4
66	S79	44	4
67	S81	40	4
68	S83	58	4
69	S84	50	4
70	S85	44	4
71	S86	46	4
72	S88	42	4
73	S89	46	4
74	S90	50	4

8	S27	52	4
9	S29	44	4
10	S40	64	4
11	S48	70	4
12	S75	42	4
13	S76	48	4
14	S80	46	4
15	S82	46	4
16	S87	52	4
17	S94	44	4
18	S96	72	4
19	S99	62	4
20	S105	54	4
21	S113	42	4
22	S117	62	4
23	S119	60	4
24	S120	60	4
25	S121	56	4
26	S124	56	4
27	S128	58	4
28	S129	56	4

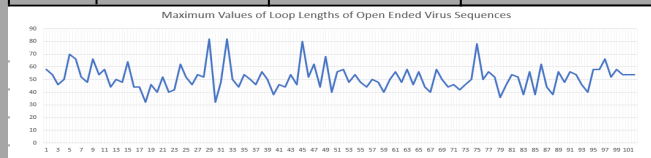


Fig. 1: Maximum Values of Loop Lengths of Open-Ended Sequences

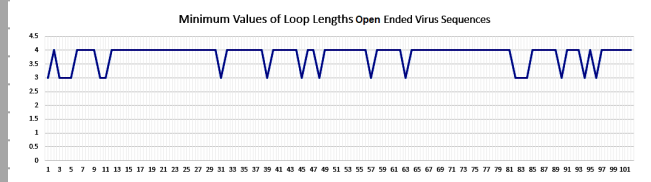


Fig. 2: Minimum Values of Loop Lengths of Open-Ended Sequences

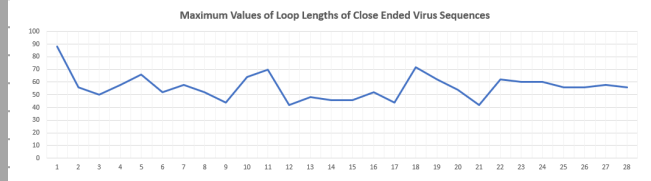


Fig. 3: Maximum Values of Loop Lengths of Close-Ended Sequences

Minimum Values of Loop Lengths of Close-Ended Sequences remains as 4 for all open-ended sequences. ‘Maximum Loop Length (MLL)’ value has been found to be a reliable quantificational measure of a virus genome, be it open ended or close ended. More the value of MLL of a sequence, less its SSI value. Every virus genome sequence would have a unique SSI value and MLL value. Given a finite set of virus genome sequences, one can find out the MLL value for every virus genome. In the case of 130 virus genome sequences, 102 MLL values have been found for open-ended sequences and 28 MLL values found for close-ended sequences. Fig. 4 shows the graph displaying 102 MLL values of open-ended genomes with four levels.

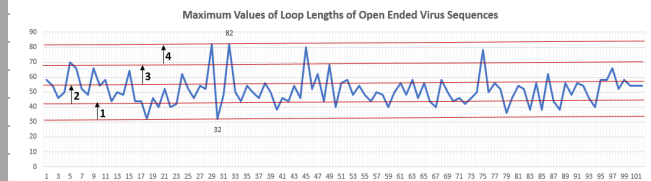


Fig. 4: MLL values of open-ended virus genomes with four levels

The genomes S29 and S32 are found to have maximum MLL value 82. Genome S30 is found to have minimum MLL value 32. The band of values from 32 to 82 is further divided into sub bands (i) 32-44, (ii)

75	S91	78	4
76	S92	50	4
77	S93	56	4
78	S95	52	4
79	S97	36	4
80	S98	46	4
81	S100	54	4
82	S101	52	3
83	S102	38	3
84	S103	56	3
85	S104	38	4
86	S106	62	4
87	S107	44	4
88	S108	38	4
89	S109	56	4
90	S110	48	3
91	S111	56	4
92	S112	54	4
93	S114	46	4
94	S115	40	3
95	S116	58	4
96	S118	58	3
97	S122	66	4
98	S123	52	4
99	S125	58	4
100	S126	54	4
101	S127	54	4
102	S130	54	4

45-57, (iii) 58-69 and (iv) 70-82. Fig. 5 shows the four sub bands and virus genomes belonging to those sub bands.

Table 2: Four sub bands and classified open-ended virus genomes

Open-Ended Virus Sequences		
Bands	MLL Ranges	Virus Sequences
1	32-44	S19, S23, S24, S25, S28, S31, S32, S39, S44, S51, S53, S59, S61, S67, S70, S79, S81, S85, S88, S97, S102, S104, S107, S108, S115
2	45-57	S2, S3, S4, S8, S9, S13, S20, S21, S26, S30, S34, S35, S36, S37, S41, S43, S45, S46, S47, S49, S50, S52, S54, S55, S57, S62, S64, S65, S66, S68, S69, S71, S72, S73, S77, S78, S84, S86, S89, S90, S92, S93, S95, S98, S100, S101, S103, S109, S110, S111, S112, S114, S123, S126, S127, S130
3	58-69	S1, S7, S11, S14, S22, S33, S58, S60, S63, S74, S83, S106, S116, S118, S122, S125,
4	70-82	S5, S38, S42, S91

Fig. 5: Shows graph displaying 28 MLL values of close-ended genomes with four levels

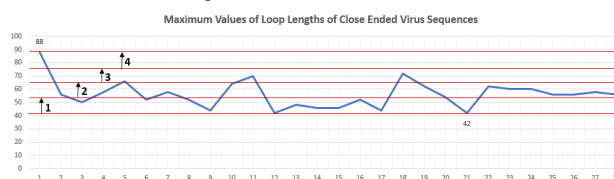


Fig. 5: MLL values of close-ended virus genomes with four levels

The band of values from 32 to 82 is further divided into sub bands (i) 42-53, (ii) 54-65, (iii) 66-76 and (iv) 77-88. Fig. 7 shows the four sub bands and virus genomes belonging to those sub bands.

Table 3: Four sub bands and classified close-ended virus genomes

Close-Ended Virus Sequences		
Bands	MLL Ranges	Virus Sequences
1	42-53	S12, S17, S27, S29, S75, S76, S80, S82, S87, S94, S113
2	54-65	S10, S15, S18, S99, S105, S117, S119, S120, S121, S124, S128, S129
3	66-76	S16, S48, S96,
4	77-88	S6

**Observations**

1. Auto Correlation Measures (ACMs) of virus sequences in band #1 are HIGH, meaning probability of they being susceptible for mutation or structural transformation is more.
2. The open-ended virus genomes S19, S23, S24, S25, S28, S31, S32, S39, S44, S51, S53, S59, S61, S67, S70, S79, S81, S85, S88, S97, S102, S104, S107, S108, S115 fall under band 1 and the are likely to undergo structural transformation.
3. The close-ended virus sequences S12, S17, S27, S29, S75, S76, S80, S82, S87, S94, S113 fall under band 1 and the are likely to undergo structural transformation.
4. The open-ended virus genomes S5, S38, S42, S91 fall under band 4 and they are likely to be stable virus genomes without undergoing any structural transformation.
5. The close-ended virus genome S6 falls under band 4 and it is likely to be a stable virus genome without undergoing any structural transformation.

**III. CLASSIFICATION OF 130 VIRUS GENOMES BASED ON CORRELATION COEFFICIENTS AND DIFFERENTIAL ERRORS**

The percentage concentration values of all four nucleotides in 130 virus genomes are calculated and subsequently the values of correlation coefficients and differential errors.

**3.1 Calculation of Percentage Concentration of Nucleotides of Viral Genome Sequences**

Analogous to the notion of pH value of a solution, the values of pA, pT, pG and pC of a genome sequence and possibly composition of these values like the proportion pA:pT:pG:pC seems to pave a way to classify and characterize genome sets. The definition of “Percentage Nucleotide Concentration” of a genome sequence is as follows. “Given a genome sequence, the number of a particular nucleotide, say A, present in that genome sequence is counted and the sum is divided by the total number of nucleotides in that genome sequence. The fraction when multiplied by 100 yields the Percentage Concentration of Adenine pA”. Similarly, one can evaluate pT, pG and pC.

### 3.1.1 Percentage Concentrations of 130 sequences

One hundred and thirty complete genomes of various categories of viruses are drawn from NCBI web site <https://www.ncbi.nlm.nih.gov/genbank/sars-cov-2-seqs/> for data analysis. The data given by NCBI is expected to be genuine. Now, pA, pT, pG and pC were calculated for all 130 virus genome data and values presented in table 4.

Table 4: Percentage Nucleotide Concentration Data of 130 Virus Genomes

Sl. No.	Seq. No.	Species	Accession ID	Total length	A	pA	T	pT	G	pG	C	pC
1	S1	Lactate dehydrogenase-elevating virus	NC_001639.1	14104	3244	23.00	3888	27.57	3646	25.85	3326	23.58
2	S2	Porcine respiratory and reproductive syndrome virus	NC_001961.1	15428	3353	21.73	3903	25.30	4047	26.23	4125	26.74
3	S3	Equine arteritis virus	NC_002532.2	12704	2692	21.19	3449	27.15	3305	26.02	3258	25.65
4	S4	Simian hemorrhagic fever virus	NC_003092.2	15717	3534	22.49	4307	27.40	3556	22.63	4314	27.45
5	S5	Mikumi yellow baboon virus	NC_025112.1	14927	3476	23.29	3853	25.81	3404	22.80	4194	28.10
6	S6	Southwest baboon virus	NC_025113.1	14851	3456	23.27	3909	26.32	3398	22.88	4088	27.53
7	S7	Forest pouched giant rat arterivirus	NC_026439.1	14953	3238	21.65	3920	26.22	3944	26.38	3851	25.75
8	S8	DeBrazzas monkey arterivirus	NC_026509.1	15684	3364	21.45	3816	24.33	3688	23.51	4816	30.71
9	S9	Pebjah virus	NC_027124.1	15478	3291	21.26	4077	26.34	3623	23.41	4487	28.99
10	S10	Kafue Kinda chacma baboon virus	NC_029053.1	14924	3311	22.19	4025	26.97	3333	22.33	4255	28.51
11	S11	Free State vervet virus	NC_029992.1	15247	3450	22.63	3929	25.77	3457	22.67	4411	28.93
12	S12	Oliver's shrew virus 1	NC_035127.1	13766	2948	21.42	3675	26.70	3438	24.97	3705	26.91
13	S13	Porcine reproductive and respiratory syndrome virus 2	NC_038291.1	15411	3345	21.71	3923	25.46	4036	26.19	4107	26.65
14	S14	Simian hemorrhagic encephalitis virus	NC_038293.1	15370	3406	22.16	3979	25.89	3502	22.78	4481	29.15
15	S15	Lelystad virus	NC_043487.1	15111	3247	21.49	3912	25.89	3815	25.25	4137	27.38
16	S16	Zambian malbrouck virus	NC_048209.1	15307	3454	22.56	4008	26.18	3431	22.41	4414	28.84
17	S17	Giardia lamblia virus	NC_003555.1	6277	1606	25.59	1530	24.37	1565	24.93	1576	25.11
18	S18	Chalara elegans RNA Virus 1	NC_005883.1	5310	1076	20.26	1441	27.14	1465	27.59	1328	25.01
19	S19	Coniothyrium minitans RNA virus	NC_007523.1	4975	1008	20.26	1024	20.58	1416	28.46	1527	30.69
20	S20	Botryotinia fuckeliana totivirus 1	NC_009224.1	5261	1232	23.42	1151	21.88	1392	26.46	1486	28.25
21	S21	Black raspberry virus F	NC_009890.1	5077	1516	29.86	1252	24.66	1336	26.31	973	19.16
22	S22	Armigeres subalbatus virus	NC_014609.1	7510	2239	29.81	1784	23.75	1672	22.26	1815	24.17
23	S23	Beauveria bassiana victorivirus	NC_024151.1	5327	961	18.04	1251	23.48	1375	25.81	1740	32.66
24	S24	Botryosphaeria dothidea victorivirus 1	NC_025214.1	5322	1145	21.51	1111	20.88	1519	28.54	1547	29.07
25	S25	Camponotus yamaokai virus	NC_027212.1	5704	1811	31.75	1359	23.83	1414	24.79	1120	19.64
26	S26	<b>Golden shiner totivirus</b>	<b>NC_030295.1</b>	7788	2902	<b>37.26</b>	2249	28.88	1436	18.44	1201	<b>15.42</b>
27	S27	Fusarium poae victorivirus 1	NC_030867.1	5124	1022	19.95	1058	20.65	1419	27.69	1625	31.71
28	S28	Australian Anopheles totivirus	NC_035674.1	6203	1570	25.31	1551	25.00	1730	27.89	1352	21.80
29	S29	Aspergillus foetidus slow virus 1	NC_038928.1	5194	1016	19.56	1172	22.56	1441	27.74	1563	30.09
30	S30	Beauveria bassiana victorivirus 1	NC_038929.1	5228	1065	20.37	1271	24.31	1418	27.12	1474	28.19
31	S31	Diatom colony associated dsRNA virus	NC_040431.1	4671	1147	24.56	1129	24.17	1401	29.99	994	21.28
32	S32	<b>Gigaspora margarita giardia-like virus</b>	<b>NC_040632.1</b>	4947	1772	<b>35.82</b>	1606	32.46	947	19.14	622	<b>12.57</b>
33	S33	Fusarium asiaticum victorivirus 1	NC_040653.1	5281	929	17.59	959	18.16	1560	29.54	1833	34.71
34	S34	Diatom colony associated dsRNA virus 13 genomic RNA	NC_040659.1	5082	1211	23.83	1188	23.38	1610	31.68	1073	21.11
35	S35	Diatom colony associated dsRNA virus 11 genomic RNA	NC_040660.1	5160	1270	24.61	1282	24.84	1587	30.76	1021	19.79
36	S36	Diatom colony associated dsRNA virus 12 genomic RNA	NC_040775.1	5941	1784	30.03	1389	23.38	1583	26.65	1185	19.95
37	S37	Alternaria arborescens victorivirus	NC_040793.1	5206	1154	22.17	1029	19.77	1486	28.54	1537	29.52
38	S38	Guangdong greater green snake arterivirus	NC_046959.1	18410	4712	25.59	5381	29.23	3886	21.11	4431	24.07
39	S39	Pothos latent virus	NC_000939.2	4415	1129	25.57	1206	27.32	1143	25.89	937	21.22
40	S40	Bovine viral diarrhea virus	NC_001461.1	12573	4049	32.20	2767	22.01	3233	25.71	2524	20.07
41	S41	O'nyong-nyong virus	NC_001512.1	11835	3676	31.06	2440	20.62	2859	24.16	2860	24.17
42	S42	Ononis yellow mosaic virus	NC_001513.1	6211	1330	21.41	1748	28.14	973	15.67	2160	34.78
43	S43	Cell fusing agent virus	NC_001564.2	10682	2618	24.51	2614	24.47	2919	27.33	2531	23.69
44	S44	Bamboo mosaic virus	NC_001642.1	6366	1925	30.24	1218	19.13	1339	21.03	1884	29.59
45	S45	Odontoglossum ringspot virus	NC_001728.1	6618	1990	30.07	2013	30.42	1441	21.77	1174	17.74
46	S46	Rupestris stem pitting associated virus	NC_001948.1	8744	2430	27.79	2560	29.28	2075	23.73	1678	19.19
47	S47	Botrytis virus F	NC_002604.1	6827	1689	24.74	1494	21.88	1405	20.58	2239	32.80
48	S48	Banana mild mosaic virus	NC_002729.1	7352	2511	34.15	2096	28.51	1571	21.37	1174	15.97
49	S49	Aconitum latent virus	NC_002795.1	8657	2306	26.64	2278	26.31	2149	24.82	1924	22.22
50	S50	Garlic latent virus	NC_003557.1	8363	2282	27.29	2413	28.85	1943	23.23	1725	20.63

51	S51	Groundnut rosette virus	NC_003603.1	4019	970	24.14	812	20.20	1176	29.26	1061	26.40
52	S52	Hibiscus chlorotic ringspot virus	NC_003608.1	3911	975	24.93	1012	25.88	1007	25.75	917	23.45
53	S53	<b>Phytophthora blight virus</b>	<b>NC_003634.1</b>	6673	1503	22.52	1588	23.80	874	13.10	2708	40.58
54	S54	Border disease virus	NC_003679.1	12333	4007	32.49	2715	22.01	3203	25.97	2408	19.52
55	S55	Obuda pepper virus	NC_003852.1	6507	1949	29.95	1867	28.69	1520	23.36	1171	18.00
56	S56	Aura virus	NC_003900.1	1824	3462	29.28	2628	22.23	2828	23.92	2906	24.58
57	S57	Grapevine rootstock stem lesion associated virus	NC_004724.1	16527	4250	25.72	4624	27.98	4224	25.56	3429	20.75
58	S58	Omsk hemorrhagic fever virus	NC_005062.1	10787	2772	25.70	2229	20.66	3361	31.16	2425	22.48
59	S59	Botrytis virus X	NC_005132.1	6966	1813	26.03	1348	19.35	1288	18.49	2517	36.13
60	S60	Olive mild mosaic virus	NC_006939.1	3683	1041	28.27	868	23.57	900	24.44	874	23.73
61	S61	Angelonia flower break virus	NC_007733.2	3962	1054	26.60	937	23.65	990	24.99	981	24.76
62	S62	Ilheus virus	NC_009028.2	10755	2903	26.99	2229	20.73	3004	27.93	2619	24.35
63	S63	Peach chlorotic mottle virus	NC_009892.1	9005	2392	26.56	2669	29.64	2042	22.68	1902	21.12
64	S64	Grapevine Algerian latent virus	NC_011535.1	4731	1243	26.27	1167	24.67	1319	27.88	1002	21.18
65	S65	Nemosa ring necrosis virus	NC_011538.1	6285	1223	19.46	1467	23.34	1210	19.25	2385	37.95
66	S66	Peach mosaic virus	NC_011552.1	7988	2439	30.53	2249	28.15	1908	23.89	1392	17.43
67	S67	Anagryis vein yellowing virus	NC_011559.1	6151	1437	23.36	1585	25.77	1154	18.76	1975	32.11
68	S68	Kedougou virus strain DakAar D1470	NC_012533.1	10723	2922	27.25	2143	19.99	3219	30.02	2439	22.75
69	S69	Bagaza virus strain DakAr B209	NC_012534.1	10941	3190	29.16	2330	21.30	3088	28.22	2333	21.32
70	S70	Bovine viral diarrhoea virus 3	NC_012812.1	12337	3945	31.98	2694	21.84	3266	26.47	2426	19.66
71	S71	Kalanchoe latent virus	NC_013006.1	8517	2381	27.96	2215	26.01	2060	24.19	1857	21.80
72	S72	Grapevine Pinot gris virus	NC_015782.2	7259	2181	30.05	2230	30.72	1582	21.79	1266	17.44
73	S73	Brassica yellows virus	NC_016038.2	5666	1625	28.68	1300	22.94	1330	23.47	1411	24.90
74	S74	Actinidia virus B	NC_016404.1	7488	2449	32.71	1558	20.81	1897	25.33	1584	21.15
75	S75	Garlic common latent virus	NC_016440.1	8638	2313	26.78	2473	28.63	2136	24.73	1716	19.87
76	S76	Ndumu virus	NC_016959.1	1724	3259	27.80	2588	22.07	3210	27.38	2632	22.45
77	S77	Pestivirus strain	NC_018713.1	12292	3933	32.00	2728	22.19	3173	25.81	2458	20.00
78	S78	Andean potato latent virus	NC_020470.1	6337	1425	22.49	1742	27.49	892	14.08	2278	35.95
79	S79	Andean potato mild mosaic virus	NC_020471.1	6226	1344	21.59	1756	28.20	970	15.58	2154	34.60
80	S80	Kama virus	NC_023439.1	10688	2449	22.91	2280	21.33	3620	33.87	2339	21.88
81	S81	Gaillardia latent virus	NC_023892.1	8659	2457	28.38	2086	24.09	2286	26.40	1829	21.12
82	S82	Pitaya virus X	NC_024458.1	6677	1800	26.96	1465	21.94	1518	22.73	1894	28.37
83	S83	Middelburg virus	NC_024887.1	11550	3125	27.06	2313	20.03	3118	27.00	2994	25.92
84	S84	Jutiapa virus	NC_026620.1	10125	2916	28.80	2599	25.67	2752	27.18	1858	18.35
85	S85	Phasey bean mild yellows virus	NC_028793.2	5851	1510	25.81	1427	24.39	1460	24.95	1413	24.15
86	S86	Grapevine Red Globe virus	NC_030693.1	6863	1277	18.61	1476	21.51	1202	17.51	2908	42.37
87	S87	Anopheles flavivirus	NC_031327.1	10588	2729	25.77	2629	24.83	2889	27.29	2341	22.11
88	S88	Potexvirus	NC_040842.1	5839	1554	26.61	1260	21.58	1257	21.53	1768	30.28
89	S89	Grapevine associated tymo-like virus	NC_040837.1	6060	1699	28.04	1942	32.05	1503	24.80	916	15.12
90	S90	Actinidia seed-borne latent virus	NC_040800.1	8192	2847	34.75	2187	26.70	1884	23.00	1274	15.55
91	S91	Kampung Karu virus	NC_040788.1	10311	2885	27.98	2271	22.03	3053	29.61	2102	20.39
92	S92	Rocio virus	NC_040776.1	10794	2976	27.57	2169	20.09	3066	28.40	2583	23.93
93	S93	Bovine viral diarrhoea virus	NC_039237.1	12513	4045	32.33	2783	22.24	3186	25.46	2499	19.97
94	S94	Kyasaur forest disease virus	NC_039218.1	10376	2509	24.18	2139	20.61	3314	31.94	2414	23.27
95	S95	Phaseolus vulgaris endornavirus	NC_039217.1	14072	4836	34.37	3986	28.33	2783	19.78	2467	17.53
96	S96	Atractylodes mottle virus	NC_038966.1	8881	2426	27.32	2429	27.35	2180	24.55	1846	20.79
97	S97	Babaco mosaic virus	NC_036587.1	6692	2080	31.08	1479	22.10	1452	21.70	1681	25.12
98	S98	Ocimum basilicum RNA virus	NC_035462.1	6930	2255	32.54	2046	29.52	1554	22.42	1075	15.51
99	S99	Actinidia virus	NC_035453.1	18848	5414	28.72	5495	29.15	4453	23.63	3486	18.50
100	S100	Lake Sinai Virus	NC_035116.1	5877	1129	19.21	1730	29.44	1417	24.11	1601	27.24
101	S101	Apis flavivirus	NC_035071.1	20414	6142	30.09	4537	22.22	4813	23.58	4922	24.11
102	S102	Agave tequilana leaf virus	NC_034833.1	6958	2140	30.76	1572	22.59	1766	25.38	1480	21.27
103	S103	Ochlerotatus caspius flavivirus	NC_034242.1	10370	2787	26.88	2704	26.08	2748	26.50	2131	20.55
104	S104	Lagenaria siceraria endornavirus	NC_034216.1	15098	5582	36.97	3901	25.84	3120	20.66	2495	16.53
105	S105	African eggplant yellowing virus	NC_034207.1	5953	1588	26.68	1380	23.18	1517	25.48	1468	24.66
106	S106	Grapevine rupestris vein feathering virus	NC_034205.1	6730	1205	17.90	1598	23.74	1424	21.16	2498	37.12
107	S107	Peach virus	NC_033828.1	6612	1287	19.46	1277	19.31	1398	21.14	2650	40.08
108	S108	Bamaga virus	NC_033725.1	10203	2866	28.09	2441	23.92	2878	28.21	2018	19.78
109	S109	Kadam virus	NC_033724.1	10215	2601	25.46	2229	21.82	3177	31.10	2208	21.62
110	S110	Gadgets Gully virus	NC_033723.1	10251	2701	26.35	2209	21.55	3207	31.28	2134	20.82
111	S111	Jugra virus	NC_033699.1	10173	2948	28.98	2319	22.80	2751	27.04	2155	21.18
112	S112	Boubou virus	NC_033693.1	10173	2940	28.90	2338	22.98	2744	26.97	2151	21.14
113	S113	New Mapoon virus	NC_032088.1	10864	2872	26.44	2395	22.05	2947	27.13	2650	24.39
114	S114	Botrytis cinerea endornavirus	NC_031752.1	11557	4006	34.66	3151	27.26	2231	19.30	2169	18.77
115	S115	Ceratobasidium endornavirus	NC_031463.1	23635	9624	40.72	6181	26.15	4217	17.84	3613	15.29
116	S116	Ceratobasidium endorna virus	NC_031462.1	15207	4392	28.88	3640	23.94	3787	24.90	3388	22.28
117	S117	Banzi virus	NC_043110.1	10182	2778	27.28	2274	22.33	2867	28.16	2263	22.23

118	S118	White bream virus	NC 008516.1	26660	8975	33.66	6143	23.04	3758	14.10	7784	29.20
119	S119	Chinook salmon bafinivirus	NC 026812.1	27004	8268	30.62	6892	25.52	3909	14.48	7929	29.36
120	S120	Fathead minnow nidovirus	NC 038295.1	27318	9234	33.80	6954	25.46	4110	15.05	7020	25.70
121	S121	Bovine nidovirus	NC 027199.1	20261	5947	29.35	6839	33.75	5208	25.70	2266	11.18
122	S122	Ball python nidovirus	NC 024709.1	33452	11117	33.23	6932	20.72	6663	19.92	8740	26.13
123	S123	Xinzhou toro-like virus	NC 033700.1	30353	9555	31.48	9194	30.29	4934	16.26	6670	21.97
124	S124	Morelia viridis nidovirus	NC 035465.1	32399	10710	33.06	7139	22.03	6186	19.09	8364	25.82
125	S125	Breda virus	NC 007447.1	28475	7690	27.01	9959	34.97	6104	21.44	4718	16.57
126	S126	Porcine torovirus	NC 022787.1	28301	8124	28.71	10267	36.28	5579	19.71	4331	15.30
127	S127	Goat torovirus	NC 034976.1	28487	7686	26.98	10072	35.36	6083	21.35	4646	16.31
128	S128	Belling River virus	NC 046956.1	30742	10484	34.10	5273	17.15	5238	17.04	9747	31.71
129	S129	Hainan hebius popei torovirus	NC_046962.1	29409	8759	29.78	1760	39.99	6561	22.31	2328	7.92
130	S130	Guangdong red-banded snake torovirus	NC_046963.1	30859	9285	30.09	8097	26.24	5722	18.54	7755	25.13

The values of pA, pT, pG and pC are studied independently as well as pairwise. Figs. 6 to 9 respectively show graphs of pA, pT, pG and pC values of all 130 virus genome sequences.

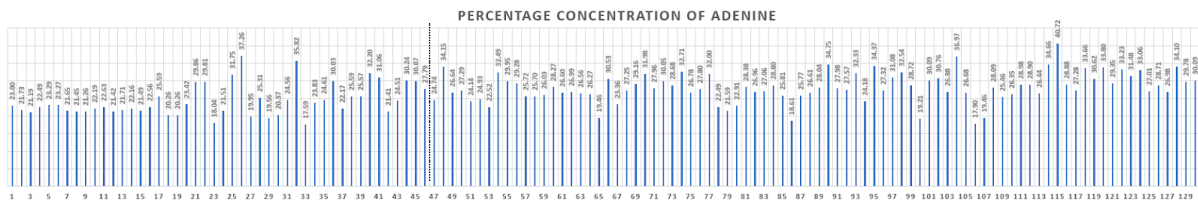


Fig. 6: pA values of all 130 virus genome sequences

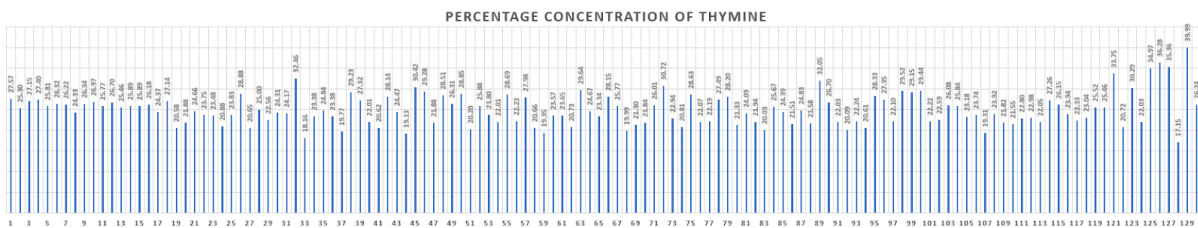


Fig. 7: pT values of all 130 virus genome sequences

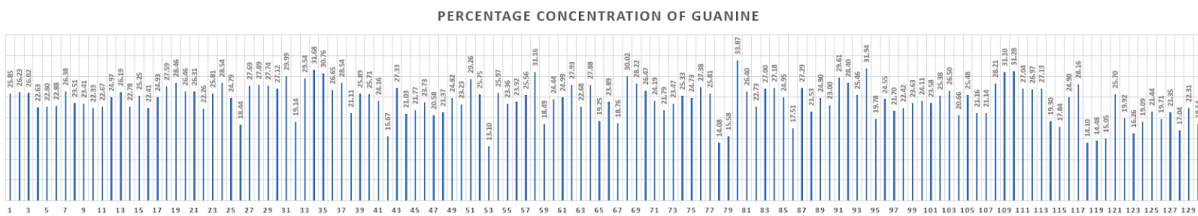


Fig. 8: pG values of all 130 virus genome sequences

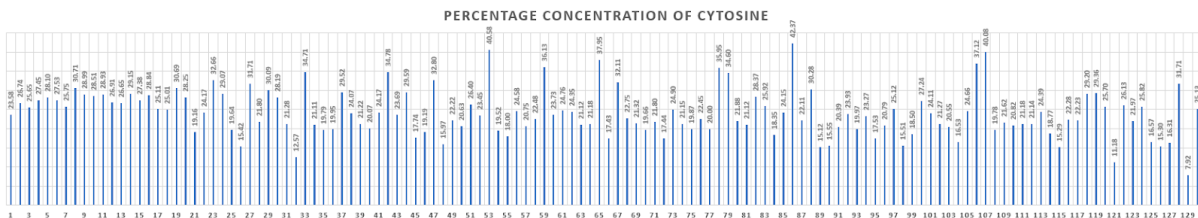


Fig. 9: pC values of all 130 virus genome sequences

Figs. 10 to 16 respectively show graphs of pA, pT, pG and pC values together, pA and pT values together, pA and pG values together, pA and pC values together, pT and pG values together, pT and pC values together, pG and pC values together of all 130 virus genomes.

Figs. 10 to shows Percentage Concentrations of Adenine, Thymine, Guanine and Cytosine

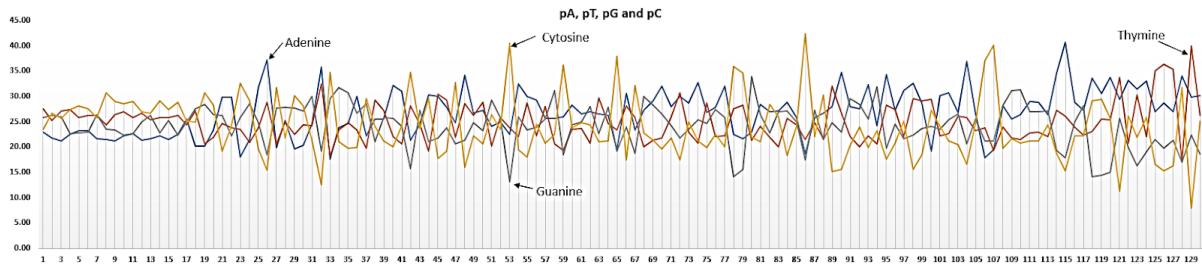


Fig. 10: pA, pT, pG and pC values of all 130 virus genome sequences

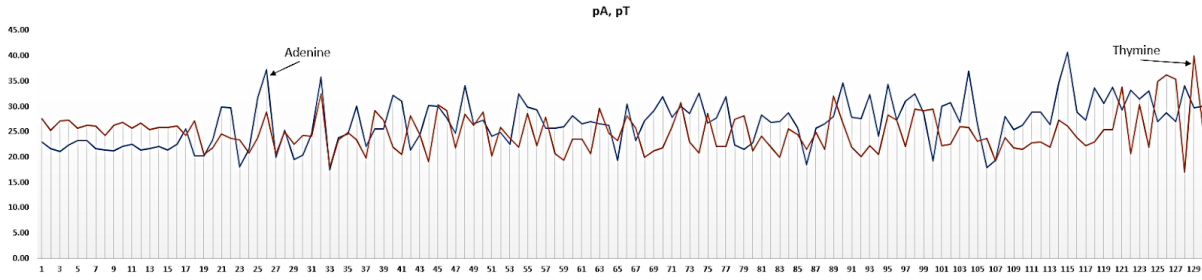


Fig. 11: pA and pT values of all 130 virus genome sequences

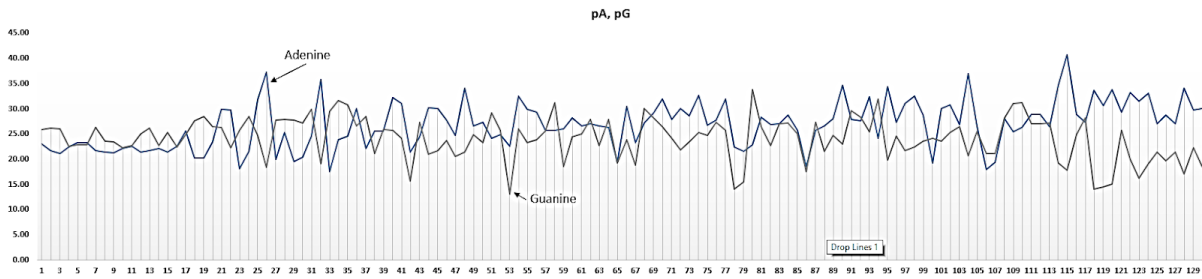


Fig. 12: pA and pG values of all 130 virus genome sequences

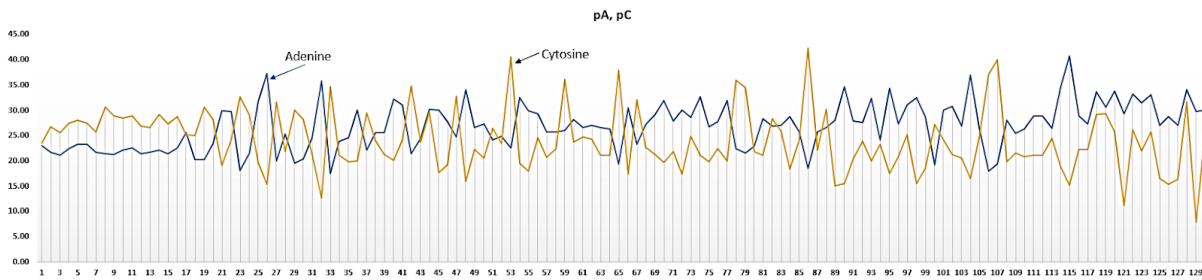


Fig. 13: pA and pC values of all 130 virus genome sequences

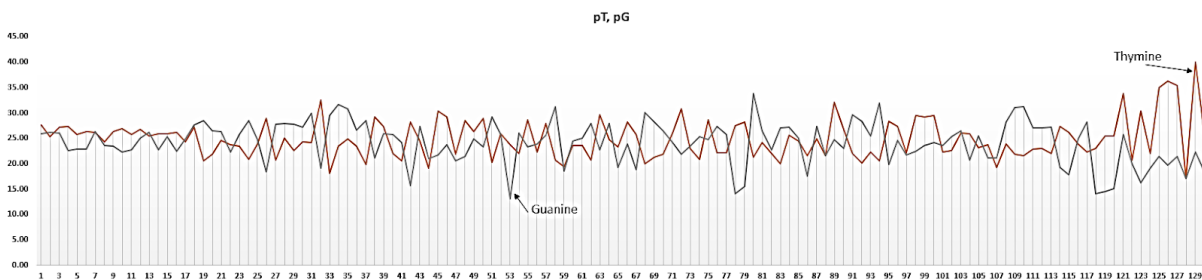


Fig. 14: pT and pG values of all 130 virus genome sequences

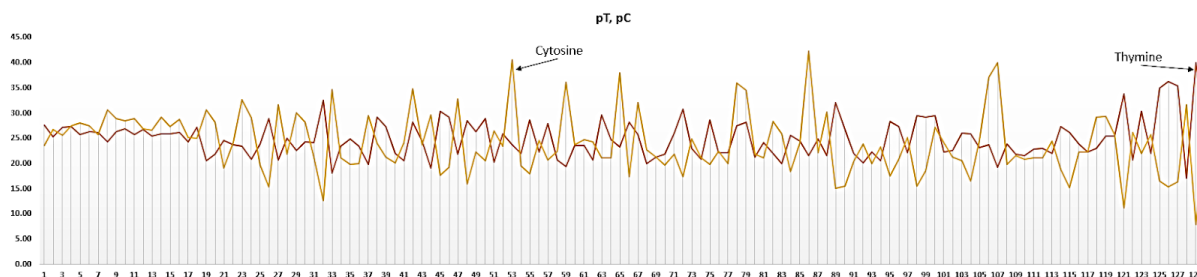


Fig. 15: pT and pC values of all 130 virus genome sequences

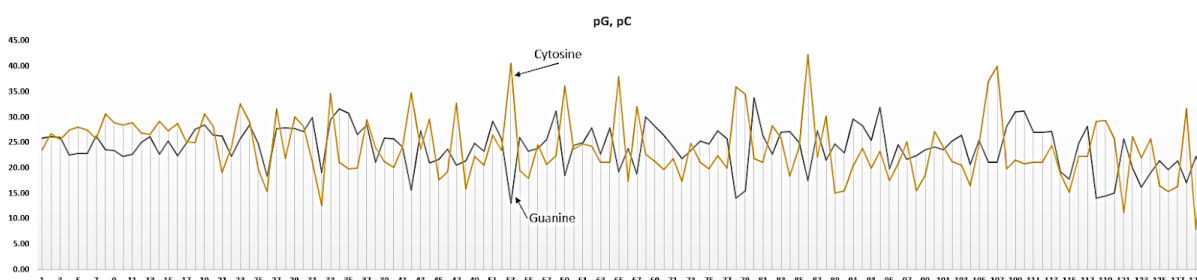


Fig. 16: pG and pC values of all 130 virus genome sequences

### 3.2 Correlation Coefficients of Pairs of Nucleotides

Now, the values of pA-pT, pA-pG, pA-pC, pT-pG, pT-pC and pG-pC are calculated and results shown in table 5. Resulting values are termed as 'errors' for convenience.

Table 5: Values of pA-pT, pA-pG, pA-pC, pT-pG, pT-pC and pG-pC

Seq. No.	pA	pT	pG	pC	pA-pT	pA-pG	pA-pC	pT-pG	pT-pC	pG-pC
S1	23.00	27.57	25.85	23.58	-4.57	-2.85	-0.58	1.72	3.98	2.27
S2	21.73	25.30	26.23	26.74	-3.56	-4.50	-5.00	-0.93	-1.44	-0.51
S3	21.19	27.15	26.02	25.65	-5.96	-4.83	-4.46	1.13	1.50	0.37
S4	22.49	27.40	22.63	27.45	-4.92	-0.14	-4.96	4.78	-0.04	-4.82
S5	23.29	25.81	22.80	28.10	-2.53	0.48	-4.81	3.01	-2.28	-5.29
S6	23.27	26.32	22.88	27.53	-3.05	0.39	-4.26	3.44	-1.21	-4.65
S7	21.65	26.22	26.38	25.75	-4.56	-4.72	-4.10	-0.16	0.46	0.62
S8	21.45	24.33	23.51	30.71	-2.88	-2.07	-9.26	0.82	-6.38	-7.19
S9	21.26	26.34	23.41	28.99	-5.08	-2.14	-7.73	2.93	-2.65	-5.58
S10	22.19	26.97	22.33	28.51	-4.78	-0.15	-6.33	4.64	-1.54	-6.18
S11	22.63	25.77	22.67	28.93	-3.14	-0.05	-6.30	3.10	-3.16	-6.26
S12	21.42	26.70	24.97	26.91	-5.28	-3.56	-5.50	1.72	-0.22	-1.94
S13	21.71	25.46	26.19	26.65	-3.75	-4.48	-4.94	-0.73	-1.19	-0.46
S14	22.16	25.89	22.78	29.15	-3.73	-0.62	-6.99	3.10	-3.27	-6.37
S15	21.49	25.89	25.25	27.38	-4.40	-3.76	-5.89	0.64	-1.49	-2.13
S16	22.56	26.18	22.41	28.84	-3.62	0.15	-6.27	3.77	-2.65	-6.42
S17	25.59	24.37	24.93	25.11	1.21	0.65	0.48	-0.56	-0.73	-0.18
S18	20.26	27.14	27.59	25.01	-6.87	-7.33	-4.75	-0.45	2.13	2.58
S19	20.26	20.58	28.46	30.69	-0.32	-8.20	-10.43	-7.88	-10.11	-2.23
S20	23.42	21.88	26.46	28.25	1.54	-3.04	-4.83	-4.58	-6.37	-1.79
S21	29.86	24.66	26.31	19.16	5.20	3.55	10.70	-1.65	5.50	7.15
S22	29.81	23.75	22.26	24.17	6.06	7.55	5.65	1.49	-0.41	-1.90
S23	18.04	23.48	25.81	32.66	-5.44	-7.77	-14.62	-2.33	-9.18	-6.85
S24	21.51	20.88	28.54	29.07	0.64	-7.03	-7.55	-7.67	-8.19	-0.53
S25	31.75	23.83	24.79	19.64	7.92	6.96	12.11	-0.96	4.19	5.15
S26	37.26	28.88	18.44	15.42	8.38	18.82	21.84	10.44	13.46	3.02
S27	19.95	20.65	27.69	31.71	-0.70	-7.75	-11.77	-7.05	-11.07	-4.02
S28	25.31	25.00	27.89	21.80	0.31	-2.58	3.51	-2.89	3.21	6.09

S29	19.56	22.56	27.74	30.09	-3.00	-8.18	-10.53	-5.18	-7.53	-2.35
S30	20.37	24.31	27.12	28.19	-3.94	-6.75	-7.82	-2.81	-3.88	-1.07
S31	24.56	24.17	29.99	21.28	0.39	-5.44	3.28	-5.82	2.89	8.71
S32	35.82	32.46	19.14	12.57	3.36	16.68	<b>23.25</b>	13.32	19.89	6.57
S33	17.59	18.16	29.54	34.71	-0.57	-11.95	-17.12	-11.38	-16.55	-5.17
S34	23.83	23.38	31.68	21.11	0.45	-7.85	2.72	-8.30	2.26	10.57
S35	24.61	24.84	30.76	19.79	-0.23	-6.14	4.83	-5.91	5.06	10.97
S36	30.03	23.38	26.65	19.95	6.65	3.38	10.08	-3.27	3.43	6.70
S37	22.17	19.77	28.54	29.52	2.40	-6.38	-7.36	-8.78	-9.76	-0.98
S38	25.59	29.23	21.11	24.07	-3.63	4.49	1.53	8.12	5.16	-2.96
S39	25.57	27.32	25.89	21.22	-1.74	-0.32	4.35	1.43	6.09	4.67
S40	32.20	22.01	25.71	20.07	10.20	6.49	12.13	-3.71	1.93	5.64
S41	31.06	20.62	24.16	24.17	10.44	6.90	6.89	-3.54	-3.55	-0.01
S42	21.41	28.14	15.67	34.78	-6.73	5.75	-13.36	12.48	-6.63	-19.11
S43	24.51	24.47	27.33	23.69	0.04	-2.82	0.81	-2.86	0.78	3.63
S44	30.24	19.13	21.03	29.59	11.11	9.21	0.64	-1.90	-10.46	-8.56
S45	30.07	30.42	21.77	17.74	-0.35	8.30	12.33	8.64	12.68	4.03
S46	27.79	29.28	23.73	19.19	-1.49	4.06	8.60	5.55	10.09	4.54
S47	24.74	21.88	20.58	32.80	2.86	4.16	-8.06	1.30	-10.91	-12.22
S48	34.15	28.51	21.37	15.97	5.64	12.79	18.19	7.14	12.54	5.40
S49	26.64	26.31	24.82	22.22	0.32	1.81	4.41	1.49	4.09	2.60
S50	27.29	28.85	23.23	20.63	-1.57	4.05	6.66	5.62	8.23	2.61
S51	24.14	20.20	29.26	26.40	3.93	-5.13	-2.26	-9.06	-6.20	2.86
S52	24.93	25.88	25.75	23.45	-0.95	-0.82	1.48	0.13	2.43	2.30
S53	22.52	23.80	13.10	40.58	-1.27	9.43	-18.06	10.70	-16.78	<b>-27.48</b>
S54	32.49	22.01	25.97	19.52	10.48	6.52	12.97	-3.96	2.49	6.45
S55	29.95	28.69	23.36	18.00	1.26	6.59	11.96	5.33	10.70	5.36
S56	29.28	22.23	23.92	24.58	7.05	5.36	4.70	-1.69	-2.35	-0.66
S57	25.72	27.98	25.56	20.75	-2.26	0.16	4.97	2.42	7.23	4.81
S58	25.70	20.66	31.16	22.48	5.03	-5.46	3.22	-10.49	-1.82	8.68
S59	26.03	19.35	18.49	36.13	6.68	7.54	-10.11	0.86	-16.78	-17.64
S60	28.27	23.57	24.44	23.73	4.70	3.83	4.53	-0.87	-0.16	0.71
S61	26.60	23.65	24.99	24.76	2.95	1.62	1.84	-1.34	-1.11	0.23
S62	26.99	20.73	27.93	24.35	6.27	-0.94	2.64	-7.21	-3.63	3.58
S63	26.56	29.64	22.68	21.12	-3.08	3.89	5.44	6.96	8.52	1.55
S64	26.27	24.67	27.88	21.18	1.61	-1.61	5.09	-3.21	3.49	6.70
S65	19.46	23.34	19.25	37.95	-3.88	0.21	-18.49	4.09	-14.61	-18.70
S66	30.53	28.15	23.89	17.43	2.38	6.65	13.11	4.27	10.73	6.46
S67	23.36	25.77	18.76	32.11	-2.41	4.60	-8.75	7.01	-6.34	-13.35
S68	27.25	19.99	30.02	22.75	7.26	-2.77	4.50	-10.03	-2.76	7.27
S69	29.16	21.30	28.22	21.32	7.86	0.93	7.83	-6.93	-0.03	6.90
S70	31.98	21.84	26.47	19.66	10.14	5.50	12.31	-4.64	2.17	6.81
S71	27.96	26.01	24.19	21.80	1.95	3.77	6.15	1.82	4.20	2.38
S72	30.05	30.72	21.79	17.44	-0.68	8.25	12.61	8.93	13.28	4.35
S73	28.68	22.94	23.47	24.90	5.74	5.21	3.78	-0.53	-1.96	-1.43
S74	32.71	20.81	25.33	21.15	11.90	7.37	11.55	-4.53	-0.35	4.18
S75	26.78	28.63	24.73	19.87	-1.85	2.05	6.91	3.90	8.76	4.86
S76	27.80	22.07	27.38	22.45	5.72	0.42	5.35	-5.31	-0.38	4.93
S77	32.00	22.19	25.81	20.00	9.80	6.18	12.00	-3.62	2.20	5.82
S78	22.49	27.49	14.08	35.95	-5.00	8.41	-13.46	13.41	-8.46	-21.87
S79	21.59	28.20	15.58	34.60	-6.62	6.01	-13.01	12.62	-6.39	-19.02
S80	22.91	21.33	33.87	21.88	1.58	<b>-10.96</b>	1.03	<b>-12.54</b>	-0.55	11.99
S81	28.38	24.09	26.40	21.12	4.28	1.97	7.25	-2.31	2.97	5.28
S82	26.96	21.94	22.73	28.37	5.02	4.22	-1.41	-0.79	-6.43	-5.63
S83	27.06	20.03	27.00	25.92	7.03	0.06	1.13	-6.97	-5.90	1.07
S84	28.80	25.67	27.18	18.35	3.13	1.62	10.45	-1.51	7.32	8.83
S85	25.81	24.39	24.95	24.15	1.42	0.85	1.66	-0.56	0.24	0.80
S86	18.61	21.51	17.51	42.37	-2.90	1.09	<b>-23.77</b>	3.99	<b>-20.87</b>	-24.86
S87	25.77	24.83	27.29	22.11	0.94	-1.51	3.66	-2.46	2.72	5.18
S88	26.61	21.58	21.53	30.28	5.04	5.09	-3.67	0.05	-8.70	-8.75
S89	28.04	32.05	24.80	15.12	-4.01	3.23	12.92	7.24	16.93	9.69
S90	34.75	26.70	23.00	15.55	8.06	11.76	19.20	3.70	11.15	7.45

S91	27.98	22.03	29.61	20.39	5.95	-1.63	7.59	-7.58	1.64	9.22
S92	27.57	20.09	28.40	23.93	7.48	-0.83	3.64	-8.31	-3.84	4.47
S93	32.33	22.24	25.46	19.97	10.09	6.86	12.36	-3.22	2.27	5.49
S94	24.18	20.61	31.94	23.27	3.57	-7.76	0.92	-11.32	-2.65	8.67
S95	34.37	28.33	19.78	17.53	6.04	14.59	16.83	8.55	10.79	2.25
S96	27.32	27.35	24.55	20.79	-0.03	2.77	6.53	2.80	6.56	3.76
S97	31.08	22.10	21.70	25.12	8.98	9.38	5.96	0.40	-3.02	-3.42
S98	32.54	29.52	22.42	15.51	3.02	10.12	17.03	7.10	14.01	6.91
S99	28.72	29.15	23.63	18.50	-0.43	5.10	10.23	5.53	10.66	5.13
S100	19.21	29.44	24.11	27.24	-10.23	-4.90	-8.03	5.33	2.19	-3.13
S101	30.09	22.22	23.58	24.11	7.86	6.51	5.98	-1.35	-1.89	-0.53
S102	30.76	22.59	25.38	21.27	8.16	5.38	9.49	-2.79	1.32	4.11
S103	26.88	26.08	26.50	20.55	0.80	0.38	6.33	-0.42	5.53	5.95
S104	36.97	25.84	20.66	16.53	11.13	16.31	20.45	5.17	9.31	4.14
S105	26.68	23.18	25.48	24.66	3.49	1.19	2.02	-2.30	-1.48	0.82
S106	17.90	23.74	21.16	37.12	-5.84	-3.25	-19.21	2.59	-13.37	-15.96
S107	19.46	19.31	21.14	40.08	0.15	-1.68	-20.61	-1.83	-20.77	-18.94
S108	28.09	23.92	28.21	19.78	4.17	-0.12	8.31	-4.28	4.15	8.43
S109	25.46	21.82	31.10	21.62	3.64	-5.64	3.85	-9.28	0.21	9.49
S110	26.35	21.55	31.28	20.82	4.80	-4.94	5.53	-9.74	0.73	10.47
S111	28.98	22.80	27.04	21.18	6.18	1.94	7.80	-4.25	1.61	5.86
S112	28.90	22.98	26.97	21.14	5.92	1.93	7.76	-3.99	1.84	5.83
S113	26.44	22.05	27.13	24.39	4.39	-0.69	2.04	-5.08	-2.35	2.73
S114	34.66	27.26	19.30	18.77	7.40	15.36	15.90	7.96	8.50	0.54
S115	40.72	26.15	17.84	15.29	14.57	<b>22.88</b>	<b>25.43</b>	8.31	10.87	2.56
S116	28.88	23.94	24.90	22.28	4.95	3.98	6.60	-0.97	1.66	2.62
S117	27.28	22.33	28.16	22.23	4.95	-0.87	5.06	-5.82	0.11	5.93
S118	33.66	23.04	14.10	29.20	10.62	19.57	4.47	8.95	-6.16	-15.10
S119	30.62	25.52	14.48	29.36	5.10	16.14	1.26	11.05	-3.84	-14.89
S120	33.80	25.46	15.05	25.70	8.35	18.76	8.10	10.41	-0.24	-10.65
S121	29.35	33.75	25.70	11.18	-4.40	3.65	18.17	8.05	22.57	<b>14.52</b>
S122	33.23	20.72	19.92	26.13	12.51	13.31	7.11	0.80	-5.40	-6.21
S123	31.48	30.29	16.26	21.97	1.19	15.22	9.50	14.03	8.32	-5.72
S124	33.06	22.03	19.09	25.82	11.02	13.96	7.24	2.94	-3.78	-6.72
S125	27.01	34.97	21.44	16.57	-7.97	5.57	10.44	13.54	18.41	4.87
S126	28.71	36.28	19.71	15.30	-7.57	8.99	13.40	16.56	20.97	4.41
S127	26.98	35.36	21.35	16.31	-8.38	5.63	10.67	14.00	19.05	5.04
S128	34.10	17.15	17.04	31.71	<b>16.95</b>	17.06	2.40	0.11	-14.55	-14.67
S129	29.78	39.99	22.31	7.92	<b>-10.20</b>	7.47	21.87	<b>17.68</b>	<b>32.07</b>	14.39
S130	30.09	26.24	18.54	25.13	3.85	11.55	4.96	7.70	1.11	-6.59

Figs. 17 to 22 respectively show graphs of pA-pT error values, pA-pG error values, pA-pC error values, pT-pG error values, pT-pC error values and pG-pC error values.

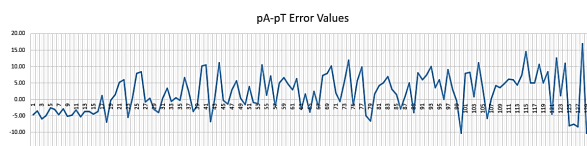


Fig. 17: Graph of pA-pT error values

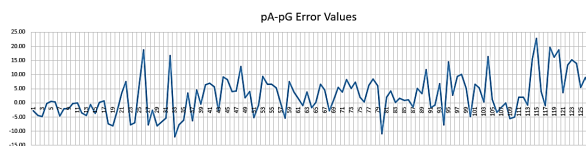


Fig. 18: Graph of pA-pG error values

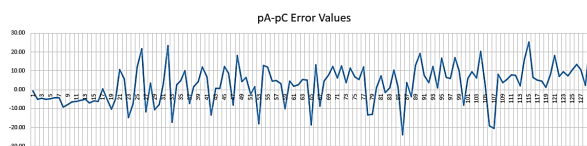


Fig. 19: Graph of pA-pC error values

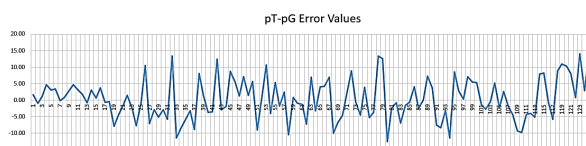


Fig. 20: Graph of pT-pG error values

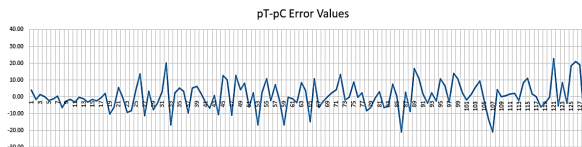


Fig. 21: Graph of pT-pC error values

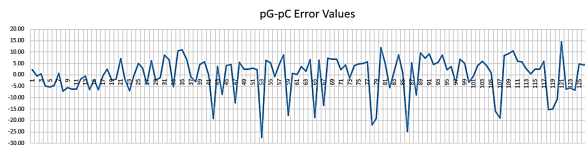


Fig. 22: Graph of pG-pC error values

### 3.2.1 Measure of Correlation Coefficients of Pairs of Nucleotides

Given a finite number of genome sequences (in this case 130), the values of pA, pT, pG and pC are evaluated. Then, the values of pA-pT, pA-pG, pA-pC, pT-pG, pT-pC and pG-pC are evaluated. Now, the Correlation Coefficient  $\tau$  of A and T, for instance, is defined as  $\tau(A,T) = \text{Max}(pA-pT) - \text{Min}(pA-pT)$ . Less the error, more the correlation coefficient  $\tau$  of A and T. This could be generalized to all other nucleotide combinations. Correlation Coefficient  $\tau$  of each of the ordered pair is calculated as

1.  $\tau(A,T) = \text{Max}(pA-pT) - \text{Min}(pA-pT) = 16.95 + 10.20 = 27.15$
2.  $\tau(A,G) = \text{Max}(pA-pG) - \text{Min}(pA-pG) = 22.88 + 10.96 = 33.84$
3.  $\tau(A,C) = \text{Max}(pA-pC) - \text{Min}(pA-pC) = 25.43 + 23.77 = 49.20$
4.  $\tau(T,G) = \text{Max}(pT-pG) - \text{Min}(pT-pG) = 17.68 + 12.54 = 30.22$
5.  $\tau(T,C) = \text{Max}(pT-pC) - \text{Min}(pT-pC) = 32.07 + 20.87 = 52.94$
6.  $\tau(G,C) = \text{Max}(pG-pC) - \text{Min}(pG-pC) = 14.52 + 27.48 = 42$

The correlation coefficients  $\tau$  of all possible pairs of nucleotides are shown in Fig. 23

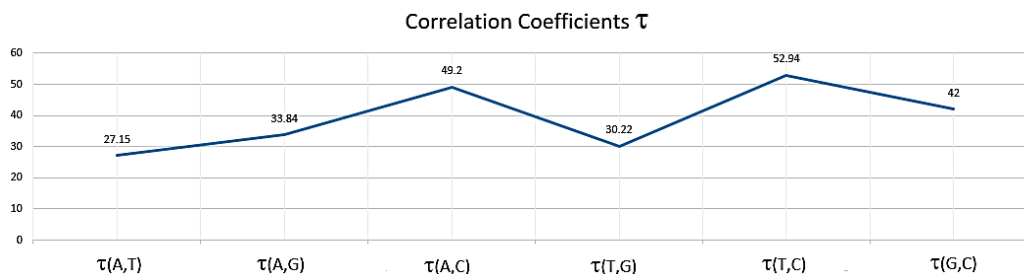


Fig. 23: Correlation coefficients of all possible pairs of nucleotides

#### Observations

1. Pairwise Correlation Coefficient  $\tau(A,T) = 27.15$
2. Pairwise Correlation Coefficient  $\tau(A,G) = 33.84$
3. Pairwise Correlation Coefficient  $\tau(A,C) = 49.2$
4. Pairwise Correlation Coefficient  $\tau(T,G) = 30.22$
5. Pairwise Correlation Coefficient  $\tau(T,C) = 52.94$
6. Pairwise Correlation Coefficient  $\tau(G,C) = 42$
7. Less the value of  $\tau$ , more the ordered couple of nucleotides is correlated.
8.  $\langle A,T \rangle$  form natural pair and highly correlated.
9.  $\langle A,G \rangle$  form unnatural pair and correlated to a certain extent.
10.  $\langle A,C \rangle$  form unnatural pair and moderately uncorrelated.
11.  $\langle T,G \rangle$  form unnatural pair and correlated to a great extent.
12.  $\langle T,C \rangle$  form unnatural pair and highly uncorrelated.
13.  $\langle G,C \rangle$  form natural pair and uncorrelated.
14. **One can select natural pairing of  $\langle A,T \rangle$  and  $\langle G,C \rangle$ . Note that the pair  $\langle A,T \rangle$  is correlated and the pair  $\langle G,C \rangle$  uncorrelated.**
15. **One can select unnatural pairing of  $\langle T,G \rangle$  and  $\langle A,C \rangle$ . Note that the pair  $\langle T,G \rangle$  is correlated and the pair  $\langle A,C \rangle$  is uncorrelated – a significant observation**

To summarize, a total of 130 virus sequences related to Corona virus family was tested in terms of percentage concentrations of nucleotides. The significant observation made after analyzing all 130 sequences, the natural pair of nucleotides Adenine and Thymine are highly correlated, whereas the natural pair of Guanine and Cytosine are uncorrelated. Similarly, the unnatural pair of Thymine and Guanine are correlated and the unnatural pair of Adenine and Cytosine are uncorrelated.

#### Inference

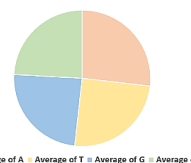
**This phenomenon seems to be a natural occurrence in all genome sequences.**

### 3.3 Differential Errors of percentage concentrations of nucleotides with respect to their averages

From entries in Table 6, one would observe that the average percentage concentrations of all nucleotides are almost same. This amounts to saying that all four nucleotides are almost equally distributed in any genome sequence. Details are given below.

**Table 6:** Percentage Concentration Averages of Nucleotides

Percentage Concentration Averages of Nucleotides	
Average of pA	26.84
Average of pT	24.92
Average of pG	24.12
Average of pC	24.11



The percentage concentration of nucleotides in all 130 virus sequences have been evaluated and the average values found as pA = 26.84, pT = 24.92, pG = 24.12 and pC = 24.11. Table 7 presents the percentage concentration values of each nucleotide subtracted from the respective average values.

**Table 7:** Values of pA, pT, pG and pC subtracted from their average values

Sequences	pA	pT	pG	pC	Avg-pA	Avg-pT	Avg-pG	Avg-pC
S1	23.00	27.57	25.85	23.58	3.84	-2.65	-1.73	0.53
S2	21.73	25.30	26.23	26.74	5.10	-0.38	-2.11	-2.62
S3	21.19	27.15	26.02	25.65	5.65	-2.23	-1.90	-1.53
S4	22.49	27.40	22.63	27.45	4.35	-2.48	1.49	-3.33
S5	23.29	25.81	22.80	28.10	3.55	-0.89	1.31	-3.98
S6	23.27	26.32	22.88	27.53	3.57	-1.40	1.24	-3.41
S7	21.65	26.22	26.38	25.75	5.18	-1.30	-2.26	-1.64
S8	21.45	24.33	23.51	30.71	5.39	0.59	0.60	-6.59
S9	21.26	26.34	23.41	28.99	5.58	-1.42	0.71	-4.88
S10	22.19	26.97	22.33	28.51	4.65	-2.05	1.78	-4.40
S11	22.63	25.77	22.67	28.93	4.21	-0.85	1.44	-4.82
S12	21.42	26.70	24.97	26.91	5.42	-1.78	-0.86	-2.80
S13	21.71	25.46	26.19	26.65	5.13	-0.54	-2.07	-2.54
S14	22.16	25.89	22.78	29.15	4.68	-0.97	1.33	-5.04
S15	21.49	25.89	25.25	27.38	5.35	-0.97	-1.13	-3.26
S16	22.56	26.18	22.41	28.84	4.27	-1.26	1.70	-4.72
S17	25.59	24.37	24.93	25.11	1.25	0.55	-0.82	-0.99
S18	20.26	27.14	27.59	25.01	6.57	-2.22	-3.47	-0.90
S19	20.26	20.58	28.46	30.69	6.58	4.34	-4.35	-6.58
S20	23.42	21.88	26.46	28.25	3.42	3.04	-2.34	-4.13
S21	29.86	24.66	26.31	19.16	-3.02	0.26	-2.20	4.95
S22	29.81	23.75	22.26	24.17	-2.98	1.17	1.85	-0.05
S23	18.04	23.48	25.81	32.66	8.80	1.44	-1.69	-8.55
S24	21.51	20.88	28.54	29.07	5.32	4.04	-4.42	-4.95
S25	31.75	23.83	24.79	19.64	-4.91	1.10	-0.67	4.48
S26	37.26	28.88	18.44	15.42	-10.42	-3.96	5.68	8.69
S27	19.95	20.65	27.69	31.71	6.89	4.27	-3.58	-7.60
S28	25.31	25.00	27.89	21.80	1.53	-0.08	-3.77	2.32
S29	19.56	22.56	27.74	30.09	7.28	2.36	-3.63	-5.98
S30	20.37	24.31	27.12	28.19	6.47	0.61	-3.01	-4.08
S31	24.56	24.17	29.99	21.28	2.28	0.75	-5.88	2.83
S32	35.82	32.46	19.14	12.57	-8.98	-7.54	4.97	11.54
S33	17.59	18.16	29.54	34.71	9.25	6.76	-5.42	-10.59
S34	23.83	23.38	31.68	21.11	3.01	1.54	-7.56	3.00
S35	24.61	24.84	30.76	19.79	2.23	0.08	-6.64	4.33
S36	30.03	23.38	26.65	19.95	-3.19	1.54	-2.53	4.17
S37	22.17	19.77	28.54	29.52	4.67	5.15	-4.43	-5.41
S38	25.59	29.23	21.11	24.07	1.24	-4.31	3.01	0.05
S39	25.57	27.32	25.89	21.22	1.27	-2.40	-1.77	2.89
S40	32.20	22.01	25.71	20.07	-5.37	2.91	-1.60	4.04
S41	31.06	20.62	24.16	24.17	-4.22	4.30	-0.04	-0.05
S42	21.41	28.14	15.67	34.78	5.42	-3.22	8.45	-10.66
S43	24.51	24.47	27.33	23.69	2.33	0.45	-3.21	0.42
S44	30.24	19.13	21.03	29.59	-3.40	5.79	3.08	-5.48
S45	30.07	30.42	21.77	17.74	-3.23	-5.50	2.34	6.37
S46	27.79	29.28	23.73	19.19	-0.95	-4.36	0.39	4.92
S47	24.74	21.88	20.58	32.80	2.10	3.04	3.54	-8.68
S48	34.15	28.51	21.37	15.97	-7.32	-3.59	2.75	8.15
S49	26.64	26.31	24.82	22.22	0.20	-1.39	-0.71	1.89

S50	27.29	28.85	23.23	20.63	-0.45	-3.93	0.88	3.49
S51	24.14	20.20	29.26	26.40	2.70	4.72	-5.14	-2.29
S52	24.93	25.88	25.75	23.45	1.91	-0.96	-1.63	0.67
S53	22.52	23.80	13.10	40.58	4.31	1.12	11.02	-16.47
S54	32.49	22.01	25.97	19.52	-5.65	2.91	-1.85	4.59
S55	29.95	28.69	23.36	18.00	-3.11	-3.77	0.76	6.12
S56	29.28	22.23	23.92	24.58	-2.44	2.69	0.20	-0.46
S57	25.72	27.98	25.56	20.75	1.12	-3.06	-1.44	3.37
S58	25.70	20.66	31.16	22.48	1.14	4.26	-7.04	1.63
S59	26.03	19.35	18.49	36.13	0.81	5.57	5.63	-12.02
S60	28.27	23.57	24.44	23.73	-1.43	1.35	-0.32	0.38
S61	26.60	23.65	24.99	24.76	0.23	1.27	-0.87	-0.65
S62	26.99	20.73	27.93	24.35	-0.15	4.20	-3.81	-0.24
S63	26.56	29.64	22.68	21.12	0.27	-4.72	1.44	2.99
S64	26.27	24.67	27.88	21.18	0.56	0.25	-3.76	2.93
S65	19.46	23.34	19.25	37.95	7.38	1.58	4.86	-13.83
S66	30.53	28.15	23.89	17.43	-3.70	-3.23	0.23	6.69
S67	23.36	25.77	18.76	32.11	3.48	-0.85	5.36	-7.99
S68	27.25	19.99	30.02	22.75	-0.41	4.94	-5.90	1.37
S69	29.16	21.30	28.22	21.32	-2.32	3.62	-4.11	2.79
S70	31.98	21.84	26.47	19.66	-5.14	3.08	-2.36	4.45
S71	27.96	26.01	24.19	21.80	-1.12	-1.09	-0.07	2.31
S72	30.05	30.72	21.79	17.44	-3.21	-5.80	2.32	6.67
S73	28.68	22.94	23.47	24.90	-1.84	1.98	0.64	-0.79
S74	32.71	20.81	25.33	21.15	-5.87	4.11	-1.22	2.96
S75	26.78	28.63	24.73	19.87	0.06	-3.71	-0.61	4.25
S76	27.80	22.07	27.38	22.45	-0.96	2.85	-3.26	1.66
S77	32.00	22.19	25.81	20.00	-5.16	2.73	-1.70	4.12
S78	22.49	27.49	14.08	35.95	4.35	-2.57	10.04	-11.83
S79	21.59	28.20	15.58	34.60	5.25	-3.28	8.54	-10.48
S80	22.91	21.33	33.87	21.88	3.92	3.59	-9.75	2.23
S81	28.38	24.09	26.40	21.12	-1.54	0.83	-2.28	2.99
S82	26.96	21.94	22.73	28.37	-0.12	2.98	1.38	-4.25
S83	27.06	20.03	27.00	25.92	-0.22	4.89	-2.88	-1.81
S84	28.80	25.67	27.18	18.35	-1.96	-0.75	-3.06	5.76
S85	25.81	24.39	24.95	24.15	1.03	0.53	-0.84	-0.04
S86	18.61	21.51	17.51	42.37	8.23	3.41	6.60	-18.26
S87	25.77	24.83	27.29	22.11	1.06	0.09	-3.17	2.00
S88	26.61	21.58	21.53	30.28	0.22	3.34	2.59	-6.16
S89	28.04	32.05	24.80	15.12	-1.20	-7.13	-0.68	9.00
S90	34.75	26.70	23.00	15.55	-7.92	-1.78	1.12	8.56
S91	27.98	22.03	29.61	20.39	-1.14	2.90	-5.49	3.73
S92	27.57	20.09	28.40	23.93	-0.73	4.83	-4.29	0.18
S93	32.33	22.24	25.46	19.97	-5.49	2.68	-1.34	4.14
S94	24.18	20.61	31.94	23.27	2.66	4.31	-7.82	0.85
S95	34.37	28.33	19.78	17.53	-7.53	-3.41	4.34	6.58
S96	27.32	27.35	24.55	20.79	-0.48	-2.43	-0.43	3.33
S97	31.08	22.10	21.70	25.12	-4.24	2.82	2.42	-1.01
S98	32.54	29.52	22.42	15.51	-5.70	-4.60	1.69	8.60
S99	28.72	29.15	23.63	18.50	-1.89	-4.23	0.49	5.62
S100	19.21	29.44	24.11	27.24	7.63	-4.52	0.01	-3.13
S101	30.09	22.22	23.58	24.11	-3.25	2.70	0.54	0.00
S102	30.76	22.59	25.38	21.27	-3.92	2.33	-1.26	2.84
S103	26.88	26.08	26.50	20.55	-0.04	-1.15	-2.38	3.56
S104	36.97	25.84	20.66	16.53	-10.13	-0.92	3.45	7.59
S105	26.68	23.18	25.48	24.66	0.16	1.74	-1.37	-0.55
S106	17.90	23.74	21.16	37.12	8.93	1.18	2.96	-13.00
S107	19.46	19.31	21.14	40.08	7.37	5.61	2.97	-15.96
S108	28.09	23.92	28.21	19.78	-1.25	1.00	-4.09	4.34
S109	25.46	21.82	31.10	21.62	1.38	3.10	-6.98	2.50
S110	26.35	21.55	31.28	20.82	0.49	3.37	-7.17	3.30

S111	28.98	22.80	27.04	21.18	-2.14	2.12	-2.93	2.93
S112	28.90	22.98	26.97	21.14	-2.06	1.94	-2.86	2.97
S113	26.44	22.05	27.13	24.39	0.40	2.88	-3.01	-0.28
S114	34.66	27.26	19.30	18.77	-7.83	-2.34	4.81	5.35
S115	40.72	26.15	17.84	15.29	-13.88	-1.23	6.27	8.83
S116	28.88	23.94	24.90	22.28	-2.04	0.98	-0.79	1.84
S117	27.28	22.33	28.16	22.23	-0.45	2.59	-4.04	1.89
S118	33.66	23.04	14.10	29.20	-6.83	1.88	10.02	-5.08
S119	30.62	25.52	14.48	29.36	-3.78	-0.60	9.64	-5.25
S120	33.80	25.46	15.05	25.70	-6.96	-0.54	9.07	-1.58
S121	29.35	33.75	25.70	11.18	-2.51	-8.83	-1.59	12.93
S122	33.23	20.72	19.92	26.13	-6.40	4.20	4.20	-2.01
S123	31.48	30.29	16.26	21.97	-4.64	-5.37	7.86	2.14
S124	33.06	22.03	19.09	25.82	-6.22	2.89	5.02	-1.70
S125	27.01	34.97	21.44	16.57	-0.17	-10.05	2.68	7.55
S126	28.71	36.28	19.71	15.30	-1.87	-11.36	4.40	8.81
S127	26.98	35.36	21.35	16.31	-0.14	-10.44	2.76	7.81
S128	34.10	17.15	17.04	31.71	-7.27	7.77	7.08	-7.59
S129	29.78	39.99	22.31	7.92	-2.95	-15.07	1.81	16.20
S130	30.09	26.24	18.54	25.13	-3.25	-1.32	5.57	-1.02

Figs. 24 to 27 respectively show the graphs of differential errors of percentage concentration values with respect to average values.



Fig. 24: Differential errors of pA values with respect to its average value

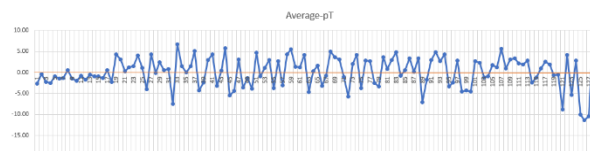


Fig. 25: Differential errors of pT values with respect to its average value



Fig. 26: Differential errors of pG values with respect to its average value

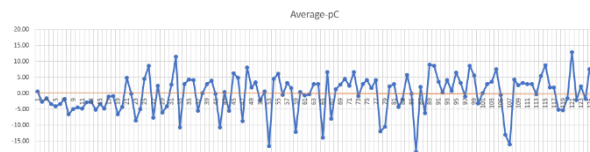


Fig. 27: Differential errors of pC values with respect to its average value

Tables 8 to 11 respectively present both positive and negative differential error values of all four nucleotides.

Table 8: Positive and negative differential error values of pA

Positive Values					Negative Values			
Sl. No.	Seq. No	pA	Avg-pA		Sl. No.	Seq. No	pA	Avg-pA
1	S1	23.00	3.84		1	S21	29.86	-3.02
2	S2	21.73	5.10		2	S22	29.81	-2.98
3	S3	21.19	5.65		3	S25	31.75	-4.91
4	S4	22.49	4.35		4	S26	37.26	-10.42
5	S5	23.29	3.55		5	S32	35.82	-8.98
6	S6	23.27	3.57		6	S36	30.03	-3.19
7	S7	21.65	5.18		7	S40	32.20	-5.37
8	S8	21.45	5.39		8	S41	31.06	-4.22
9	S9	21.26	5.58		9	S44	30.24	-3.40
10	S10	22.19	4.65		10	S45	30.07	-3.23
11	S11	22.63	4.21		11	S46	27.79	-0.95
12	S12	21.42	5.42		12	S48	34.15	-7.32
13	S13	21.71	5.13		13	S50	27.29	-0.45
14	S14	22.16	4.68		14	S54	32.49	-5.65
15	S15	21.49	5.35		15	S55	29.95	-3.11
16	S16	22.56	4.27		16	S56	29.28	-2.44
17	S17	25.59	1.25		17	S60	28.27	-1.43

18	S18	20.26	6.57		18	S62	26.99	-0.15
19	S19	20.26	6.58		19	S66	30.53	-3.70
20	S20	23.42	3.42		20	S68	27.25	-0.41
21	S23	18.04	8.80		21	S69	29.16	-2.32
22	S24	21.51	5.32		22	S70	31.98	-5.14
23	S27	19.95	6.89		23	S71	27.96	-1.12
24	S28	25.31	1.53		24	S72	30.05	-3.21
25	S29	19.56	7.28		25	S73	28.68	-1.84
26	S30	20.37	6.47		26	S74	32.71	-5.87
27	S31	24.56	2.28		27	S76	27.80	-0.96
28	S33	17.59	9.25		28	S77	32.00	-5.16
29	S34	23.83	3.01		29	S81	28.38	-1.54
30	S35	24.61	2.23		30	S82	26.96	-0.12
31	S37	22.17	4.67		31	S83	27.06	-0.22
32	S38	25.59	1.24		32	S84	28.80	-1.96
33	S39	25.57	1.27		33	S89	28.04	-1.20
34	S42	21.41	5.42		34	S90	34.75	-7.92
35	S43	24.51	2.33		35	S91	27.98	-1.14
36	S47	24.74	2.10		36	S92	27.57	-0.73
37	S49	26.64	0.20		37	S93	32.33	-5.49
38	S51	24.14	2.70		38	S95	34.37	-7.53
39	S52	24.93	1.91		39	S96	27.32	-0.48
40	S53	22.52	4.31		40	S97	31.08	-4.24
41	S57	25.72	1.12		41	S98	32.54	-5.70
42	S58	25.70	1.14		42	S99	28.72	-1.89
43	S59	26.03	0.81		43	S101	30.09	-3.25
44	S61	26.60	0.23		44	S102	30.76	-3.92
45	S63	26.56	0.27		45	S103	26.88	-0.04
46	S64	26.27	0.56		46	S104	36.97	-10.13
47	S65	19.46	7.38		47	S108	28.09	-1.25
48	S67	23.36	3.48		48	S111	28.98	-2.14
49	S75	26.78	0.06		49	S112	28.90	-2.06
50	S78	22.49	4.35		50	S114	34.66	-7.83
51	S79	21.59	5.25		51	S115	40.72	-13.88
52	S80	22.91	3.92		52	S116	28.88	-2.04
53	S85	25.81	1.03		53	S117	27.28	-0.45
54	S86	18.61	8.23		54	S118	33.66	-6.83
55	S87	25.77	1.06		55	S119	30.62	-3.78
56	S88	26.61	0.22		56	S120	33.80	-6.96
57	S94	24.18	2.66		57	S121	29.35	-2.51
58	S100	19.21	7.63		58	S122	33.23	-6.40
59	S105	26.68	0.16		59	S123	31.48	-4.64
60	S106	17.90	8.93		60	S124	33.06	-6.22
61	S107	19.46	7.37		61	S125	27.01	-0.17
62	S109	25.46	1.38		62	S126	28.71	-1.87
63	S110	26.35	0.49		63	S127	26.98	-0.14
64	S113	26.44	0.40		64	S128	34.10	-7.27
					65	S129	29.78	-2.95
					66	S130	30.09	-3.25

Table 9: Positive and negative differential error values of pT

Positive Values				Negative Values			
Sl. No.	Seq. No	pT	Avg-pT	Sl. No.	Seq. No	pT	Avg-pT
1	S8	24.33	0.59	1	S1	27.57	-2.65
2	S17	24.37	0.55	2	S2	25.30	-0.38
3	S19	20.58	4.34	3	S3	27.15	-2.23
4	S20	21.88	3.04	4	S4	27.40	-2.48
5	S21	24.66	0.26	5	S5	25.81	-0.89
6	S22	23.75	1.17	6	S6	26.32	-1.40
7	S23	23.48	1.44	7	S7	26.22	-1.30
8	S24	20.88	4.04	8	S9	26.34	-1.42
9	S25	23.83	1.10	9	S10	26.97	-2.05

10	S27	20.65	4.27		10	S11	25.77	-0.85
11	S29	22.56	2.36		11	S12	26.70	-1.78
12	S30	24.31	0.61		12	S13	25.46	-0.54
13	S31	24.17	0.75		13	S14	25.89	-0.97
14	S33	18.16	6.76		14	S15	25.89	-0.97
15	S34	23.38	1.54		15	S16	26.18	-1.26
16	S35	24.84	0.08		16	S18	27.14	-2.22
17	S36	23.38	1.54		17	S26	28.88	-3.96
18	S37	19.77	5.15		18	S28	25.00	-0.08
19	S40	22.01	2.91		19	S32	32.46	-7.54
20	S41	20.62	4.30		20	S38	29.23	-4.31
21	S43	24.47	0.45		21	S39	27.32	-2.40
22	S44	19.13	5.79		22	S42	28.14	-3.22
23	S47	21.88	3.04		23	S45	30.42	-5.50
24	S51	20.20	4.72		24	S46	29.28	-4.36
25	S53	23.80	1.12		25	S48	28.51	-3.59
26	S54	22.01	2.91		26	S49	26.31	-1.39
27	S56	22.23	2.69		27	S50	28.85	-3.93
28	S58	20.66	4.26		28	S52	25.88	-0.96
29	S59	19.35	5.57		29	S55	28.69	-3.77
30	S60	23.57	1.35		30	S57	27.98	-3.06
31	S61	23.65	1.27		31	S63	29.64	-4.72
32	S62	20.73	4.20		32	S66	28.15	-3.23
33	S64	24.67	0.25		33	S67	25.77	-0.85
34	S65	23.34	1.58		34	S71	26.01	-1.09
35	S68	19.99	4.94		35	S72	30.72	-5.80
36	S69	21.30	3.62		36	S75	28.63	-3.71
37	S70	21.84	3.08		37	S78	27.49	-2.57
38	S73	22.94	1.98		38	S79	28.20	-3.28
39	S74	20.81	4.11		39	S84	25.67	-0.75
40	S76	22.07	2.85		40	S89	32.05	-7.13
41	S77	22.19	2.73		41	S90	26.70	-1.78
42	S80	21.33	3.59		42	S95	28.33	-3.41
43	S81	24.09	0.83		43	S96	27.35	-2.43
44	S82	21.94	2.98		44	S98	29.52	-4.60
45	S83	20.03	4.89		45	S99	29.15	-4.23
46	S85	24.39	0.53		46	S100	29.44	-4.52
47	S86	21.51	3.41		47	S103	26.08	-1.15
48	S87	24.83	0.09		48	S104	25.84	-0.92
49	S88	21.58	3.34		49	S114	27.26	-2.34
50	S91	22.03	2.90		50	S115	26.15	-1.23
51	S92	20.09	4.83		51	S119	25.52	-0.60
52	S93	22.24	2.68		52	S120	25.46	-0.54
53	S94	20.61	4.31		53	S121	33.75	-8.83
54	S97	22.10	2.82		54	S123	30.29	-5.37
55	S101	22.22	2.70		55	S125	34.97	-10.05
56	S102	22.59	2.33		56	S126	36.28	-11.36
57	S105	23.18	1.74		57	S127	35.36	-10.44
58	S106	23.74	1.18		58	S129	39.99	-15.07
59	S107	19.31	5.61		59	S130	26.24	-1.32
60	S108	23.92	1.00					
61	S109	21.82	3.10					
62	S110	21.55	3.37					
63	S111	22.80	2.12					
64	S112	22.98	1.94					
65	S113	22.05	2.88					
66	S116	23.94	0.98					
67	S117	22.33	2.59					
68	S118	23.04	1.88					
69	S122	20.72	4.20					
70	S124	22.03	2.89					
71	S128	17.15	7.77					

Table 10: Positive and negative differential error values of pG

Positive Values					Negative Values			
Sl. No.	Seq. No	pG	Avg-pG		Sl. No.	Seq. No	pG	Avg-pG
1	S4	22.63	1.49		1	S1	25.85	-1.73
2	S5	22.80	1.31		2	S2	26.23	-2.11
3	S6	22.88	1.24		3	S3	26.02	-1.90
4	S8	23.51	0.60		4	S7	26.38	-2.26
5	S9	23.41	0.71		5	S12	24.97	-0.86
6	S10	22.33	1.78		6	S13	26.19	-2.07
7	S11	22.67	1.44		7	S15	25.25	-1.13
8	S14	22.78	1.33		8	S17	24.93	-0.82
9	S16	22.41	1.70		9	S18	27.59	-3.47
10	S22	22.26	1.85		10	S19	28.46	-4.35
11	S26	18.44	5.68		11	S20	26.46	-2.34
12	S32	19.14	4.97		12	S21	26.31	-2.20
13	S38	21.11	3.01		13	S23	25.81	-1.69
14	S42	15.67	8.45		14	S24	28.54	-4.42
15	S44	21.03	3.08		15	S25	24.79	-0.67
16	S45	21.77	2.34		16	S27	27.69	-3.58
17	S46	23.73	0.39		17	S28	27.89	-3.77
18	S47	20.58	3.54		18	S29	27.74	-3.63
19	S48	21.37	2.75		19	S30	27.12	-3.01
20	S50	23.23	0.88		20	S31	29.99	-5.88
21	S53	13.10	11.02		21	S33	29.54	-5.42
22	S55	23.36	0.76		22	S34	31.68	-7.56
23	S56	23.92	0.20		23	S35	30.76	-6.64
24	S59	18.49	5.63		24	S36	26.65	-2.53
25	S63	22.68	1.44		25	S37	28.54	-4.43
26	S65	19.25	4.86		26	S39	25.89	-1.77
27	S66	23.89	0.23		27	S40	25.71	-1.60
28	S67	18.76	5.36		28	S41	24.16	-0.04
29	S72	21.79	2.32		29	S43	27.33	-3.21
30	S73	23.47	0.64		30	S49	24.82	-0.71
31	S78	14.08	10.04		31	S51	29.26	-5.14
32	S79	15.58	8.54		32	S52	25.75	-1.63
33	S82	22.73	1.38		33	S54	25.97	-1.85
34	S86	17.51	6.60		34	S57	25.56	-1.44
35	S88	21.53	2.59		35	S58	31.16	-7.04
36	S90	23.00	1.12		36	S60	24.44	-0.32
37	S95	19.78	4.34		37	S61	24.99	-0.87
38	S97	21.70	2.42		38	S62	27.93	-3.81
39	S98	22.42	1.69		39	S64	27.88	-3.76
40	S99	23.63	0.49		40	S68	30.02	-5.90
41	S100	24.11	0.01		41	S69	28.22	-4.11
42	S101	23.58	0.54		42	S70	26.47	-2.36
43	S104	20.66	3.45		43	S71	24.19	-0.07
44	S106	21.16	2.96		44	S74	25.33	-1.22
45	S107	21.14	2.97		45	S75	24.73	-0.61
46	S114	19.30	4.81		46	S76	27.38	-3.26
47	S115	17.84	6.27		47	S77	25.81	-1.70
48	S118	14.10	10.02		48	S80	33.87	-9.75
49	S119	14.48	9.64		49	S81	26.40	-2.28
50	S120	15.05	9.07		50	S83	27.00	-2.88
51	S122	19.92	4.20		51	S84	27.18	-3.06
52	S123	16.26	7.86		52	S85	24.95	-0.84
53	S124	19.09	5.02		53	S87	27.29	-3.17
54	S125	21.44	2.68		54	S89	24.80	-0.68
55	S126	19.71	4.40		55	S91	29.61	-5.49
56	S127	21.35	2.76		56	S92	28.40	-4.29
57	S128	17.04	7.08		57	S93	25.46	-1.34
58	S129	22.31	1.81		58	S94	31.94	-7.82

59	S130	18.54	5.57		59	S96	24.55	-0.43
					60	S102	25.38	-1.26
					61	S103	26.50	-2.38
					62	S105	25.48	-1.37
					63	S108	28.21	-4.09
					64	S109	31.10	-6.98
					65	S110	31.28	-7.17
					66	S111	27.04	-2.93
					67	S112	26.97	-2.86
					68	S113	27.13	-3.01
					69	S116	24.90	-0.79
					70	S117	28.16	-4.04
					71	S121	25.70	-1.59

Table 11: Positive and negative differential error values of pC

Positive Values				Negative Values			
Sl. No.	Seq. No	pC	Avg-pC	Sl. No.	Seq. No	pC	Avg-pC
1	S1	23.58	0.53	1	S2	26.74	-2.62
2	S21	19.16	4.95	2	S3	25.65	-1.53
3	S25	19.64	4.48	3	S4	27.45	-3.33
4	S26	15.42	8.69	4	S5	28.10	-3.98
5	S28	21.80	2.32	5	S6	27.53	-3.41
6	S31	21.28	2.83	6	S7	25.75	-1.64
7	S32	12.57	11.54	7	S8	30.71	-6.59
8	S34	21.11	3.00	8	S9	28.99	-4.88
9	S35	19.79	4.33	9	S10	28.51	-4.40
10	S36	19.95	4.17	10	S11	28.93	-4.82
11	S38	24.07	0.05	11	S12	26.91	-2.80
12	S39	21.22	2.89	12	S13	26.65	-2.54
13	S40	20.07	4.04	13	S14	29.15	-5.04
14	S43	23.69	0.42	14	S15	27.38	-3.26
15	S45	17.74	6.37	15	S16	28.84	-4.72
16	S46	19.19	4.92	16	S17	25.11	-0.99
17	S48	15.97	8.15	17	S18	25.01	-0.90
18	S49	22.22	1.89	18	S19	30.69	-6.58
19	S50	20.63	3.49	19	S20	28.25	-4.13
20	S52	23.45	0.67	20	S22	24.17	-0.05
21	S54	19.52	4.59	21	S23	32.66	-8.55
22	S55	18.00	6.12	22	S24	29.07	-4.95
23	S57	20.75	3.37	23	S27	31.71	-7.60
24	S58	22.48	1.63	24	S29	30.09	-5.98
25	S60	23.73	0.38	25	S30	28.19	-4.08
26	S63	21.12	2.99	26	S33	34.71	-10.59
27	S64	21.18	2.93	27	S37	29.52	-5.41
28	S66	17.43	6.69	28	S41	24.17	-0.05
29	S68	22.75	1.37	29	S42	34.78	-10.66
30	S69	21.32	2.79	30	S44	29.59	-5.48
31	S70	19.66	4.45	31	S47	32.80	-8.68
32	S71	21.80	2.31	32	S51	26.40	-2.29
33	S72	17.44	6.67	33	S53	40.58	-16.47
34	S74	21.15	2.96	34	S56	24.58	-0.46
35	S75	19.87	4.25	35	S59	36.13	-12.02
36	S76	22.45	1.66	36	S61	24.76	-0.65
37	S77	20.00	4.12	37	S62	24.35	-0.24
38	S80	21.88	2.23	38	S65	37.95	-13.83
39	S81	21.12	2.99	39	S67	32.11	-7.99
40	S84	18.35	5.76	40	S73	24.90	-0.79
41	S87	22.11	2.00	41	S78	35.95	-11.83
42	S89	15.12	9.00	42	S79	34.60	-10.48
43	S90	15.55	8.56	43	S82	28.37	-4.25
44	S91	20.39	3.73	44	S83	25.92	-1.81

45	S92	23.93	0.18		45	S85	24.15	-0.04
46	S93	19.97	4.14		46	S86	42.37	-18.26
47	S94	23.27	0.85		47	S88	30.28	-6.16
48	S95	17.53	6.58		48	S97	25.12	-1.01
49	S96	20.79	3.33		49	S100	27.24	-3.13
50	S98	15.51	8.60		50	S105	24.66	-0.55
51	S99	18.50	5.62		51	S106	37.12	-13.00
52	S101	24.11	0.00		52	S107	40.08	-15.96
53	S102	21.27	2.84		53	S113	24.39	-0.28
54	S103	20.55	3.56		54	S118	29.20	-5.08
55	S104	16.53	7.59		55	S119	29.36	-5.25
56	S108	19.78	4.34		56	S120	25.70	-1.58
57	S109	21.62	2.50		57	S122	26.13	-2.01
58	S110	20.82	3.30		58	S124	25.82	-1.70
59	S111	21.18	2.93		59	S128	31.71	-7.59
60	S112	21.14	2.97		60	S130	25.13	-1.02
61	S114	18.77	5.35					
62	S115	15.29	8.83					
63	S116	22.28	1.84					
64	S117	22.23	1.89					
65	S121	11.18	12.93					
66	S123	21.97	2.14					
67	S125	16.57	7.55					
68	S126	15.30	8.81					
69	S127	16.31	7.81					
70	S129	7.92	16.20					

**Observations**

Number of Positive Values of Avg-pA = 64 Number of Negative Values of Avg-pA = 66	Positive and negative values of Avg-pT are distributed in the ratio <b>64:66</b> (almost equally distributed). This property has been observed in all 130 sequences.
Number of Positive Values of Avg-pT = 71 Number of Negative Values of Avg-pT = 59	Positive and negative values of Avg-pT are distributed in the ratio <b>71:59</b> . This property has been observed in all 130 sequences.
Number of Positive Values of Avg-pG = 59 Number of Negative Values of Avg-pG = 71	Positive and negative values of Avg-pG are distributed in the ratio <b>59:71</b> . This property has been observed in all 130 sequences.
Number of Positive Values of Avg-pC = 70 Number of Negative Values of Avg-pC = 60	Positive and negative values of Avg-pC are distributed in the ratio <b>70:60</b> . This property has been observed in all 130 sequences.

From tables 7 to 10, it is clear that there is a pattern of distribution of nucleotides in all 130 virus sequences. This observation paves a new avenue for further research in bio-informatics.

**3.4 Virus Genomes with Dominant Nucleotides**

From the given set of 130 virus genome sequences, the subset of sequences consisting of maximum number of Adenines are separated, which are called A-Dominant sequences. Similarly, T-Dominant sequences, G-Dominant sequences and C-Dominant sequences are separated. There are 52 A-Dominant sequences, 21 T-Dominant sequences, 22 G-Dominant sequences and 35 C-Dominant sequences in the set of 130 virus genomes. They are given in tables 12 to 15.

*Table 12: A-Dominant Virus Sequences (52)*

Sl. No.	Seq. No.	Species	Accession ID	Total length	A	pA	T	pT	G	pG	C	pC
1	S17	Giardia lamblia virus	NC_003555.1	6277	1606	25.59	1530	24.37	1565	24.93	1576	25.11
2	S21	Black raspberry virus F	NC_009890.1	5077	1516	29.86	1252	24.66	1336	26.31	973	19.16
3	S22	Armigeres subalbatus virus	NC_014609.1	7510	2239	29.81	1784	23.75	1672	22.26	1815	24.17
4	S25	Camponotus yamaokai virus	NC_027212.1	5704	1811	31.75	1359	23.83	1414	24.79	1120	19.64
5	S26	Golden shiner totivirus	NC_030295.1	7788	2902	37.26	2249	28.88	1436	18.44	1201	15.42
6	S32	Gigaspora margarita giardia-like virus	NC_040632.1	4947	1772	35.82	1606	32.46	947	19.14	622	12.57
7	S36	Diatom colony associated dsRNA virus 12 genomic RNA	NC_040775.1	5941	1784	30.03	1389	23.38	1583	26.65	1185	19.95
8	S40	Bovine viral diarrhea virus	NC_001461.1	12573	4049	32.20	2767	22.01	3233	25.71	2524	20.07
9	S41	O'nyong-nyong virus	NC_001512.1	11835	3676	31.06	2440	20.62	2859	24.16	2860	24.17
10	S44	Bamboo mosaic virus,	NC_001642.1	6366	1925	30.24	1218	19.13	1339	21.03	1884	29.59
11	S48	Banana mild mosaic virus	NC_002729.1	7352	2511	34.15	2096	28.51	1571	21.37	1174	15.97
12	S49	Aconitum latent virus	NC_002795.1	8657	2306	26.64	2278	26.31	2149	24.82	1924	22.22

13	S54	Border disease virus	NC_003679.1	12333	4007	32.49	2715	22.01	3203	25.97	2408	19.52
14	S55	Obuda pepper virus	NC_003852.1	6507	1949	29.95	1867	28.69	1520	23.36	1171	18.00
15	S56	Aura virus	NC_003900.1	11824	3462	29.28	2628	22.23	2828	23.92	2906	24.58
16	S60	Olive mild mosaic virus	NC_006939.1	3683	1041	28.27	868	23.57	900	24.44	874	23.73
17	S61	Angelonia flower break virus	NC_007733.2	3962	1054	26.60	937	23.65	990	24.99	981	24.76
18	S66	Peach mosaic virus	NC_011552.1	7988	2439	30.53	2249	28.15	1908	23.89	1392	17.43
19	S69	Bagaza virus strain DakAr B209	NC_012534.1	10941	3190	29.16	2330	21.30	3088	28.22	2333	21.32
20	S70	Bovine viral diarrhea virus 3	NC_012812.1	12337	3945	31.98	2694	21.84	3266	26.47	2426	19.66
21	S71	Kalanchoe latent virus	NC_013006.1	8517	2381	27.96	2215	26.01	2060	24.19	1857	21.80
22	S73	Brassica yellows virus	NC_016038.2	5666	1625	28.68	1300	22.94	1330	23.47	1411	24.90
23	S74	Actinidia virus B	NC_016404.1	7488	2449	32.71	1558	20.81	1897	25.33	1584	21.15
24	S76	Ndumu virus	NC_016959.1	11724	3259	27.80	2588	22.07	3210	27.38	2632	22.45
25	S77	Pestivirus strain	NC_018713.1	12292	3933	32.00	2728	22.19	3173	25.81	2458	20.00
26	S81	Gaillardia latent virus	NC_023892.1	8659	2457	28.38	2086	24.09	2286	26.40	1829	21.12
27	S83	Middelburg virus	NC_024887.1	11550	3125	27.06	2313	20.03	3118	27.00	2994	25.92
28	S84	Jutiapa virus	NC_026620.1	10125	2916	28.80	2599	25.67	2752	27.18	1858	18.35
29	S85	Phasey bean mild yellows virus	NC_028793.2	5851	1510	25.81	1427	24.39	1460	24.95	1413	24.15
30	S90	Actinidia seed-borne latent virus	NC_040800.1	8192	2847	34.75	2187	26.70	1884	23.00	1274	15.55
31	S93	Bovine viral diarrhea virus	NC_039237.1	12513	4045	32.33	2783	22.24	3186	25.46	2499	19.97
32	S95	Phaseolus vulgaris endornavirus	NC_039217.1	14072	4836	34.37	3986	28.33	2783	19.78	2467	17.53
33	S97	Babaco mosaic virus	NC_036587.1	6692	2080	31.08	1479	22.10	1452	21.70	1681	25.12
34	S98	Ocimum basilicum RNA virus	NC_035462.1	6930	2255	32.54	2046	29.52	1554	22.42	1075	15.51
35	S101	Apis flavivirus	NC_035071.1	20414	6142	30.09	4537	22.22	4813	23.58	4922	24.11
36	S102	Agave tequilana leaf virus	NC_034833.1	6958	2140	30.76	1572	22.59	1766	25.38	1480	21.27
37	S103	Ochlerotatus caspius flavivirus	NC_034242.1	10370	2787	26.88	2704	26.08	2748	26.50	2131	20.55
38	S104	Lagenaria siceraria endornavirus	NC_034216.1	15098	5582	36.97	3901	25.84	3120	20.66	2495	16.53
39	S105	African eggplant yellowing virus	NC_034207.1	5953	1588	26.68	1380	23.18	1517	25.48	1468	24.66
40	S111	Jugra virus	NC_033699.1	10173	2948	28.98	2319	22.80	2751	27.04	2155	21.18
41	S112	Bouboui virus	NC_033693.1	10173	2940	28.90	2338	22.98	2744	26.97	2151	21.14
42	S114	Botrytis cinerea endornavirus	NC_031752.1	11557	4006	34.66	3151	27.26	2231	19.30	2169	18.77
43	S115	Ceratobasidium endornavirus	NC_031463.1	23635	9624	40.72	6181	26.15	4217	17.84	3613	15.29
44	S116	Ceratobasidium endorna virus	NC_031462.1	15207	4392	28.88	3640	23.94	3787	24.90	3388	22.28
45	S118	White bream virus	NC_008516.1	26660	8975	33.66	6143	23.04	3758	14.10	7784	29.20
46	S119	Chinook salmon bafinivirus	NC_026812.1	27004	8268	30.62	6892	25.52	3909	14.48	7929	29.36
47	S120	Fathead minnow nidovirus	NC_038295.1	27318	9234	33.80	6954	25.46	4110	15.05	7020	25.70
48	S122	Ball python nidovirus	NC_024709.1	33452	11117	33.23	6932	20.72	6663	19.92	8740	26.13
49	S123	Xinzhou toro-like virus	NC_033700.1	30353	9555	31.48	9194	30.29	4934	16.26	6670	21.97
50	S124	Morelia viridis nidovirus	NC_035465.1	32399	10710	33.06	7139	22.03	6186	19.09	8364	25.82
51	S128	Bellinger River virus	NC_046956.1	30742	10484	34.10	5273	17.15	5238	17.04	9747	31.71
52	S130	Guangdong red-banded snake torovirus	NC_046963.1	30859	9285	30.09	8097	26.24	5722	18.54	7755	25.13

Fig. 28 shows graph of Percentage Concentrations in virus sequences of dominant adenines and other nucleotides

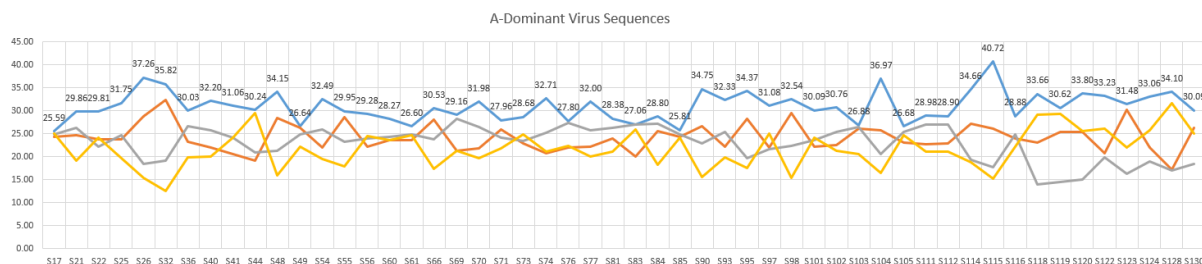


Fig. 28: Percentage Concentrations in virus sequences of dominant adenines and other nucleotides

13	S80	Kama virus	NC_023439.1	10688	2449	22.91	2280	21.33	3620	33.87	2339	21.88
14	S87	Anopheles flavivirus	NC_031327.1	10588	2729	25.77	2629	24.83	2889	27.29	2341	22.11
15	S91	Kampung Karu virus	NC_040788.1	10311	2885	27.98	2271	22.03	3053	29.61	2102	20.39
16	S92	Rocio virus	NC_040776.1	10794	2976	27.57	2169	20.09	3066	28.40	2583	23.93
17	S94	Kyasanur forest disease virus	NC_039218.1	10376	2509	24.18	2139	20.61	3314	31.94	2414	23.27
18	S108	Bamaga virus	NC_033725.1	10203	2866	28.09	2441	23.92	2878	28.21	2018	19.78
19	S109	Kadam virus	NC_033724.1	10215	2601	25.46	2229	21.82	3177	31.10	2208	21.62
20	S110	Gadgets Gully virus	NC_033723.1	10251	2701	26.35	2209	21.55	3207	31.28	2134	20.82
21	S113	New Mapoon virus	NC_032088.1	10864	2872	26.44	2395	22.05	2947	27.13	2650	24.39
22	S117	Banzi virus	NC_043110.1	10182	2778	27.28	2274	22.33	2867	28.16	2263	22.23

Table 13: T-Dominant Virus Sequences (21)

Sl. No.	Seq. No.	Species	Accession ID	Total length	A	pA	T	pT	G	pG	C	pC
1	S1	Lactate dehydrogenase-elevating virus	NC_001639.1	14104	3244	23.00	3888	27.57	3646	25.85	3326	23.58
2	S3	Equine arteritis virus	NC_002532.2	12704	2692	21.19	3449	27.15	3305	26.02	3258	25.65
3	S38	Guangdong greater green snake arterivirus	NC_046959.1	18410	4712	25.59	5381	29.23	3886	21.11	4431	24.07
4	S39	Pothos latent virus	NC_000939.2	4415	1129	25.57	1206	27.32	1143	25.89	937	21.22
5	S45	Odontoglossum ringspot virus,	NC_001728.1	6618	1990	30.07	2013	30.42	1441	21.77	1174	17.74
6	S46	Rupestris stem pitting associated virus	NC_001948.1	8744	2430	27.79	2560	29.28	2075	23.73	1678	19.19
7	S50	Garlic latent virus	NC_003557.1	8363	2282	27.29	2413	28.85	1943	23.23	1725	20.63
8	S52	Hibiscus chlorotic ringspot virus	NC_003608.1	3911	975	24.93	1012	25.88	1007	25.75	917	23.45
9	S57	Grapevine rootstock stem lesion associated virus	NC_004724.1	16527	4250	25.72	4624	27.98	4224	25.56	3429	20.75
10	S63	Peach chlorotic mottle virus	NC_009892.1	9005	2392	26.56	2669	29.64	2042	22.68	1902	21.12
11	S72	Grapevine Pinot gris virus	NC_015782.2	7259	2181	30.05	2230	30.72	1582	21.79	1266	17.44
12	S75	Garlic common latent virus	NC_016440.1	8638	2313	26.78	2473	28.63	2136	24.73	1716	19.87
13	S89	Grapevine associated tymo-like virus	NC_040837.1	6060	1699	28.04	1942	32.05	1503	24.80	916	15.12
14	S96	Atractylodes mottle virus	NC_038966.1	8881	2426	27.32	2429	27.35	2180	24.55	1846	20.79
15	S99	Actinidia virus	NC_035453.1	18848	5414	28.72	5495	29.15	4453	23.63	3486	18.50
16	S100	Lake Sinai Virus	NC_035116.1	5877	1129	19.21	1730	29.44	1417	24.11	1601	27.24
17	S121	Bovine nidovirus	NC_027199.1	20261	5947	29.35	6839	33.75	5208	25.70	2266	11.18
18	S125	Breda virus	NC_007447.1	28475	7690	27.01	9959	34.97	6104	21.44	4718	16.57
19	S126	Porcine torovirus	NC_022787.1	28301	8124	28.71	10267	36.28	5579	19.71	4331	15.30
20	S127	Goat torovirus	NC_034976.1	28487	7686	26.98	10072	35.36	6083	21.35	4646	16.31
21	S129	Hainan hebivirus popei torovirus	NC_046962.1	29409	8759	29.78	11760	39.99	6561	22.31	2328	7.92

Fig. 29 shows graph of Percentage Concentrations in virus sequences of dominant thymines and other nucleotides

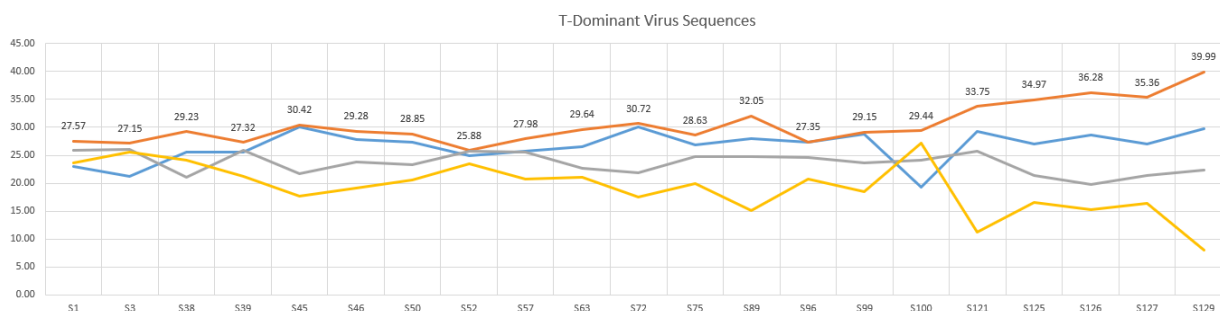


Fig. 29: Percentage Concentrations in virus sequences of dominant thymines and other nucleotides

Table 14: G-Dominant Virus Sequences (22)

Sl. No.	Seq. No.	Species	Accession ID	Total length	A	pA	T	pT	G	pG	C	pC
1	S7	Forest pouched giant rat arterivirus	NC_026439.1	14953	3238	21.65	3920	26.22	3944	26.38	3851	25.75
2	S18	Chalara elegans RNA Virus 1	NC_005883.1	5310	1076	20.26	1441	27.14	1465	27.59	1328	25.01
3	S28	Australian Anopheles totivirus	NC_035674.1	6203	1570	25.31	1551	25.00	1730	27.89	1352	21.80

4	S31	Diatom colony associated dsRNA virus	NC_040431.1	4671	1147	24.56	1129	24.17	1401	29.99	994	21.28
5	S34	Diatom colony associated dsRNA virus 13 genomic RNA	NC_040659.1	5082	1211	23.83	1188	23.38	1610	31.68	1073	21.11
6	S35	Diatom colony associated dsRNA virus 11 genomic RNA	NC_040660.1	5160	1270	24.61	1282	24.84	1587	30.76	1021	19.79
7	S43	Cell fusing agent virus	NC_001564.2	10682	2618	24.51	2614	24.47	2919	27.33	2531	23.69
8	S51	Groundnut rosette virus	NC_003603.1	4019	970	24.14	812	20.20	1176	29.26	1061	26.40
9	S58	Omsk hemorrhagic fever virus	NC_005062.1	10787	2772	25.70	2229	20.66	3361	31.16	2425	22.48
10	S62	Ilheus virus	NC_009028.2	10755	2903	26.99	2229	20.73	3004	27.93	2619	24.35
11	S64	Grapevine Algerian latent virus	NC_011535.1	4731	1243	26.27	1167	24.67	1319	27.88	1002	21.18
12	S68	Kedougou virus strain DakAar D1470	NC_012533.1	10723	2922	27.25	2143	19.99	3219	30.02	2439	22.75
13	S80	Kama virus	NC_023439.1	10688	2449	22.91	2280	21.33	3620	33.87	2339	21.88
14	S87	Anopheles flavivirus	NC_031327.1	10588	2729	25.77	2629	24.83	2889	27.29	2341	22.11
15	S91	Kampung Karu virus	NC_040788.1	10311	2885	27.98	2271	22.03	3053	29.61	2102	20.39
16	S92	Rocio virus	NC_040776.1	10794	2976	27.57	2169	20.09	3066	28.40	2583	23.93
17	S94	Kyasanur forest disease virus	NC_039218.1	10376	2509	24.18	2139	20.61	3314	31.94	2414	23.27
18	S108	Bamaga virus	NC_033725.1	10203	2866	28.09	2441	23.92	2878	28.21	2018	19.78
19	S109	Kadam virus	NC_033724.1	10215	2601	25.46	2229	21.82	3177	31.10	2208	21.62
20	S110	Gadgets Gully virus	NC_033723.1	10251	2701	26.35	2209	21.55	3207	31.28	2134	20.82
21	S113	New Mapoon virus	NC_032088.1	10864	2872	26.44	2395	22.05	2947	27.13	2650	24.39
22	S117	Banzi virus	NC_043110.1	10182	2778	27.28	2274	22.33	2867	28.16	2263	22.23

Fig. 30 shows graph of Percentage Concentrations in virus sequences of dominant guanines and other nucleotides

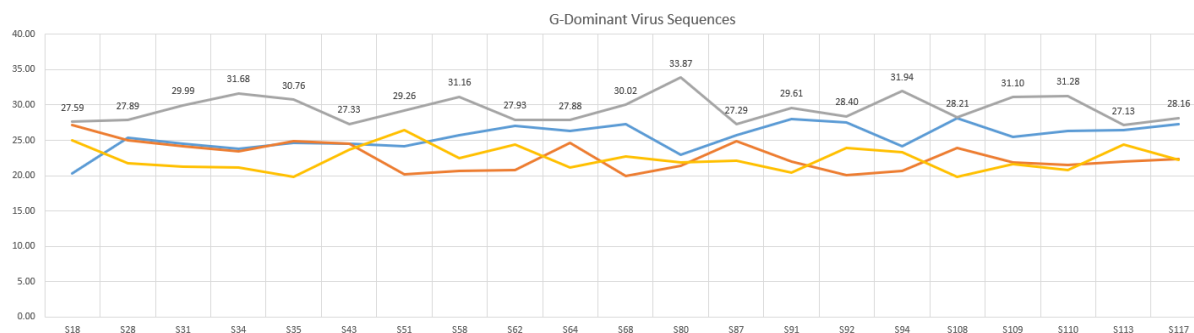


Fig. 30: Percentage Concentrations in virus sequences of dominant guanines and other nucleotides

Table 15: C-Dominant Virus Sequences (35)

Sl. No.	Seq. No.	Species	Accession ID	Total length	A	pA	T	pT	G	pG	C	pC
1	S2	Porcine respiratory and reproductive syndrome virus	NC_001961.1	15428	3353	21.73	3903	25.30	4047	26.23	4125	26.74
2	S4	Simian hemorrhagic fever virus	NC_003092.2	15717	3534	22.49	4307	27.40	3556	22.63	4314	27.45
3	S5	Mikumi yellow baboon virus	NC_025112.1	14927	3476	23.29	3853	25.81	3404	22.80	4194	28.10
4	S6	Southwest baboon virus	NC_025113.1	14851	3456	23.27	3909	26.32	3398	22.88	4088	27.53
5	S8	DeBrazzas monkey arterivirus	NC_026509.1	15684	3364	21.45	3816	24.33	3688	23.51	4816	30.71
6	S9	Pebjah virus	NC_027124.1	15478	3291	21.26	4077	26.34	3623	23.41	4487	28.99
7	S10	Kafue Kinda chacma baboon virus	NC_029053.1	14924	3311	22.19	4025	26.97	3333	22.33	4255	28.51
8	S11	Free State vervet virus	NC_029992.1	15247	3450	22.63	3929	25.77	3457	22.67	4411	28.93
9	S12	Olivier's shrew virus 1	NC_035127.1	13766	2948	21.42	3675	26.70	3438	24.97	3705	26.91
10	S13	Porcine reproductive and respiratory syndrome virus 2	NC_038291.1	15411	3345	21.71	3923	25.46	4036	26.19	4107	26.65
11	S14	Simian hemorrhagic encephalitis virus	NC_038293.1	15370	3406	22.16	3979	25.89	3502	22.78	4481	29.15
12	S15	Lelystad virus	NC_043487.1	15111	3247	21.49	3912	25.89	3815	25.25	4137	27.38
13	S16	Zambian malbrouck virus	NC_048209.1	15307	3454	22.56	4008	26.18	3431	22.41	4414	28.84
14	S19	Coniothyrium minitans RNA virus	NC_007523.1	4975	1008	20.26	1024	20.58	1416	28.46	1527	30.69
15	S20	Botryotinia fuckeliana totivirus 1	NC_009224.1	5261	1232	23.42	1151	21.88	1392	26.46	1486	28.25
16	S23	Beauveria bassiana victorivirus	NC_024151.1	5327	961	18.04	1251	23.48	1375	25.81	1740	32.66
17	S24	Botryosphaeria dothidea victorivirus 1	NC_025214.1	5322	1145	21.51	1111	20.88	1519	28.54	1547	29.07
18	S27	Fusarium poae victorivirus 1	NC_030867.1	5124	1022	19.95	1058	20.65	1419	27.69	1625	31.71

19	S29	Aspergillus foetidus slow virus I	NC_038928.1	5194	1016	19.56	1172	22.56	1441	27.74	1563	30.09
20	S30	Beauveria bassiana victorivirus I	NC_038929.1	5228	1065	20.37	1271	24.31	1418	27.12	1474	28.19
21	S33	Fusarium asiaticum victorivirus I	NC_040653.1	5281	929	17.59	959	18.16	1560	29.54	1833	34.71
22	S37	Alternaria arborescens victorivirus	NC_040793.1	5206	1154	22.17	1029	19.77	1486	28.54	1537	29.52
23	S42	Ononis yellow mosaic virus	NC_001513.1	6211	1330	21.41	1748	28.14	973	15.67	2160	34.78
24	S47	Botrytis virus F	NC_002604.1	6827	1689	24.74	1494	21.88	1405	20.58	2239	32.80
25	S53	Physalis mottle virus	NC_003634.1	6673	1503	22.52	1588	23.80	874	13.10	2708	40.58
26	S59	Botrytis virus X	NC_005132.1	6966	1813	26.03	1348	19.35	1288	18.49	2517	36.13
27	S65	Nemesia ring necrosis virus	NC_011538.1	6285	1223	19.46	1467	23.34	1210	19.25	2385	37.95
28	S67	Anagryis vein yellowing virus	NC_011559.1	6151	1437	23.36	1585	25.77	1154	18.76	1975	32.11
29	S78	Andean potato latent virus	NC_020470.1	6337	1425	22.49	1742	27.49	892	14.08	2278	35.95
30	S79	Andean potato mild mosaic virus	NC_020471.1	6226	1344	21.59	1756	28.20	970	15.58	2154	34.60
31	S82	Pitaya virus X	NC_024458.1	6677	1800	26.96	1465	21.94	1518	22.73	1894	28.37
32	S86	Grapevine Red Globe virus	NC_030693.1	6863	1277	18.61	1476	21.51	1202	17.51	2908	42.37
33	S88	Potexvirus	NC_040842.1	5839	1554	26.61	1260	21.58	1257	21.53	1768	30.28
34	S106	Grapevine rupestris vein feathering virus	NC_034205.1	6730	1205	17.90	1598	23.74	1424	21.16	2498	37.12
35	S107	Peach virus	NC_033828.1	6612	1287	19.46	1277	19.31	1398	21.14	2650	40.08

Fig. 31 shows graph of Percentage Concentrations in virus sequences of dominant cytosines and other nucleotides

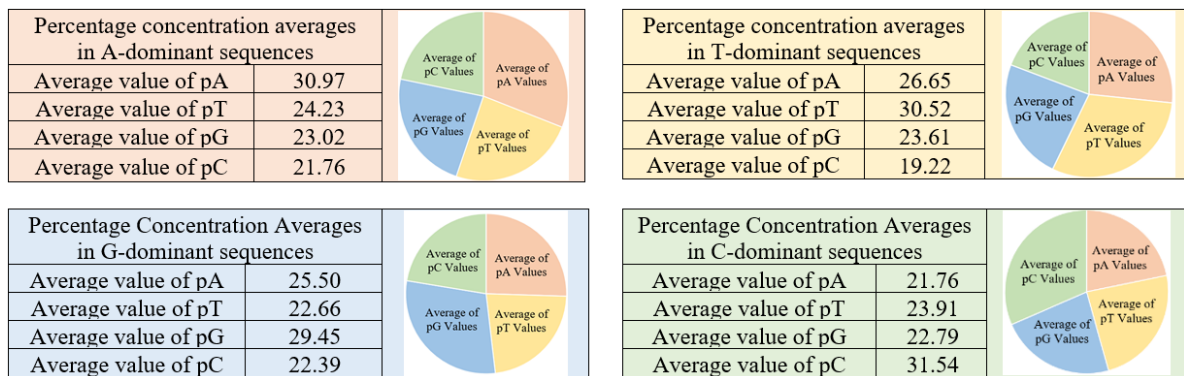
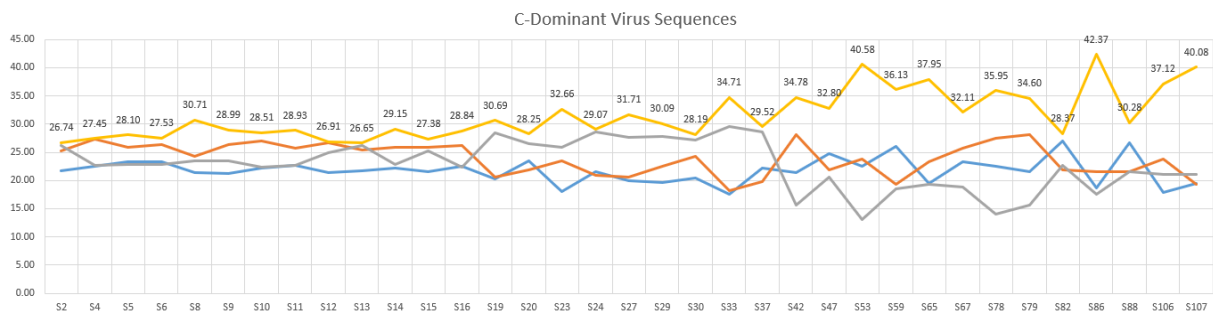


Fig. 32: Percentage concentration averages of all dominant sequences

Thus, classification of all 130 virus genomes have been carried out based on correlation coefficients and differential errors.

The next quantificational measure used to classify 130 genome sequences is based on the fundamental notion of 'Golden Ratio' defined for certain geometric figures and numerical sequences.

#### IV. CLASSIFICATION OF 130 VIRUS GENOMES BASED ON JCP GOLDEN RATIO VALUES

The term 'Golden Ratio (GR)' refers to a number mostly observed while taking ratios of distances in simple geometric figures such as pentagon, pentagram, decagon and dodecahedron. It is usually denoted by the symbol  $\phi$ .

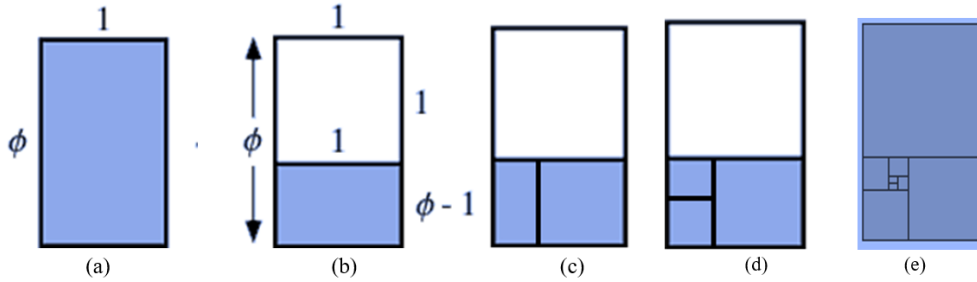


Fig. 33: Rectangle that exhibits Golden Ratio Property (GRP)

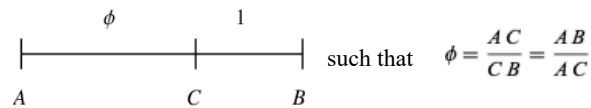
Fig. 33(a) shows a rectangle having sides in the ratio  $1: x$ . Now, for a unique value of  $x$ , which is denoted as  $\phi$ , partitioning the given rectangle as a square and a small rectangle as shown in Fig. 33(b), the smaller rectangle will have its sides in the same ratio  $1: x$ . One can further partition the small rectangle into a square and a smaller rectangle. This kind of partitioning could be carried out indefinitely. One can observe smaller rectangles have their sides in the ratio  $1: x$  in every partitioning. Now the rectangle shown in Fig. 33(a) is called a 'Golden Rectangle'. Such a rectangle and successive points dividing a golden rectangle into squares lie on a logarithmic spiral, giving a figure known as a whirling square. This property of rectangles having sides in the ratio  $1: x$  is called Golden Ratio Property (GRP). The unique number is  $x = \phi = 1.618$ .

Based on this definition one would obtain the equation:

$$\frac{\phi}{1} = \frac{1}{\phi - 1}, \quad \phi^2 - \phi - 1 = 0.$$

#### Euclid's Definition

Consider a line segment AB divided as AC and CB in the ratio  $\phi:1$ .



$$\frac{\phi + 1}{\phi} = \phi, \text{ and now one obtains } \phi^2 - \phi - 1 = 0,$$

Solving this quadratic equation, one obtains

$$\phi = \frac{1}{2}(1 + \sqrt{5})$$

1.618033988749894848204586834365638117720 ...

#### Fibonacci Series

Another fascinating connection with the Fibonacci numbers is given by the series

$$\phi = 1 + \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{F_n F_{n+1}}.$$

One can represent the series as a nested radical such as

$$\phi = \sqrt{1 + \sqrt{1 + \sqrt{1 + \sqrt{1 + \dots}}}}$$

This is equivalent to the recurrence equation  $a_n^2 = a_{n-1} + 1$ , with  $a_1 = 1$ , giving  $\lim_{n \rightarrow \infty} a_n = \phi$ .  $\phi$  is the "worst" real number for rational approximation because its continued fraction representation  $\phi = [1, 1, 1, 1, \dots]$

$$1 + \frac{1}{1 + \frac{1}{1 + \dots}}$$

Let  $x_n = p_n/q_n$  converge according to the recurrence relation  $x_n = 1 + \frac{1}{x_{n-1}}$  with  $x_1 = 1$

Now, this has the solution  $x_n = \frac{F_{n+1}}{F_n}$ , where  $F_n$  is the  $n^{\text{th}}$  Fibonacci number

Let  $x_n = p_n/q_n$  converge according to the recurrence relation with  $x_1 = 1$

As a result, one obtains the equation:

$$\phi = \lim_{n \rightarrow \infty} x_n = \lim_{n \rightarrow \infty} \frac{F_n}{F_{n-1}}$$

Golden Ratio is also obtained using the recurrence relation  $\phi^n = \phi^{n-1} + \phi^{n-2}$ .

The question that arises here is whether it is possible to attribute or extend the notion of ‘Golden Ratio’ to arbitrary geometrical figures and arbitrary sequences and arrays. Mainly, the objective of this paper is to classify virus genome sequences based on a quantificational measure ‘Golden Ratio’.

### Golden Ratios of Virus Genome Sequences due to Jean Claude Perez

Jean Claude Perez proposed a method of attributing a quantificational measure on the lines of ‘Golden Ratio’. Golden Ratio (GR) of a viral sequence may be treated as the solution of rational polynomial  $(x1A + x2T) / (x3G + x4C)$ , where  $x1$ ,  $x2$ ,  $x3$  and  $x4$  denote certain coefficients; A denotes the number of Adenines in the sequence; T denotes the number of Thymines in the sequence; G denotes the number of Guanines in the sequence; C denotes the number of Cytosines in the sequence. In order to honor Jean Claude Perez, we denote this Golden Ratio as ‘JCP Golden Ratio’.

#### 4.1 JCP Golden Ratio (GR) Calculations

Since life began on earth, four types of bases A, G, C, and T/U form two sets of natural base pairs  $A \leftrightarrow T$  and  $G \leftrightarrow C$  that have remained unchanged as the components of nucleic acids that replicate and transfer genetic information. Throughout evolution, except for the U to T modification, the four base structures have not changed. Recently, researchers have developed new artificial pairs of nucleobases (unnatural base pairs) that function alongside the natural base pairs. Some unnatural base pairs in duplex DNA can be efficiently and faithfully amplified in a polymerase chain reaction (PCR) using thermostable DNA polymerases. The addition of unnatural base pair systems could expand the genetic alphabet of DNA, thus providing a new mechanism for the generation novel biopolymers by the site-specific incorporation of functional components into nucleic acids and proteins. Moreover, the process of unnatural base pair development might provide clues to the origin of the natural base pairs in a primordial soup on the early Earth. In this Account, we describe the development of four representative types of unnatural base pairs that function as pairs of nucleobases in PCR and reconsider the origin of the natural nucleic acids. They are  $A \leftrightarrow C$ ,  $T \leftrightarrow G$ ,  $A \leftrightarrow G$  and  $T \leftrightarrow C$ . This section of the paper proposes classification of a set of 130 virus genomes based on the extended notion of ‘Goldent Ratios’. Two types of pairing are considered here: (i) natural pairing  $A \leftrightarrow T$  and  $G \leftrightarrow C$  and unnatural pairing  $A \leftrightarrow C$ ,  $T \leftrightarrow G$ . The rational polynomial used for calculating Golden Ratio of sequences obeying natural pairing is  $(x1A + x2T) / (x3G + x4C)$  and (ii) the rational polynomial used for calculating Golden Ratio of sequences obeying natural pairing is  $(x1A + x2C) / (x3T + x4G)$ . Now, Golden Ratios (GRs) of all 130 virus genome sequences are calculated for all eight assignments A1 to A8 in the context of ‘Natural Pairing’. Eight assignments of values, for instance, 1, 2, 3 and 4 to coefficients are considered here such that A1:  $x1=1$ ,  $x2=2$ ,  $x3=3$  and  $x4=4$ . One can cyclically permute the sequence of assigned values 1,2,3,4 as 4,1,2,3 and further permute as 3,4,1,2, and 2,3,4,1. One may have the graphical inverse of A1=1,2,3,4 as A5=4,3,2,1 and permute it as A6=1,4,3,2, A7=2,1,4,3 and A8=3,2,1,4. Golden Ratios are obtained for all the eight assignments and results studied. One can repeat this for sequences obeying unnatural pairing also.

<p><b>Golden Ratio (GR)</b> of Viral Sequence obeying <b>Natural Pairing</b> may be treated as the solution of rational polynomial <math>(x1A + x2T) / (x3G + x4C)</math>, where <math>x1</math>, <math>x2</math>, <math>x3</math> and <math>x4</math> denote the coefficients; A denotes the number of Adenines in the sequence; T denotes the number of Thymines in the sequence; G denotes the number of Guanines in the sequence; C denotes the number of Cytosines in the sequence. The term ‘<b>Natural Pairing</b>’ refers to pairing of (i) Adenine and Thymine and (ii) Guanine and Cytosine. The first assignment A1 refers to the numerical values of the coefficients in the rational polynomial <math>x1=1</math>; <math>x2=2</math>; <math>x3=3</math> and <math>x4=4</math>. Assignments A2, A3 and A4 refer to the cyclically permuted values. Assignments A5, A6, A7 and A8 refer to the graphical inverse and their cyclic permutations.</p>		x	x	x	x
		1	2	3	4
	A1	1	2	3	4
	A2	4	1	2	3
	A3	3	4	1	2
	A4	2	3	4	1
	A5	4	3	2	1
	A6	1	4	3	2
	A7	2	1	4	3
A8	3	2	1	4	

The results are given below in tables 16 to 28 along with bar graphs.

Table 16: Golden Ra of Virus Genomes S1 to S10 (Natural pairing)

Seq. No.	Accession ID	Total length	A	T	G	C	Golden Ratios for Natural Pairing							
							A1	A2	A3	A4	A5	A6	A7	A8
S1	NC_001639.1	14104	3244	3888	3646	3326	0.454	0.976	2.455	1.013	2.32	1.068	0.422	1.032
S2	NC_001961.1	15428	3353	3903	4047	4125	0.389	0.845	2.087	0.906	2.055	0.93	0.371	0.869
S3	NC_002532.2	12704	2692	3449	3305	3258	0.417	0.867	2.227	0.954	2.139	1.003	0.384	0.916
S4	NC_003092.2	15717	3534	4307	3556	4314	0.435	0.919	2.284	1.078	2.368	1.075	0.418	0.923
S5	NC_025112.1	14927	3476	3853	3404	4194	0.414	0.915	2.191	1.039	2.314	1.015	0.412	0.898
S6	NC_025113.1	14851	3456	3909	3398	4088	0.424	0.93	2.246	1.054	2.347	1.039	0.418	0.92
S7	NC_026439.1	14953	3238	3920	3944	3851	0.406	0.867	2.18	0.929	2.105	0.968	0.38	0.907
S8	NC_026509.1	15684	3364	3816	3688	4816	0.362	0.791	1.903	0.928	2.042	0.9	0.361	0.772
S9	NC_037124.1	15478	3291	4077	3623	4487	0.397	0.832	2.078	0.991	2.164	0.987	0.381	0.835
S10	NC_029053.1	14924	3311	4025	3333	4255	0.42	0.888	2.198	1.063	2.318	1.048	0.407	0.883

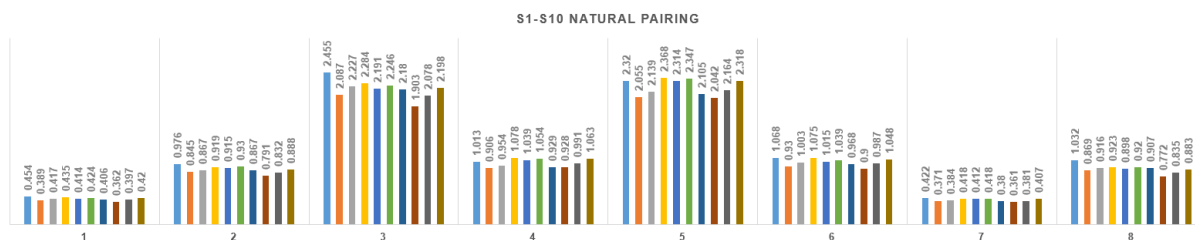


Table 17: Golden Ratios of Virus Genomes S11 to S20 (Natural pairing)

Seq. No.	Accession ID	Total length	A	T	G	C	Golden Ratios for Natural Pairing							
							A1	A2	A3	A4	A5	A6	A7	A8
S11	NC_029992.1	15247	3450	3929	3457	4411	0.403	0.879	2.122	1.024	2.259	0.998	0.4	0.862
S12	NC_035127.1	13766	2948	3675	3438	3705	0.409	0.859	2.17	0.969	2.156	0.995	0.384	0.886
S13	NC_038291.1	15411	3345	3923	4036	4107	0.392	0.848	2.1	0.911	2.064	0.936	0.372	0.873
S14	NC_038293.1	15370	3406	3979	3502	4481	0.399	0.86	2.096	1.014	2.225	0.992	0.393	0.848
S15	NC_043487.1	15111	3247	3912	3815	4137	0.395	0.843	2.1	0.939	2.101	0.958	0.376	0.862
S16	NC_048209.1	15307	3454	4008	3431	4414	0.41	0.886	2.153	1.043	2.291	1.019	0.404	0.871
S17	NC_003555.1	6277	1606	1530	1565	1576	0.424	1.012	2.318	0.995	2.34	0.984	0.431	1.001
S18	NC_005883.1	5310	1076	1441	1465	1328	0.407	0.83	2.181	0.9	2.026	0.97	0.364	0.901
S19	NC_007523.1	4975	1008	1024	1416	1527	0.295	0.682	1.592	0.707	1.629	0.698	0.296	0.674
S20	NC_009224.1	5261	1232	1151	1392	1486	0.349	0.839	1.901	0.838	1.962	0.816	0.36	0.817

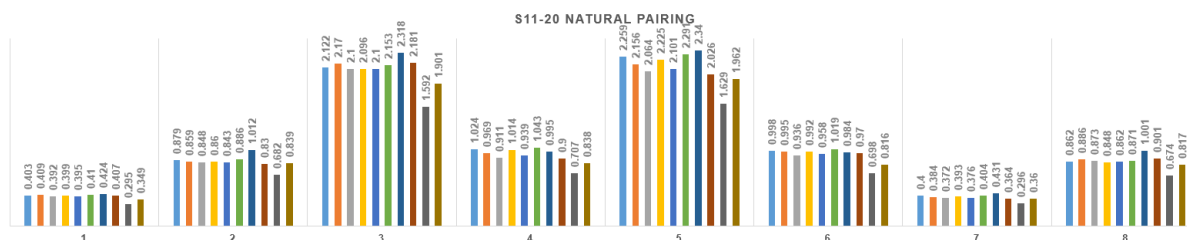
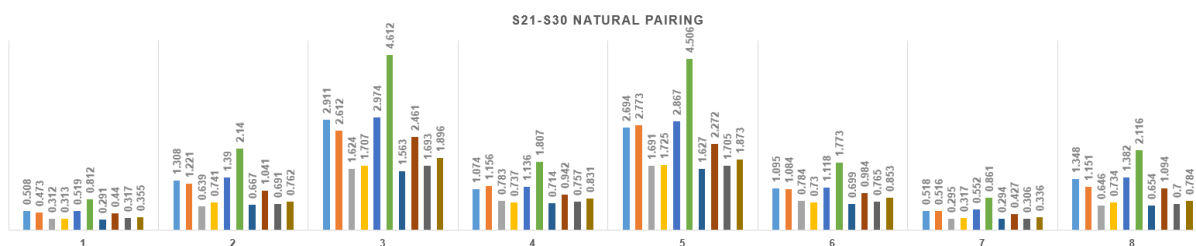


Table 18: Golden Ratios of Virus Genomes S21 to S30 (Natural pairing)

Seq. No.	Accession ID	Total length	A	T	G	C	Golden Ratios for Natural Pairing							
							A1	A2	A3	A4	A5	A6	A7	A8
S21	NC_009890.1	5077	1516	1252	1336	973	0.508	1.308	2.911	1.074	2.694	1.095	0.518	1.348
S22	NC_014609.1	7510	2239	1784	1672	1815	0.473	1.221	2.612	1.156	2.773	1.084	0.516	1.151
S23	NC_024151.1	5327	961	1251	1375	1740	0.312	0.639	1.624	0.783	1.691	0.784	0.295	0.646
S24	NC_025214.1	5322	1145	1111	1519	1547	0.313	0.741	1.707	0.737	1.725	0.73	0.317	0.734
S25	NC_027212.1	5704	1811	1359	1414	1120	0.519	1.39	2.974	1.136	2.867	1.118	0.552	1.382
S26	NC_030295.1	7788	2902	2249	1436	1201	0.812	2.14	4.612	1.807	4.506	1.773	0.861	2.116
S27	NC_030867.1	5124	1022	1058	1419	1625	0.291	0.667	1.563	0.714	1.627	0.699	0.294	0.654
S28	NC_035674.1	6203	1570	1551	1730	1352	0.44	1.041	2.461	0.942	2.272	0.984	0.427	1.094
S29	NC_038928.1	5194	1016	1172	1441	1563	0.317	0.691	1.693	0.757	1.705	0.765	0.306	0.7
S30	NC_038929.1	5228	1065	1271	1418	1474	0.355	0.762	1.896	0.831	1.873	0.853	0.336	0.784



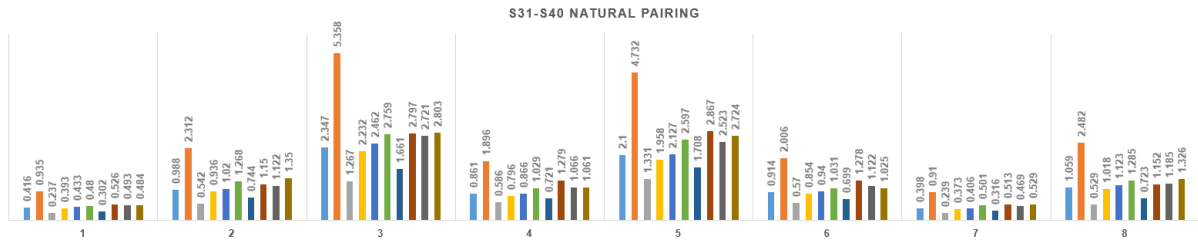


Table 20: Golden Ratios of Virus Genomes S41 to S50 (Natural pairing)

Seq. No.	Accession ID	Total length	A	T	G	C	Golden Ratios for Natural Pairing							
							A1	A2	A3	A4	A5	A6	A7	A8
S41	NC_001512.1	11835	3676	2440	2859	2860	0.427	1.199	2.423	1.026	2.567	0.939	0.489	1.112
S42	NC_001513.1	6211	1330	1748	973	2160	0.417	0.838	2.074	1.306	2.572	1.149	0.424	0.778
S43	NC_001564.2	10682	2618	2614	2919	2531	0.415	0.974	2.294	0.92	2.188	0.946	0.407	1.002
S44	NC_001642.1	6366	1925	1218	1339	1884	0.377	1.07	2.084	1.036	2.488	0.873	0.46	0.925
S45	NC_001728.1	6618	1990	2013	1441	1174	0.667	1.557	3.7	1.444	3.451	1.505	0.645	1.628
S46	NC_001948.1	8744	2430	2560	2075	1678	0.583	1.337	3.227	1.256	2.985	1.322	0.556	1.412
S47	NC_002604.1	6827	1689	1494	1405	2239	0.355	0.865	1.877	1	2.225	0.881	0.394	0.777
S48	NC_002729.1	7352	2511	2096	1571	1174	0.712	1.821	4.061	1.516	3.784	1.542	0.725	1.87
S49	NC_002795.1	8657	2306	2278	2149	1924	0.485	1.142	2.673	1.088	2.58	1.109	0.479	1.165
S50	NC_003557.1	8363	2282	2413	1943	1725	0.558	1.273	3.059	1.242	2.916	1.286	0.538	1.319

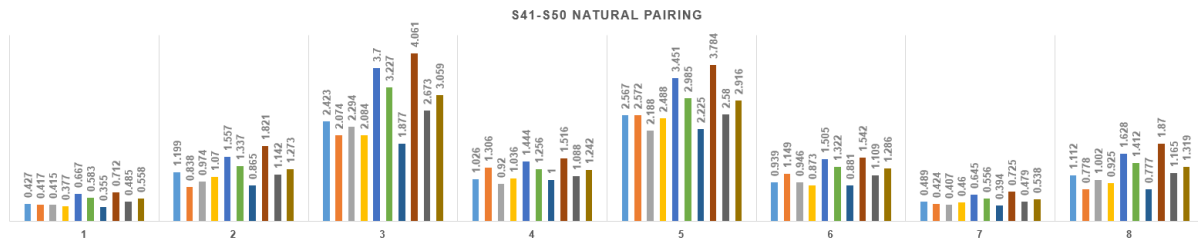


Table 21: Golden Ratios of Virus Genomes S51 to S60 (Natural pairing)

Seq. No.	Accession ID	Total length	A	T	G	C	Golden Ratios for Natural Pairing							
							A1	A2	A3	A4	A5	A6	A7	A8
S51	NC_003603.1	4019	970	812	1176	1061	0.333	0.847	1.867	0.759	1.85	0.746	0.348	0.836
S52	NC_003608.1	3911	975	1012	1007	917	0.448	1.03	2.454	1.008	2.366	1.034	0.436	1.058
S53	NC_003634.1	6673	1503	1588	874	2708	0.347	0.769	1.726	1.252	2.418	0.977	0.395	0.656
S54	NC_003679.1	12333	4007	2715	3203	2408	0.49	1.375	2.853	1.061	2.742	1.03	0.535	1.359
S55	NC_003852.1	6507	1949	1867	1520	1171	0.614	1.474	3.447	1.31	3.181	1.364	0.6	1.544
S56	NC_003900.1	11824	3462	2628	2828	2906	0.433	1.146	2.418	1.041	2.538	0.977	0.476	1.082
S57	NC_004724.1	16527	4250	4624	4224	3429	0.511	1.154	2.819	1.1	2.599	1.164	0.482	1.226
S58	NC_005062.1	10787	2772	2229	3361	2425	0.365	0.951	2.098	0.77	1.943	0.782	0.375	0.978
S59	NC_005132.1	6966	1813	1348	1288	2517	0.323	0.849	1.713	1	2.217	0.809	0.391	0.716
S60	NC_006939.1	3683	1041	868	900	874	0.448	1.137	2.49	1.047	2.531	1.014	0.474	1.105

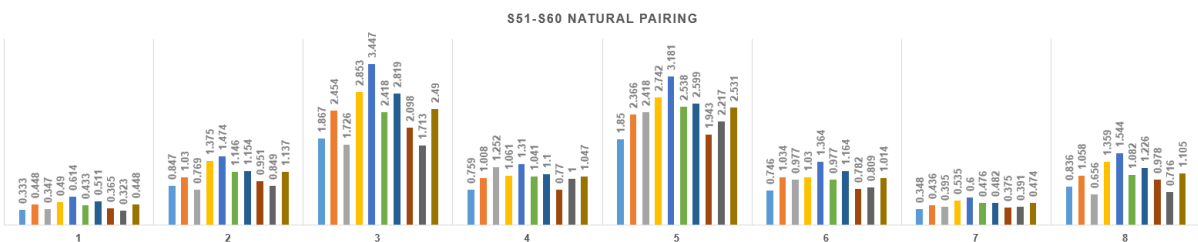


Table 22: Golden Ratios of Virus Genomes S61 to S70 (Natural pairing)

Seq. No.	Accession ID	Total length	A	T	G	C	Golden Ratios for Natural Pairing							
							A1	A2	A3	A4	A5	A6	A7	A8
S61	NC_007733.2	3962	1054	937	990	981	0.424	1.046	2.34	0.995	2.373	0.973	0.441	1.024
S62	NC_009028.2	10755	2903	2229	3004	2619	0.377	0.998	2.138	0.853	2.121	0.829	0.404	0.976
S63	NC_009892.1	9005	2392	2669	2042	1902	0.562	1.249	3.053	1.27	2.936	1.316	0.537	1.296
S64	NC_011535.1	4731	1243	1167	1319	1002	0.449	1.087	2.526	0.953	2.327	0.991	0.441	1.138
S65	NC_011538.1	6285	1223	1467	1210	2385	0.315	0.664	1.594	0.947	1.934	0.844	0.326	0.614
S66	NC_011552.1	7988	2439	2249	1908	1392	0.614	1.502	3.476	1.288	3.168	1.344	0.603	1.58
S67	NC_011559.1	6151	1437	1585	1154	1975	0.405	0.89	2.086	1.157	2.452	1.049	0.423	0.826
S68	NC_012533.1	10723	2922	2143	3219	2439	0.371	1.005	2.141	0.801	2.04	0.79	0.395	1.005
S69	NC_012534.1	10941	3190	2330	3088	2333	0.422	1.145	2.436	0.91	2.321	0.898	0.45	1.145
S70	NC_012812.1	12337	3945	2694	3266	2426	0.478	1.337	2.785	1.031	2.663	1	0.52	1.327

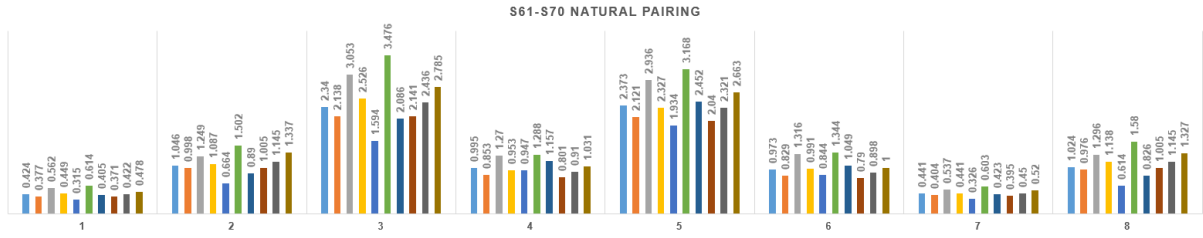


Table 23: Golden Ratios of Virus Genomes S71 to S80 (Natural pairing)

Seq. No.	Accession ID	Total length	A	T	G	C	Golden Ratios for Natural Pairing							
							A1	A2	A3	A4	A5	A6	A7	A8
S71	NC_013006.1	8517	2381	2215	2060	1857	0.5	1.211	2.771	1.129	2.705	1.136	0.505	1.219
S72	NC_015782.2	7259	2181	2230	1582	1266	0.676	1.573	3.758	1.455	3.479	1.525	0.65	1.655
S73	NC_016038.2	5666	1625	1300	1330	1411	0.438	1.131	2.426	1.062	2.554	1.001	0.476	1.071
S74	NC_016404.1	7488	2449	1558	1897	1584	0.462	1.328	2.68	1.043	2.69	0.979	0.523	1.27
S75	NC_016440.1	8638	2313	2473	2136	1716	0.546	1.244	3.022	1.173	2.784	1.24	0.518	1.32
S76	NC_016959.1	11724	3259	2588	3210	2632	0.418	1.091	2.375	0.923	2.297	0.913	0.439	1.088
S77	NC_018713.1	12292	3933	2728	3173	2458	0.485	1.345	2.807	1.059	2.716	1.028	0.527	1.326
S78	NC_020470.1	6337	1425	1742	892	2278	0.416	0.863	2.063	1.381	2.689	1.16	0.441	0.775
S79	NC_020471.1	6226	1344	1756	970	2154	0.421	0.848	2.094	1.318	2.599	1.159	0.429	0.786
S80	NC_023439.1	10688	2449	2280	3620	2339	0.346	0.847	1.984	0.697	1.736	0.744	0.333	0.917

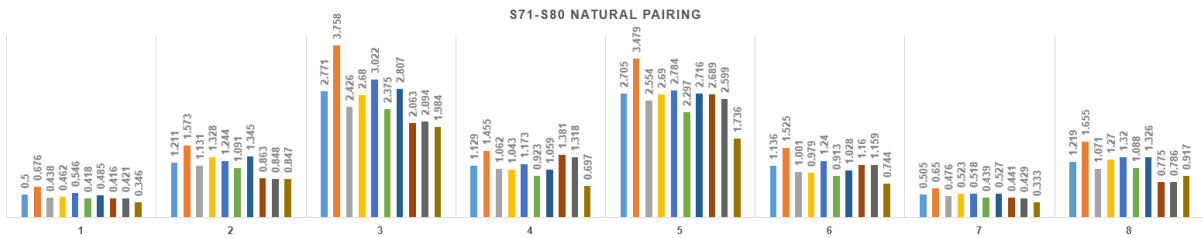


Table 24: Golden Ratios of Virus Genomes S81 to S90 (Natural pairing)

Seq. No.	Accession ID	Total length	A	T	G	C	Golden Ratios for Natural Pairing							
							A1	A2	A3	A4	A5	A6	A7	A8
S81	NC_023892.1	8659	2457	2086	2286	1829	0.467	1.184	2.643	1.018	2.513	1.027	0.478	1.202
S82	NC_024458.1	6677	1800	1465	1518	1894	0.389	0.993	2.122	1.003	2.351	0.918	0.43	0.915
S83	NC_024887.1	11550	3125	2313	3118	2994	0.363	0.973	2.045	0.852	2.106	0.806	0.399	0.927
S84	NC_026620.1	10125	2916	2599	2752	1858	0.517	1.287	2.959	1.059	2.643	1.111	0.508	1.369
S85	NC_028793.2	5851	1510	1427	1460	1413	0.435	1.043	2.388	1.006	2.381	1.001	0.441	1.038
S86	NC_030693.1	6863	1277	1476	1202	2908	0.277	0.591	1.387	0.904	1.795	0.762	0.297	0.528
S87	NC_031327.1	10588	2729	2629	2889	2341	0.442	1.058	2.47	0.96	2.315	0.992	0.435	1.097
S88	NC_040842.1	5839	1554	1260	1257	1768	0.375	0.956	2.024	1.013	2.334	0.902	0.422	0.862
S89	NC_040837.1	6060	1699	1942	1503	916	0.683	1.518	3.857	1.331	3.218	1.492	0.609	1.738
S90	NC_040800.1	8192	2847	2187	1884	1274	0.671	1.788	3.900	1.391	3.559	1.414	0.693	1.850

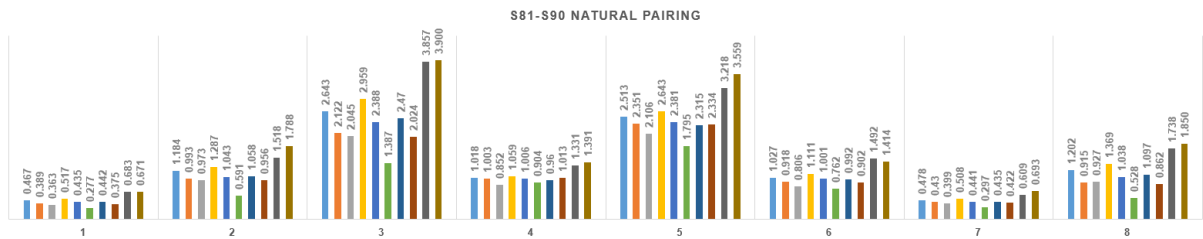


Table 25: Golden Ratios of Virus Genomes S91 to S100 (Natural pairing)

Seq. No.	Accession ID	Total length	A	T	G	C	Golden Ratios for Natural Pairing							
							A1	A2	A3	A4	A5	A6	A7	A8
S91	NC_040788.1	10311	2885	2271	3053	2102	0.422	1.112	2.444	0.879	2.235	0.895	0.434	1.151
S92	NC_040776.1	10794	2976	2169	3066	2583	0.374	1.013	2.138	0.839	2.112	0.811	0.405	0.990
S93	NC_039237.1	12513	4045	2783	3186	2499	0.491	1.367	2.842	1.078	2.765	1.042	0.537	1.342
S94	NC_039218.1	10376	2509	2139	3314	2414	0.346	0.877	1.975	0.729	1.819	0.749	0.349	0.910
S95	NC_039217.1	14072	4836	3986	2783	2467	0.703	1.799	3.946	1.590	3.896	1.564	0.736	1.776
S96	NC_038966.1	8881	2426	2429	2180	1846	0.523	1.225	2.894	1.148	2.737	1.186	0.510	1.268
S97	NC_036587.1	6692	2080	1479	1452	1681	0.454	1.233	2.525	1.147	2.782	1.036	0.519	1.125
S98	NC_035462.1	6930	2255	2046	1554	1075	0.708	1.747	4.035	1.460	3.623	1.532	0.694	1.854
S99	NC_035453.1	18848	5414	5495	4453	3486	0.600	1.402	3.345	1.282	3.077	1.347	0.577	1.480
S100	NC_035116.1	5877	1129	1730	1417	1601	0.430	0.817	2.231	1.024	2.188	1.079	0.380	0.875

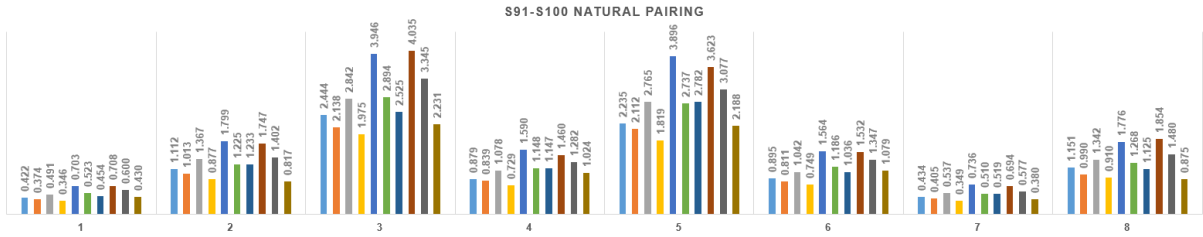


Table 26: Golden Ratios of Virus Genomes S101 to S110 (Natural pairing)

Seq. No.	Accession ID	Total length	A	T	G	C	Golden Ratios for Natural Pairing							
							A1	A2	A3	A4	A5	A6	A7	A8
S101	NC_035071.1	20414	6142	4537	4813	4922	0.445	1.193	2.495	1.071	2.624	1.000	0.494	1.122
S102	NC_034833.1	6958	2140	1572	1766	1480	0.471	1.270	2.688	1.052	2.648	1.020	0.508	1.244
S103	NC_034242.1	10370	2787	2704	2131	0.488	1.165	2.735	1.042	2.525	1.087	0.476	1.221	
S104	NC_034216.1	15098	5582	3901	3120	2495	0.692	1.911	3.988	1.527	3.895	1.476	0.754	1.873
S105	NC_034207.1	5953	1588	1380	1517	1468	0.417	1.039	2.309	0.970	2.330	0.949	0.435	1.018
S106	NC_034205.1	6730	1205	1598	1424	2498	0.308	0.620	1.558	0.879	1.798	0.819	0.303	0.596
S107	NC_033828.1	6612	1287	1277	1398	2650	0.259	0.597	1.339	0.777	1.648	0.673	0.284	0.534
S108	NC_033725.1	10203	2866	2441	2878	2018	0.463	1.177	2.655	0.964	2.416	0.996	0.465	1.231
S109	NC_033724.1	10215	2601	2229	3177	2208	0.384	0.973	2.201	0.797	1.996	0.825	0.384	1.020
S110	NC_033723.1	10251	2701	2209	3207	2134	0.392	1.015	2.266	0.803	2.039	0.830	0.395	1.066

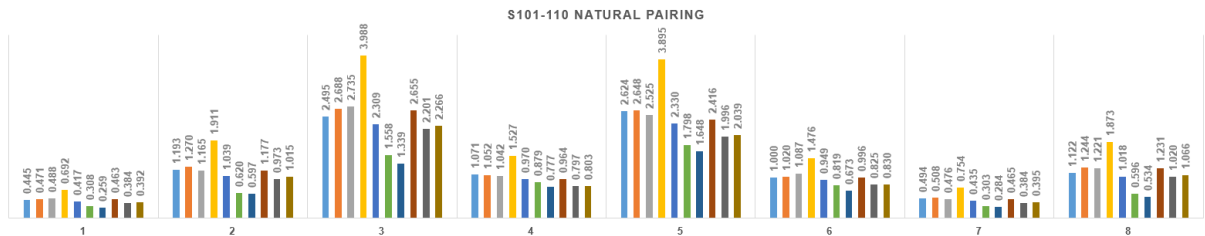


Table 27: Golden Ratios of Virus Genomes S111 to S120 (Natural pairing)

Seq. No.	Accession ID	Total length	A	T	G	C	Golden Ratios for Natural Pairing							
							A1	A2	A3	A4	A5	A6	A7	A8
S111	NC_033699.1	10173	2948	2319	2751	2155	0.449	1.179	2.566	0.976	2.448	0.973	0.470	1.185
S112	NC_033693.1	10173	2940	2338	2744	2151	0.452	1.180	2.579	0.982	2.457	0.980	0.471	1.189
S113	NC_032088.1	10864	2872	2395	2947	2650	0.394	1.002	2.206	0.895	2.185	0.880	0.412	0.989
S114	NC_031752.1	11557	4006	3151	2231	2169	0.670	1.748	3.748	1.574	3.842	1.505	0.723	1.679
S115	NC_031463.1	23635	9624	6181	4217	3613	0.811	2.318	4.683	1.845	4.734	1.728	0.917	2.208
S116	NC_031462.1	15207	4392	3640	3787	3388	0.468	1.195	2.625	1.063	2.598	1.044	0.490	1.179
S117	NC_043110.1	10182	2778	2274	2867	2263	0.415	1.068	2.357	0.901	2.242	0.904	0.428	1.080
S118	NC_008516.1	26660	8975	6143	3758	7784	0.501	1.362	2.664	1.594	3.55	1.249	0.627	1.123
S119	NC_026812.1	27004	8268	6892	3909	7929	0.507	1.264	2.649	1.579	3.413	1.299	0.594	1.083
S120	NC_038295.1	27318	9234	6954	4110	7020	0.572	1.498	3.058	1.676	3.792	1.405	0.677	1.292

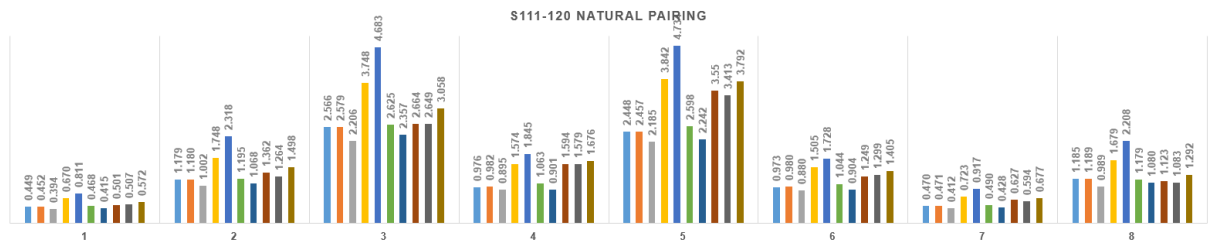
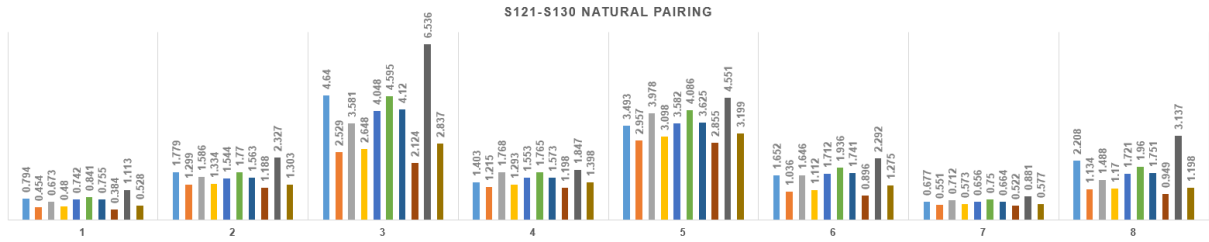


Table 28: Golden Ratios of Virus Genomes S121 to S130 (Natural pairing)

Seq. No.	Accession ID	Total length	A	T	G	C	Golden Ratios for Natural Pairing							
							A1	A2	A3	A4	A5	A6	A7	A8
S121	NC_027199.1	20261	5947	6839	5208	2266	0.794	1.779	4.64	1.403	3.493	1.652	0.677	2.208
S122	NC_024709.1	33452	11117	6932	6663	8740	0.454	1.299	2.529	1.215	2.957	1.036	0.551	1.134
S123	NC_033700.1	30353	9555	9194	4934	6670	0.673	1.586	3.581	1.768	3.978	1.646	0.712	1.488
S124	NC_035465.1	32399	10710	7139	6186	8364	0.48	1.334	2.648	1.293	3.098	1.112	0.573	1.17
S125	NC_007447.1	28475	7690	9959	6104	4718	0.742	1.544	4.048	1.553	3.582	1.712	0.656	1.721
S126	NC_022787.1	28301	8124	10267	5579	4331	0.841	1.77	4.595	1.765	4.086	1.936	0.75	1.96
S127	NC_034976.1	28487	7686	10072	6083	4646	0.755	1.563	4.12	1.573	3.625	1.741	0.664	1.751
S128	NC_046956.1	30742	10484	5273	5238	9747	0.384	1.188	2.124	1.198	2.855	0.896	0.522	0.949
S129	NC_046962.1	29409	8759	11760	6561	2328	1.113	2.327	6.536	1.847	4.551	2.292	0.881	3.137
S130	NC_046963.1	30859	9285	8097	5722	7755	0.528	1.303	2.837	1.398	3.199	1.275	0.577	1.198



Now, the GR values of all 130 virus genomes are analyzed based on ‘Assignments’ of values to the coefficients in the rational polynomial  $(x1A + x2T) / (x3G + x4C)$ . It has been observed (i) GR values of all 130 virus genomes remain almost the same for the assignments of A1 and A7 [Ref. Fig. 34], (ii) GR values of all 130 virus genomes remain almost the same but different from previous values for the assignments of A2 and A8 [Ref. Fig. 35], (iii) GR values of all 130 virus genomes remain almost the same but different from previous values for the assignments of A3 and A5 [Ref. Fig. 36] and (iv) GR values of all 130 virus genomes remain almost the same but different from previous values for the assignments of A4 and A6 [Ref. Fig. 37].

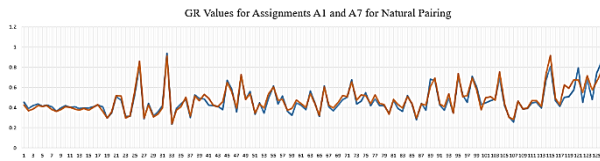


Fig. 34: GR values for assignments A1 and A7

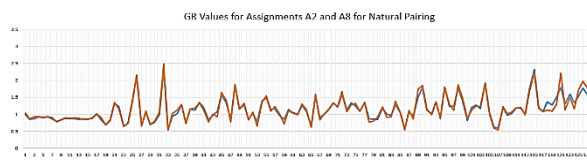


Fig. 35: GR values for assignments A2 and A8

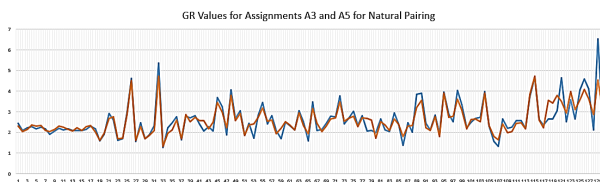


Fig. 36: GR values for assignments A3 and A5

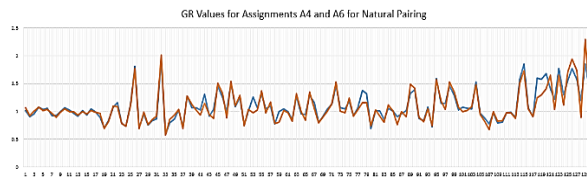


Fig. 37: GR values for assignments A4 and A6

Now, Golden Ratios of all 130 virus genome sequences are calculated for all eight assignments A1 to A8 in the context of ‘Unnatural Pairing’.

<b>Golden Ratio (GR)</b> of Viral Sequence obeying <b>Unnatural Pairing</b> may be treated as the solution of rational polynomial $(x1A + x4C) / (x2T + x3G)$ , where $x1, x2, x3$ and $x4$ denote the coefficients; A denotes the number of Adenines in the sequence; T denotes the number of Thymines in the sequence; G denotes the number of Guanines in the sequence; C denotes the number of Cytosines in the sequence. The term ‘ <b>Unnatural Pairing</b> ’ refers to pairing of (i) Adenine and Cytosine and (ii) Thymine and Guanine. The first assignment A1 refers to the numerical values of the coefficients in the rational polynomial $x1=1; x2=2; x3=3$ and $x4=4$ . Assignments A2, A3 and A4 refer to the cyclically permuted values. Assignments A5, A6, A7 and A8 refer to the graphical inverse and their cyclic permutations.	$x1$	$x2$	$x3$	$x4$	
	A1	1	2	3	4
	A2	4	1	2	3
	A3	3	4	1	2
	A4	2	3	4	1
	A5	4	3	2	1
	A6	1	4	3	2
	A7	2	1	4	3
	A8	3	2	1	4

Eight assignments of values, for instance, 1, 2, 3 and 4 to coefficients are considered here such that A1:  $x1=1, x2=2, x3=3$  and  $x4=4$ . One can cyclically permute the sequence of assigned values 1,2,3,4 as 4,1,2,3 and further permute as 3,4,1,2, and 2,3,4,1. One may have the graphical inverse of  $A1=1,2,3,4$  as  $A5=4,3,2,1$  and permute it as  $A6=1,4,3,2, A7=2,1,4,3$  and  $A8=3,2,1,4$ . Golden Ratios are obtained for all the eight assignments and results studied. The results are given below in tables 29 to 41 along with bar graphs.

Table 29: Golden Ratios of Virus Genomes S1 to S10 (Unnatural pairing)

Seria 1 No.	Accession ID	Total length	A	T	G	C	Golden Ratios for Unnatural Pairing							
							A1	A2	A3	A4	A5	A6	A7	A8
S1	NC_001639.1	14104	3244	3888	3646	3326	0.884	2.053	0.853	0.373	0.859	0.373	0.891	2.016
S2	NC_001961.1	15428	3353	3903	4047	4125	0.995	2.149	0.931	0.388	0.885	0.418	0.949	2.24
S3	NC_002532.2	12704	2692	3449	3305	3258	0.935	2.042	0.853	0.366	0.827	0.388	0.909	2.068
S4	NC_003092.2	15717	3534	4307	3556	4314	1.078	2.371	0.925	0.419	0.92	0.435	1.079	2.289
S5	NC_025112.1	14927	3476	3853	3404	4194	1.13	2.484	1	0.442	0.985	0.463	1.118	2.448
S6	NC_025113.1	14851	3456	3909	3398	4088	1.099	2.436	0.974	0.434	0.967	0.45	1.095	2.382
S7	NC_026439.1	14953	3238	3920	3944	3851	0.947	2.075	0.887	0.375	0.855	0.397	0.915	2.131
S8	NC_026509.1	15684	3364	3816	3688	4816	1.21	2.493	1.04	0.44	0.97	0.493	1.14	2.593
S9	NC_027124.1	15478	3291	4077	3623	4487	1.116	2.351	0.945	0.414	0.906	0.451	1.079	2.362
S10	NC_029053.1	14924	3311	4025	3333	4255	1.126	2.432	0.949	0.428	0.933	0.452	1.116	2.367

<b>Golden Ratio (GR)</b> of Viral Sequence obeying <b>Unnatural Pairing</b> may be treated as the solution of rational polynomial $(x1A + x4C) / (x2T + x3G)$ , where $x1, x2, x3$ and $x4$ denote the coefficients; A denotes the number of Adenines in the sequence; T denotes the number of Thymines in the sequence; G denotes the number of Guanines in the sequence; C denotes the number of Cytosines in the sequence. The term ' <b>Unnatural Pairing</b> ' refers to pairing of (i) Adenine and Cytosine and (ii) Thymine and Guanine. The first assignment A1 refers to the numerical values of the coefficients in the rational polynomial $x1=1; x2=2; x3=3$ and $x4=4$ . Assignments A2, A3 and A4 refer to the cyclically permuted values. Assignments A5, A6, A7 and A8 refer to the graphical inverse and their cyclic permutations.		$x1$	$x2$	$x3$	$x4$
	A1	1	2	3	4
	A2	4	1	2	3
	A3	3	4	1	2
	A4	2	3	4	1
	A5	4	3	2	1
	A6	1	4	3	2
	A7	2	1	4	3
	A8	3	2	1	4

Eight assignments of values, for instance, 1, 2, 3 and 4 to coefficients are considered here such that A1:  $x1=1, x2=2, x3=3$  and  $x4=4$ . One can cyclically permute the sequence of assigned values 1,2,3,4 as 4,1,2,3 and further permute as 3,4,1,2, and 2,3,4,1. One may have the graphical inverse of A1=1,2,3,4 as A5=4,3,2,1 and permute it as A6=1,4,3,2, A7=2,1,4,3 and A8=3,2,1,4. Golden Ratios are obtained for all the eight assignments and results studied. The results are given below in tables 29 to 41 along with bar graphs.

Table 29: Golden Ratios of Virus Genomes S1 to S10 (Unnatural pairing)

Serial No.	Accession ID	Total length	A	T	G	C	Golden Ratios for Unnatural Pairing							
							A1	A2	A3	A4	A5	A6	A7	A8
S1	NC_001639.1	14104	3244	3888	3646	3326	0.884	2.053	0.853	0.373	0.859	0.373	0.891	2.016
S2	NC_001961.1	15428	3353	3903	4047	4125	0.995	2.149	0.931	0.388	0.885	0.418	0.949	2.24
S3	NC_002532.2	12704	2692	3449	3305	3258	0.935	2.042	0.853	0.366	0.827	0.388	0.909	2.068
S4	NC_003092.2	15717	3534	4307	3556	4314	1.078	2.371	0.925	0.419	0.92	0.435	1.079	2.289
S5	NC_025112.1	14927	3476	3853	3404	4194	1.13	2.484	1	0.442	0.985	0.463	1.118	2.448
S6	NC_025113.1	14851	3456	3909	3398	4088	1.099	2.436	0.974	0.434	0.967	0.45	1.095	2.382
S7	NC_026439.1	14953	3238	3920	3944	3851	0.947	2.075	0.887	0.375	0.855	0.397	0.915	2.131
S8	NC_026509.1	15684	3364	3816	3688	4816	1.21	2.493	1.04	0.44	0.97	0.493	1.14	2.593
S9	NC_027124.1	15478	3291	4077	3623	4487	1.116	2.351	0.945	0.414	0.906	0.451	1.079	2.362
S10	NC_029053.1	14924	3311	4025	3333	4255	1.126	2.432	0.949	0.428	0.933	0.452	1.116	2.367

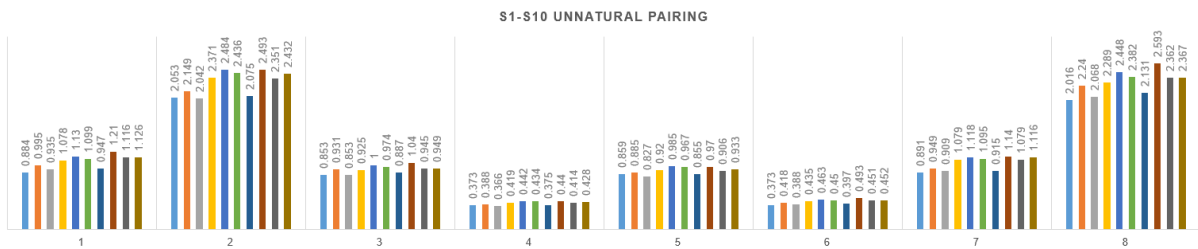


Table 30: Golden Ratios of Virus Genomes S11 to S20 (Unnatural pairing)

Serial No.	Accession ID	Total length	A	T	G	C	Golden Ratios for Unnatural Pairing							
							A1	A2	A3	A4	A5	A6	A7	A8
S11	NC_029992.1	15247	3450	3929	3457	4411	1.157	2.493	0.999	0.441	0.973	0.47	1.133	2.474
S12	NC_035127.1	13766	2948	3675	3438	3705	1.005	2.171	0.896	0.387	0.865	0.414	0.976	2.193
S13	NC_038291.1	15411	3345	3923	4036	4107	0.99	2.142	0.925	0.386	0.881	0.415	0.947	2.227
S14	NC_038293.1	15370	3406	3979	3502	4481	1.155	2.464	0.987	0.435	0.955	0.468	1.126	2.455
S15	NC_043487.1	15111	3247	3912	3815	4137	1.027	2.2	0.925	0.393	0.884	0.425	0.986	2.258
S16	NC_048209.1	15307	3454	4008	3431	4414	1.152	2.489	0.985	0.439	0.965	0.466	1.136	2.447
S17	NC_003555.1	6277	1606	1530	1565	1576	1.019	2.393	1.037	0.441	1.036	0.439	1.019	2.404
S18	NC_005883.1	5310	1076	1441	1465	1328	0.877	1.896	0.813	0.341	0.776	0.367	0.84	1.964
S19	NC_007523.1	4975	1008	1024	1416	1527	1.13	2.233	1.102	0.405	0.941	0.486	0.986	2.636
S20	NC_009224.1	5261	1232	1151	1392	1486	1.107	2.385	1.112	0.437	1.028	0.478	1.03	2.609

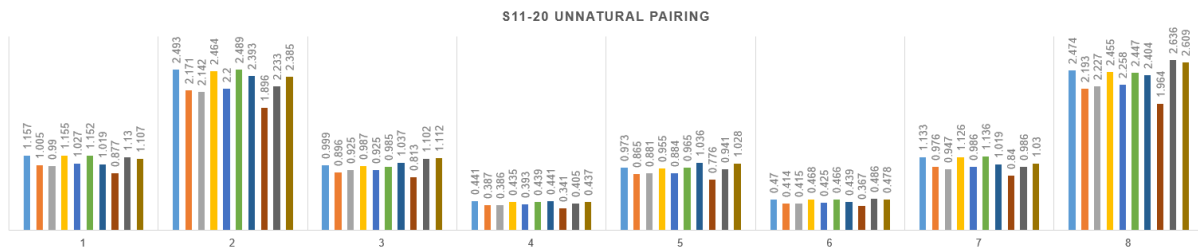


Table 31: Golden Ratios of Virus Genomes S21 to S30 (Unnatural pairing)

Serial No.	Accession ID	Total length	A	T	G	C	Golden Ratios for Unnatural Pairing							
							A1	A2	A3	A4	A5	A6	A7	A8
S21	NC_009890.1	5077	1516	1252	1336	973	0.83	2.289	1.023	0.44	1.094	0.383	0.902	2.197
S22	NC_014609.1	7510	2239	1784	1672	1815	1.106	2.808	1.174	0.522	1.238	0.482	1.171	2.667
S23	NC_024151.1	5327	961	1251	1375	1740	1.195	2.265	0.997	0.395	0.858	0.486	1.057	2.538
S24	NC_025214.1	5322	1145	1111	1519	1547	1.081	2.222	1.094	0.407	0.961	0.47	0.964	2.572
S25	NC_027212.1	5704	1811	1359	1414	1120	0.903	2.532	1.12	0.487	1.211	0.418	0.995	2.399
S26	NC_030295.1	7788	2902	2249	1436	1201	0.875	2.97	1.064	0.56	1.331	0.398	1.176	2.276
S27	NC_030867.1	5124	1022	1058	1419	1625	1.18	2.3	1.117	0.414	0.95	0.503	1.027	2.706
S28	NC_035674.1	6203	1570	1551	1730	1352	0.841	2.062	0.934	0.388	0.94	0.375	0.849	2.093
S29	NC_038928.1	5194	1016	1172	1441	1563	1.09	2.159	1.007	0.387	0.879	0.459	0.969	2.457
S30	NC_038929.1	5228	1065	1271	1418	1474	1.024	2.113	0.944	0.379	0.862	0.429	0.943	2.295

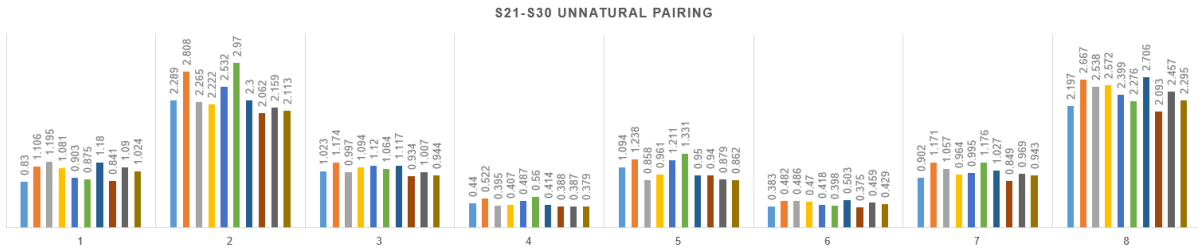


Table 32: Golden Ratios of Virus Genomes S31 to S40 (Unnatural pairing)

Serial No.	Accession ID	Total length	A	T	G	C	Golden Ratios for Unnatural Pairing							
							A1	A2	A3	A4	A5	A6	A7	A8
S31	NC_040431.1	4671	1147	1129	1401	994	0.792	1.925	0.917	0.365	0.901	0.359	0.783	2.027
S32	NC_040632.1	4947	1772	1606	947	622	0.703	2.558	0.889	0.484	1.148	0.325	1.002	1.876
S33	NC_040653.1	5281	929	959	1560	1833	1.252	2.259	1.195	0.404	0.925	0.539	1.021	2.909
S34	NC_040659.1	5082	1211	1188	1610	1073	0.763	1.829	0.908	0.349	0.872	0.35	0.739	1.988
S35	NC_040660.1	5160	1270	1282	1587	1021	0.73	1.827	0.871	0.349	0.869	0.334	0.734	1.901
S36	NC_040775.1	5941	1784	1389	1583	1185	0.866	2.347	1.081	0.452	1.134	0.403	0.922	2.314
S37	NC_040793.1	5206	1154	1029	1486	1537	1.12	2.306	1.166	0.425	1.015	0.493	0.992	2.711
S38	NC_046959.1	18410	4712	5381	3886	4431	1	2.443	0.905	0.437	0.973	0.409	1.085	2.175
S39	NC_000939.2	4415	1129	1206	1143	937	0.834	1.268	0.881	0.39	0.923	0.363	0.715	2
S40	NC_001461.1	12573	4049	2767	3233	2524	0.928	1.513	1.202	0.5	1.267	0.438	0.837	2.537

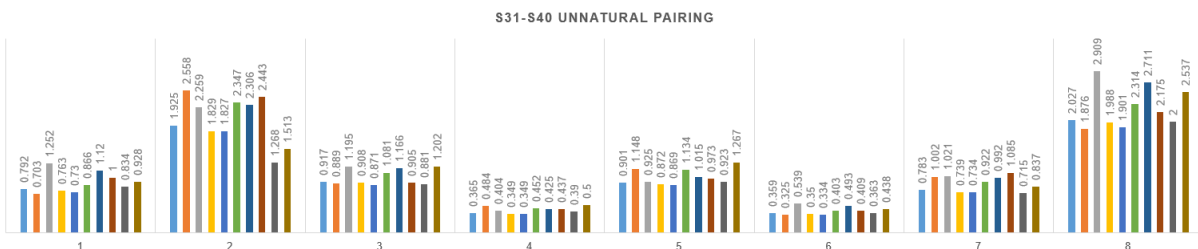
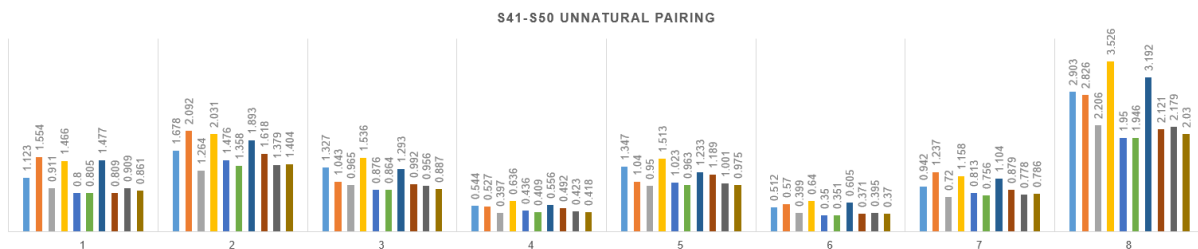


Table 33: Golden Ratios of Virus Genomes S31 to S50 (Unnatural pairing)

Serial No.	Accession ID	Total length	A	T	G	C	Golden Ratios for Unnatural Pairing							
							A1	A2	A3	A4	A5	A6	A7	A8
S41	NC_001512.1	11835	3676	2440	2859	2860	1.123	1.678	1.327	0.544	1.347	0.512	0.942	2.903
S42	NC_001513.1	6211	1330	1748	973	2160	1.554	2.092	1.043	0.527	1.04	0.57	1.237	2.826
S43	NC_001564.2	10682	2618	2614	2919	2531	0.911	1.264	0.965	0.397	0.95	0.399	0.72	2.206
S44	NC_001642.1	6366	1925	1218	1339	1884	1.466	2.031	1.536	0.636	1.513	0.64	1.158	3.526
S45	NC_001728.1	6618	1990	2013	1441	1174	0.8	1.476	0.876	0.436	1.023	0.35	0.813	1.95
S46	NC_001948.1	8744	2430	2560	2075	1678	0.805	1.358	0.864	0.409	0.963	0.351	0.756	1.946
S47	NC_002604.1	6827	1689	1494	1405	2239	1.477	1.893	1.293	0.556	1.233	0.605	1.104	3.192
S48	NC_002729.1	7352	2511	2096	1571	1174	0.809	1.618	0.992	0.492	1.189	0.371	0.879	2.121
S49	NC_002795.1	8657	2306	2278	2149	1924	0.909	1.379	0.956	0.423	1.001	0.395	0.778	2.179
S50	NC_003557.1	8363	2282	2413	1943	1725	0.861	1.404	0.887	0.418	0.975	0.37	0.786	2.03



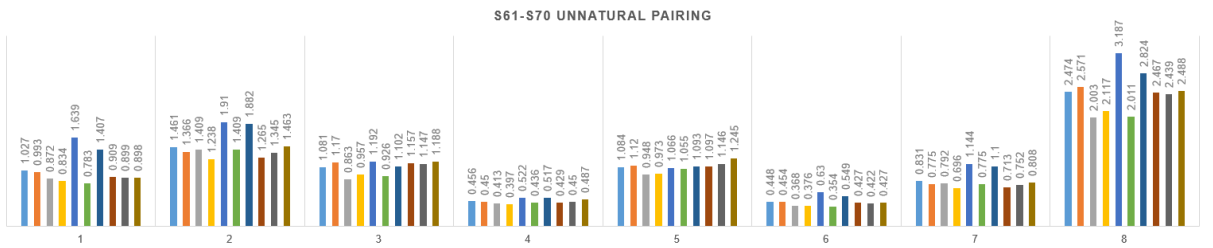
**Table 34:** Golden Ratios of Virus Genomes S51 to S60 (Unnatural pairing)

Serial No.	Accession ID	Total length	A	T	G	C	Golden Ratios for Unnatural Pairing							
							A1	A2	A3	A4	A5	A6	A7	A8
S51	NC_003603.1	4019	970	812	1176	1061	1.012	1.28	1.137	0.42	1.031	0.456	0.736	2.555
S52	NC_003608.1	3911	975	1012	1007	917	0.92	1.319	0.941	0.405	0.953	0.397	0.75	2.175
S53	NC_003634.1	6673	1503	1588	874	2708	2.127	2.78	1.373	0.691	1.339	0.771	1.656	3.787
S54	NC_003679.1	12333	4007	2715	3203	2408	0.906	1.497	1.197	0.497	1.266	0.431	0.826	2.508
S55	NC_003852.1	6507	1949	1867	1520	1171	0.799	1.423	0.911	0.433	1.037	0.356	0.785	0.004
S56	NC_003900.1	11824	3462	2628	2828	2906	1.097	1.618	1.214	0.512	1.237	0.488	0.913	2.722
S57	NC_004724.1	16527	4250	4624	4224	3429	0.819	1.267	0.863	0.387	0.915	0.356	0.713	1.964
S58	NC_005062.1	10787	2772	2229	3361	2425	0.857	1.171	1.072	0.395	1.007	0.401	0.663	2.304
S59	NC_005132.1	6966	1813	1348	1288	2517	1.811	2.277	1.567	0.668	1.475	0.739	1.332	3.892
S60	NC_006939.1	3683	1041	868	900	874	1.022	1.518	1.114	0.476	1.143	0.451	0.857	2.511



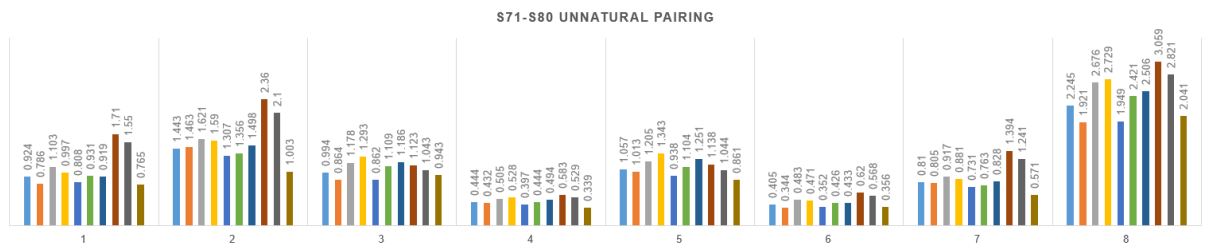
**Table 35:** Golden Ratios of Virus Genomes S61 to S70 (Unnatural pairing)

Serial No.	Accession ID	Total length	A	T	G	C	Golden Ratios for Unnatural Pairing							
							A1	A2	A3	A4	A5	A6	A7	A8
S61	NC_007733.2	3962	1054	937	990	981	1.027	1.461	1.081	0.456	1.084	0.448	0.831	2.474
S62	NC_009028.2	10755	2903	2229	3004	2619	0.993	1.366	1.17	0.45	1.12	0.454	0.775	2.571
S63	NC_009892.1	9005	2392	2669	2042	1902	0.872	1.409	0.863	0.413	0.948	0.368	0.792	2.003
S64	NC_011535.1	4731	1243	1167	1319	1002	0.834	1.238	0.957	0.397	0.973	0.376	0.696	2.117
S65	NC_011538.1	6285	1223	1467	1210	2385	1.639	1.91	1.192	0.522	1.066	0.63	1.144	3.187
S66	NC_011552.1	7988	2439	2249	1908	1392	0.783	1.409	0.926	0.436	1.055	0.354	0.775	2.011
S67	NC_011559.1	6151	1437	1585	1154	1975	1.407	1.882	1.102	0.517	1.093	0.549	1.1	2.824
S68	NC_012533.1	10723	2922	2143	3219	2439	0.909	1.265	1.157	0.429	1.097	0.427	0.713	2.467
S69	NC_012534.1	10941	3190	2330	3088	2333	0.899	1.345	1.147	0.45	1.146	0.422	0.752	2.439
S70	NC_012812.1	12337	3945	2694	3266	2426	0.898	1.463	1.188	0.487	1.245	0.427	0.808	2.488



**Table 36:** Golden Ratios of Virus Genomes S71 to S80 (Unnatural pairing)

Serial No.	Accession ID	Total length	A	T	G	C	Golden Ratios for Unnatural Pairing							
							A1	A2	A3	A4	A5	A6	A7	A8
S71	NC_013006.1	8517	2381	2215	2060	1857	0.924	1.443	0.994	0.444	1.057	0.405	0.81	2.245
S72	NC_015782.2	7259	2181	2230	1582	1266	0.786	1.463	0.864	0.432	1.013	0.344	0.805	1.921
S73	NC_016038.2	5666	1625	1300	1330	1411	1.103	1.621	1.178	0.505	1.205	0.483	0.917	2.676
S74	NC_016404.1	7488	2449	1558	1897	1584	0.997	1.59	1.293	0.528	1.343	0.471	0.881	2.729
S75	NC_016440.1	8638	2313	2473	2136	1716	0.808	1.307	0.862	0.397	0.938	0.352	0.731	1.949
S76	NC_016959.1	11724	3259	2588	3210	2632	0.931	1.356	1.109	0.444	1.104	0.426	0.763	2.421
S77	NC_018713.1	12292	3933	2728	3173	2458	0.919	1.498	1.186	0.494	1.251	0.433	0.828	2.506
S78	NC_020470.1	6337	1425	1742	892	2278	1.71	2.36	1.123	0.583	1.138	0.62	1.394	3.059
S79	NC_020471.1	6226	1344	1756	970	2154	1.55	2.1	1.043	0.529	1.044	0.568	1.241	2.821
S80	NC_023439.1	10688	2449	2280	3620	2339	0.765	1.003	0.943	0.339	0.861	0.356	0.571	2.041



S71-S80 UNNATURAL PAIRING

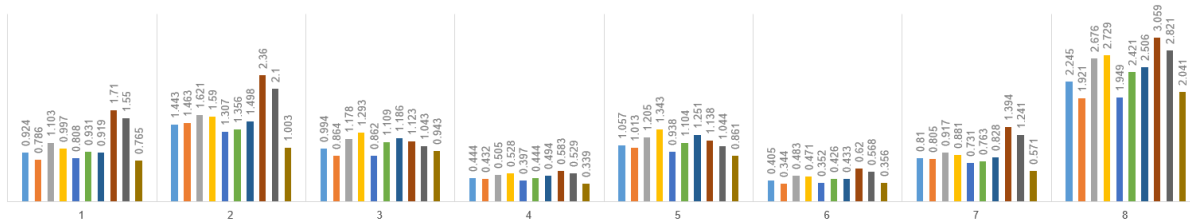


Table 37: Golden Ratios of Virus Genomes S81 to S90 (Unnatural pairing)

Serial No.	Accession ID	Total length	A	T	G	C	Golden Ratios for Unnatural Pairing							
							A1	A2	A3	A4	A5	A6	A7	A8
S81	NC_023892.1	8659	2457	2086	2286	1829	0.886	1.363	1.037	0.437	1.076	0.402	0.763	2.274
S82	NC_024458.1	6677	1800	1465	1518	1894	1.252	1.709	1.245	0.524	1.223	0.536	0.98	2.917
S83	NC_024887.1	11550	3125	2313	3118	2994	1.08	1.452	1.241	0.476	1.176	0.489	0.827	2.757
S84	NC_026620.1	10125	2916	2599	2752	1858	0.769	1.266	0.947	0.408	1.016	0.355	0.701	2.035
S85	NC_028793.2	5851	1510	1427	1460	1413	0.99	1.414	1.026	0.438	1.034	0.429	0.804	2.36
S86	NC_030693.1	6863	1277	1476	1202	2908	1.968	2.201	1.357	0.591	1.173	0.745	1.331	3.722
S87	NC_031327.1	10588	2729	2629	2889	2341	0.868	1.264	0.96	0.401	0.97	0.386	0.714	2.154
S88	NC_040842.1	5839	1554	1260	1257	1768	1.371	3.052	1.301	0.553	1.268	0.577	1.337	3.106
S89	NC_040837.1	6060	1699	1942	1503	916	0.638	1.928	0.747	0.364	0.873	0.287	0.772	1.626
S90	NC_040800.1	8192	2847	2187	1884	1274	0.792	2.554	1.042	0.494	1.225	0.374	0.978	2.179

S81-S90 UNNATURAL PAIRING

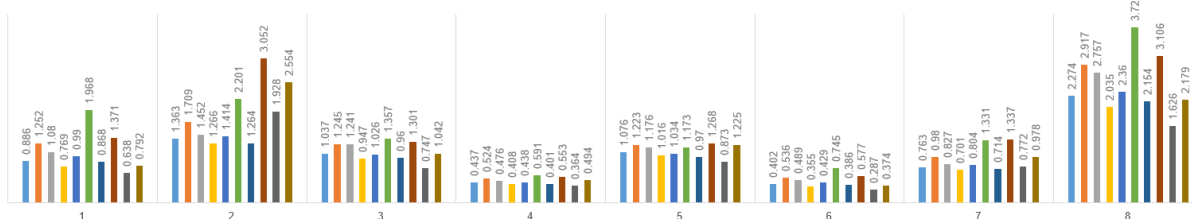


Table 38: Golden Ratios of Virus Genomes S91 to S100 (Unnatural pairing)

Serial No.	Accession ID	Total length	A	T	G	C	Golden Ratios for Unnatural Pairing							
							A1	A2	A3	A4	A5	A6	A7	A8
S91	NC_040788.1	10311	2885	2271	3053	2102	0.824	2.130	1.059	0.413	1.055	0.388	0.833	2.246
S92	NC_040776.1	10794	2976	2169	3066	2583	0.983	2.367	1.200	0.454	1.146	0.455	0.949	2.601
S93	NC_039237.1	12513	4045	2783	3186	2499	0.928	2.586	1.196	0.502	1.268	0.437	1.003	2.528
S94	NC_039218.1	10376	2509	2139	3314	2414	0.855	1.970	1.040	0.377	0.954	0.396	0.796	2.263
S95	NC_039217.1	14072	4836	3986	2783	2467	0.900	2.799	1.038	0.525	1.244	0.402	1.129	2.266
S96	NC_038966.1	8881	2426	2429	2180	1846	0.860	2.245	0.922	0.418	0.991	0.376	0.931	2.083
S97	NC_036587.1	6692	2080	1479	1452	1681	1.203	3.048	1.303	0.57	1.362	0.529	1.262	2.939
S98	NC_035462.1	6930	2255	2046	1554	1075	0.748	2.375	0.915	0.452	1.091	0.342	0.936	1.959
S99	NC_035453.1	18848	5414	5495	4453	3486	0.795	2.229	0.878	0.417	0.990	0.350	0.913	1.954
S100	NC_035116.1	5877	1129	1730	1417	1601	0.976	2.041	0.790	0.355	0.762	0.387	0.954	2.007

S91-S100 UNNATURAL PAIRING

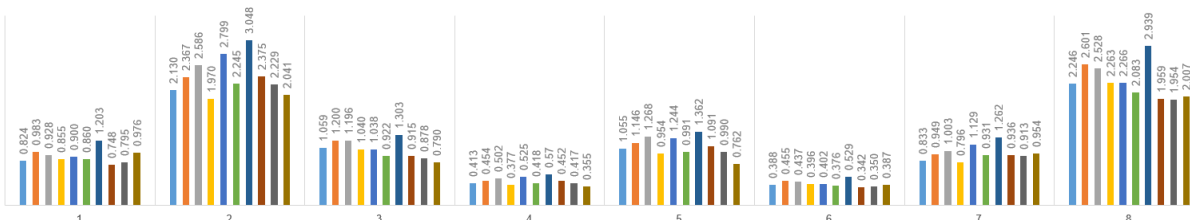


Table 39: Golden Ratios of Virus Genomes S101 to S110 (Unnatural pairing)

Serial No.	Accession ID	Total length	A	T	G	C	Golden Ratios for Unnatural Pairing							
							A1	A2	A3	A4	A5	A6	A7	A8
S101	NC_035071.1	20414	6142	4537	4813	4922	1.098	2.777	1.231	0.523	1.269	0.490	1.137	2.744
S102	NC_034833.1	6958	2140	1572	1766	1480	0.954	2.547	1.164	0.488	1.217	0.440	1.009	2.513
S103	NC_034242.1	10370	2787	2704	2748	2131	0.828	2.139	0.930	0.403	0.975	0.369	0.873	2.070
S104	NC_034216.1	15098	5582	3901	3120	2495	0.906	2.939	1.160	0.564	1.383	0.423	1.138	2.446
S105	NC_034207.1	5953	1588	1380	1517	1468	1.020	2.436	1.094	0.454	1.090	0.449	1.017	2.486
S106	NC_034205.1	6730	1205	1598	1424	2498	1.499	2.769	1.101	0.467	0.957	0.581	1.357	2.945
S107	NC_033828.1	6612	1287	1277	1398	2650	1.761	3.215	1.408	0.554	1.176	0.708	1.532	3.659
S108	NC_033725.1	10203	2866	2441	2878	2018	0.809	2.137	0.999	0.411	1.030	0.375	0.844	2.148
S109	NC_033724.1	10215	2601	2229	3177	2208	0.817	1.983	1.010	0.382	0.967	0.380	0.791	2.178
S110	NC_033723.1	10251	2701	2209	3207	2134	0.800	1.995	1.027	0.387	0.992	0.377	0.784	2.182

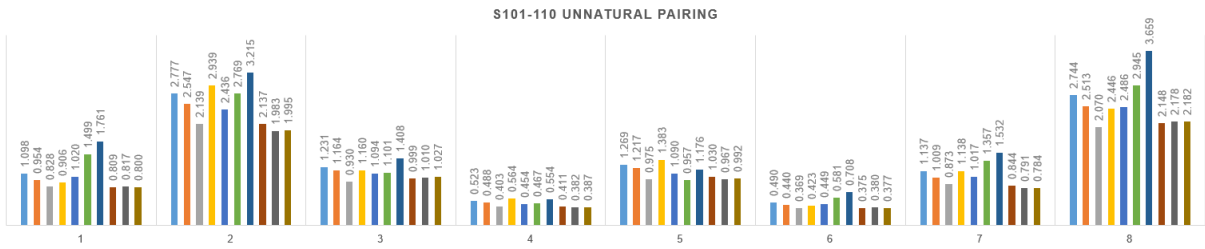


Table 40: Golden Ratios of Virus Genomes S111 to S120 (Unnatural pairing)

Serial No.	Accession ID	Total length	A	T	G	C	Golden Ratios for Unnatural Pairing							
							A1	A2	A3	A4	A5	A6	A7	A8
S111	NC_033699.1	10173	2948	2319	2751	2155	0.897	2.334	1.093	0.448	1.119	0.414	0.927	2.363
S112	NC_033693.1	10173	2940	2338	2744	2151	0.894	2.327	1.084	0.446	1.112	0.411	0.926	2.348
S113	NC_032088.1	10864	2872	2395	2947	2650	0.988	2.345	1.110	0.442	1.080	0.443	0.965	2.483
S114	NC_031752.1	11557	4006	3151	2231	2169	0.975	2.959	1.102	0.554	1.307	0.432	1.202	2.425
S115	NC_031463.1	23635	9624	6181	4217	3613	0.962	3.375	1.247	0.645	1.56	0.450	1.305	2.613
S116	NC_031462.1	15207	4392	3640	3787	3388	0.962	2.472	1.087	0.466	1.133	0.430	1.008	2.415
S117	NC_043110.1	10182	2778	2274	2867	2263	0.899	2.235	1.074	0.427	1.065	0.412	0.898	2.344
S118	NC_008516.1	26660	8975	6143	3758	7784	1.702	4.337	1.499	0.769	1.683	0.684	1.95	3.618
S119	NC_026812.1	27004	8268	6892	3909	7929	1.567	3.865	1.291	0.673	1.438	0.613	1.789	3.194
S120	NC_038295.1	27318	9234	6954	4110	7020	1.422	3.822	1.307	0.683	1.511	0.579	1.689	3.095

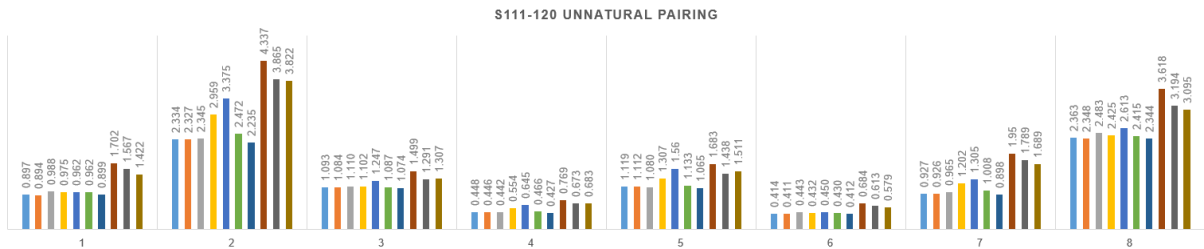
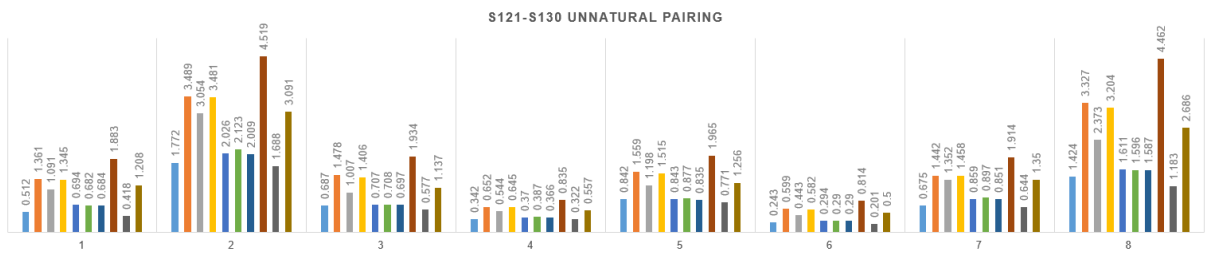


Table 41: Golden Ratios of Virus Genomes S121 to S130 (Unnatural pairing)

Serial No.	Accession ID	Total length	A	T	G	C	Golden Ratios for Unnatural Pairing							
							A1	A2	A3	A4	A5	A6	A7	A8
S121	NC_027199.1	20261	5947	6839	5208	2266	0.512	1.772	0.687	0.342	0.842	0.243	0.675	1.424
S122	NC_024709.1	33452	11117	6932	6663	8740	1.361	3.489	1.478	0.652	1.559	0.599	1.442	3.327
S123	NC_033700.1	30353	9555	9194	4934	6670	1.091	3.054	1.007	0.544	1.198	0.443	1.352	2.373
S124	NC_035465.1	32399	10710	7139	6186	8364	1.345	3.481	1.406	0.645	1.515	0.582	1.458	3.204
S125	NC_007447.1	28475	7690	9959	6104	4718	0.694	2.026	0.707	0.37	0.843	0.294	0.859	1.611
S126	NC_022787.1	28301	8124	10267	5579	4331	0.682	2.123	0.708	0.387	0.877	0.29	0.897	1.596
S127	NC_034976.1	28487	7686	10072	6083	4646	0.684	2.009	0.697	0.366	0.835	0.29	0.851	1.587
S128	NC_046956.1	30742	10484	5273	5238	9747	1.883	4.519	1.934	0.835	1.965	0.814	1.914	4.462
S129	NC_046962.1	29409	8759	11760	6561	2328	0.418	1.688	0.577	0.322	0.771	0.201	0.644	1.183
S130	NC_046963.1	30859	9285	8097	5722	7755	1.208	3.091	1.137	0.557	1.256	0.5	1.35	2.686



Now, the GR values of all 130 virus genomes are analyzed based on ‘Assignments’ of values to the coefficients in the rational polynomial  $(x1A + x2C) / (x3T + x4G)$ . It has been observed (i) GR values of all 130 virus genomes remain almost the same for the assignments of A1 and A7 [Ref. Fig. 38], (ii) GR values of all 130 virus genomes remain almost the same but different from previous values for the assignments of A2 and A8 [Ref. Fig. 39], (iii) GR values of all 130 virus genomes remain almost the same but different from previous values for the assignments of A3 and A5 [Ref. Fig. 40] and (iv) GR values of all 130 virus genomes remain almost the same but different from previous values for the assignments of A4 and A6 [Ref. Fig. 41].

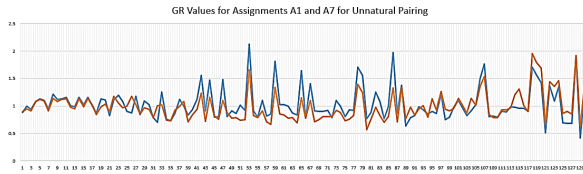


Fig. 38: GR values for assignments A1 and A7

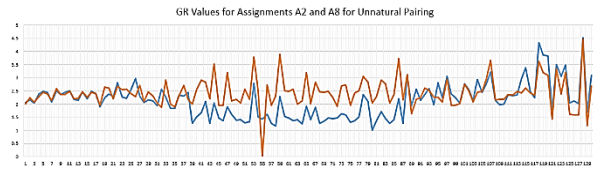


Fig. 39: GR values for assignments A2 and A8

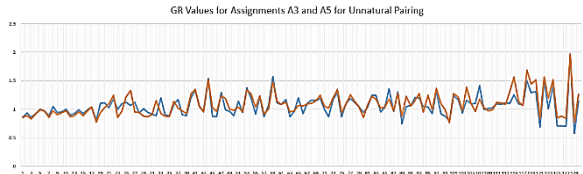


Fig. 40: GR values for assignments A3 and A5

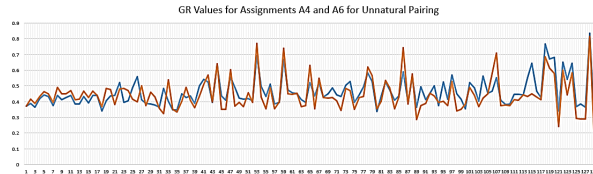


Fig. 41: GR values for assignments A4 and A6

This exercise of evaluating JCP Golden Ratios for all 130 virus genome sequences for A1=1,3,5,7, A2=7,1,3,5, A3=5,7,1,3, A4=3,5,7,1, A5=7,5,3,1, A6=1,7,5,3, A7=3,1,7,5 and A8=5,3,1,7 has been repeated and results compared with the previous ones. Due to space constraints, the results of this exercise are not provided.

**As per the definition of 'JCP Golden Ratios', their values are expected to remain within certain limits irrespective of the numerical values assigned to the coefficients in the formula  $(x1A + x2T) / (x3G + x4C)$ .** Table 42 presents the error values obtained using the formula  $A(1,2,3,4) - A(1,3,5,7)$ .

Table 42: Error values obtained using the formula  $A(1,2,3,4) - A(1,3,5,7)$

Sequences	A(1,2,3,4) – A(1,3,5,7)							
	A1	A2	A3	A4	A5	A6	A7	A8
S1	0.095	0.011	-0.733	0.002	-0.635	-0.012	0.099	-0.004
S2	0.082	0.010	-0.598	-0.005	-0.588	-0.011	0.086	0.004
S3	0.085	0.016	-0.648	-0.005	-0.601	-0.017	0.092	0.004
S4	0.092	0.018	-0.614	-0.022	-0.721	-0.021	0.097	0.017
S5	0.090	0.011	-0.583	-0.021	-0.712	-0.014	0.093	0.015
S6	0.091	0.013	-0.604	-0.019	-0.715	-0.014	0.095	0.014
S7	0.085	0.012	-0.635	-0.003	-0.590	-0.013	0.089	0.002
S8	0.078	0.012	-0.497	-0.024	-0.642	-0.014	0.082	0.016
S9	0.084	0.018	-0.556	-0.023	-0.664	-0.021	0.089	0.016
S10	0.089	0.018	-0.581	-0.027	-0.720	-0.022	0.094	0.019
S11	0.087	0.013	-0.559	-0.024	-0.704	-0.016	0.091	0.016
S12	0.085	0.016	-0.611	-0.011	-0.627	-0.018	0.090	0.009
S13	0.083	0.011	-0.601	-0.005	-0.590	-0.012	0.086	0.004
S14	0.085	0.015	-0.553	-0.025	-0.693	-0.018	0.090	0.017
S15	0.083	0.014	-0.588	-0.011	-0.613	-0.015	0.088	0.009
S16	0.088	0.015	-0.566	-0.026	-0.716	-0.018	0.092	0.018
S17	0.095	-0.004	-0.660	0.000	-0.673	0.003	0.094	-0.001
S18	0.082	0.017	-0.658	-0.001	-0.549	-0.017	0.088	-0.001
S19	0.065	0.002	-0.444	-0.005	-0.479	-0.003	0.065	0.004
S20	0.079	-0.003	-0.529	-0.004	-0.578	0.002	0.078	0.002
S21	0.117	-0.029	-0.930	0.027	-0.693	0.024	0.110	-0.043
S22	0.113	-0.018	-0.716	-0.001	-0.827	0.017	0.107	0.000
S23	0.065	0.017	-0.432	-0.021	-0.522	-0.019	0.069	0.015
S24	0.070	-0.001	-0.485	-0.001	-0.498	0.001	0.070	0.000
S25	0.124	-0.036	-0.915	0.026	-0.764	0.032	0.114	-0.037
S26	0.193	-0.048	-1.392	0.034	-1.223	0.044	0.179	-0.044
S27	0.064	0.004	-0.426	-0.009	-0.489	-0.005	0.066	0.007
S28	0.096	-0.008	-0.772	0.016	-0.593	0.006	0.095	-0.023
S29	0.067	0.009	-0.474	-0.008	-0.499	-0.010	0.070	0.006
S30	0.075	0.011	-0.539	-0.007	-0.538	-0.012	0.078	0.005
S31	0.091	-0.010	-0.765	0.020	-0.531	0.008	0.089	-0.032
S32	0.210	-0.042	-1.788	0.055	-1.169	0.034	0.199	-0.098
S33	0.053	0.003	-0.342	-0.009	-0.404	-0.005	0.052	0.006
S34	0.086	-0.012	-0.744	0.020	-0.484	0.009	0.083	-0.037
S35	0.094	-0.011	-0.833	0.023	-0.519	0.009	0.092	-0.044
S36	0.113	-0.032	-0.869	0.026	-0.678	0.028	0.105	-0.040

S37	0.069	-0.006	-0.467	0.000	-0.498	0.005	0.068	0.000
S38	0.113	0.015	-0.767	-0.018	-0.855	-0.017	0.118	0.014
S39	0.106	-0.001	-0.842	0.012	-0.668	-0.001	0.107	-0.018
S40	0.119	-0.044	-0.863	0.028	-0.727	0.038	0.106	-0.040
S41	0.107	-0.032	-0.677	0.010	-0.750	0.032	0.096	-0.011
S42	0.088	0.032	-0.460	-0.113	-0.982	-0.047	0.098	0.039
S43	0.091	-0.004	-0.692	0.009	-0.593	0.003	0.091	-0.012
S44	0.096	-0.023	-0.512	-0.018	-0.828	0.027	0.088	0.011
S45	0.146	-0.007	-1.144	0.020	-0.914	0.006	0.145	-0.027
S46	0.126	-0.002	-1.003	0.016	-0.787	0.001	0.126	-0.023
S47	0.083	0.001	-0.450	-0.038	-0.764	-0.003	0.082	0.020
S48	0.165	-0.038	-1.285	0.036	-0.982	0.032	0.154	-0.055
S49	0.108	-0.004	-0.796	0.009	-0.709	0.004	0.106	-0.011
S50	0.121	0.001	-0.917	0.008	-0.796	-0.002	0.121	-0.011
S51	0.077	-0.014	-0.550	0.009	-0.514	0.012	0.073	-0.011
S52	0.098	0.000	-0.728	0.006	-0.652	-0.001	0.098	-0.007
S53	0.078	0.020	-0.345	-0.158	-1.046	-0.033	0.085	0.037
S54	0.120	-0.046	-0.891	0.030	-0.722	0.040	0.107	-0.046
S55	0.136	-0.015	-1.086	0.025	-0.828	0.013	0.132	-0.035
S56	0.104	-0.021	-0.674	0.005	-0.743	0.021	0.097	-0.005
S57	0.109	0.001	-0.876	0.013	-0.685	-0.002	0.110	-0.018
S58	0.085	-0.023	-0.672	0.020	-0.499	0.019	0.079	-0.032
S59	0.080	-0.004	-0.380	-0.056	-0.828	0.005	0.077	0.023
S60	0.105	-0.016	-0.710	0.007	-0.721	0.015	0.100	-0.007
S61	0.097	-0.010	-0.667	0.003	-0.681	0.009	0.095	-0.004
S62	0.090	-0.022	-0.634	0.013	-0.584	0.020	0.084	-0.017
S63	0.120	0.008	-0.900	0.003	-0.812	-0.008	0.124	-0.004
S64	0.101	-0.013	-0.798	0.019	-0.603	0.011	0.097	-0.027
S65	0.068	0.020	-0.364	-0.067	-0.708	-0.026	0.074	0.027
S66	0.138	-0.021	-1.114	0.029	-0.811	0.019	0.132	-0.045
S67	0.089	0.017	-0.496	-0.060	-0.856	-0.022	0.095	0.029
S68	0.089	-0.029	-0.669	0.021	-0.537	0.025	0.081	-0.032
S69	0.102	-0.033	-0.762	0.024	-0.609	0.029	0.092	-0.036
S70	0.117	-0.045	-0.874	0.030	-0.698	0.034	0.105	-0.046
S71	0.113	-0.010	-0.820	0.010	-0.746	0.009	0.110	-0.012
S72	0.147	-0.006	-1.168	0.022	-0.913	0.006	0.146	-0.029
S73	0.104	-0.017	-0.670	0.001	-0.756	0.016	0.099	-0.002
S74	0.116	-0.046	-0.801	0.024	-0.738	0.041	0.103	-0.033
S75	0.117	-0.001	-0.940	0.015	-0.729	0.001	0.118	-0.021
S76	0.098	-0.024	-0.723	0.018	-0.619	0.020	0.092	-0.024
S77	0.119	-0.043	-0.868	0.028	-0.722	0.037	0.106	-0.040
S78	0.090	0.030	-0.437	-0.143	-1.083	-0.046	0.100	0.041
S79	0.089	0.032	-0.464	-0.115	-0.993	-0.046	0.099	0.039
S80	0.077	-0.014	-0.668	0.020	-0.427	0.011	0.073	-0.037
S81	0.108	-0.020	-0.814	0.020	-0.666	0.019	0.102	-0.026
S82	0.092	-0.010	-0.551	-0.013	-0.739	0.010	0.088	0.009
S83	0.088	-0.021	-0.584	0.008	-0.602	0.020	0.081	-0.010
S84	0.117	-0.024	-0.975	0.030	-0.659	0.020	0.111	-0.050
S85	0.098	-0.004	-0.686	0.003	-0.674	0.005	0.097	-0.003
S86	0.061	0.017	-0.297	-0.085	-0.707	-0.025	0.066	0.027
S87	0.098	-0.009	-0.763	0.015	-0.615	0.008	0.096	-0.020
S88	0.016	-0.259	-1.375	-0.036	-0.984	-0.124	0.035	-0.333
S89	0.076	-0.876	-2.210	-0.473	-2.824	-0.164	-0.128	-0.520
S90	0.160	0.010	-0.915	-0.170	-1.383	-0.061	0.120	0.158
S91	0.120	0.094	-0.412	-0.005	-0.556	0.029	0.109	0.149
S92	0.028	-0.194	-1.240	-0.121	-0.997	-0.146	0.032	-0.230
S93	0.148	0.161	-0.518	0.123	-0.331	0.094	0.165	0.126
S94	0.045	-0.163	-1.039	-0.050	-0.728	-0.060	0.038	-0.199
S95	0.407	0.806	1.028	0.815	1.396	0.755	0.435	0.720
S96	0.166	0.024	-0.621	0.209	-0.292	0.210	0.145	-0.004
S97	0.254	0.644	0.881	0.323	0.533	0.352	0.296	0.611

S98	0.464	1.149	2.106	0.530	1.198	0.685	0.462	1.282
S99	0.279	0.350	0.372	0.317	0.082	0.409	0.235	0.455
S100	-0.093	-1.151	-2.974	-0.466	-2.753	-0.346	-0.222	-1.049
S101	0.065	0.018	-1.099	0.047	-0.559	-0.078	0.124	-0.126
S102	0.114	-0.034	-0.808	0.020	-0.722	0.030	0.104	-0.027
S103	0.150	-0.052	-0.456	-0.023	-0.868	0.111	0.082	0.092
S104	0.346	1.124	1.132	0.481	1.064	0.362	0.468	1.014
S105	-0.052	-0.370	-2.086	-0.291	-1.551	-0.392	-0.012	-0.491
S106	-0.241	-1.158	-3.798	-0.543	-2.739	-0.689	-0.240	-1.322
S107	-0.084	-0.659	-1.856	-0.374	-1.989	-0.338	-0.132	-0.588
S108	0.055	-0.054	-1.119	-0.172	-1.058	-0.186	0.069	-0.054
S109	-0.154	-0.856	-2.912	-0.772	-2.976	-0.711	-0.197	-0.782
S110	0.125	0.120	-0.340	0.092	-0.248	0.096	0.121	0.128
S111	0.078	-0.231	-1.150	-0.074	-1.054	-0.031	0.041	-0.196
S112	0.168	0.140	-0.200	0.160	-0.232	0.193	0.148	0.178
S113	0.071	-0.140	-1.033	0.043	-0.616	0.009	0.069	-0.206
S114	0.157	-0.091	-1.428	0.227	-0.615	0.134	0.174	-0.245
S115	0.271	0.797	-0.508	0.551	0.755	0.239	0.451	0.400
S116	0.182	0.233	0.098	0.026	-0.503	0.149	0.154	0.332
S117	0.097	-0.023	-0.730	0.019	-0.595	0.019	0.090	-0.027
S118	0.127	-0.012	-0.577	-0.097	-1.358	0.016	0.120	0.036
S119	0.121	0.003	-0.586	-0.100	-1.285	-0.005	0.121	0.039
S120	0.140	-0.011	-0.710	-0.070	-1.345	0.013	0.134	0.033
S121	0.162	-0.019	-1.824	0.059	-0.745	0.013	0.161	-0.177
S122	0.116	-0.032	-0.637	-0.013	-0.958	0.034	0.105	0.008
S123	0.153	0.006	-0.914	-0.043	-1.278	-0.008	0.154	0.028
S124	0.121	-0.026	-0.662	-0.020	-1.013	0.029	0.112	0.012
S125	0.151	0.022	-1.291	0.017	-0.918	-0.021	0.158	-0.025
S126	0.172	0.021	-1.462	0.020	-1.050	-0.020	0.179	-0.030
S127	0.153	0.023	-1.321	0.018	-0.925	-0.022	0.161	-0.027
S128	0.105	-0.033	-0.467	-0.048	-1.063	0.041	0.092	0.020
S129	0.216	-0.006	-2.775	0.084	-0.906	0.003	0.220	-0.323
S130	0.123	-0.004	-0.720	-0.031	-1.034	0.003	0.121	0.020

Figs. 42 to 57 show the values of  $A(1,2,3,4)$  and  $A(1,3,5,7)$  and their differences

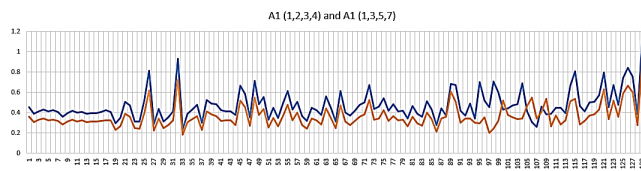


Fig. 42:  $A1(1,2,3,4)$  and  $A1(1,3,5,7)$

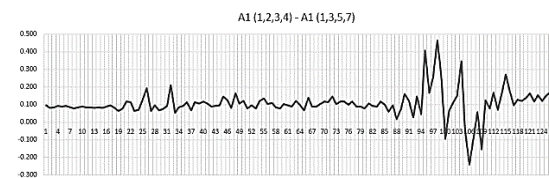


Fig. 43:  $A1(1,2,3,4) - A1(1,3,5,7)$

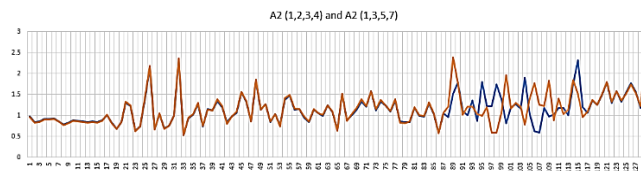


Fig. 44:  $A2(1,2,3,4)$  and  $A2(1,3,5,7)$

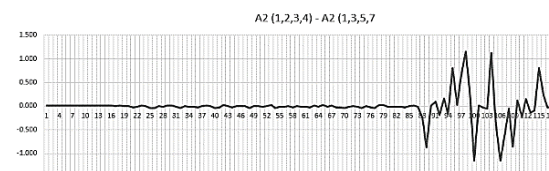


Fig. 45:  $A2(1,2,3,4) - A2(1,3,5,7)$

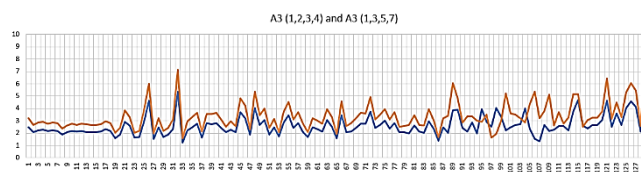


Fig. 46:  $A3(1,2,3,4)$  and  $A3(1,3,5,7)$

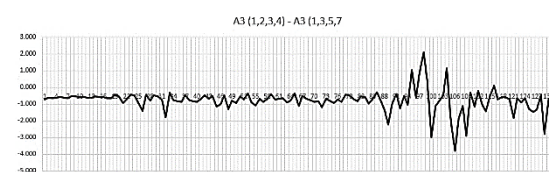


Fig. 47:  $A3(1,2,3,4) - A3(1,3,5,7)$

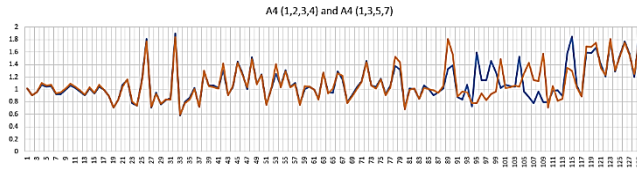


Fig. 48: A4(1,2,3,4) and A4(1,3,5,7)

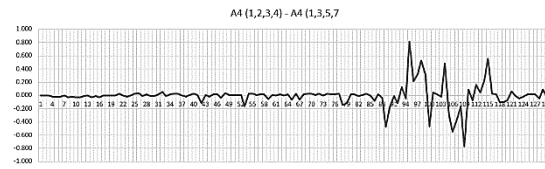


Fig. 49: A4(1,2,3,4) – A4(1,3,5,7)

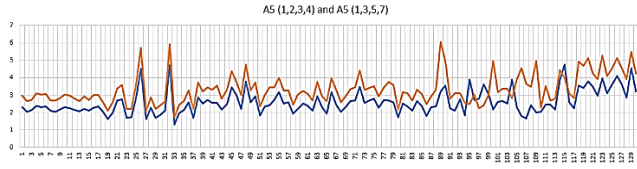


Fig. 50: A5(1,2,3,4) and A5(1,3,5,7)

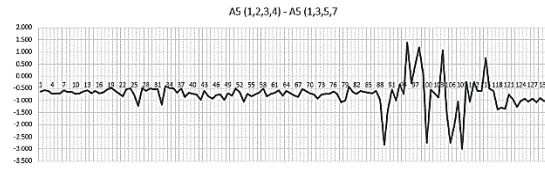


Fig. 51: A5(1,2,3,4) – A5(1,3,5,7)

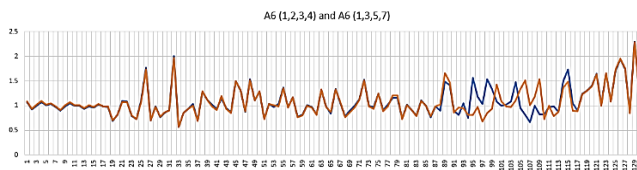


Fig. 52: A6(1,2,3,4) and A6(1,3,5,7)

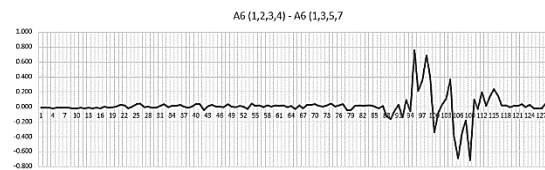


Fig. 53: A6(1,2,3,4) – A6(1,3,5,7)

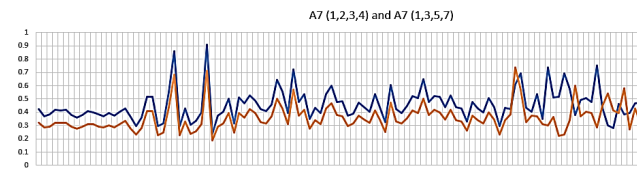


Fig. 54: A7(1,2,3,4) and A7(1,3,5,7)

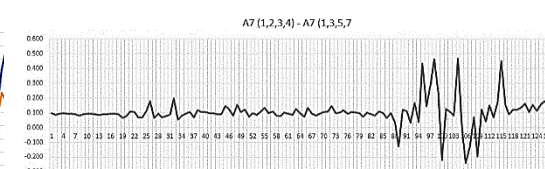


Fig. 55: A7(1,2,3,4) – A7(1,3,5,7)

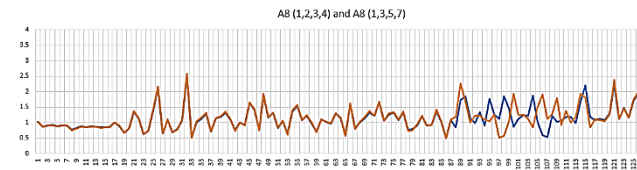


Fig. 56: A8(1,2,3,4) and A8(1,3,5,7)

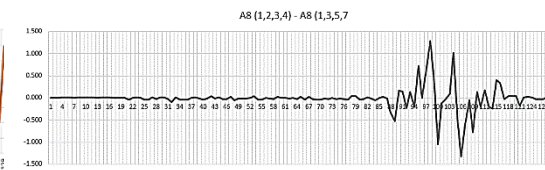


Fig. 57: A8(1,2,3,4) – A8(1,3,5,7)

With reference to table 42 and Figs. 42 to 57, one would observe that virus sequences S26 and S32 show considerable errors in their GR values  $A1(1,2,3,4) - A1(1,3,5,7)$ ,  $A3(1,2,3,4) - A3(1,3,5,7)$ ,  $A5(1,2,3,4) - A5(1,3,5,7)$  and  $A7(1,2,3,4) - A7(1,3,5,7)$ . These error values are given below highlighted in red color.

S26	0.193	-0.048	-1.392	0.034	-1.223	0.044	0.179	-0.044
S32	0.210	-0.042	-1.788	0.055	-1.169	0.034	0.199	-0.098

Once again with reference to table 42 and Figs. 42 to 57, one would observe that many of the virus sequences in the band S88 and S130 show considerable errors in their GR values. The gray colored rows in table 43 do not show considerable errors, whereas the remaining sequences exhibit large errors.

Table 43: Maximum error values obtained using the formula  $A(1,2,3,4) - A(1,3,5,7)$

Sequences	$A(1,2,3,4) - A(1,3,5,7)$							
	A1	A2	A3	A4	A5	A6	A7	A8
S88	0.016	-0.259	-1.375	-0.036	-0.984	-0.124	0.035	-0.333
S89	0.076	-0.876	-2.21	-0.473	-2.824	-0.164	-0.128	-0.52
S90	0.16	0.01	-0.915	-0.17	-1.383	-0.061	0.12	0.158
S91	0.12	0.094	-0.412	-0.005	-0.556	0.029	0.109	0.149
S92	0.028	-0.194	-1.24	-0.121	-0.997	-0.146	0.032	-0.23
S93	0.148	0.161	-0.518	0.123	-0.331	0.094	0.165	0.126
S94	0.045	-0.163	-1.039	-0.05	-0.728	-0.06	0.038	-0.199
S95	0.407	0.806	1.028	0.815	1.396	0.755	0.435	0.72

S96	0.166	0.024	-0.621	0.209	-0.292	0.21	0.145	-0.004
S97	0.254	0.644	0.881	0.323	0.533	0.352	0.296	0.611
S98	0.464	1.149	2.106	0.53	1.198	0.685	0.462	1.282
S99	0.279	0.35	0.372	0.317	0.082	0.409	0.235	0.455
S100	-0.093	-1.151	-2.974	-0.466	-2.753	-0.346	-0.222	-1.049
S101	0.065	0.018	-1.099	0.047	-0.559	-0.078	0.124	-0.126
S102	0.114	-0.034	-0.808	0.02	-0.722	0.03	0.104	-0.027
S103	0.15	-0.052	-0.456	-0.023	-0.868	0.111	0.082	0.092
S104	0.346	1.124	1.132	0.481	1.064	0.362	0.468	1.014
S105	-0.052	-0.37	-2.086	-0.291	-1.551	-0.392	-0.012	-0.491
S106	-0.241	-1.158	-3.798	-0.543	-2.739	-0.689	-0.24	-1.322
S107	-0.084	-0.659	-1.856	-0.374	-1.989	-0.338	-0.132	-0.588
S108	0.055	-0.054	-1.119	-0.172	-1.058	-0.186	0.069	-0.054
S109	-0.154	-0.856	-2.912	-0.772	-2.976	-0.711	-0.197	-0.782
S110	0.125	0.12	-0.34	0.092	-0.248	0.096	0.121	0.128
S111	0.078	-0.231	-1.15	-0.074	-1.054	-0.031	0.041	-0.196
S112	0.168	0.14	-0.2	0.16	-0.232	0.193	0.148	0.178
S113	0.071	-0.14	-1.033	0.043	-0.616	0.009	0.069	-0.206
S114	0.157	-0.091	-1.428	0.227	-0.615	0.134	0.174	-0.245
S115	0.271	0.797	-0.508	0.551	0.755	0.239	0.451	0.4
S116	0.182	0.233	0.098	0.026	-0.503	0.149	0.154	0.332
S117	0.097	-0.023	-0.73	0.019	-0.595	0.019	0.09	-0.027
S118	0.127	-0.012	-0.577	-0.097	-1.358	0.016	0.12	0.036
S119	0.121	0.003	-0.586	-0.1	-1.285	-0.005	0.121	0.039
S120	0.14	-0.011	-0.71	-0.07	-1.345	0.013	0.134	0.033
S121	0.162	-0.019	-1.824	0.059	-0.745	0.013	0.161	-0.177
S122	0.116	-0.032	-0.637	-0.013	-0.958	0.034	0.105	0.008
S123	0.153	0.006	-0.914	-0.043	-1.278	-0.008	0.154	0.028
S124	0.121	-0.026	-0.662	-0.02	-1.013	0.029	0.112	0.012
S125	0.151	0.022	-1.291	0.017	-0.918	-0.021	0.158	-0.025
S126	0.172	0.021	-1.462	0.02	-1.05	-0.02	0.179	-0.03
S127	0.153	0.023	-1.321	0.018	-0.925	-0.022	0.161	-0.027
S128	0.105	-0.033	-0.467	-0.048	-1.063	0.041	0.092	0.02
S129	0.216	-0.006	-2.775	0.084	-0.906	0.003	0.22	-0.323
S130	0.123	-0.004	-0.72	-0.031	-1.034	0.003	0.121	0.02

Figs. 58 to 65 show the error values of 43 virus sequences from A88(1,2,3,4) - A88(1,2,3,4) to A130(1,2,3,4) - A130(1,2,3,4).

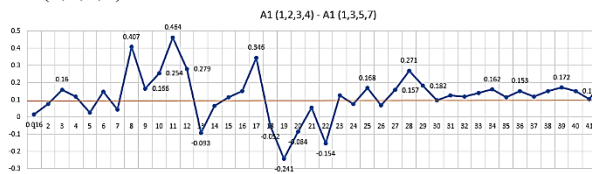


Fig. 58: A1(1,2,3,4) – A1(1,3,5,7)

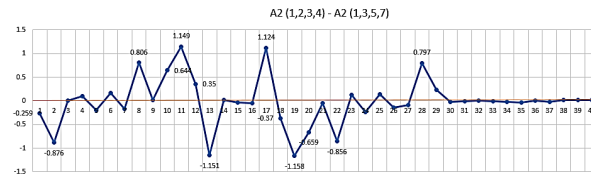


Fig. 59: A2(1,2,3,4) – A2(1,3,5,7)

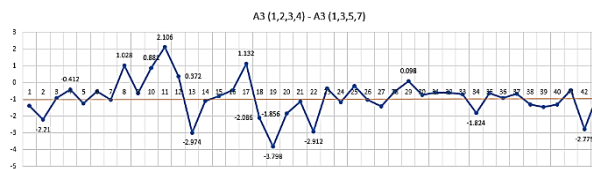


Fig. 60: A3(1,2,3,4) – A3(1,3,5,7)

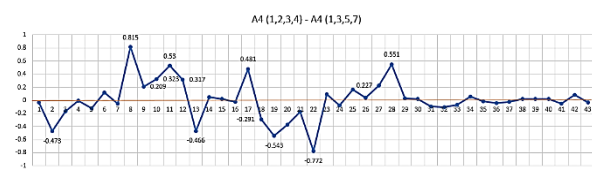


Fig. 61: A4(1,2,3,4) – A4(1,3,5,7)

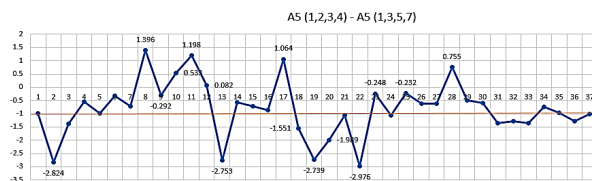


Fig. 62: A5(1,2,3,4) – A5(1,3,5,7)

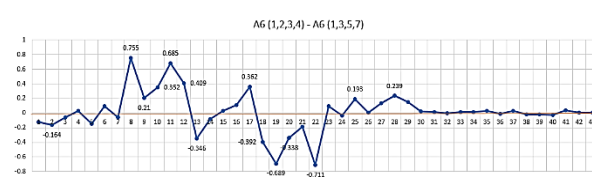


Fig. 63: A6(1,2,3,4) – A6(1,3,5,7)

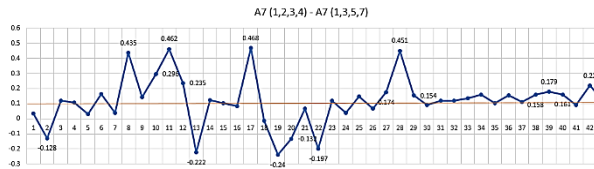


Fig. 64: A7(1,2,3,4) – A7(1,3,5,7)

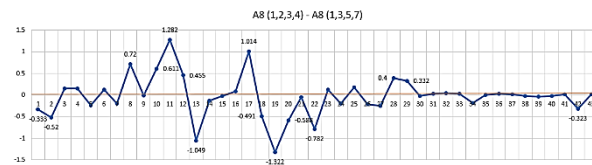


Fig. 65: A8(1,2,3,4) – A8(1,3,5,7)

**Observations**

1. Among the 43 sequences, the following 17 sequences do not show wide differences:  
S92, S93, S94, S101, S102, S103, S111, S113, S117, S118, S119, S120, S122, S124, S125, S128, S130
2. Among the 43 sequences, the following 26 sequences show wide differences:  
S88, S89, S90, S91, S95, S96, S97, S98, S99, S100, S104, S105, S106, S107, S108, S109, S110, S112, S114, S115, S116, S121, S123, S126, S127, S129
3. Among the remaining 87 sequences, the following 2 sequences show wide differences  
S26, S32.
4. Among the 130 sequences, the following 85 sequences do not show any difference:  
S1, S2, S3, S4, S5, S6, S7, S8, S9, S10, S11, S12, S13, S14, S15, S16, S17, S18, S19, S20, S21, S22, S23, S24, S25, S27, S28, S29, S30, S31, S33, S34, S35, S36, S37, S38, S39, S40, S41, S42, S43, S44, S45, S46, S47, S48, S49, S50, S51, S52, S53, S54, S55, S56, S57, S58, S59, S60, S61, S62, S63, S64, S65, S66, S67, S68, S69, S70, S71, S72, S73, S74, S75, S76, S77, S78, S79, S80, S81, S82, S83, S84, S85, S86, S87

Based on above observations on GR values, one could classify all 130 virus sequences in four categories. Table 44 presents these four categories.

Table 44: Classification of 130 Virus Sequences Based on JCP Golden Ratios

Category	Virus Sequences	Remarks
1	S1, S2, S3, S4, S5, S6, S7, S8, S9, S10, S11, S12, S13, S14, S15, S16, S17, S18, S19, S20, S21, S22, S23, S24, S25, S27, S28, S29, S30, S31, S33, S34, S35, S36, S37, S38, S39, S40, S41, S42, S43, S44, S45, S46, S47, S48, S49, S50, S51, S52, S53, S54, S55,	All these 85 sequences satisfy Golden Ratio Property (GRP). Different value assignments to nucleotides have not affected GRP. This means that these virus sequences have not changed their structures due to value assignments and hence
	S56, S57, S58, S59, S60, S61, S62, S63, S64, S65, S66, S67, S68, S69, S70, S71, S72, S73, S74, S75, S76, S77, S78, S79, S80, S81, S82, S83, S84, S85, S86, S87	<b>mutation possibility among these 85 sequences is very remote.</b>
2	S88, S89, S90, S91, S95, S96, S97, S98, S99, S100, S104, S105, S106, S107, S108, S109, S110, S112, S114, S115, S116, S121, S123, S126, S127, S129	All these 26 sequences do not satisfy Golden Ratio Property (GRP). Different value assignments to nucleotides have affected GRP. This means that these virus sequences have changed their structures due to value assignments and hence <b>mutation possibility among these 26 sequences is likely.</b>
3	S92, S93, S94, S101, S102, S103, S111, S113, S117, S118, S119, S120, S122, S124, S125, S128, S130	All these 17 sequences do not satisfy Golden Ratio Property (GRP). Different value assignments to nucleotides have affected GRP. This means that these virus sequences have changed their structures due to value assignments and hence <b>mutation possibility among these 17 sequences is likely.</b>
4	S26, S32	Both of these 2 sequences do not satisfy Golden Ratio Property (GRP). Different value assignments to nucleotides have affected GRP. This means that these virus sequences have changed their structures due to value assignments and hence <b>mutation possibility among these 2 sequences is likely.</b>

**Final Observation**

From the results presented above, it is clear that one could classify any given set of virus genomes based on (i) Auto Correlation Measures, (ii) Percentage Concentration of Nucleotides and Pairwise Correlation Coefficients and Golden Ratio (GR) values of the sequences.

**Inference**

After analyzing more than 70,000 virus genome sequences provided in the NCBI web site, most of the patterns of results so far obtained indicate the possibility of creating any desired virus apart from natural ones and also the possibility of creating engineered versions of natural viruses – *A Subjective Understanding by the Research Team*

## V. SONIC ATTACK – A FEASIBLE SOLUTION TO BREAK THE STRUCTURE OF A VIRUS GENOME

DNA molecule is the most evolved and most complex molecule created by nature and its primary role is long-term storage of genetic information. **Genetic modification could be carried out under the influence of acoustic, electromagnetic, and scalar waves, meaning genetic code of a DNA could be read and possibly rewritten.**

### 5.1 Natural Frequencies of DNA

Mobin Marvi & Majid Ghadiri have proposed ‘A Mathematical model for Vibration Behavior Analysis of DNA and Using a Resonant frequency of DNA for Genome engineering’ ([www.nature.com/scientificreports](http://www.nature.com/scientificreports)). In this model, the mass of the nucleobases in the DNA structure, the effects of the fluid surrounding the DNA (nucleoplasm) and the effects of temperature changes are also considered. A set of governing equations have been derived from Hamilton’s principle and by solving these equations using the generalized differential quadrature method (GDQM), the frequency and mode shape of the DNA has been obtained. An interesting but important topic called “**wave genome**” has been mooted by some Russian scientists, which states that DNA could be affected by acoustic, electromagnetic, and scalar waves. **Under the influence of these waves, the genetic code can be read or rewritten.** Konstantin Meyl has adapted the scalar waves described by ‘Nicola Tesla’ to biology and proposed a relationship between the scalar waves and DNA. Braddon et al investigated the manipulability of DNA from human emotions also. **Rein and Mccraty studied the impact of music on the DNA. Another study was carried out on the effect of sound waves on the synthesis and genes of chrysanthemum.** Peter Garjajev and his research group proved that DNA can be reprogrammed by words and using the correct resonant frequencies of DNA. Russian quantum biologist Poponin tried to prove that human DNA has a direct effect on the physical world and he also found out that DNA can cause disturbing patterns in the vacuum, thus producing magnetized microscopic wormholes. Nobel Laureate Luc Montagnier known for his study on HIV and AIDS, has demonstrated that DNA can be generated by teleportation through quantum imprint and also showed that DNA emits electromagnetic signals that teleport the DNA to other places, such as water molecules, a shocking revelation. Mobin Marvi & Majid Ghadiri has advocated the PBD (Peyrard-Bishop-Dauxois) model for describing DNA. Fig. 66 shows the shape of a DNA before reaching its resonant frequency and its shape distortions after reaching its resonant frequency.

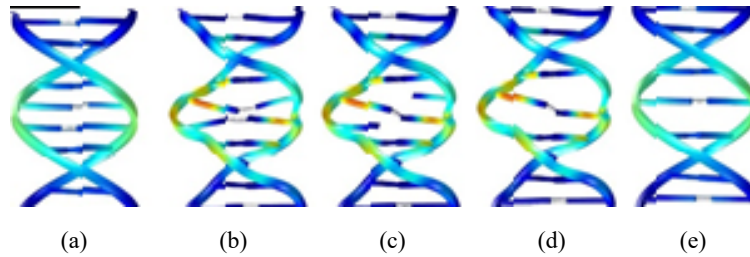


Fig. 66: Original DNA and engineered DNA

#### Legend:

- (a) Shape of DNA before reaching resonant frequency
- (b) Shape after reaching resonant frequency
- (c) Formation of hydrogen bond between a nucleobase to adjacent nucleobase
- (d) Removal of certain nucleobases
- (e) New shape for the DNA – DNA engineered

The mathematical model for vibration behavior analysis of DNA and using a resonant frequency of DNA due to Mobin Marvi and Majid Ghadiri is given by two governing equations of DNA considering the effects of the cellular fluid and temperature change. Appropriate details about the formulation of these two governing equations could be obtained from [www.nature.com/scientificreports](http://www.nature.com/scientificreports).

$$\begin{aligned}
 & \left( \begin{matrix} \mathbf{e}_{01}^T \\ \mathbf{e}_{11} \\ \mathbf{e}_{21} \end{matrix} \begin{matrix} \mathbf{Y}_{e_0A} \\ \mathbf{Y}_{e_1I_2} \\ \mathbf{Y}_{e_2I_1} \end{matrix} - \begin{matrix} \mathbf{e}_{02}^T \\ \mathbf{e}_{12} \\ \mathbf{e}_{22} \end{matrix} \begin{matrix} \mathbf{Y}_{e_0A} \\ \mathbf{Y}_{e_1I_2} \\ \mathbf{Y}_{e_2I_1} \end{matrix} - (e_0a)^2 \left( \begin{matrix} \mathbf{e}_{01}^T \\ \mathbf{e}_{11} \\ \mathbf{e}_{21} \end{matrix} \begin{matrix} \mathbf{Y}_{e_0A} \\ \mathbf{Y}_{e_1I_2} \\ \mathbf{Y}_{e_2I_1} \end{matrix} + 2 \begin{matrix} \mathbf{e}'_{01} \\ \mathbf{e}'_{11} \\ \mathbf{e}'_{21} \end{matrix} \begin{matrix} \mathbf{Y}_{e_0A} \\ \mathbf{Y}_{e_1I_2} \\ \mathbf{Y}_{e_2I_1} \end{matrix} \right) \begin{matrix} \mathbf{r}_1 \\ \mathbf{r}_1 \end{matrix} \right) \\
 & - \left( \begin{matrix} \mathbf{e}_{02}^T \\ \mathbf{e}_{12} \\ \mathbf{e}_{22} \end{matrix} \begin{matrix} \mathbf{Y}_{e_0A} \\ \mathbf{Y}_{e_1I_2} \\ \mathbf{Y}_{e_2I_1} \end{matrix} + 2(e_0a)^2 \begin{matrix} \mathbf{e}'_{01} \\ \mathbf{e}'_{11} \\ \mathbf{e}'_{21} \end{matrix} \begin{matrix} \mathbf{Y}_{e_0A} \\ \mathbf{Y}_{e_1I_2} \\ \mathbf{Y}_{e_2I_1} \end{matrix} \right) \begin{matrix} \mathbf{r}'_1 \\ \mathbf{r}'_1 \end{matrix} \\
 & + (\mathbf{K}(\mathbf{r}_1 - \mathbf{r}_2) - (e_0a)^2 \mathbf{K}(\mathbf{r}'_1 - \mathbf{r}'_2)) \\
 & - ((\mathbf{O}_2^T \mathbf{Y}_{A_s_2} \alpha \Delta \mathbf{T} \mathbf{O}_2) \mathbf{r}_1 - (e_0a)^2 (\mathbf{O}_2^T \mathbf{Y}_{A_s_2} \alpha \Delta \mathbf{T} \mathbf{O}_2) \mathbf{r}'_2) \\
 & - \rho \omega^2 (\mathbf{r}_1 \mathbf{A} - (e_0a)^2 \mathbf{r}'_1 \mathbf{A}) - m_K \omega^2 (\mathbf{r}_1 - (e_0a)^2 \mathbf{r}'_1) \\
 & + \omega (\mathbf{C}_a (\mathbf{r}_1 - \mathbf{r}_2) - (e_0a)^2 \mathbf{C}_a (\mathbf{r}'_1 - \mathbf{r}'_2)) \\
 & + \omega (\rho_n \mathbf{r}_1 (\mathbf{A}_{s_1} + \mathbf{A}_{s_2}) - (e_0a)^2 \rho_n \mathbf{r}'_1 (\mathbf{A}_{s_1} + \mathbf{A}_{s_2})) = 0
 \end{aligned}$$

... Equation #1

$$\begin{aligned}
 & \left( \begin{matrix} \mathbf{e}_{01}^T \\ \mathbf{e}_{11} \\ \mathbf{e}_{21} \end{matrix} \begin{matrix} \mathbf{Y}_{e_0A} \\ \mathbf{Y}_{e_1I_2} \\ \mathbf{Y}_{e_2I_1} \end{matrix} - \begin{matrix} \mathbf{e}_{02}^T \\ \mathbf{e}_{12} \\ \mathbf{e}_{22} \end{matrix} \begin{matrix} \mathbf{Y}_{e_0A} \\ \mathbf{Y}_{e_1I_2} \\ \mathbf{Y}_{e_2I_1} \end{matrix} - (e_0a)^2 \left( \begin{matrix} \mathbf{e}_{01}^T \\ \mathbf{e}_{11} \\ \mathbf{e}_{21} \end{matrix} \begin{matrix} \mathbf{Y}_{e_0A} \\ \mathbf{Y}_{e_1I_2} \\ \mathbf{Y}_{e_2I_1} \end{matrix} + 2 \begin{matrix} \mathbf{e}'_{01} \\ \mathbf{e}'_{11} \\ \mathbf{e}'_{21} \end{matrix} \begin{matrix} \mathbf{Y}_{e_0A} \\ \mathbf{Y}_{e_1I_2} \\ \mathbf{Y}_{e_2I_1} \end{matrix} \right) \begin{matrix} \mathbf{r}_2 \\ \mathbf{r}_2 \end{matrix} \right) \\
 & - \left( \begin{matrix} \mathbf{e}_{02}^T \\ \mathbf{e}_{12} \\ \mathbf{e}_{22} \end{matrix} \begin{matrix} \mathbf{Y}_{e_0A} \\ \mathbf{Y}_{e_1I_2} \\ \mathbf{Y}_{e_2I_1} \end{matrix} + 2(e_0a)^2 \begin{matrix} \mathbf{e}'_{01} \\ \mathbf{e}'_{11} \\ \mathbf{e}'_{21} \end{matrix} \begin{matrix} \mathbf{Y}_{e_0A} \\ \mathbf{Y}_{e_1I_2} \\ \mathbf{Y}_{e_2I_1} \end{matrix} \right) \begin{matrix} \mathbf{r}'_2 \\ \mathbf{r}'_2 \end{matrix} \\
 & - (\mathbf{K}(\mathbf{r}_1 - \mathbf{r}_2) - (e_0a)^2 \mathbf{K}(\mathbf{r}'_1 - \mathbf{r}'_2)) \\
 & - ((\mathbf{O}_1^T \mathbf{Y}_{A_s_1} \alpha \Delta \mathbf{T} \mathbf{O}_1) \mathbf{r}_2 - (e_0a)^2 (\mathbf{O}_1^T \mathbf{Y}_{A_s_1} \alpha \Delta \mathbf{T} \mathbf{O}_1) \mathbf{r}'_2) \\
 & - \rho \omega^2 (\mathbf{r}_2 \mathbf{A} - (e_0a)^2 \mathbf{r}'_2 \mathbf{A}) - m_K \omega^2 (\mathbf{r}_2 - (e_0a)^2 \mathbf{r}'_2) \\
 & - \omega (\mathbf{C}_a (\mathbf{r}_1 - \mathbf{r}_2) - (e_0a)^2 \mathbf{C}_a (\mathbf{r}'_1 - \mathbf{r}'_2)) \\
 & + \omega (\rho_n \mathbf{r}_2 (\mathbf{A}_{s_1} + \mathbf{A}_{s_2}) - (e_0a)^2 \rho_n \mathbf{r}'_2 (\mathbf{A}_{s_1} + \mathbf{A}_{s_2})) = 0
 \end{aligned}$$

... Equation #2

Based on these governing equations, Mobin Marvi and Majid Ghadiri have evaluated the mechanical and geometrical properties of DNA, which are presented in Fig. 67.

Mechanical or thermal properties	Symbol	Quantity	Reference
Young modulus	$E$	0.3 (GPa)	
Poisson's ratio	$\nu$	0.5	
Shear modulus	$G$	0.1 (GPa)	$G = \frac{E}{2(1+\nu)}$
Density	$\rho$	1.7 (g/m <sup>3</sup> )	
Mass of Adenine	$M$	226 * 10 <sup>-27</sup> (Kg)	
Mass of Thymine	$M$	211 * 10 <sup>-27</sup> (Kg)	
Mass of Guanine	$M$	252 * 10 <sup>-27</sup> (Kg)	
Mass of Cytosine	$M$	185 * 10 <sup>-27</sup> (Kg)	
Hydrogen bond strength	Between Adenine-Thymine ( $k_{AT}$ )	19.5 (N/m)	
	Between Guanine-Cytosine ( $k_{GC}$ )	56.3 (N/m)	
Damper constant	$C$	0.05 (N s/m)	
Density of nucleoplasm	$\rho_n$	0.14 (g/cm <sup>3</sup> )	
Viscosity of nucleoplasm	$\mu_n$	1.35 (cP) or 0.135 * 10 <sup>-2</sup> ( $\frac{N \cdot s}{m^2}$ )	
Osmotic pressure	$P$	4 (atm)	
Thermal conductivity	$k_t$	$k = 150 \frac{W}{mK}$	
Specific heat capacity	$c_p$	$C_p = 40 \frac{cal}{molK}$ or $2.56 * 10^{-4} \frac{KJ}{gK}$	
Thermal diffusivity	$\alpha$	$\alpha = 3.44 * 10^{-8} \frac{m^2}{s}$ $\alpha = \frac{k}{\rho c_p}$	

Geometry	Quantity	Reference
Pitch/turn of helix	34 (Å)	
Rise/bp along axis	3.4 (Å)	
Radius	10 (Å)	
Rotation/bp	34.3°	
Curvature	0.008 * 10 <sup>10</sup>	$\kappa = \frac{radius}{\sqrt{radius^2 + pitch^2}}$
Twist	0.027 * 10 <sup>10</sup>	$\tau = \frac{pitch}{\sqrt{radius^2 + pitch^2}}$
Straight and open length	7.14 * 10 <sup>-9</sup>	$L = \sqrt{(2\pi r)^2 + (pitch)^2}$
Length of Adenine	2 * 5.8 (Å)	
Length of Thymine	2 * 4.8 (Å)	
Length of Guanine	2 * 5.7 (Å)	
Length of Cytosine	2 * 4.7 (Å)	
Distance between Adenine-Thymine	2.83 (Å)	

Thickness	Width
0.98 (Å)	3.17 (Å)

Cross-section of DNA strands

Adenine	Thymine	Guanine	Cytosine
0.53	0.5	0.55	0.45

Sectional radius of nucleobases (Å)

Fig. 67: Mechanical and geometrical properties of DNA

Courtesy: www.nature.com/scientificreports

Moreover, the authors have estimated the natural frequencies of DNA in terms of its length. The estimated values are shown in Fig. 68.

Number of Nucleobases	Number of turn of helix	Natural frequency (Hz)		
		First mode	Second mode	Third mode
10	1	5.0198 * 10 <sup>9</sup>	7.0444 * 10 <sup>9</sup>	9.6939 * 10 <sup>9</sup>
20	2	1.4375 * 10 <sup>9</sup>	2.6497 * 10 <sup>9</sup>	3.8139 * 10 <sup>9</sup>
30	3	5.2032 * 10 <sup>8</sup>	1.1533 * 10 <sup>9</sup>	1.8835 * 10 <sup>9</sup>
40	4	6.6750 * 10 <sup>8</sup>	6.8998 * 10 <sup>8</sup>	1.3232 * 10 <sup>9</sup>
50	5	5.3081 * 10 <sup>8</sup>	5.4125 * 10 <sup>8</sup>	1.0504 * 10 <sup>9</sup>
60	6	4.3337 * 10 <sup>8</sup>	4.5284 * 10 <sup>8</sup>	8.6065 * 10 <sup>8</sup>
70	7	3.7059 * 10 <sup>8</sup>	3.8425 * 10 <sup>8</sup>	7.3720 * 10 <sup>8</sup>
80	8	3.2702 * 10 <sup>8</sup>	3.3004 * 10 <sup>8</sup>	6.5024 * 10 <sup>8</sup>

Fig. 68: Variations in natural frequencies with respect to DNA lengths

Based on the details given above, one would wonder whether it is possible to develop a formal method of dealing with Corona viruses using sonic frequencies with the idea of causing a physical disruption in the virus structure. In this context, this paper first points out to a work carried out by Mark D. Temple, for generating sonic frequencies artificially in order to express Corona viruses. Subsequently, a novel technique for artificially generating sonic bursts is proposed for the purpose of physically dismembering Corona viruses, for that matter any real time viral genome.

### 5.2 Sonic frequencies artificially generated to express corona viruses

“Mark D. Temple describes in his paper (<https://bmcbioinformatics.biomedcentral.com/articles/10.1186/s12859-020-03760-7>), a real-time audio and visual display of the Coronavirus genome wherein he uses a combination of sonification and an animated display to inquire into the SARS-CoV-2 genome. The audio data is generated in real time from a variety of RNA motifs that are known to be important in the functioning of RNA. Additionally, metadata relating to RNA translation and transcription has been used to shape the auditory and visual displays. Together these tools provide a unique approach to further understand the metabolism of the viral RNA genome. This audio signal provides a further means to represent the function of the RNA in addition to traditional written and visual approaches. Sonification of the SARS-CoV-2 genomic RNA sequence results in a complex auditory stream composed of up to 12 individual audio tracks. Each auditory motive is derived from the actual RNA sequence or from metadata. This approach has been used to represent transcription or translation of the viral RNA genome.

Observations:

1. As the DNA length increases, the natural frequency decreases. This need not be true in all cases because the natural frequencies depend on various factors like helical turns, viscosity, density and temperature of the fluid surrounding the RNA.
2. It is expected that the natural frequency associated with a Corona virus RNA could be of the order of kilo hertz.
3. Frequency varies between various modes need not be uniform.

The display highlights the real-time interaction of functional RNA elements. The sonification of codons derived from all three reading frames of the viral RNA sequence in combination with sonified metadata provide the framework for this display. Functional RNA motifs such as transcription regulatory sequences and stem loop regions have also been sonified. Using the tool, audio can be generated in real-time from either genomic or sub-genomic representations of the RNA. Given the large size of the viral genome, a collection of interactive buttons has been provided to navigate to regions of interest, such as cleavage regions in the polyprotein, untranslated regions or each gene. These tools are available through an internet browser and the user can interact with the data display in real time. Each nucleotide generates a note on every beat whereas each di-nucleotide generates a note every second beat. Each codon (in an ORF) generates a note every third beat. Together these notes are syncopated to create a characteristic sound during peptide translation that is distinct from the surrounding untranslated region. Audio from the GC tracks are only triggered when the GC ratio changes by an increment of 0.1. If a note sequence has identical adjacent notes then the length of the note is extended rather than being repeated. This creates space and clarity for other notes layered in the auditory display. Table xxx provides scale degrees and instrumentation used for the RNA features that are sonified". **Reproduced verbatim from <https://bmcbioinformatics.biomedcentral.com/articles/10.1186/s12859-020-03760-7>**

Table 45: Scale degrees and instrumentation of the RNA features being sonified

Sonified motif	Instrument	Pan	Translation Scale Bb aeolian mode		Transcription Scale C Lydian mode		
			Scale degrees		Octave	Scale degrees	Octave
			Nucleotide	Synth	L	1, 3	2, 3
Di-nucleotide	Synth	R	1, 4, 5, 6	1, 2, 3, 4	1, 3, 5	1	
GC Content (10 bp)	AM synth + delay	L	1, 3, 6, 7	2,3	1, 3, 5, 7	4, 5	
GC Content (100 bp)	AM synth + delay	R	1, 3, 6, 7	2, 3	1, 3, 5, 7	4, 5	
3 bp repeat	Synth	L	1, 3	4	1, 4, 5	6	
Codon Frame 1 (translation)	FM synth + distortion	L	1, 3, 4, 5, 7	2, 3, 4, 5	-	-	
Codon Frame 2 (translation)	FM synth + distortion	C	1, 3, 4, 5, 7	2, 3, 4, 5	-	-	
Codon Frame 3 (translation)	FM synth + distortion	R	1, 3, 4, 5, 7	2, 3, 4, 5	-	-	
Tri-nucleotide (transcription)	FM synth + distortion	L	-	-	1, 3, 4, 5, 7	3, 4, 5	
Untranslated regions	AM synth	R	1, 2, 3	5	1, 4, 6, 7	3	
Transcription regulating sequences (TRS)	AM synth	L	1, 2, 4, 5, 6	5	1, 2, 3, 4, 5, 6, 7	6	
Cleavage sites in the polyprotein	AM synth + distortion	L	1, 6, 7	4	1, 2, 3	6	
Stem and loop regions (SL)	AM synth + delay + distortion	R	1, 2, 6, 7	5	1, 4, 5, 7		
Nucleotide	Synth	L	1, 3	2, 3	1, 5	2, 3	
Di-nucleotide	Synth	R	1, 4, 5, 6	1, 2, 3, 4	1, 3, 5	1	
GC Content (10 bp)	AM synth + delay	L	1, 3, 6, 7	2,3	1, 3, 5, 7	4, 5	
GC Content (100 bp)	AM synth + delay	R	1, 3, 6, 7	2, 3	1, 3, 5, 7	4, 5	
3 bp repeat	Synth	L	1, 3	4	1, 4, 5	6	
Codon Frame 1 (translation)	FM synth + distortion	L	1, 3, 4, 5, 7	2, 3, 4, 5	-	-	
Codon Frame 2 (translation)	FM synth + distortion	C	1, 3, 4, 5, 7	2, 3, 4, 5	-	-	
Codon Frame 3 (translation)	FM synth + distortion	R	1, 3, 4, 5, 7	2, 3, 4, 5	-	-	
Tri-nucleotide (transcription)	FM synth + distortion	L	-	-	1, 3, 4, 5, 7	3, 4, 5	
Untranslated regions	AM synth	R	1, 2, 3	5	1, 4, 6, 7	3	
Transcription regulating sequences (TRS)	AM synth	L	1, 2, 4, 5, 6	5	1, 2, 3, 4, 5, 6, 7	6	
Cleavage sites in the polyprotein	AM synth + distortion	L	1, 6, 7	4	1, 2, 3	6	
Stem and loop regions (SL)	AM synth + delay + distortion	R	1, 2, 6, 7	5	1, 4, 5, 7		

(Courtesy: <https://bmcbioinformatics.biomedcentral.com/articles/10.1186/s12859-020-03760-7>),

### 5.3 Sonic frequencies artificially generated to dismember corona viruses

What is intended in this section is to explore the possibility of generating sonic frequencies in order to dismember corona viruses. The basic idea that is used here is that one can assign four of the heptonic scales to the four nucleotides. The seven svaras of the saptak are the fundamentals of heptatonic scales or melakarta ragas and thaats in Carnatic and Hindustani classical music. The seven svaras are Shadja (षड्ज), Rishabh (ऋषभ), Gandhara (गान्धार), Madhyama (मध्यम), Panchama (पंचम), Dhaivata (धैवत) and Nishada (निषाद). The frequencies corresponding to the svaras are determined as shown here. Shadja is realized as the fundamental frequency  $f_s$ , which may be in any suitable scale. Then the remaining svaras are realized as the frequencies given by  $f_r = f_s + 33\text{Hz}$ ,  $f_g = f_s + 69\text{Hz}$ ,  $f_m = f_s + 88\text{Hz}$ ,  $f_p = f_s + 131\text{Hz}$ ,  $f_d = f_s + 179\text{Hz}$ ,  $f_n = f_s + 233\text{Hz}$ , and the Shadja of the next octave  $f_{ss} = f_s + 254\text{Hz}$ . For example, if one chooses 50Hz for  $f_s$ , then  $f_r = 83\text{Hz}$ ,  $f_g = 119\text{Hz}$ ,  $f_m = 138\text{Hz}$ ,  $f_p = 181\text{Hz}$ ,  $f_d = 229\text{Hz}$ ,  $f_n = 283\text{Hz}$ , and the Shadja of the next octave  $f_{ss} = 304\text{Hz}$ . For convenience, four frequencies such as  $f_s$ ,  $f_m$ ,  $f_p$  and  $f_{ss}$  are chosen and allotted to the four nucleotides such as Adenine (A), Thymine (T), Guanine (G) and Cytosine (C) respectively. Table 46 shows frequency allocation to the nucleotides and a sample frequency allocation.

Table 46: Frequency allocation to nucleotides

Nucleotides	A	T	G	C	Nucleotides	A	T	G	C
Frequencies	$f_s$	$f_m$	$f_p$	$f_{ss}$	Sample Frequencies	$f_s = 50\text{Hz}$	$f_m = 138\text{Hz}$	$f_p = 181\text{Hz}$	$f_{ss} = 304\text{Hz}$

For example, let us consider the first 30 nucleotides in S1: >NC\_001639.1 Lactate dehydrogenase-elevating virus, complete genome. Then the frequency allocation for the nucleotides is given in Table 47.

Table 47: Frequency allocation to 30 nucleotides ( $s1$  to  $s30$ ) in the sequence S1: >NC\_001639.1

Sample Number	$s1$	$s2$	$s3$	$s4$	$s5$	$s6$	$s7$	$s8$	$s9$	$s10$
Nucleotide	c	t	a	c	c	c	a	g	g	a
Frequency Allotted	$f_{SS} = 304\text{Hz}$	$f_M = 138\text{Hz}$	$f_S = 50\text{Hz}$	$f_{SS} = 304\text{Hz}$	$f_{SS} = 304\text{Hz}$	$f_{SS} = 304\text{Hz}$	$f_S = 50\text{Hz}$	$f_P = 181\text{Hz}$	$f_P = 181\text{Hz}$	$f_S = 50\text{Hz}$
Sample Number	$s11$	$s12$	$s13$	$s14$	$s15$	$s16$	$s17$	$s18$	$s19$	$s20$
Nucleotide	a	a	a	g	c	c	a	a	c	c
Frequency Allotted	$f_S = 50\text{Hz}$	$f_S = 50\text{Hz}$	$f_S = 50\text{Hz}$	$f_P = 181\text{Hz}$	$f_{SS} = 304\text{Hz}$	$f_{SS} = 304\text{Hz}$	$f_S = 50\text{Hz}$	$f_S = 50\text{Hz}$	$f_{SS} = 304\text{Hz}$	$f_{SS} = 304\text{Hz}$
Sample Number	$s21$	$s22$	$s23$	$s24$	$s25$	$s26$	$s27$	$s28$	$s29$	$s30$
Nucleotide	a	a	c	c	t	c	g	a	t	c
Frequency Allotted	$f_S = 50\text{Hz}$	$f_S = 50\text{Hz}$	$f_{SS} = 304\text{Hz}$	$f_{SS} = 304\text{Hz}$	$f_M = 138\text{Hz}$	$f_{SS} = 304\text{Hz}$	$f_P = 181\text{Hz}$	$f_S = 50\text{Hz}$	$f_M = 138\text{Hz}$	$f_{SS} = 304\text{Hz}$

Fig. 69 shows the pure tones generated programmatically.

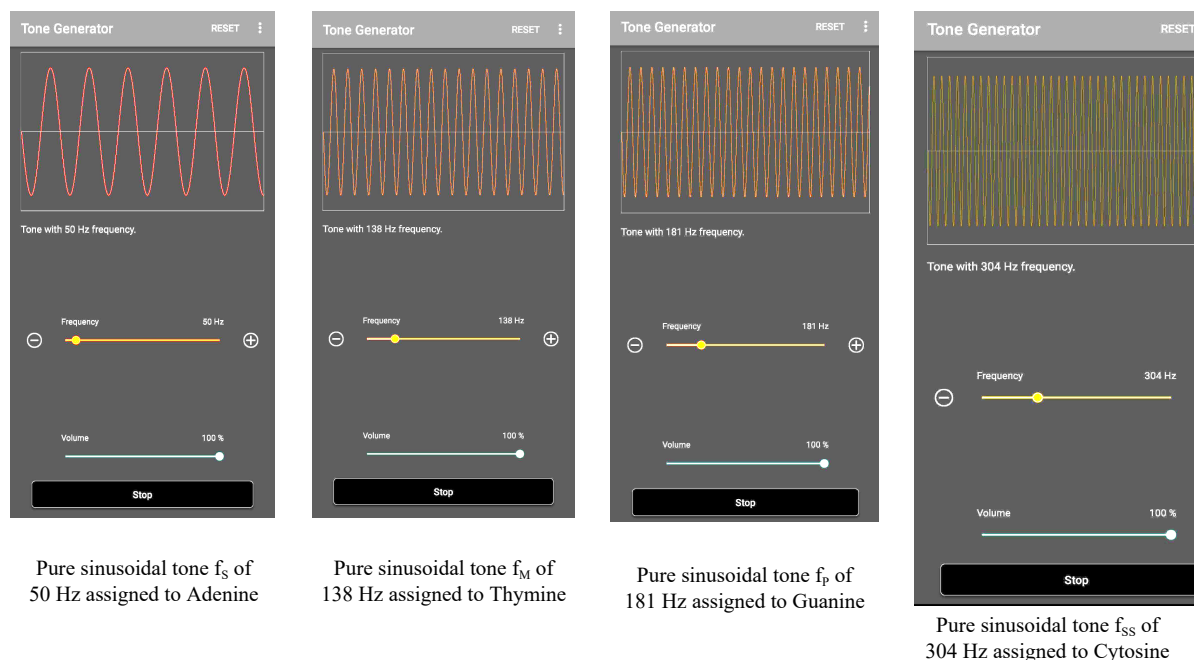


Fig. 69: Pure tones generated programmatically

Now let us consider the first thirty nucleotides in the viral genome S1: >NC\_001639.1 and allot the four frequencies to the respective nucleotides. Fig. 70 shows the audio frequency bursts for first 30 nucleotides of S1: >NC\_001639.1 virus sequence.

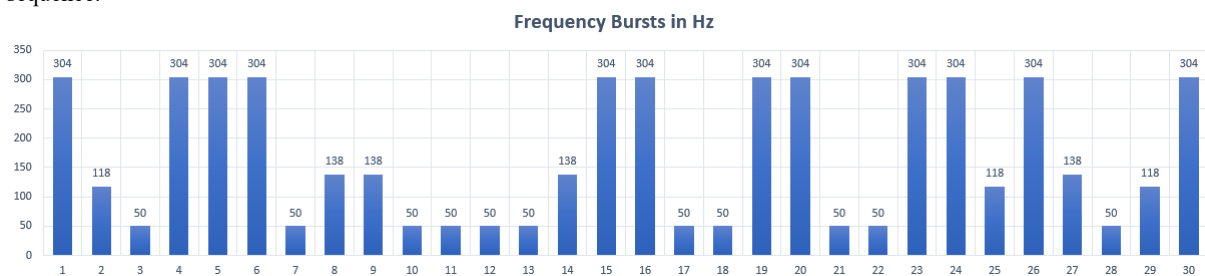


Fig. 70: Audio frequency bursts for first 30 nucleotides of S1: >NC\_001639.1 virus sequence

The waveform of the audio signal is described mathematically as  $s(t) = A(t) \cdot \sin[2\pi f(t) \cdot t + \phi(t)]$ . In our case,  $A(t)$  is a constant and the audio burst is on for time  $\tau$ . The function of time is determined by the pulse repetition frequency and the duty cycle. The frequency spectrum of such a sequence of rectangular pulses is represented by a  $(\sin x)/x$  function. Fig. 71 shows a sequence of rectangular amplitude modulated pulses and its frequency spectrum.

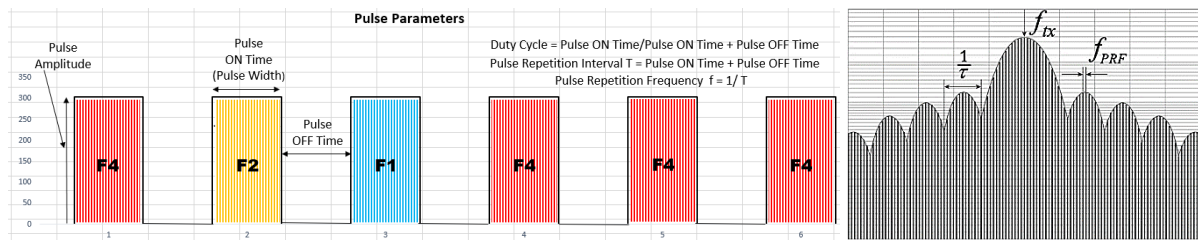


Fig. 71: Frequency spectrum of a sequence of rectangular amplitude modulated pulses

The essential parts of the audio transmission power are in a region  $B_{HF} = 2/\tau$  in the vicinity of the pulse frequency  $f_c$ . The audio burst bandwidth  $B_{HF}$  increases with decreasing pulse width like  $B_{HF} = \tau^{-1}$ . The shortening of the pulses limits the maximum range in the case of simple pulse modulation. Under these conditions, the pulse energy  $E_p$  is increased only by the pulse power  $P_s$  at a required virus target range. For the maximum range, the pulse energy is crucial, and not its pulse power:

$$E_p = P_s \cdot \tau = P_{av} \cdot T = \frac{P_{av}}{f_{PRF}} \quad \text{where} \quad \begin{array}{l} E_p = \text{energy content of the pulse} \\ P_s = \text{transmission pulse power} \\ P_{av} = \text{average power} \end{array}$$

Significant improvements could be achieved with the internal modulation of the transmitted pulse (intra-pulse modulation). The intra-pulse modulation is carried out in this case, using the frequencies  $f_s$ ,  $f_M$ ,  $f_p$  and  $f_{SS}$ . Due to intra-pulse modulation, the frequency spectrum would keep changing and the changes occur as per the nucleotides present in a virus genome. For example, let us consider the meta data of S1: >NC\_001639.1 Lactate dehydrogenase-elevating virus, complete genome. Then the bursts of sonic frequencies generated for the virus metadata mentioned above, are shown below as a sequence of frequency components separated with a timing symbol  $\sim$  that denotes the burst OFF time duration. This pulse OFF time duration could be varied from pulse to pulse, if needed. As outlined earlier, four discrete frequencies of  $f_s = 50\text{Hz}$ ,  $f_M = 138\text{Hz}$ ,  $f_p = 181\text{Hz}$ , and  $f_{SS} = 304\text{Hz}$  corresponding to the nucleotides Adenine, Thymine, Guanine and Cytosine are used here. Each pulse ON time and OFF time could be varied depending on prevailing context. In what is given below, the basic frequency of 50 Hz is assigned to Adenine. However, one can as well change the basic frequency as per the need. In fact, significant improvements have been achieved with this internal pulse modulation using the frequencies  $f_s$ ,  $f_M$ ,  $f_p$  and  $f_{SS}$ . The frequency spectrum changes occur as per the nucleotides present in the target virus genome S1: >NC\_001639.1 and the sequence of frequencies shown above is meant for dismembering it. These frequencies resonate with the natural frequencies of nucleotides and cause considerable physical change in the structure of the entire genome. **The efficacy of this technique of deactivating virus genome by sonic attack has to be checked, real time, by transmitting generated pulses to the actual target genome and by observing the changes in the shape of the genome.** The control parameters are (i) pulse width, (ii) pulse OFF time, (iii) amplitude of pulse and (iv) frequencies embedded inside the pulses.

### 5.3.1 Use of Ultrasound for deactivating virus genomes

What has been shown above is just an example to explain the 'Sonic Attack' technique. In fact, one should avoid audio signals of frequencies less than 10,000Hz because these sounds may cause damage to ears. Moreover, these signals are subharmonic, whereas the actual frequencies to be used to destroy a virus genome of length more than 20,000 bases should be near or ultrasonic frequencies above 10,000Hz. We have found that the base frequency  $f_s = 20,000\text{Hz}$  together with  $f_M = 20,088\text{Hz}$ ,  $f_p = 20,131\text{Hz}$  and  $f_{SS} = 20,254\text{Hz}$  would do a perfect job of dismembering the structure of virus genome. It is to be noted that one would require different frequency sequences for different virus genomes like F1 for S1, F2 for S2, F3 for S3 and so on, up to F130 for S130.

### 5.3.2 Pulse to Pulse Frequency Agile System

Alternatively, one could develop a random frequency sequence generator to produce only four discrete frequencies of  $f_s$ ,  $f_M$ ,  $f_p$  and  $f_{SS}$  and tune the generator to deactivate any real time virus genome. Fig. 72 shows a hypothetical block diagram that is self explanatory as to how to produce random frequency sequences. Apart from this, Phase Shift Keying (PSK) could also be incorporated as a part of intra-pulse modulation scheme to improve efficacy of the system.



























2. Pairwise Spectral Correlation of SARS-Corona Viruses, London Journal of Research in Computer Science and Technology, London Journals Press, Volume 20 | Issue 2 | Compilation 1.0 Year 2020, pp 81-148

In this paper, a set of 130 virus genomes obtained from NCBI web site <https://www.ncbi.nlm.nih.gov/genbank/sars-cov-2-seqs/> has been analyzed and classified based on three quantificational measures (i) Four Level Auto Correlation Measures (ACMs), (ii) Percentage Concentration of Nucleotides and Pairwise Correlation Coefficients and (iii) JCP Golden Ratio (GR) values of the sequences. Moreover, a novel technique called '**Sonic Attack**' has been introduced which is expected to destroy the shape and structure of any kind of virus genome, be it inside the body or in the environment.

SARS-Corona like virus has been described in ancient ayurvedic scriptures about 5,000 years ago in a chapter called 'Krimi' i.e., infections, in the '**Charak Samhita**' that describes an infection exactly like the novel coronavirus. In the chapter 'Krimi', there is mention of a term called '**Sleshma Krimi**', which is described as something that cannot be seen by the naked eye as they are '**Maha Sukshma**' which means they are so minute that one needs a special instrument to see that.

ॐ त्र्यम्बकं यजामहे सुगन्धिं पुष्टिवर्धनम् ।  
उर्वारुकमिव बन्धनान् मृत्योर्मुक्षीय मामृतात् ॥

Om, We Worship the Tryambaka (the Three-Eyed One, Lord Shiva)  
Who is Fragrant (as the Spiritual Essence), Increasing the Nourishment (of our Spiritual Core);  
From these many Bondages (of Samsara) similar to Cucumbers (tied to their Creepers),  
May I be Liberated from Death (Attachment to Perishable Things), So that I am not separated from the perception  
of Immortality (Immortal Essence pervading everywhere)

ओम शांति शांति शांति ॥