



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 20 | 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 20, 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 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



1

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



11

- 1. A Review on Honeypot Deployment
pg. 1-10
 - 2. Pairwise Spatial Correlation of SARS-Corona Viruses
pg. 11-78
 - 3. Future Automobile Market and Role of Lithium Ion...
pg. 79-82
 - 4. A Review on Face Recognition Technique...
pg. 83-89
-



79

- V. London Journals Press Memberships



Scan to know paper details and
author's profile

A Review on Honeypot Deployment

Dr. Apurva Saxena Verma & Dr. Anubha Dubey

Rabindranath Tagore University

ABSTRACT

A honeypot is a source which is proposed to be attacked and cooperated to gain more information about the attacker and its used tools. It can also be installed to draw and divert an attacker from their existent targets. The honeypots are planned and placed in a cloud network to gather the strange presence as well as known occurrences happened in cloud computing. Honeypots are a computer software tool designed to help, learn the purpose and methods of the hacker community. Honeypot describes in-depth the concepts of their involvement in the field of security. The present review paper proposes and designs the importance of honeypot deployment technique in the protection of data based on the in- frastructures of honeypot, its advantages and disadvantages. The authors also try to offer possible solutions for issues that are generated while using honeypot.

Keywords: honeypots, attacker, information, technique.

Classification: C.2.m

Language: English



London
Journals Press

LJP Copyright ID: 975841
Print ISSN: 2514-863X
Online ISSN: 2514-8648

London Journal of Research in Computer Science and Technology



Volume 20 | Issue 1 | Compilation 1.0

© 2020. Dr. Apurva Saxena Verma & Dr. Anubha Dubey. This is a research/review paper, distributed under the terms of the Creative Commons Attribution-Noncom-mercial 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.

A Review on Honeypot Deployment

Dr. Apurva Saxena Verma^α & Dr. Anubha Dubey^σ

ABSTRACT

A honeypot is a source which is proposed to be attacked and cooperated to gain more information about the attacker and its used tools. It can also be installed to draw and divert an attacker from their existent targets. The honeypots are planned and placed in a cloud network to gather the strange presence as well as known occurrences happened in cloud computing. Honeypots are a computer software tool designed to help, learn the purpose and methods of the hacker community. Honeypot describes in-depth the concepts of their involvement in the field of security. The present review paper proposes and designs the importance of honeypot deployment technique in the protection of data based on the infrastructures of honeypot, its advantages and disadvantages. The authors also try to offer possible solutions for issues that are generated while using honeypot.

Keywords: honeypots, attacker, information, technique.

Author α: Research Scholar, Computer Science Engineering, Rabindranath Tagore University, Bhopal, India.

σ: Independent Researcher and analyst, Bioinformatics, MANIT, Bhopal, India.

I. INTRODUCTION

A honeypot is a program, machine, or system that is putting on a network to attract the attackers. The idea is to mislead the intruder by making the honeypot, which seems like a genuine system. Honeypots are naturally virtual machines that follow real machines by simulate running services and open ports, services which one might find on a typical machine on a network. This means, that

a honeypot is expected to get attacked and potentially subjugated [2]. Honeypots do not fix anything. A honeypot is a security resource whose importance lies in being surveyed, attacked. In simple terms, "A honeypot is a software system on the Internet that is definitely set up to attract and "trap" people who endeavor to infiltrate other people's computer systems [3].

Maintaining a honeypot is said to require a substantial amount of attention and may propose as its highest value nothing more than a learning experience. The area of computer security, a honeypot is considered to capture all traffic and movement directed to the system. Honeypots differ from regular network systems in that considerably greater emphasis is placed on logging all activity to the site, either by the honeypot itself or through the use of a network/packet sniffer [3]. It is intended to look like something a trespasser can attack to gain contact to a given system. According to the usage honeypots are classified as Production honeypots and Research honeypots.

- Production honeypots are used to diminish the risks in the business/production environment and thus are mostly deployed in organizations [3].
- Research honeypots are intended to congregate as much information as possible. While research honeypots do not append security value to an organization, but they can help a lot in understanding the attacker's society and their intentions [3].

Every traffic from and to a honeypot is apprehensive because no productive systems positioned on this resource. All data collected by a honeypot is consequently fascinating data. Data

collected by a honeypot is of high assessment and can lead to a better perceptive and acquaintance which in turn can help to amplify overall network security [6]. One can also argue that a honeypot can be used for hindrance because it can discourage attackers from attacking other systems by engaging them lengthy enough and combine their resources. According to their level of involvement, they categorized into three types:

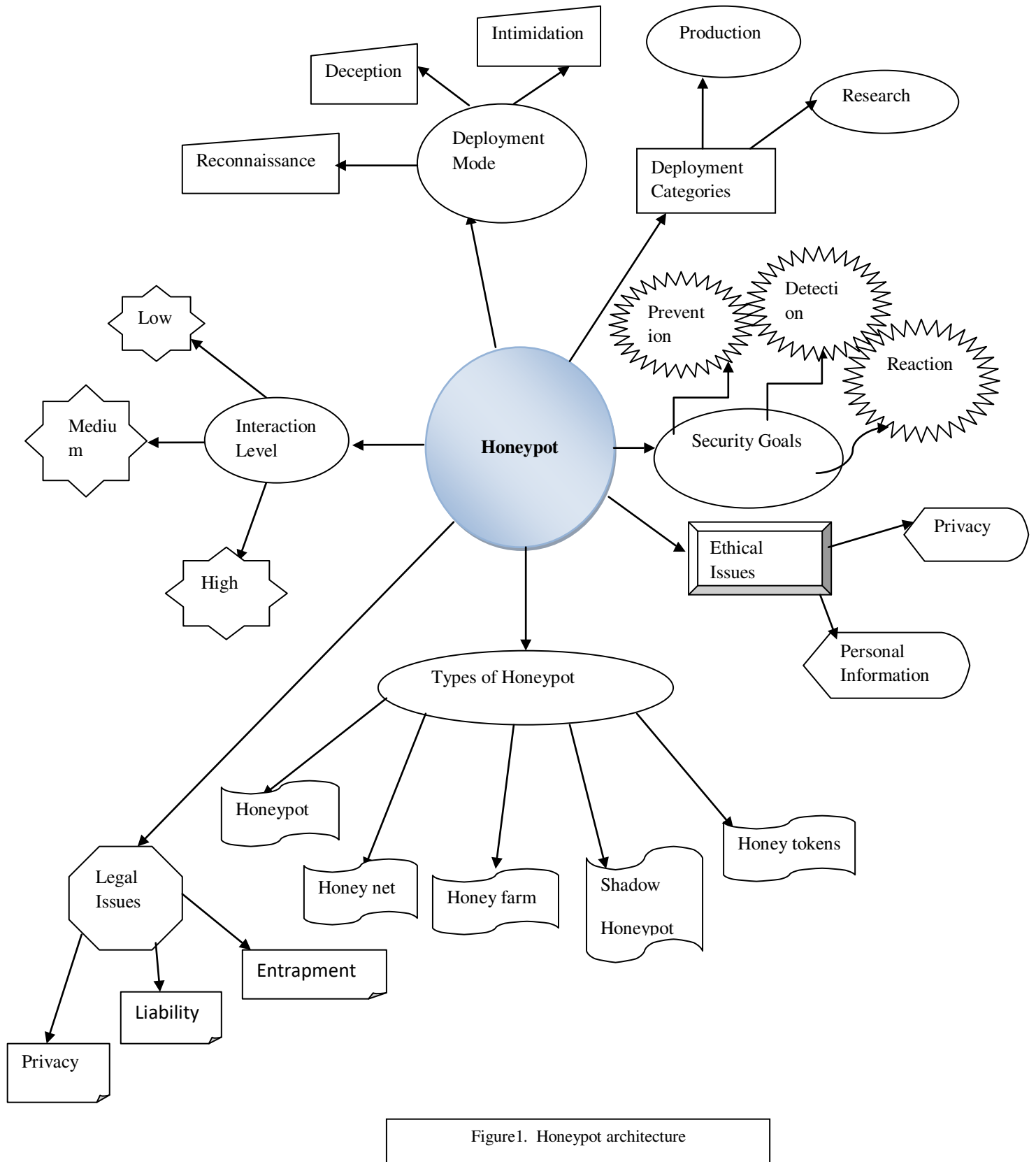
- *Low involvement:* It is the level in which the honeypots imitate simple services, and the freedom given to attackers is least. They are inactive in approach so attackers cannot use them to attack other systems. Thus they are well suitable for organizations, and many production honeypots are low involvement honeypots.
- *Mid involvement:* This honeypot grants more services than low level but doesn't provide a real operating system [3]. The risks also rise with the level of emulation they provide to attackers.
- *High involvement:* This honeypot gives an excellent operating system to attack. Expose the system to sufficient risk and complexity. At the same time, the possibility to build up information about the attack as well as the magnetism of the honeypot increases a lot, so they primarily used for research purposes.

Figure 1 shows all levels of honeypot at a glance, types of a honeypot, their deployment mode which further categorized into three different ways. Legal and ethical issues are also tried to incorporate with the architecture.

Contribution of authors: In this review paper authors are trying to compile working of different honeypots. Besides advantages and disadvantages, an effort is trying to design a proposed solutions for issues obtained during applying honeypot. And ethics of using honeypot in cloud are also discussed.

Paper organization: The remainder of this paper is organized as follows. Related works are summarized in Section II. The advantages of

honeypot are explained in Section III. In section IV the detail description of drawbacks of honeypot and its issues with proposed solutions. The significance of honeypot in various fields is describes in the section V. Ethics of Honeypot is explained in section VI. Finally, we conclude the paper with its future scope in Section VII.



II. RELATED WORK

Paul. A.J, et al. 2007 presented security in a cloud computing environment mostly uses this infrastructure as a service for the research work. They provide opportunities for offering capital cost and focused on core competencies. The purpose of this paper is to use Honey pot as a security tool in a cloud environment.

Joshi Ashay Mukundrao et al. 2011 discussed cloud computing is the newly emerging field because of its performance, high availability, least cost. In cloud computing, the data get stored in service providers. This paper has been written to focus on the problem of data security.

Stephen Brown et al. 2012 conducted a study using honeypots within various cloud computing platforms such as Amazon EC2, Windows Azure etc. to learn more about what kind of packets they receive. They used multiple honeypots such as Dionaea, Kippo, and Amun on the cloud instances. They gathered data about where attacks came from, what kinds of attacks made, and various cloud instances.

Nithin Chandra et.al, 2012 presented a Cloud Security using Honey pots. The concepts were first introduced by several icons in computer security, specifically Cliff Stoll in the book "The Cuckoo's Egg", and Bill Cheswick's paper "An Evening with Berferd." This paper explained that honeypots are used for securing cloud systems, their advantages and disadvantages etc.

Michael Beham et.al, 2013 explained that in many organizations reasonably attractive towards the services which are used in a cloud environment. This study of the security in the cloud environment is assessed by deploying and running Dionaea honeypots for a few months in the cloud provide networks.

Hwan-Seok Yang, 2013 explained that in cloud computing, intrusion detection and prevention systems are one such measure to lessen the attacks. Different researchers have proposed

different IDSs time to time. Some of these IDS's combine features of two or more IDSs which are called as hybrid intrusion detection Systems. For a signature based IDS if an attacker attacks slowly and organized, the attack may go undetected through the IDS. Thus, signature-based IDS fail to detect unknown attacks. Hybrid Intrusion Detection System (HIDS) combines the positive features of two different detection methodologies -Honeypot methodology and anomaly based intrusion detection methodology.

Navneet Kambow et.al, 2014 said that network forensics is used to notice attacker's activity and analyze their behavior. This assessment paper is based upon the preface to honeypots, their significance in network security, types of honeypots, their advantages, disadvantages and legal issues connected to honeypots.

Huseyin Ulusoy,et.al 2015 gave many research projects from the past, and built intrusion detection systems and honeypot architectures based on virtual machine introspection (VMI). These systems directly provide benefit from the use of virtualization technology. They compare the performance of existing nested-virtualization solutions and analyze the impact of the performance overhead on VMI-based intrusion detection and honey pot systems.

Ramya. R, cloud 2015 explained in cloud computing, accessing the data from data centers reduces the chances of eavesdropping and storage cost. A honey pot is a system that intentionally designs to invite malicious users to enter in the network. Honeypots is used in various cloud computing platforms (such as Amazon EC2, Windows Azure etc.) with the objective of learning more about what kind of packets they receive. Malicious activities of the attacker are used to educate the network from the attack and design network with the novel security tools.

Chaimae Saadi et.al 2016 explained that cloud computing security has developed into a basic necessity. It attains knowledge about vulnerabilities, attacks, activities of attackers and

tools to secure it. This work suggests new cloud infrastructure architecture, which combines IDS based on mobile agent and using three types of honeypots in order to detect attacks, to study the behavior of attackers, augment the added value of Honeypot and IDS based mobile agents, to solve systems boundaries intrusion detection, recover knowledge bases IDS. Therefore boost the detection pace in the cloud environment.

Manoj Agnihotri 2017 has explained that cloud computing provide contact to large pools of high end resources through diversity of interfaces that are comparable to the move toward in HPC computing resource management. The honeypots are planned and placed in a cloud network that help us to collect the unknown as well as recognized incidents occurred in cloud computing. The purpose of this paper is to show the significance of honeypot in the protection of cloud based infrastructures.

Michail Tsikerdekis et.al, 2018 explained that honeypots has been used lengthily for over two

decades. They also provide recommendations for future honeypot software which is more flexible, modular and include a dynamic intelligence design.

Christos Dalamagkas et. al, 2019 compared honeypot with the smart grid technology to distract the attackers interest and investigate the intruder strategies to protect the real data. Conpot as honeypot used to support many smart grid cases was also shown.

III. WORKING OF HONEYPOT

Authors are tried to compile the different honeypots and their working in the different environment of operating system in table 1. These involve different interaction level of honeypots, their log files generated, services, and their requirements on the basis of available literature [3].

Table 1: Compare among Various Honeypots.

Honeypots	Interaction Level	Types of Honeypot	Software Type	Log Files	Services	OS	Requirements
BOF(Back Office Friendly)	Low Interaction Level	--	Freeware	NO	7 (Telnet,FTP,HTTP, POP3,imap2)	Win32,Unix	--
Specter(Spec o3)	High Interaction Level	Production Honeypot	Retail Software	YES	14 (SMTP,FTP,POP3, HTTP,SSH,DNSimap4,Telnet)	WindowsNT ,2000,XP	--
Decoy Server (Decoo2)	High Interaction Level	--	NO	Syslog	Unlimited	Windows (9X,2000,NT),Solaris	Java Runtime
Honeyd(provo3)	Medium Interaction Level	--	YES	YES	Unlimited	Unix	Libdnet, Libevent
KF Sensor	Medium Interaction Level	Production Honeypot	NO	NO	Unlimited	Win32	HTTP
Dionea	Medium Interaction Level	Production Honeypot	YES	YES	FTP,HTTP,MYSQL	Ubuntu	--

IV. ADVANTAGES OF HONEYPOTS

Honeypots have several advantages exclusive to technology. Most notable are explained here [11].

- **Data rate:** The complete quantity of information is irresistible, building it

enormously complicated to obtain some value from the data. Honeypots can give us the accurate data we needed in a rapid and easy-to-recognize design. This makes the study easier and response time is much faster.

- **Resources:** One more challenge for the security mechanisms is its limitations. But honeypot does not need the newest cutting-edge technology and vast amounts of RAM. In minimal number of computers it deploys easily and performs better. This means honeypot is relatively cheaper in terms of availing the resources [11].
- **Accessibility:** The researchers believe in simplicity, is the prime advantage of honeypots. Here no algorithms to develop and no databases to sustain. Whereas some honeypots, particularly research honeypots, are involved on the basic principle.
- **Return on the stake:** As firewalls fruitfully remain intruders away, they turn into sufferers of their accomplishment. Administration start inquiry about the return on deal, by differentiating a threat gets disappear: Investments has new protection technologies, such as robust validation, encryption, etc.
- **Reduced False Positives:** Honeypots help in dipping false positives. The better the probability that a security resource produce false positives or misleading alerts, the fewer likely the technology will be installed [19].
- **Catching False negatives:** Catching false negatives with the help of honeypots is calm comfortable because every connection made to a honeypot is considered unauthorized. Conven-

tional attacks detecting tools become failing in detecting new attacks like signature- based detection tools [18].

- **Encryption:** Honeypots can detain malicious activity if it is in an encrypted form.

V. DRAWBACKS OF HONEYPOTS

Amid amazing recompense, honeypot consists of numerous obstacles because they work with security mechanism and enhance the overall architecture of security.

- **Narrow Field of outlook:** They observe activity against the intruder [8]. An invader splits the network and attacks on systems.
- **Fingerprinting:** In this technique, an attacker identifies the accurate uniqueness of a honeypot as it contains exceptionality [9].
- **Risk:** They launch threat in our surroundings. Several honeypots have unique levels of risk. Low interaction honeypots has low chances, but high interaction honeypots introduce high risks in whole possible platform for the attacker. The simpler the honeypot, minor will be a risk.

VI. HONEYPOT: ISSUES AND PROPOSED SOLUTIONS

There are some issues generated while using honeypot, which is tried to solve through the proposed possible solutions are given in table2.

Table2: Honeypot on different OS, its issues and proposed possible Solutions

Honeypot	Linux	Windows	Issues	Possible Solution
Open Source	NO	NO	This tool is Simulate	➤ Used in Virtual environment with proper Security facility
Services	Web Server(HTT P,FTP,SSH)	IIS,FTP	Easy to Deploy	➤ Honeypot can be used in cloud environment where security gets enhanced. ➤ Alert system generated ➤ Log files can be formed ➤ Continuous system monitoring possible.
Database	MYSQL	MYSQL	On real System	➤ Security get increased on cloud environment ➤ Another tool can be added like Kerberos which increases the security. ➤ Here risk gets decreases and chances of gathering information about attacker get increases.

Modifying Syslog	Syslogd, Klogd	-		--
Logging	Bash or Script Session	Perl Script, PHP Scripts		--
Port No.	8080	8080		--
Integrity Checking	Tripwire	-		--

The flaws can be removed [22] if the above solutions are used as per requirement. It is also tried to explain the working of honeypot on the different operating system.

VII. SIGNIFICANCE OF HONEYPOT

- **Honeypots in Educational Resource:** Jeremiah K. Jones & Gordon W. Romney [3] discussed the aspects of using the honeynets in an educational area. A lab of honeypot can form for tracing the malicious traffic in the network[22].The honeypot is implemented at Brigham Young University observe the certain benefits such as it inform about the new threats, securing the lab at a higher level, learning the network, basics of security closely identifies the flaws.
- **Virtual Honeynet in Teaching and Research:** An additional way of implementing the honeypot in the educational areas can be done by executing real or virtual honeynet for better understanding [23]. In the real honeynets, all the connections have to be measured very generously to eliminate any possibility of fault.
- **Honeypot with IDS:** An Intrusion Detection System (IDS) distinguishes between the traffic coming from various customers and the attackers, to concurrently ease the problems of throughput, latency and security of the Network. IDS can be further categorizes into Network based and Host-based [24]. Honeypots can either be the host and Network based, but generally, they are not network-based as all interface operations characteristically performed over a network link.
- **Honeypot uses in cyber security:** Honeypot can be used in Cyber security to stop the malicious attack from the attacker. Deception technology is a future concept of cyber security defense. Its

related application can detect, analyze, and defend against minimal time period and advanced attacks in real time [24]. As a technique honeypot can be attach with different technologies to empower the security on the basis of the some examples of honeypot such as Spam honeypot. It is used to attract spam and unauthorized emails, malware honeypot is designed as a simulator to get the malware attacks, spider honeypot is designed as a false web-page to detect and analyze on the basis of the unusual activity and many more.

- **IoT (Internet of Things) with honeypots:** The Internet of things is defined as the series of multiple technologies, real-time analytic, machine learning, commodity sensors, and embedded systems. The Honeypot tool that is mostly using in IOT is:
- **Honey Thing:** This honeypot act as a modem or router run on the web server and support TR-069(Technical Report 069) protocol [23]. It offers an easy web-based interface.
- **Database honeypots:** It is designed to find different attacking techniques through SQL (Structured Query Language) service. One of the tools given below:
- **Elastic Honey:** This honeypot is created for database which catches the malicious request from the attacker.
- **Honeypot with Cloud provider:** To enhance the security of the data with the cloud services we can append the honeypot with it. There are several honeypot available which can be used as per the need [25].
- **Honeypot with Kerberos:** It can also be attached with the tools or software like Kerberos to increase the security and make the data continuous monitoring from the attacker [26].

VIII. ETHICS OF HONEYPOT

Cloud computing brings so many profits to organizations that are mostly unfeasible to resist migrating to it [22]. The Cloud has transformed the business but also has quickly turned into a hacking quicksand, who miscalculates the power and importance of proper cloud security. Various technical criteria have authority on arising of ethical issues in Cloud. Along with them, security, privacy, compliance and performance metrics have a better impact on moral matters.

Subsequently, we tried to demonstrate how each of these technological criteria can have a direct impact on ethical considerations in Cloud:

- a) Privacy and Security: When illegal access to your responsive data tenant in a Cloud is gained, either by a hacker or by the Cloud provider itself. Hence privacy and security mechanisms are essential to avoid such ethical issues.
- b) Compliance: A significant part of privacy and security mechanisms created as a set of principles. A cloud service should obey with a subset of the standards concerning the request of the service. When a Cloud-based appliance in the SaaS model with some privacy or security necessities is going to be launched in the market, it should fulfill the predefined standards [23]. Additionally, it is one of the critical parameters to choose a cloud supplier to get either IaaS or PaaS services.
- c) Performance metrics: The expected performances of the provided services are predetermined in Service Level Agreement (SLA), which is a part of the general terms and conditions, focused on the performance metrics [22-23]. Availability and application response time is a pair of examples of SLA metrics. The violation of the performance metrics of SLA is a punishable offence.
- d) Security is the primary factor preventing organizations from migrating to the cloud. With the central security pressure like unauthorized access, hijacking, and malevolent insiders. The first step is to intensify cloud

security consciousness among staff to diminish the human factor in security occurrences.

IX CONCLUSION AND FUTURE SCOPE

A honeypot is a valuable resource, mainly to assemble information about events of attackers as well as their deployed tools. However, there have been numerous research offerings to make honeypot technologies more secure, reliable and risk-free.

In this paper, we have given an overview of the classification of honeypots technologies with their advantages and disadvantages. We recommended the use of an appropriate kind of honeypots for certain specific applications which depends on the interaction and design of the system. In the future we can try to find out possible solution of syslog, logging, integrity checking. By the use of deception technology examples, the level of security will get enhanced and made the honeypot much more than the software tool for research. This technology makes the honeypot more advanced at the level of authenticity, ease to operate, interactive, scalable in physical, virtual and cloud environment. Flexible deployment is one of the properties of the deception technology which is used in machine learning.

A combination of traditional network security monitoring and recent advancements in honeypots would be like:

- Cloud environment: In the cloud, honeypot can be effective tool for information gathering.
- Grid computing: It is used to secure high interaction honeypot by secure protection language (SPL).SPL to enforce security properties which cannot manage with other approach.
- Edge computing: Here honeypot can be used as active cyber defense techniques of a specific organization.

REFERENCES

1. Dhamija, et.al, "A user study using images for authentication," in Proc. 9th USINEX Security Symp., Denver, CO, Aug. 2000, pp. 45– 58.
2. Reto Baumann et.al 2002 White Paper: Honeypots .
3. Fabien Pouget, Marc Dacier et.al, 2003 White Paper: "Honeypot, Honeytoken, Honeytoken: Terminological issues" Institute Eurécom 2229, Route des Crêtes ; BP 193 06904 Sophia Antipolis Cedex ; France.
4. X. Suo,et.al,"Graphical passwords: A survey," in Proc. 21stAnnu.Comput. Security Appl.Conf., Dec. 5–9, 2005, pp. 463–472.
5. Paul. A.J, et.al 2007 "A Fast and Secure Encryption Algorithm For Message Communication", IET-UK International Conference on Information and Communication Technology in Electrical Sciences (ICTES 2007), Dr. M.G.R. University, Chennai, Tamil Nadu, India. pp. 629-634.
6. Joshi Ashay Mukundrao, et.al, "Enhancing Security in Cloud Computing" Information and Knowledge Management ,ISSN 2224-5758 (Paper) ISSN 2224-896X,Vol 1, No.1, 2011.
7. Stephen Brown,et.al, "Honeypots in the Cloud" University of Wisconsin – Madison, December 19, 2012.
8. Nithin Chandra, et.al , "Cloud Security using Honeypot Systems", International Journal of Scientific & Engineering Research Volume 3, Issue 3, March -2012 1 ISSN 2229-5518 IJSER © 2012.
9. Michael Beham, et.al ,2013, "Intrusion detection and Honey pots in nested virtualization environments", DSN, 43rd Annual IEEE/IFIP International Conference on Dependable Systems and Networks Budapest June 24, 2013 to June 27, 2013 pp 1-6.
10. Hwan-Seok Yang, "A study on attack information collection using virtualization technology" 74:8791–8799 DOI 10.1007/s11042-013-1487-8, Springer Science + Business Media New York 2013.
11. Navneet Kambow et.al 2014 Honeypots: The Need of Network Security (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 5 (5) , 2014, 6098-6101.
12. Ramya. R, "Securing the system using honeypot in cloud computing environment", International Journal of Multidisciplinary Research and Development ,Volume: 2, Issue: 4, 172-176 April 2015.
13. Chaimae Saadi et.al, 2016 Cloud Computing Security Using IDS-AM-Clust, Honeyd, Honeywall and Honeycomb International Conference on Computational Modeling and Security (CMS 2016) doi: 10.1016/ j.procs. 2016.05.189
14. Sultan Aldossary, et.al, "Data Security, Privacy, Availability and Integrity in Cloud Computing: Issues and Current Solutions" (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 7, No. 4, 2016.
15. Yunfei CI, et.al, "Design and Implementation of the Components of the Symmetric Cryptographic Algorithm" IEEE Second International Conference on Data Science in Cyberspace 978-1-5386-1600-0/17 \$31.00 © 2017 IEEE DOI 10.1109/DSC.2017.23
16. Liangxuan Zhang, et.al, "Privacy-Preserving Attribute-Based Encryption Supporting Expressive Access Structures" 2017 IEEE Second International Conference on Data Science in Cyberspace 978-1-5386-1600-0/17 \$31.00 © 2017 IEEE DOI 10.1109/ DSC. 2017.61.
17. Manoj Agnihotri et.al 2017 Analysis of Cloud Security through Honeypots International Journal of Innovations in Engineering and Technology (IJIET) <http://dx.doi.org/10.21172/ijiet.82.034>.
18. Nooreen Fatima Khan et.al 2018, Cloud security using self-acting spontaneous honeypots International Journal of Engineering & Technology 7 (2.8) (2018) 243-247
19. Michail Tsikerdekis et.al, 2018 Approaches for Preventing Honeypot Detection and Compromise 978-1-5386-5150-6/18/\$31.00 © 2018 IEEE.

20. Ronald Mitchell Campbell, 2013 The legal and ethical issues of deploying Honeypots Honor .project University of South Africa.
21. Fahien Pouget et.al, 2003 Whitepapers: Honeypot, Honey net: A Comparative Survey, Eurecom issues September14 2003, France.
22. Aman Sachan, et.al, 2016 Honeypots: Sweet OR Sour spot in Network Security? International Journal of Current Engineering and Technology Vol.6, No.3 (June 2016).
23. Yogendra Kumar Jain et.al, 2011 Honeypot based Secure Network System, International Journal on Computer Science and Engineering (IJCSE) Vol. 3 No. 2 Feb 2011 ISSN : 0975-339.
24. Christos Dalamagkas, et. al, 2019“A Survey On Honeypots, Honeynets And Their Applications On Smart Grid” IEEE NetSoft 2019 - 1st Workshop on Cyber-Security Threats, Trust and Privacy Management in Software-Defined and Virtualized Infrastructures (SecSoft).
25. Apurva Saxena, et.al, 2019 Kerberos based Data security in Research & Production Honey Pot. EPH - International Journal of Science and Engineering Volume-5, Issue-1, Jan.
26. Apurva Saxena, et.al,2019 “KERBEROS AUTHENTICATION MODEL FOR DATA SECURITY IN CLOUD COMPUTING USING HONEY-POT” GLOBAL JOURNAL OF ENGINEERING SCIENCE AND RESEARCHES [Saxena, 6(8): August 2019] ISSN 2348 – 8034 DOI- 10.5281/zenodo. 3367428.



Scan to know paper details and
author's profile

Pairwise Spatial Correlation of SARS-Corona Viruses

Ethirajan Govindarajan

University of Petroleum

ABSTRACT

A set of complete genome data of nine SARS-Corona type viruses are analyzed for possible common features among them. A novel genetic features called '**Percentage Concentration of Nucleotides**' denoted as pA, pT, pG and pC are evaluated for each genome data and cross verified with other data whether all of them possess the same genetic features or not. It was found to be approximately the same for all genomes. This means, all SARS-Corona type viruses have the same structural genetic property, in the system biological sense. On the other hand, the genetic functional properties of SARS-Corona type viruses are different from one another. The scope of this paper is limited to study of spatial correlation among the genome data sets under consideration. Adjoints of a genome data are four independent binary sequences corresponding to the nucleotides of adenine, thymine, guanine and cytosine. For example, the adjoint of adenine of a genome sequence is a binary sequence consisting of 1's in the place of adenine in the genome sequence and 0's in all other places. Two-dimensional array of an adjoint of period 128 is obtained by dividing the binary sequence into segments of length 128 and placing all segments one below the other. One can try with periods of 64, 32, and so on and obtain respective two-dimensional arrays. This two-dimensional array is referred here as adjoint array. Adjoint arrays of all nine genome data are pair wise correlated and results reported in this paper.

Keywords: SARS-Corona Viruses, Artificial Intelligence, Machine Learning

Classification: J.3

Language: English



LJP Copyright ID: 975842
Print ISSN: 2514-863X
Online ISSN: 2514-8648

London Journal of Research in Computer Science and Technology



Volume 20 | Issue 1 | Compilation 1.0

© 2020. 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.

Pairwise Spatial Correlation of SARS-Corona Viruses

Ethirajan Govindarajan

ABSTRACT

A set of complete genome data of nine SARS-Corona type viruses are analyzed for possible common features among them. A novel genetic features called 'Percentage Concentration of Nucleotides' denoted as pA, pT, pG and pC are evaluated for each genome data and cross verified with other data whether all of them possess the same genetic features or not. It was found to be approximately the same for all genomes. This means, all SARS-Corona type viruses have the same structural genetic property, in the system biological sense. On the other hand, the genetic functional properties of SARS-Corona type viruses are different from one another. The scope of this paper is limited to study of spatial correlation among the genome data sets under consideration. Adjoints of a genome data are four independent binary sequences corresponding to the nucleotides of adenine, thymine, guanine and cytosine. For example, the adjoint of adenine of a genome sequence is a binary sequence consisting of 1's in the place of adenine in the genome sequence and 0's in all other places. Two-dimensional array of an adjoint of period 128 is obtained by dividing the binary sequence into segments of length 128 and placing all segments one below the other. One can try with periods of 64, 32, and so on and obtain respective two-dimensional arrays. This two-dimensional array is referred here as adjoint array. Adjoint arrays of all nine genome data are pair wise correlated and results reported in this paper.

Curtain Raiser

A Verse from Taittiriya Upanishad

आकाशाद्वायुः । वायुर्ओरग्निः । अग्नेरापः ।
अद्भ्यः पृथिवी । पृथिव्या ओषधयः । ओषधीभ्योन्नम् । अन्नात्पुरुषः ।

The Life Cycle

The Forward Process

Gaseous Forms emanated from Space.
Heat Forms emanated from Gaseous Forms.
Fluid Forms emanated from Heat Forms.
Solid Forms emanated from Fluid Forms.
Plants and Trees emanated from Solid Forms.
Consumables emanated from Plants and Trees.
Living Forms emanated from Consumables

The Reverse Process

Living Forms become Consumables.
Consumables become Plants and Trees.
Plants and Trees become Solid Forms.
Solid Forms become Fluid Forms.
Fluid Forms become Heat Forms.
Heat Forms become Gaseous Forms.
Gaseous Forms become Space.

Universe appears and disappears as said above. Universal Consciousness plays major role in this cyclic process. Universal Consciousness is an inseparable blend of Knowledge and Energy. All Entities, both sentient and insentient play active roles in these processes and Universal Consciousness is indeed the passive spectator and is closed under all processes.

RNAs and DNAs are the fundamental units used in these processes. Life Forms are constructs made up of these units.

RNAs are termed as ‘Viruses’. They are constructive as well as destructive. Constructive Viruses act in the Forward Process and the destructive Viruses act in the Reverse Process.

SARS-Corona type viruses are destructive Viruses and they do their job. It is up to Humans to understand them, their role and delay in becoming consumables used in the Reverse Process.

This report is a result of an intensive research carried out by devoted Scientists who work with passion.

The team acknowledges **Mr. Rahul Sharma** (Senior Data Scientist, Pentagram Research Centre, Hyderabad, Telangana State, India) for his untiring effort in making this report.

Keywords: SARS-Corona Viruses, Artificial Intelligence, Machine Learning

Author: Adjunct Professor, Department of Cybernetics, School of Computer Science University of Petroleum and Energy Studies, Dehradun, Uttarakhand, India Director, PRC Global Technologies Inc., Ontario, Canada President, Pentagram Research Centre Pvt. Ltd., Hyderabad, India.

I. DATA USED FOR ANALYSIS

Nine complete genomes of viruses as given in the NCBI web site <https://www.ncbi.nlm.nih.gov/genbank/sars-cov-2-seqs/> are considered here for data analysis. These genome data are assumed to be genuine and analysis carried out on them.

1. SARS coronavirus SZ16, complete genome

GenBank: **AY304488.1**

[FASTA Graphics](#)

LOCUS: AY304488; 29731 bp RNA; linear VRL 05-NOV-2003

DEFINITION: SARS coronavirus SZ16, complete genome.

ACCESSION AY304488

VERSION AY304488.1

KEYWORDS .

SOURCE: Civet SARS CoV SZ16/2003

ORGANISM: Civet SARS CoV SZ16/2003

Viruses; Riboviria; Nidovirales; Coronaviridae; Orthocoronavirinae; Betacoronavirus; Sarbecovirus.

REFERENCE 1 (bases 1 to 29731)

AUTHORS Guan,Y., Zheng,B.J., He,Y.Q., Liu,X.L., Zhuang,Z.X., Cheung,C.L.,

Luo,S.W., Li,P.H., Zhang,L.J., Guan,Y.J., Butt,K.M., Wong,K.L.,

Chan,K.W., Lim,W., Shortridge,K.F., Yuen,K.Y., Peiris,J.S. and Poon,L.L.

TITLE: Isolation and characterization of viruses related to the SARS coronavirus from animals in southern China

JOURNAL Science 302 (5643), 276-278 (2003)

PUBMED [12958366](#)

REFERENCE 2 (bases 1 to 29731)

AUTHORS: Guan,Y. and Zheng,B.J.

2. SARS coronavirus ZS-C, complete genome

GenBank: **AY395003.1**

[FASTA Graphics](#)

LOCUS: AY395003; 29647 bp RNA; linear VRL 29-JAN-2004

DEFINITION: SARS coronavirus ZS-C, complete genome.

ACCESSION AY395003

VERSION AY395003.1

KEYWORDS .

SOURCE: SARS coronavirus ZS-C
 ORGANISM: [SARS coronavirus ZS-C](#)
 Viruses; Riboviria; Nidovirales; Coronaviridae; Orthocoronavirinae; Betacoronavirus; Sarbecovirus.
 REFERENCE 1 (bases 1 to 29647)
 CONSRTM The SARS epidemiology consortium of Guangdong
 TITLE: From independent foci of epidemic outbreak to large genomic alteration in late phase viruses: evolution of the SARS-coronavirus
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 29647)
 CONSRTM: The SARS epidemiology consortium of Guangdong
 TITLE Direct Submission
 JOURNAL Submitted (19-SEP-2003) Guangdong, China

3. Bat SARS coronavirus Rm1, complete genome

GenBank: **DQ412043.1**
[FASTA Graphics](#)
 LOCUS: DQ412043; 29749 bp RNA; linear VRL 13-JUL-2006
 DEFINITION: Bat SARS coronavirus Rm1, complete genome.
 ACCESSION: DQ412043 DQ071612 DQ159957
 VERSION: DQ412043.1
 KEYWORDS .
 SOURCE: Bat SARS CoV Rm1/2004
 ORGANISM: [Bat SARS CoV Rm1/2004](#)
 Viruses; Riboviria; Nidovirales; Coronaviridae; Orthocoronavirinae; Betacoronavirus; Sarbecovirus.
 REFERENCE 1 (bases 1 to 29749)
 AUTHORS Li,W., Shi,Z., Yu,M., Ren,W., Smith,C., Epstein,J.H., Wang,H., Crameri,G., Hu,Z., Zhang,H., Zhang,J., McEachern,J., Field,H., Daszak,P., Eaton,B.T., Zhang,S. and Wang,L.F.
 TITLE: Bats are natural reservoirs of SARS-like coronaviruses
 JOURNAL Science 310 (5748), 676-679 (2005)
 PUBMED [16195424](#)
 REFERENCE 2 (bases 1 to 29749)
 AUTHORS: Ren,W.Z., Li,W.D., Yu,M., Zhou,P., Hao,P., Zhang,S.Y., Zhao,G.P., Zhong,Y., Wang,S.Y., Wang,L.-F. and Shi,Z.L.
 TITLE: Genetic diversity of SARS-like coronaviruses in horseshoe bats
 JOURNAL Unpublished
 REFERENCE 3 (bases 1 to 29749)
 AUTHORS Li,W. and Shi,Z.
 TITLE Direct Submission
 JOURNAL Submitted (23-MAY-2005) Chinese Academy of Science, Wuhan Institute of Virology, China
 REFERENCE 4 (bases 1 to 29749)
 AUTHORS: Yu,M. and Wang,L.-F.

4. Bat SARS-like coronavirus isolate Rs4084, complete genome

GenBank: **KY417144.1**
[FASTA Graphics](#)
 LOCUS: KY417144; 29770 bp RNA; linear VRL 18-DEC-2017
 DEFINITION: Bat SARS-like coronavirus isolate Rs4084, complete genome.
 ACCESSION: KY417144
 VERSION: KY417144.1
 KEYWORDS .
 SOURCE: Bat SARS-like coronavirus
 ORGANISM: [Bat SARS-like coronavirus](#)
 Viruses; ssRNA viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales; Coronaviridae; Coronavirinae; Betacoronavirus.
 REFERENCE 1 (bases 1 to 29770)
 AUTHORS: Hu,B., Zeng,L.P., Yang,X.L., Ge,X.Y., Zhang,W., Li,B., Xie,J.Z., Shen,X.R., Zhang,Y.Z., Wang,N., Luo,D.S., Zheng,X.S., Wang,M.N., Daszak,P., Wang,L.F., Cui,J. and Shi,Z.L.
 TITLE: Discovery of a rich gene pool of bat SARS-related coronaviruses provides new insights into the origin of SARS coronavirus
 JOURNAL: PLoS Pathog. 13 (11), e1006698 (2017)
 PUBMED: [29190287](#)
 REMARK: Publication Status: Online-Only
 REFERENCE 2 (bases 1 to 29770)

AUTHORS: Hu,B., Zeng,L.P., Yang,X.L., Ge,X.Y., Zhang,W., Li,B., Luo,D.S.,
Zhang,Y.Z., Wang,M.N., Daszak,P., Wang,L., Cui,J. and Shi,Z.L.

TITLE: Direct Submission

JOURNAL: Submitted (30-DEC-2016) CAS Key Laboratory of Special Pathogens and Biosafety and Center for Emerging Infectious Diseases,

Wuhan Institute of Virology, Chinese Academy of Sciences, No. 44 Xiao, Hong Shan, Wuhan, Hubei 430071, China

5. Wuhan seafood market pneumonia virus isolate Wuhan-Hu-1, complete genome

GenBank: [MN908947.1](#)

COMMENT: [WARNING] On Jan 14, 2020 this sequence was replaced by [MN908947.2](#). [FASTA Graphics](#)

LOCUS: MN908947; 30473 bp ss-RNA; linear VRL 12-JAN-2020

DEFINITION: Wuhan seafood market pneumonia virus isolate Wuhan-Hu-1, complete genome.

ACCESSION: MN908947

VERSION: MN908947.1

KEYWORDS

SOURCE: Wuhan seafood market pneumonia virus

ORGANISM: [Wuhan seafood market pneumonia virus](#)

Viruses; Riboviria; Nidovirales; Cornidovirineae; Coronaviridae; Orthocoronavirinae; Betacoronavirus; unclassified Betacoronavirus.

REFERENCE 1 (bases 1 to 30473)

AUTHORS: Zhang,Y.-Z., Wu,F., Chen,Y.-M., Pei,Y.-Y., Xu,L., Wang,W., Zhao,S., Yu,B., Hu,Y., Tao,Z.-W.,
Song,Z.-G., Tian,J.-H., Zhang,Y.-L., Liu,Y., Zheng,J.-J., Dai,F.-H., Wang,O.-M., She,J.-L. and Zhu,T.-Y.

TITLE: A novel coronavirus associated with a respiratory disease in Wuhan of Hubei province, China

JOURNAL: Unpublished

REFERENCE 2 (bases 1 to 30473)

AUTHORS: Zhang,Y.-Z., Wu,F., Chen,Y.-M., Pei,Y.-Y., Xu,L., Wang,W., Zhao,S., Yu,B., Hu,Y., Tao,Z.-W.,
Song,Z.-G., Tian,J.-H., Zhang,Y.-L., Liu,Y., Zheng,J.-J., Dai,F.-H., Wang,O.-M., She,J.-L. and Zhu,T.-Y.

TITLE: Direct Submission

JOURNAL: Submitted (05-JAN-2020) Department of Zoonoses, National Institute of Communicable Disease Control and Prevention, Chinese

Center for Disease Control and Prevention, Changping Liuzi 5, Beijing 102206, China

6. Wuhan seafood market pneumonia virus isolate Wuhan-Hu-1, complete genome

GenBank: [MN908947.2](#)

[WARNING] On Jan 17, 2020 this sequence was replaced by [MN908947.3](#). On Jan 14, 2020 this sequence version replaced [MN908947.1](#).

[FASTA Graphics](#)

LOCUS: MN908947; 29875 bp ss-RNA; linear VRL 14-JAN-2020

DEFINITION: Wuhan seafood market pneumonia virus isolate Wuhan-Hu-1, complete genome.

ACCESSION: MN908947

VERSION: MN908947.2

KEYWORDS

SOURCE: Wuhan seafood market pneumonia virus

ORGANISM: [Wuhan seafood market pneumonia virus](#)

Viruses; Riboviria; Nidovirales; Cornidovirineae; Coronaviridae; Orthocoronavirinae; Betacoronavirus; unclassified Betacoronavirus.

REFERENCE 1 (bases 1 to 29875)

AUTHORS: Zhang,Y.-Z., Wu,F., Chen,Y.-M., Pei,Y.-Y., Xu,L., Wang,W., Zhao,S., Yu,B., Hu,Y., Tao,Z.-W.,
Song,Z.-G., Tian,J.-H., Zhang,Y.-L., Liu,Y., Zheng,J.-J., Dai,F.-H., Wang,Q.-M., She,J.-L. and Zhu,T.-Y.

TITLE: A novel coronavirus associated with a respiratory disease in Wuhan of Hubei province, China

JOURNAL: Unpublished

REFERENCE 2 (bases 1 to 29875)

AUTHORS: Zhang,Y.-Z., Wu,F., Chen,Y.-M., Pei,Y.-Y., Xu,L., Wang,W., Zhao,S., Yu,B., Hu,Y., Tao,Z.-W.,
Song,Z.-G., Tian,J.-H., Zhang,Y.-L., Liu,Y., Zheng,J.-J., Dai,F.-H., Wang,Q.-M., She,J.-L. and Zhu,T.-Y.

TITLE: Direct Submission

JOURNAL: Submitted (05-JAN-2020) Department of Zoonoses, National Institute of Communicable Disease Control and Prevention, Chinese

Center for Disease Control and Prevention, Changping Liuzi 5, Beijing 102206, China

7. Severe acute respiratory syndrome coronavirus 2 isolate Wuhan-Hu-1, complete genome

GenBank: [MN908947.3](#)

COMMENT: On Jan 17, 2020 this sequence version replaced [MN908947.2](#).

[FASTA Graphics](#)

LOCUS: MN908947; 29903 bp ss-RNA; linear VRL 18-MAR-2020
 DEFINITION: Severe acute respiratory syndrome coronavirus 2 isolate Wuhan-Hu-1, complete genome.
 ACCESSION: MN908947
 VERSION: MN908947.3
 KEYWORDS .
 SOURCE: Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)
 ORGANISM: [Severe acute respiratory syndrome coronavirus 2](#)
 Viruses; Riboviria; Nidovirales; Coronaviridae; Orthocoronavirinae; Betacoronavirus; Sarbecovirus.
 REFERENCE 1 (bases 1 to 29903)
 AUTHORS: Wu,F., Zhao,S., Yu,B., Chen,Y.M., Wang,W., Song,Z.G., Hu,Y., Tao,Z.W., Tian,J.H., Pei,Y.Y., Yuan,M.L., Zhang,Y.L., Dai,F.H., Liu,Y., Wang,Q.M., Zheng,J.J., Xu,L., Holmes,E.C. and Zhang,Y.Z.
 TITLE: A new coronavirus associated with human respiratory disease in China
 JOURNAL: Nature 579 (7798), 265-269 (2020)
 PUBMED: [32015508](#)
 REFERENCE 2 (bases 1 to 29903)
 AUTHORS: Wu,F., Zhao,S., Yu,B., Chen,Y.-M., Wang,W., Hu,Y., Song,Z.-G., Tao,Z.-W., Tian,J.-H., Pei,Y.-Y., Yuan,M.L., Zhang,Y.-L., Dai,F.-H., Liu,Y., Wang,Q.-M., Zheng,J.-J., Xu,L., Holmes,E.C. and Zhang,Y.-Z.
 TITLE: Direct Submission
 JOURNAL: Submitted (05-JAN-2020) Shanghai Public Health Clinical Center & School of Public Health, Fudan University, Shanghai, China
 ORGANISM: [Bat SARS-like coronavirus](#)
 Viruses; Riboviria; Nidovirales; Coronaviridae; Orthocoronavirinae; Betacoronavirus; Sarbecovirus.
 REFERENCE 1 (bases 1 to 29802)
 AUTHORS: Hu,D., Zhu,C., Ai,L., He,T., Wang,Y., Ye,F., Yang,L., Ding,C., Zhu,X., Lv,R., Zhu,J., Hassan,B., Feng,Y., Tan,W. and Wang,C.
 TITLE: Genomic characterization and infectivity of a novel SARS-like coronavirus in Chinese bats
 JOURNAL: Emerg Microbes Infect 7 (1), 154 (2018)
 PUBMED: [30209269](#)
 REMARK: Publication Status: Online-Only
 REFERENCE 2 (bases 1 to 29802)
 AUTHORS: Hu,D.
 TITLE: Direct Submission
 JOURNAL: Submitted (05-JAN-2018) Institute of Military Medicine Nanjing Command, Nanjing, Institute of Military Medicine Nanjing Command, Nanjing, NO. 293 East Zhongshan Road, Nanjing, Jiangsu 210002, China

9. Bat SARS-like coronavirus isolate bat-SL-CoVZXC21, complete genome

GenBank: **MG772934.1**
[FASTA Graphics](#)
 LOCUS: MG772934; 29732 bp RNA; linear VRL 05-FEB-2020
 DEFINITION: Bat SARS-like coronavirus isolate bat-SL-CoVZXC21, complete genome.
 ACCESSION: MG772934
 VERSION: MG772934.1
 KEYWORDS .
 SOURCE: Bat SARS-like coronavirus
 ORGANISM: [Bat SARS-like coronavirus](#)
 Viruses; Riboviria; Nidovirales; Coronaviridae; Orthocoronavirinae; Betacoronavirus; Sarbecovirus.
 REFERENCE 1 (bases 1 to 29732)
 AUTHORS: Hu,D., Zhu,C., Ai,L., He,T., Wang,Y., Ye,F., Yang,L., Ding,C., Zhu,X., Lv,R., Zhu,J., Hassan,B., Feng,Y., Tan,W. and Wang,C.
 TITLE: Genomic characterization and infectivity of a novel SARS-like coronavirus in Chinese bats
 JOURNAL: Emerg Microbes Infect 7 (1), 154 (2018)
 PUBMED: [30209269](#)
 REMARK: Publication Status: Online-Only
 REFERENCE 2 (bases 1 to 29732)
 AUTHORS: Hu,D.
 TITLE: Direct Submission
 JOURNAL: Submitted (05-JAN-2018) Institute of Military Medicine Nanjing Command, Nanjing, Institute of Military Medicine Nanjing Command, Nanjing, NO. 293 East Zhongshan Road, Nanjing, Jiangsu 210002, China

Actual genome data of the above nine viruses, referenced by NCBI are (i) AY304488.1, (ii) AY395003, (iii) DQ412043, (iv) KY417144.1, (v) MN908947.1, (vi) MN908947.2, (vii) MN908947.3, (viii) Mg772933 and (ix) 9 Mg772934 are not presented as such in this paper. Genome data is a characteristic sequence consisting of symbols A, T, G and C and it is numerically coded as a sequence consisting of the numbers 1, 2, 3 and 4 (i.e., A ↔ 1, T ↔ 2, G ↔ 3 C ↔ 4). For example, the Genome 1 (AY304488.1) and its numerical equivalent are given below.

Genome data of AY304488.1

[illegible]

[illegible]

Genome data of AY304488.1 Numerical Sequence

4214441331 1113441144 1144243124 242321312 4232242421 1143114222 1111242323 2134232434 2433423412 3442132341 4421434132 1211141212 1211122221 4232432231 4113111431 3211142434 4424242344 1314234221 4332224332 4332234314 2431241241 3411241121 3222442344 3323231443 1113321131 2331131424 2322422332 3241143131 1114141432 4411424132 2234423244 2441322213 1143234231 3234323242 2433331424 2223311313 3444212433 1334143231 1414424111 1123341422 3233242132 1313423311 1113343214 2344441442 2311441341 2323242241 121143224 2312344221 1344144124 1433441441 3324322313 4432322431 1112331433 4122413214 3324321343 3212114124 3331321424 3234414124 3233443122 23334431114 4441122341 2143431123 2442424243 2131114332 1121133331 443233241 244212334 144212334 1121 1324221231 4221323314 3142233412 3241444421 3131121214 311411142 3213114121 341234132 3323414244 3231142414 2432313424 123313323 41212343414 4212343414 11141122442 3233444131 2333214442 4223122341 4411131222 2424341434 3433341132 4112323414 2422244311 4114223122 1412431324 3113131332 3442142342 3443324144 2314312311 1223442332 4424231343 424231133 1342143134 1441314144 4224311122 1113234441 311122231 411222411 3333112344 4111322232 2234442422 1142411113 2411134212 4411441432 322311131 111314231 323224123 3334321214 3442432321 4442223241 2424414133 1323211411 4123414223 242142231 2311123211 242234312 33113224412 33113222412 3341314323 4314224423 1113441422 3231141223 33414231 1122213221 2311133144 2142141231 3323144214 1234 321323111 1234112324 4234442341 134441133 1223314423 133112214 141141241 11412231 1142431424 434113331 3213121423 1232222331 3342323323 2234421232 23234234212 1121134323 21432332 2444234342 132434231 21 2233424121 3441412434 1221423323 14112323314 312334124 3123312142 31312424 2241 31312424 2241 31312142 3132432311 4323221141 221142232 2334312222 4122231123 1133132234 4124122223 3412422242 4234424214 1132442222 1223141421 211313242 312121413 242224111 4412232321 3244234332 1142121113 22114431 33311344431 1113323422 3314112233 1411413131 3412322211 4144142323 2332222444 2414133423 4233232212 413124112 2222434341 4142231234 134114414 2411224423 1222341113 134134234 111332414223 1233212224 2311413241 2214324223 441434412 332212142 2413144234 2414411 23241221 1233412123 2114435233 242232411 4131422424 1232323234 2114322233 3341421423 223111142 4133442124 222312222 3341243241 1333241141 1413322242 2341121141 214323123 2323111124 2441223123 232221141 334142431 123241324 241242334 341132324 3121414241 422132331 3242241243 424111341 3331422214 42342132321 3313414243 414214241 2344424221 33414411 311321144 2242231133 2124241412 3141421421 3214242421 333123222 42441131 1443 3231142341 134142431

1434443223 1213422414 112331342 1243223341 4144132423 2321112334 4241234242 2131312211 3314111311 4112142343 4122324244 2332221423 3421411141 1232422243 4221111333
3323414411 2211133232 1144222331 3113121423 2223331132 2411332214 1131123231 3112414122 2313422312 3114323223 1413132342 212311113 2342423242 1414232231 1424332144
3113221423 1322234123 2322321341 3133423223 2311314222 1411441322 2423124244 2214411412 3332122312 4223123132 3313232134 2141224214 2212223123 1234233231 1311114222
2412414321 2321223224 4222214442 4413123133 1131131331 4312341313 2323133113 1131122231 2311144232 3114123132 1433214131 3312312212 4113324244 4242331122 2332344241
3423111413 2243132231 3311313121 3133113142 3342331231 2142142313 4112413121 2231344131 1441311442 1414423113 1144132211 3412221422 3322122211 1142214231 4112223244
1221112323 4231412432 2113313341 4111323421 1244212332 3122321112 3423241141 2141442311 1412332332 332313413 3234142411 4113341144 1123323441 2341111331 3132312312
2141221134 2111233444 2422141321 3313332422 3222342224 2331412112 4223421131 1332342341 2322322331 4421144211 1234133231 3314124413 4224221133 4134121231 1112224112
2414133141 4422142234 1441223223 2413413341 2122233234 2111421422 4132422214 1132323432 3413143322 4321414133 2221212234 1324112314 1113424222 1231341332 2324123312
2124223121 1442311344 2131323311 3414421114 1131331344 1441114141 3113122441 1114231331 3111242324 3214131134 4232431232 3113441111 1221133442 3412231231 3322144141
1414233113 111421322 2422144112 1132214242 2322234231 2124112332 1134222144 1231224241 3114123422 1313323113 1212324222 4422313113 3123414422 1412332133 2312322124
1421323323 1212414223 2322321121 4442441111 1334233233 4142142313 1234242411 3134222311 3111323441 3223123132 121214414 3214442331 4113312323 4233221214 1422313311
3421131423 4242211311 1234111242 3412222123 2142144224 1311341442 1123421133 1131312242 1331142321 2442331122 2313131112 3422342412 3423113131 4113111122 1123442121
2341233123 2213134412 1123341144 1244114321 1321211133 1122111122 4113133341 2432231421 2332324431 2242242222 1214213211 1313442321 3422421221 2214311342 3114242421
1123134434 2232414112 3441122332 2123231414 1233222211 2422311313 3423443442 3212343224 2422111342 4423443213 2324132124 1241441312 3423221421 4121211233 1214442142
2432412411 1314124231 3313414222 3213111413 2224222334 2334242214 1314232141 4212214221 3413432141 3132213323 2231122242 2113322232 3141111223 2321441414 2423313134
4443243132 2241242231 4332313322 4222414223 1411142111 3132424221 2444234333 1332211314 2121111323 2241411423 2331411414 2112424414 1414134223 2331212324 2123141212
3314134132 2232344114 1214223312 3323423123 221411112 2111442421 3211124123 1332311314 2224422321 4214421323 1231414142 1432133231 3422243132 1421441214 2422312313
1222224223 3213321412 3242342221 1444141411 1311123311 1222442411 3223233222 2142424112 211233342 3121141122 3221222324 2132322221 2213414224 1141312321 1324111224
1123414413 4142241311 33442212212 13134444323 4233231234 2342114222 2323414241 4124344221 4132121111 1423223343 1342233231 2342133131 1421231323 2422424141 3412342112
2233112423 4111343132 2422112323 3232321114 1223233241 3111142142 1442211433 3232131134 2323123212 1233321424 2124221231 2112422113 1413323222 4412244123 2323232332
4323123421 4141121242 1321411411 3132422422 2322212312 3242341441 44423423132 1211122141 3411332141 2422212323 4311231321 4142333212 2124132323 3241221414 2412121142
1414233113 4442421243 2122314331 3424144221 4111312324 1313214111 3414413231 4231232222 4214113311 1412422141 4214114412 1413442323 243212114 2431233132 2142214141
3131223114 4111122331 2333212212 1111133121 1234221421 2141313413 4421231344 2231444114 2411441221 4411123431 3222231211 2224111424 1412322421 1414111122 2342312312
2211124111 2314134422 4141113441 3422414313 1342124232 4141224244 4413142231 1333431232 133342122 3142121314 1421224134 3122241321 1422243132 1421242341 1221423421
3222334141 2211441334 2141144113 1411432241 1444111414 2233232221 423242221 3313214111 3441231312 1422411122 412223132 2423341321 3113141414 1133123341 4112423422
2323111334 14114441442423131 3213233111 1244214412 1413113311 324123132 3231432311 1412143311 3223213341 1234212142 2111441241 312313323 221132114 34113321
3324123133 142212334 2342212323 3111141411 3412214412 2113111442 1123134222 4142134422 1332221111 1411223441 4241233212 2342341122 1121323224 4223313211 1122222342
2123241114 4122422133 1411341341 1221411412 41111223434 2113131221 3414114232 2322214111 2212123442 2123232221 4122122322 4411223232 1422221421 1113214411 2242131122
1313422414 2144214114 2122342211 1112323221 1313232234 2111221232 2233134343 3412211221 2323113241 4441112222 4211123232 4141123442 1232323421 2322323213 2122244222
3222424211 2421232114 2342342222 33223214242 2142411222 22233242442 2422122321 1233432213 1311223212 4221122432 4211432214 2142123312 224232313 322422244 2234134122
2322411323 32131424 4422312242 2144413424 2231114412 2413323143 1224412432 1411342131 4223141122 2213324233 4434231323 3222223341 212123232 2414111122 4222212221
2213324222 4134212112 3413323224 2223342122 2234213241 2224124134 1122422334 2124221322 214212132 1223214111 2334144432 2242341123 3221331232 1412422422 2344224224
214214212 3311313421 2322412124 1232123322 3414424224 3142234123 1234324121 343411243 2344141434 3223132321 4114212232 211233423 1131313423 14242123421 2341112331
3344323342 2423411314 24141112233 1122324241 1223234241 3232132412 2412212323 3311322342 432312232 4142441322 211131441 1241144421 4231441324 1232121222
3223121323 2234242323 11112334343 4224144424 1422231441 3342332411 1131442123 1313141244 3424444412 2223241122 2131411222 3131342114 1141421113 3224142344 2122112324
1213222223 1233411324 41111234314 3132423422 4211324234 2242323214 2141324134 2312323441 1442122423 2234223144 1134242232 1241314322 3313121321 4231132224 2422113123
2231232422 1232431412 44222241341 1422221323 2244212331 1111422113 341422233 4214134241 4134313221 3411133323 2134222131 2323234422 2421412243 2342413424 4431411332
3222323121 44431232231 4141113312 3221223112 3242411142 2241412414 2423142213 1132314133 1441313222 1141122241 2342414421 211213322 311114321 4344413131 2422334341
2321223142 3211234113 3412124112 3444113213 4111113241 112322241 4241242331 1232111131 4214123242 2212423114 1342343211 1411122432 132423441 1313114111 4121422222
1314211422 3234214114 2131413322 3224112341 2114214211 1124241424 113323321 1312232213 1422232222 1114221234 2211334414 1221223234 3224223423 4122332223 2214124322
1234413214 1214122324 1124412312 3221414111 1231112412 2332214111 3441224133 1233232414 2432314124 1222421423 1231223222 2341112111 4123423322 2231434123 3222134413
3223232322 4121411111 2314111134 2344443213 2134234212 4122141131 3131223322 2412132344 23344221443 3321423234 2313134112 4112332314 2242234122 2242144243 2322222132
3423223341 1241223421 4141442244 1141442243 1141442243 123212331 2222342144 2423423444 2442342344 2313232141 1422221133 1234212333 4111442323 142122322 12314121 2223421313
3322421222 4221213231 3422432441 3141424322 1232342212 3312332244 1242124132 2244211414 2214423313 3322423221 3132132114 1142223212 3423132142 3213141233 2142124311
1324213113 41133212223 4421242144 1232321312 3332422211 2112313412 2141313424 2124133242 224232332 3223123431 2311242412 1342111424 2221424424 2232341144 2323332342
2213123232 4234242132 1323342332 3321221223 4412122332 3142232342 3442142142 2212311122 4131432342 2223233132 1411441232 2322342342 1123414222 2322222312 3242224142
1214242324 2334144134 2214134222 4234433313 2421424132 4222214223 2142231412 2421222414 4112312332 2412242233 4241442241 1233222344 1232222424 212223344 2222333121
1413411242 1232122423 2122242423 1134142344 1223232422 211414212 4423231111 1132412232 212331322 1412221321 4422241331 3442342323 2321444222 2342411411 3311123214
4211112234 3213431314 1423232441 4221414132 1211413321 2422342421 2121411312 1411321222 4132331344 2213121421 4413421243 2311341342 2342344142 2134111334 24221112314
2221341412 4133242321 2322442414 4114414414 131424112 4142242342 3222434131 3232224131 1111232341 2244432413 3411132231 1333242342 3214113211 4244323314 2141142214
1123312232 3323231231 4141321214 2324411314 1232412223 4141341311 3141234221 1244211232 2311312423 4241224341 1124411441 2134224222 3224134243 3411123241 1422432322
1223444212 4212341111 2234242222 1334221113 2231214224 214442113 121142231 1222142231 442324111 4122224132 2421341234 214112332 4144124233 2322212413
2323441231 3144211241 2144122111 3322422244 2211233124 1232332132 3223222221 1412231221 2312234323 2422242342 1212341241 2123313422 4411413313 2141434233 2142314221
1313322112 2421232324 1222232314 1314111423 4141334234 1332141314 1411441211 4122111232 2224234123 4232123423 4232212411 2323212133 2332242221 1231322414 4142142233
1123142221 1442232334 1123113214 1144212314 4222314141 1312412322 3144212233 3144224234 2342411411 3311223443 2422131212 3232342342 2231113134 2342341311 2321334211
3324321421 2442232313 4142122221 3113123132 2214144122 2312322322 1314112342 4232323214 4224411332 1132241131 1112232211 3334142412 4122331234 2222114222 4223141241
4231223422 2234241113 2141413233 2414232222 2422232221 4313112342 2244232221 2214242233 2122123341 1223423412 3234212342 342232213 4121131413 4124242324 4232222423
2214422424 2234114132 2342214222 121232342 141234234 2134233323 1234321241 2141233423 2311223342 3141421342 2324233221 2134422113 3412213232 2321234224 2122242321
2234221224 2412314134 2434142322 2123123123 4234213143 3222333141 42323123 241321423 232214142 232221411 3242142123 3211234222 1312411342 1222442122 3334422132 21422242321
1442421142 1224233232 4322143142 1241232222 2134213134 2121323222 3232323223 1321221444 1223221222 1221423341 1414422141 3232124123 4223222122 3222422133 4212232234
2342342142 2233442222 4232221421 1144322142 2413342214 2422323222 2123142142 2332442414 1411311222 1332121231 1424441333 3422223442 4421131321 3212231234 2224113422
1141221132 2322323212 2331332114 4112321241 1332231241 2321413242 1111232423 1423111323 4141242323 3214234242 4332424241 1411422131 3213241242 4224211122 3223344111
2323214114 2441411231 2122424222 3411113141 4114231313 2224313113 1233222424 2222324232 2223421244 123413323 4232131412 2112133223 2343133111 2342431211 4432342142
4224133421 2234224113 1222132242 2214412412 1234434221 2344142344 4133133444 1223134134 2321342112 3323122423 1132423242 4111113221 113112422 2311232334 2111242313
2223144323 1234234412 3411434113 2233111131 2334131241 3442123144 4111232141 1411334113 1242313314 1131333411 1132114213 234212341 1111234242 2414212342 2133113422
3121123123 4142211441 4122124114 1123443423 1233223232 2441424114 1242144212 2341241434 1344114424 1233223223 2444231221 2332144214 1131144421 1424222141
2123412423 4142423331 1124413411 3223223123 4331213411 3122322411 4221323111 2211412331 4112241441 1122233422 3344242212 2322141342 4211313441 1424134232 2111421413
1121123114 2313244132 1341421431 4131232442 3234334233 2144141411 1413422321 4231231411 2341422344 2142121141 1224311333 1332133222 3234233412 2142124131 4414411312
4241112333 4213122444 2113132312 3321413321 4112221414 1311423311 4414422321 3222322214 1314141441 1113334421 1132311121 4223214224 1241113342 2114114422 1112131332
1233234233 3413222134 2342141321 4324224133 42333111234 2141311321 44234441 122 414232342 2244224232 3422223413 2131444234 2111341212 1133122144 2134113233 1331411441
1241441142 3232311312 3223232141 4141423321 4133414334 1122142321 1414413113 4211412331 4411313244 2223323323 4224123223 2423212232 1312344141 2231441244 1112442111
3312242323 1422311133 2113214324 4111224421 4414223234 2112314441 3233322221 4142231114 4141324232 1443242343 11232331 1332212334 2321232323 2241144243 411442433
1234132423 4331232414 1143222221 1143332223 4323221132 3413444324 2214144323 4334141334 1441321423 1232434221 4133342222 3121222141 1431111112 3240232322 3411123244
2111114211 2234232434 2244133131 1331231331 1334112221 2213142422 1422223132 2113133412 1421232421 1421441141 2311313142 1222121142 2332211131 2232441343 3223423244
1231422222 4113222131 3213123323 1423232144 141212241 4324134324 2114211121 4141123342 31222

2322331412 3212244312 1234211121 1231211414 4241433214 2333114423 1322221231 3342123214 1414414121 4132422341 3342321332 3422323212 2323411224 1413142241 42243232343
3234423212 2133131441 2244212322 3411323423 4212314412 3241222411 4124141411 1221323223 2423221124 4421232223 4112344441 3322323123 2414231232 3141411423 21242123313
3212313421 2212234113 2414121134 4244412213 222441221 2323421123 3241332222 2332221214 111141412 3232133413 2314112324 1423142241 1434312134 1141232312 2331421123
4233431221 4121422344 1141422321 4231313142 4113422224 3413413111 4342441134 4142313311 1412221134 2324121233 2122344142 3214343113 2142424231 4131311223 4124222412
3331332233 1111442131 4414412231 1413111421 2324222142 3322144323 2114221111 2131211321 4131223313 3214144221 2311111332 3141422232 3214431332 1242143142
1411322311 2322332312 2142223232 2314124241 4142321123 4414221323 4144214242 1323441411 3134142123 2313112214 2334223214 4411414241 1412424131 2313222242 1341123223
4111221241 1113324334 1234111132 1424214142 4411331441 4423321423 3211313241 2222344124 3314223424 2421221444 1424324234 1213232121 4334123424 2412341342 3223123444
2123231111 3341221111 2122234441 2132111123 2132131124 1214423434 3234343432 1313232222 3121112241 1132311224 1141421311 4132123222 2423414232 1112341223 4413111411
4234231412 2321324222 3123111242 4212334214 2112212314 2231323223 2411234213 1422432341 1114414213 2421212233 4312442342 4112214413 4444443414 1223423144 1113341414
2131144131 1212222112 2413232341 3142212311 1141121332 4413141232 2442233114 232434432 2324423423 1112232231 4142323132 3422213222 1231411211 3421111341 4141133131
1324134241 1234224111 1232242141 1133232212 2141412312 3222412423 4112411413 1442411121 3343223211 3131122242 2141434112 4423422331 3111134232 2222124241 4422121122
4141311434 2321342241 1111242213 3122344214 3413142322 3122412414 1333224231 1212314212 3241212241 4141114214 2311141341 4142422321 1232411443 4224112323 3421241411
3334111112 2334122223 2341211232 4231213131 2422212314 1114234112 2214113242 1311121441 4324341123 2334214122 1411341311 1123211423 3142222221 3314232132 1131241221
4233242241 2442141413 3414421414 1442413432 2312121113 2241131423 1133122123 2323214121 4413341214 4111331412 3144214432 1314241242 4212312333 2224111123 1122144113
2411233221 4442112123 2221241444 3231131134 2122432414 3224323432 3122323421 3312321313 3423242143 4114213131 2342323332 1421144214 4242441342 331222242 1413232221
1422132134 232144312 3322123223 1414231111 2114413131 2241441313 2211234111 1442244141 3323144132 2211141242 2121441424 1232121141 3423444323 3112321323 43221221131
2132141112 3424132312 1414231113 3122324131 413124323 2243244222 3334341233 4222313422 1412411231 1321422232 4113122331 4423111311 4322323242 3232314111 4323411422
3422224214 2241241312 1422123442 3423311241 2242323332 2223142123 2422211444 1222123231 3123224134 1323333422 2143332114 4224131321 1441231441 4142234413 3214123311
1234141232 3342132232 3123421324 213213123 2221341324 4123132342 2322211343 4322312233 2423221112 1444212212 1331312311 342322321 1224424223 4131111321 414141233
2323211324 2341232422 3423121132 2244132442 2412314122 3311124411 1334212411 3232323442 4133423113 2131123311 3224214312 3424134412 3213231411 1342214111 1213133114
2422412231 2212342141 4124143121 1123412423 2332222331 2332222331 4413441123 4112232323 2133222314 14113313242 2422211422 3141313242 2422211422 3141313242 2322323123
3232313222 3212323112 1134123412 2441414244 1342224312 1111232412 2214211222 1113411223 4422442222 1421224231 2132442232 312424123 3411141132 1323243312 1223122123
2244142441 1242342143 2321221414 3123411222 1332332342 3222343134 1441234111 442423411 4132144431 4132144132 2211141242 2121441424 1232121141 3423444323 3112321323 43221221131
4221211442 3233112141 2221441332 2141313222 1311112323 3422121123 2232211211 1331414222 3123314143 4433431134 1442322244 1241221121 1234232221 4141113321 3123321223
1232331312 4222331111 1121131411414 2244232211 2322241222 3134222333 4211343211 4122111441 323441131 2211312142 411212223 3323232312 2434242421 2142321124 2333142141
1113131134 4441341412 2124214111 2132323423 4141123142 3141223441 1311144124 213132342 232242414 2214232422 3222312332 3141332131 3141332131 4141234123
1233232222 1121141311 3322413241 1133242114 1442241113 3314413414 1134213432 4112331324 1412211223 3131124132 1111141412 222142142 221311132 1314334122 1224114132
2344231114 4214222142 4131341313 1422131331 2222113444 1312414111 2331114231 4222424313 4243421233 1231122412 1413431212 1134243133 3421234422 4311414124 2222123341
1222413241 2331411422 3343324224 1222112312 1334221344 1134342414 1131224144 1422111221 3133122221 2444212331 4134141323 1111122142 241214131 2343411141 3322412411
1123232323 224232122 3124222142 2231314222 2324313121 1211132414 1131222324 1323122241 1113232241 133224112 2314212342 2111222412 2412342223 3422113332 3314123233
1114224241 4441111421 4113411324 1134233441 1441332322 3431234421 1422234141 3123411131 1254424223 1111323231 4422413112 2123233111 1234232212 1441113321 1212312321
1232434111 321214241 4232344112 1422111214 1422142221 3423214442 1411412313 1322122414 2233232423 3424231211 1331322341 4413321413 4232142413 4112323223 44114233411
4142142232 4331423142 4221123142 21234244431 4341312242 1422211223 3131423234 1141321414 1433421211 1212331442 212212134 321232123 1444411314 4111412323 1411113131
1231424211 1311333222 2241422124 2332333122 2121113411 1114213444 233233224 2121342321 1131211413 1341224223 3112342314 4222141134 2212334441 2224241233 2331413422
2323241111 2321112341 2414412433 1412412222 1122333342 1142124223 3411344311 3311411122 3123342141 4412341234 2114214122 2242331331 1411111244 2124413223 2422442122
142422231 412314111 2224442422 1122113133 1142342321 1232424221 133131241 112412312 123122212 4242242331 111332133 4242242331 111332133 4141313223 3222411323
2121224232 2114114211 1431141232 221222422 1221222422 1424241421 3233213231 4422314433 2341441422 2231231232 2411342442 1122141424 1141214224 1242123133 3332221422
1244231231 11212222131 2413141424 2221222114 2413312221 222424412 2222122411 2322141333 2224124231 2221414221 2211241214 3222314114 4423241214 4222211331 2332122212 2324234341
4131311124 1112323244 4323232323 2222233224 2144123114 1141132414 1324323212 2122122411 1124221421 1232323212 1431341232 1142223112 21232323141 4442224222 3422224221
1144412333 2141413141 4124123121 2124231111 2341222112 2341422243 132412124 2312344222 2434223123 2224131111 3241332112 2221114142 2143131322 2323222111 1122113123
3322242421 2322212113 3342124114 4121213123 1324232312 4214422423 3222211414 2223111442 1222221132 2344242233 2122111122 1411122221 3134412242 2141344222 2214423424
1121412223 3334143241 3423413442 1222232233 4212221113 4411421112 2212342411 3212312311 1123321411 2414131234 2322312232 2424111124 4142234231 1424111234 2432321131
3422231231 231213131 1231122214431 4424211222 4133223232 4442413313 1232232313 1224442112 1221411142 2323244222 2331313222 2221123421 12112244 2422324212 3412333131
3311111212 2242112232 32234234212 1424232342 4214114241 1412422222 4114422211 3234212334 3222423441 4211323231 2312422234 2424441123 242123431 2422223231 3241133313
1231233123 141112343 4413314111 11423323212 2342312212 112121112 2344113121 2224123342 2322444223 4223311214 2133114122 3123421422 4114233211 2212112212 112121332
1422213141 233413422 1334442223 1313131412 1242112323 442242444 4231233411 1442234144 44142423424 2211223221 22334441221 1113122123 3222221414 4142132324 1223421444
114421413 1322321321 4224224223 1142222111 2341443344 1433222323 3144111122 1444142314 422122131 1441323232 4112221212 2221123314 2414232124 2332323221 1422424222
4111313122 2411441222 4114112223 3443231232 2242312124 1423122443 2243131244 2111141242 3111212213 1412224144 2234242222 3333323211 3232112214 1442331141 1123422412
4231132234 232421212 4113123221 1233414231 2322442141 3411224123 4131241142 4141441342 23344341212 1224214233 1114112321 2244131424 1134133423 24422121331 3423134123
2431141224 2212131324 3141224421 2233144231 242223242 1322144121 4132224224 1221432132 1421344111 1124212232 3442212142 123422213 323423123 2241122342 2142421121
141412234 2121442142 1142222211 2213412214 2141311321 1234423222 4212334211 1144244341 312232121 2321412423 4331312242 1423112323 4211222342 2444411212 3321342222
343414112 112432341 4244413321 2234232231 14131312434 1141414323 1132322434 2413132411 4111232411 1144444114 2223111312 2223233222 221122224 1411122121 4423144242
211344114 211313242 2221223133 1422342422 2112113323 1414243423 1234234242 4123134411 213344112 3442133231 2122112342 1313124241 2223423441 311324241 3113242143
14423144 2442342142 3123121231 2234234421 4142342342 4213221323 2134214414 2342331233 14142223323 4233434234 2424411121 44213332 32342422 2343122421 2341112334 1212133224 1123341223
3132234441 1112322424 2123131144 111141112 4344114411 2221141133 4312213241 1122411311 2414221411 4114124114 2341223334 1134234113 1432232211 4413112342 4113412211
1414142232 2111411422 1342421122 223232412 22241132323 4211123121 2442242443 1422312111 3243134343 1332141112 2314133221 1224143341 3142241131 4422411144 2123211414
1141142112 1133342342 3311241332 2422342411 2422342342 1421111232 4231323232 2422331231 2441111313 2321422233 4331113334 2144144221 2324242444 1413413431 2233122323
232422442 1412324143 2123342412 4441331313 3114224144 1413434413 4112223441 2311334111 3412142244 4243231133 2322222323 222123341 4242232322 2122141413 1331142242
412414141 1121122142 1413141121 4122232424 1331112232 3123243221 2233412412 211411441 3222123124 4242341144 2313422314 2313422314 222111313 1313422314 1111142121
412414431 2322312422 3343141222 41341221 4342242324 3241141224 1131131112 2314434424 1231332423 4211111222 1112311241 4241223144 22411332 122 3331111212 3134112121
2121123344 2233212322 2332423432 1421122344 1243241233 224112341 412112422 3422232234 12342132414 1421344124 4221423434 2424113332 3412342422 333242223 4243113222 312313123
1424231344 1322424113 3323241112 2141221414 1211143114 2212331222 3222123131 1222221424 2233124112 2142341413 4413211111 2231411234 2242442341 1321423224 1234214134
1143121443 4214113442 4142444222 4331233422 3221223343 2234122242 2342322443 42131343421 411112112 2343424112 1111312334 1342134442 2212113334 2244132241 2213414122
1422234241 222321441 2421224212 2422323422 3243423413 3212331334 34122222 2144242123 4422312121 222421411 2341241143 413321312 2122212312 2322334222 3243211323
411124413 1144412214 2221231234 4114212222 3222342334 1414141211 4212314214 232124412 1211413232 4141312141 1223243221 4231133231 4334122241 1444411114 2411131313
4214411122 3323222122 4231331213 3414241332 3221113142 1232432232 1412334212 2241443113 2221421441 3422313242 1414111221 4214131414 2332122311 1123421412 2422412422
211413422 3221113144 4144311232 3411121414 4111243143 3424224133 1322342112 4413411233 1244112221 2312313443 1431431421 4213423444 2223211341 4113111223 13214131
2123214241 2423222423 1133111413 2143211121 3221121343 2142242222 242324224 323212242 3242132414 1421344124 4221423434 2424123334 22412323 234321424 2341112323 212142313
2221321111 4411433222 1432421424 343232211 1124231142 4224231133 1322442312 4224233242 1114311421 1421221221 2212242322 2331142221 1412234221 2412334131 4114132142
1221443223 1331342211 1411424423 3114112331 1442132112 1332224421 2244213442 3312212322 1421411222 3442122421 1243311213 3222223214 1211211134 223222442 4233424223
2334413211 41422234223 4222323422 3423423242 1413112211 223323142 3343331223 4312234112 3342232122 3213342231 2323342213 4214224322 3422442421 3342322234 2432144434
2411232332 4122411444 1311141114 1224224241 1232344242 4433333141 1223231441 3144342412 3311133311 4223241223 3234232312 4122432332 4142234331 2334433141 42444421333
4342323141 2211331442 3441111313 1241423233 4214124143 1143422242 2122141112 2133134324 3413432321 3341423122 4133222234 2341214114 4342144321 2233111421 211221112
1413144143 4433213411 4314112122 3422234213 2141321132 3141141312 3222412422 3223142244 1332214112 1341313121 2231221441 2212313314 222413312

II. ON THE NOTION OF PERCENTAGE NUCLEOTIDE CONCENTRATION OF GENOME

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. Now, pA, pT, pG and pC were calculated for all nine virus genome data (i) AY304488.1, (ii) AY395003, (iii) DQ412043, (iv) KY417144.1, (v) MN908947.1, (vi) MN908947.2, (vii) MN908947.3, (viii) Mg772933 and (ix) Mg772934 and the results reported here. Basic details of all the nine virus genomes are given below.

Genome title	GenBank ID	NCBI Archive Date	Sequence length	Number of Adenines	Number of Adenines	Number of Adenines	Number of Adenines
SARS coronavirus SZ16 (Genome 1)	AY304488.1	05-NOV-2003	29731	8461	9131	6191	5948
SARS coronavirus ZS-C (Genome 2)	AY395003.1	29-JAN-2004	29647	8454	9106	6167	5920
Bat SARS coronavirus Rml (Genome 3)	DQ412043.1	13-JUL-2006	29749	8499	9019	6242	5989
Bat SARS-like coronavirus isolate Rs4084 (Genome 4)	KY417144.1	18-DEC-2017	29770	8487	9094	6149	5995
Wuhan seafood market pneumonia virus isolate Wuhan-Hu-1 (Genome 5)	MN908947.1	14-JAN-2020	30473	9115	9802	5951	5605
Wuhan seafood market pneumonia virus isolate Wuhan-Hu-1 (Genome 6)	MN908947.2	18-MAR-2020	29875	8921	9592	5860	5502
Severe acute respiratory syndrome coronavirus 2 isolate Wuhan-Hu-1 (Genome 7)	MN908947.3	18-MAR-2020	29903	8954	9594	5863	5492
Bat SARS-like coronavirus isolate bat-SL-CoVZC45 (Genome 8)	MG772933.1	05-FEB-2020	29802	8740	9468	6020	5574
Bat SARS-like coronavirus isolate bat-SL-CoVZXC21 (Genome 9)	MG772934.1	05-FEB-2020	29732	8734	9455	5976	5567

Fig. 1: shows the tabular and graphical display of the percentage nucleotide concentrations of nine viruses.

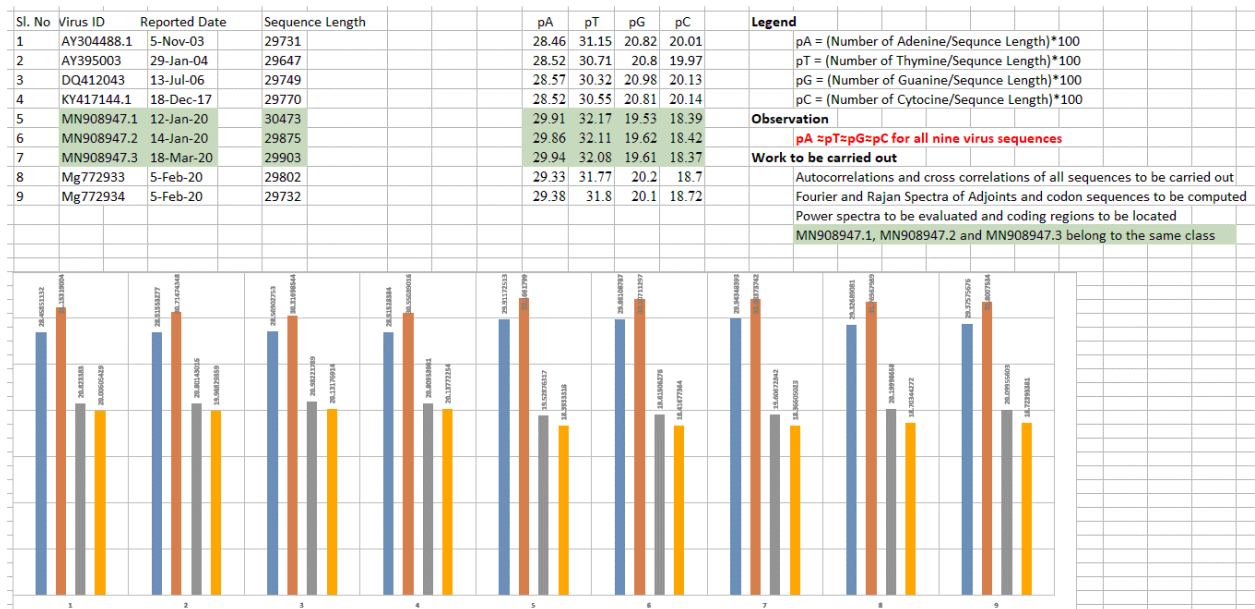


Fig. 1: Tabular and graphical display of the percentage nucleotide concentrations of nine viruses

Important Observations

1. pA, pT, pG and pC are approximately the same for all nine virus genome sequences
2. AY304488.1, AY395003, DQ412043 and KY417144.1 virus genomes form one class
3. MN908947.1, MN908947.2 and MN908947.3 virus genomes form one class
4. Mg772933 and Mg772934 virus genome form one class
5. T

III. ON THE NOTION OF ADJOINTS OF GENOME SEQUENCES AND THEIR ARRAY REPRESENTATIONS

Adjoint of a particular nucleotide in a genome sequence is the binary sequence obtained by substituting the particular nucleotides in the genome sequence by 1's and the others by 0's. For example, let us consider a sample characteristic sequence of GAATGATTACCAAGGCCAAGCTCAAGCTCTCCTTCCTTGG. Now the four adjoints of this sample sequence of length 40 are (i) $A(n) = 0110010010011000011000011000000000000000$, (ii) $T(n) = 00010011000000000000001000001010011001100$, (iii) $G(n) = 1000100000000110000100000100000000000011$ and (iv) $C(n) = 0000000001100001100010100010101100110000$. Adjoints of a genome sequence play a significant role in studying the structural and functional properties of genomes. The biological aspects of these sequences are out of the scope of this paper and hence not discussed. Figs. 2-10 respectively portray the array images of the genomes AY304488.1, AY395003, DQ412043 and KY417144.1, MN908947.1, MN908947.2, MN908947.3, Mg772933 and Mg772934 and their adjoints arrays. All the arrays are of width of 128 and height of (LENGTH/128)+required number of 0's to complete the last row. For example, the genome 1: (AY304488.1) has a LENGTH of 29731. Its corresponding sequence array will have a width of 128 and height of 238. Four subarrays of the adjoints A(n), T(n), G(n) and C(n) of size 128x238 are also shown in Figs. 2-10. A color code has been assumed here as red for adenine, green for thymine, yellow for guanine and blue for cytosine. That is how a genome array, AY304488.1 for instance, is a multicolored array of size 128x238, and its adjoints arrays are in respective colors. Color coding has been given to arrays in order to discriminate one from the rest. (Adenine is represented in red color, Thymine in green, Guanine in yellow and Cytosine is represented in blue color.)

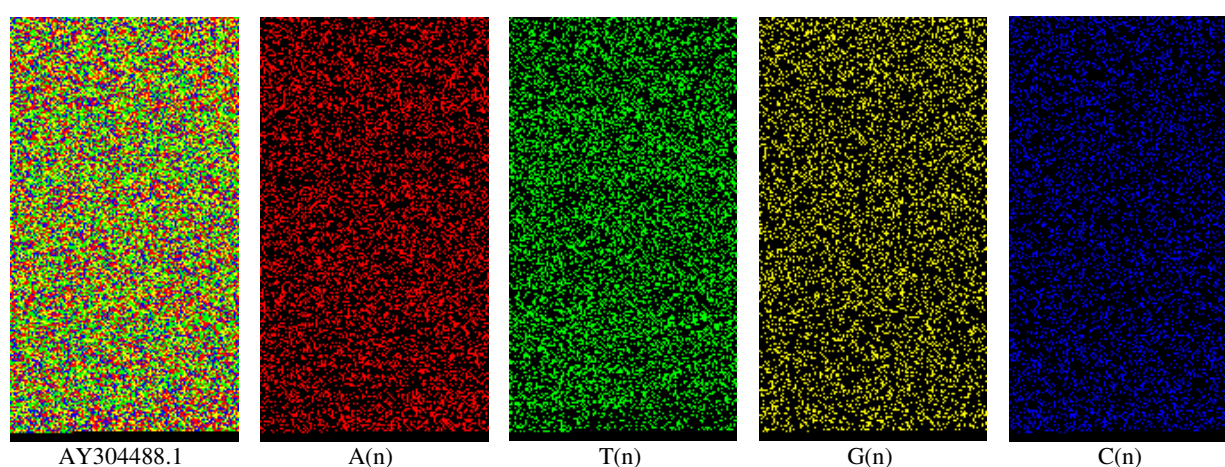


Fig. 2: Genome 1: AY304488.1 and its adjoints

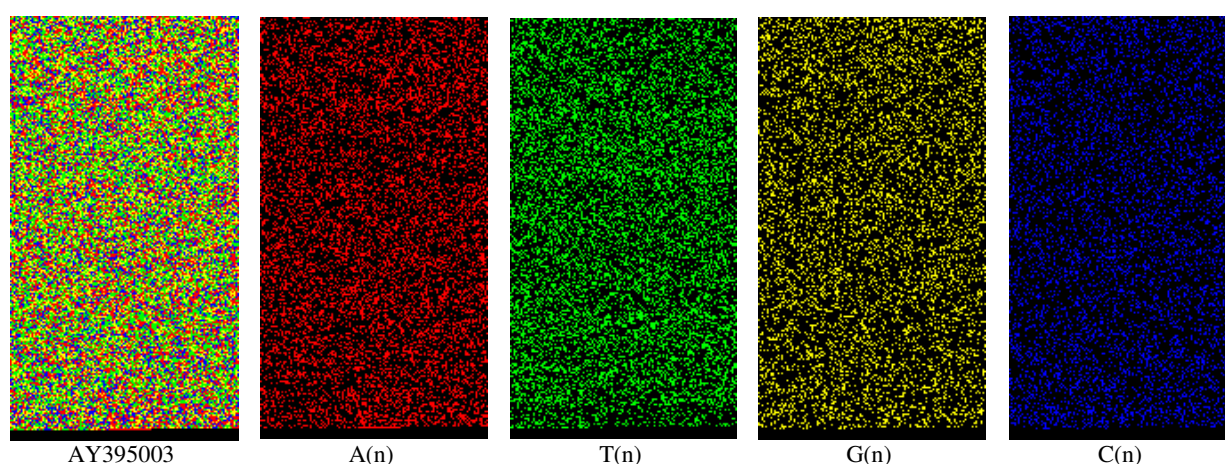


Fig. 3: Genome 2: AY395003 and its adjoints

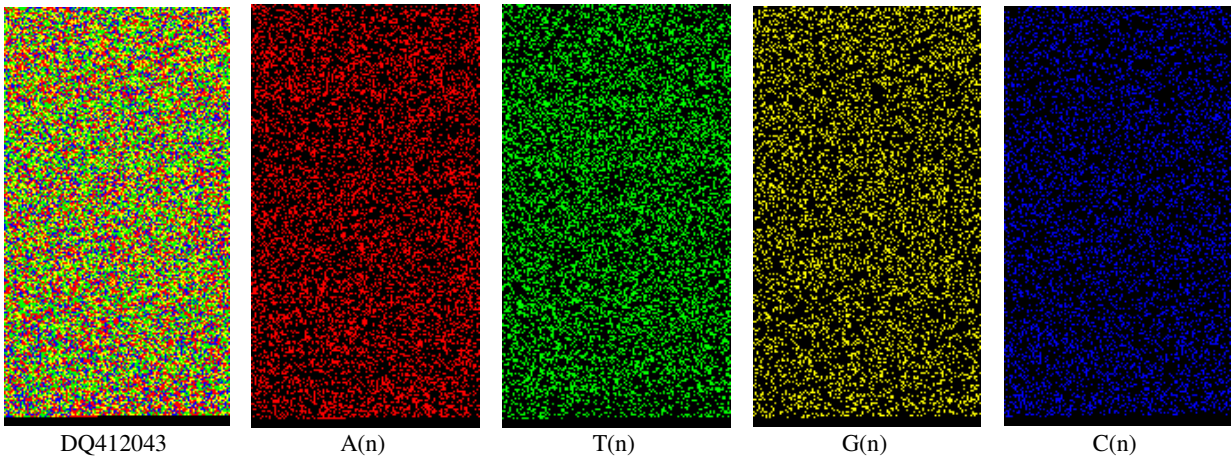


Fig. 4: Genome 3: DQ412043 and its adjoints

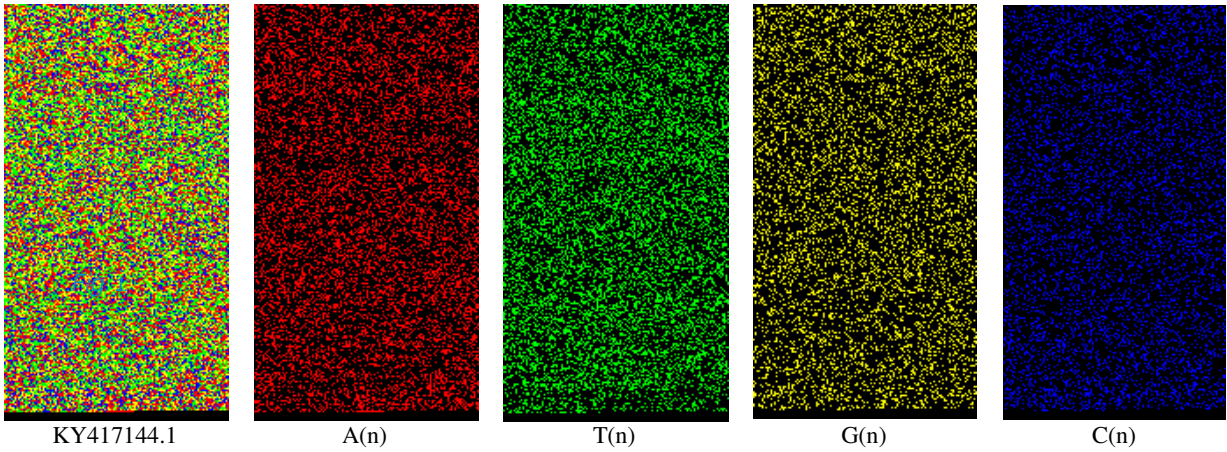


Fig. 5: Genome 4: KY417144.1 and its adjoints

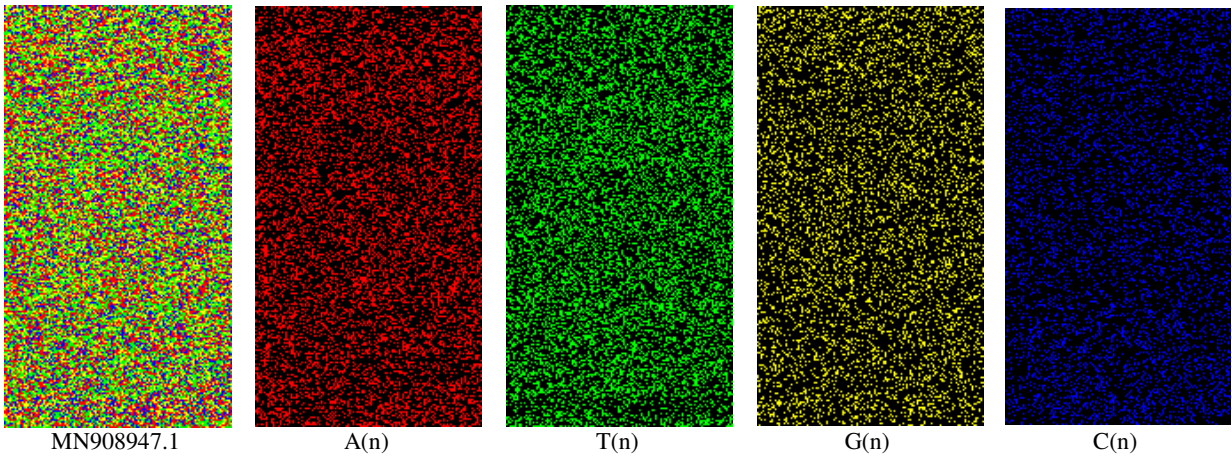


Fig. 6: Genome 5: MN908947.1 and its adjoints

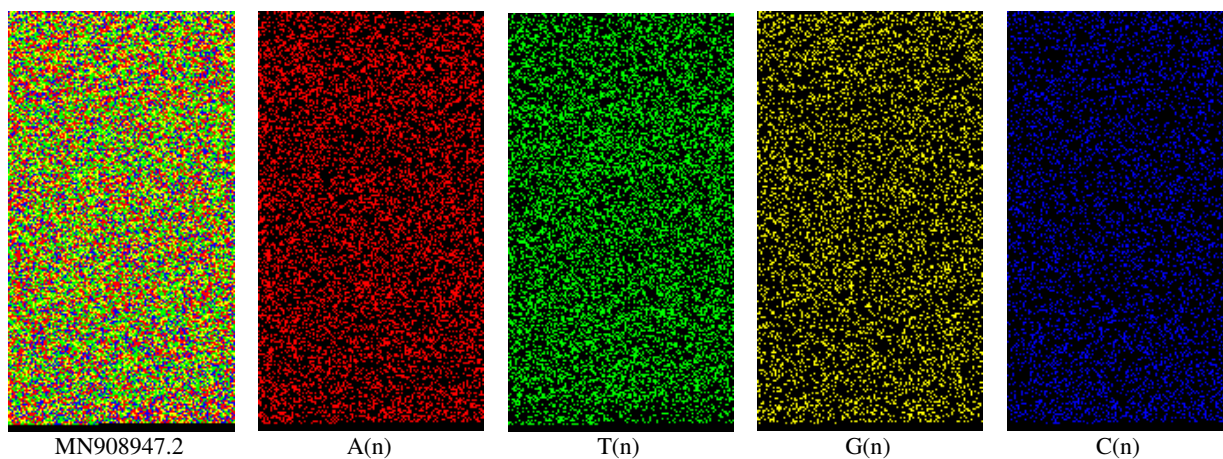


Fig. 7: Genome 6: MN908947.2 and its adjoints

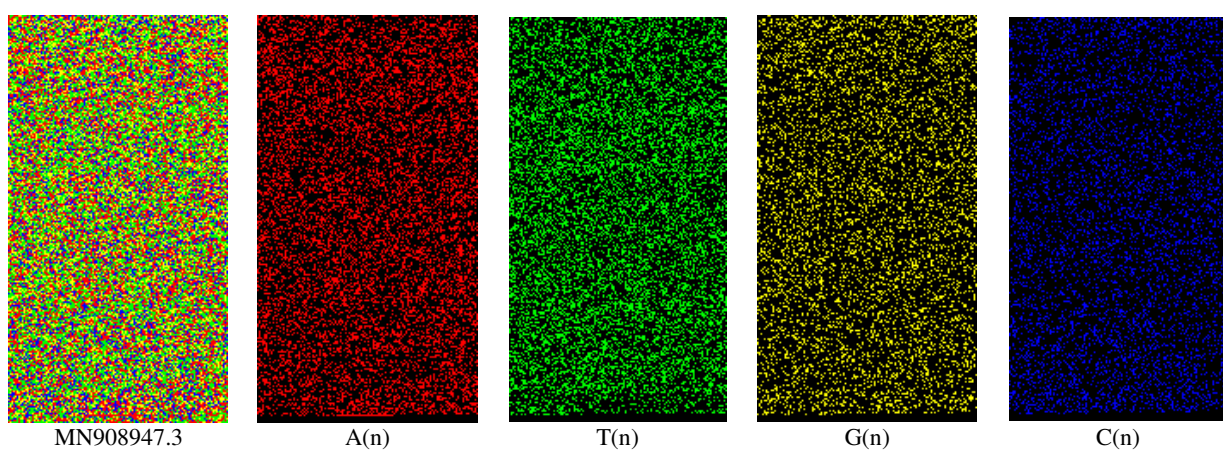


Fig. 8: Genome 7: MN908947.3 and its adjoints

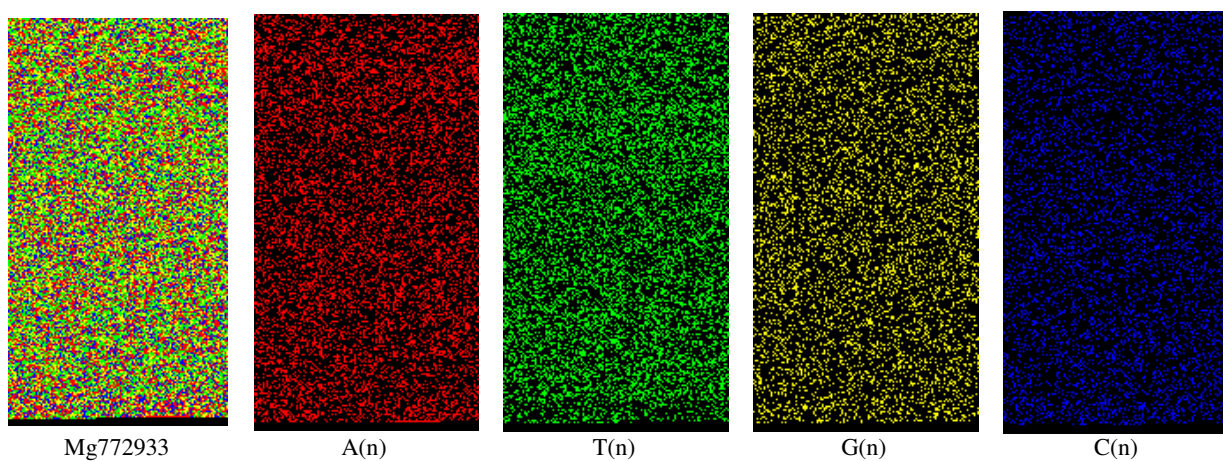


Fig. 9: Genome 8: Mg772933 and its adjoints

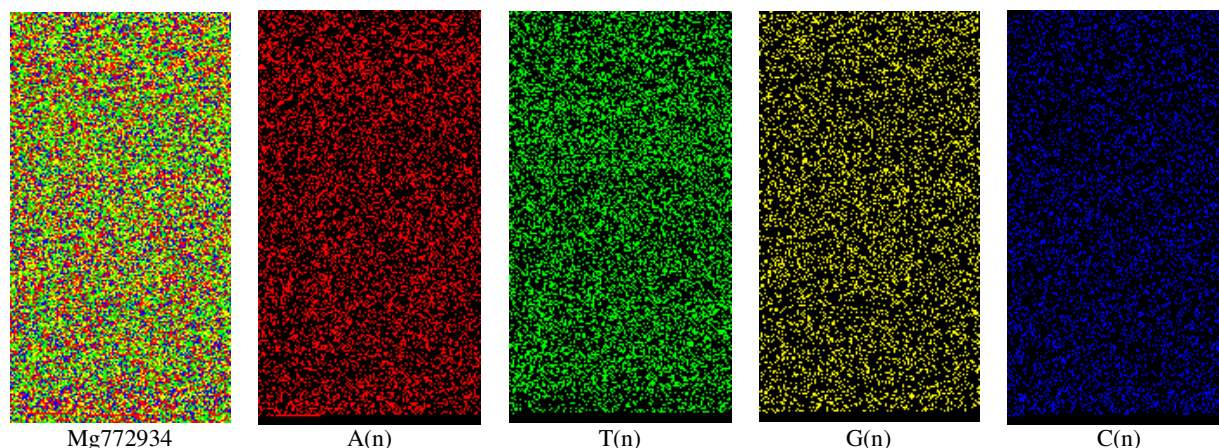


Fig. 10: Genome 9: Mg772934 and its adjoints

Total number of nucleotides in each genome

The total number of nucleotides in the genome is the total number of A, T, G, and C respectively. These are listed below along with the graph.

Genome	Genome ID	Total no. of A	Total no. of T	Total no. of G	Total no. of C
1	AY304488.1	8461	9131	6191	5948
2	AY395003	8454	9106	6167	5920
3	DQ412043	8499	9019	6242	5989
4	KY417144.1	8487	9094	6149	5995
5	MN908947.1	9155	9802	5951	5605
6	MN908947.2	8921	9592	5860	5502
7	MN908947.3	8954	9594	5863	5492
8	Mg772933	8740	9468	6020	5574
9	Mg772934	8734	9455	5976	5567

Percentage Nucleotide Concentration

Percentage nucleotide concentration is the percentage of each nucleotide in the entire genome.

Genome	Genome ID	p(A)	p(T)	p(G)	p(C)
1	AY304488.1	28.45	30.71	20.82	20.00
2	AY395003	28.51	30.71	20.80	19.96
3	DQ412043	28.56	30.31	20.98	20.13
4	KY417144.1	28.50	30.54	20.65	20.13
5	MN908947.1	30.04	32.16	19.52	18.39
6	MN908947.2	29.86	32.10	19.61	18.41
7	MN908947.3	29.94	32.08	19.60	18.36
8	Mg772933	29.32	31.76	20.19	18.70
9	Mg772934	29.37	31.80	20.09	18.72

Fig. 11 shows number of nucleotides and their % concentration in all nine genomes.

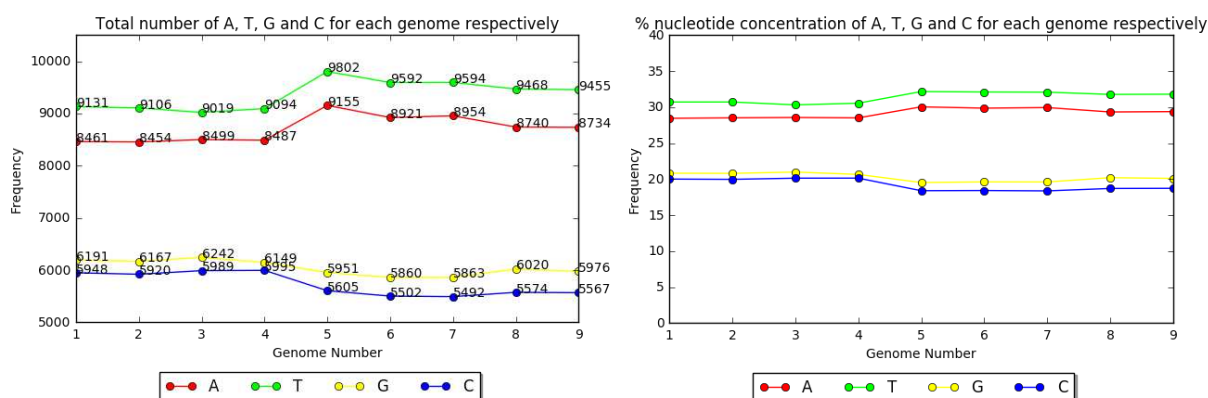


Fig. 11: Total number of nucleotides in nine genome sequences and their % concentration

From the above graph and data, it is evident that the % concentration of the nucleotides is quite similar for all genomes. Since the line graphs for percentage nucleotide concentration are fairly straight lines almost parallel to the x-axis, a relationship between the data exists. This relationship can establish the evolutionary sequence of the family of coronaviruses. **An important observation made from the above graph and data is the following. The nine % concentrations of the base pairs A and T and the % concentration of the base pairs G and C form a lateral symmetry with respect to the average value, in this case it is 25%. The % concentrations of A and those of G seem to form a dual pair with respect to the average value. Similarly, the % concentrations of T and those of C seem to form a dual pair with respect to the average value. This observation calls for genetic interpretation by peers.**

Changes in percentage nucleotide concentrations

Let us see for each genome, what is the change in percentage nucleotide concentration has occurred. The change in percentage nucleotide concentration can give a statistical overview of how the percentage distribution of nucleotides changes from one genome to another. A positive value in the following graphs depicts an increase in the concentration from Genome X to Genome Y. A negative value depicts a decrease in the concentration from Genome X to Genome Y

Differences in % nucleotide concentrations with respect to Genome 1

Genome (X, Y)	$p(A_x) - p(A_y)$	$p(T_x) - p(T_y)$	$p(G_x) - p(G_y)$	$p(C_x) - p(C_y)$
(1, 1)	0	0	0	0
(1, 2)	0.05702145	0.00269209	-0.02195284	-0.0377607
(1, 3)	0.11051621	-0.39506595	0.15883489	0.12571485
(1, 4)	0.05005435	-0.16452032	-0.16836117	0.13166825
(1, 5)	1.58447756	1.4541285	-1.29461983	-1.61272248
(1, 6)	1.40257655	1.39506158	-1.20832024	-1.58931789
(1, 7)	1.48497261	1.37168602	-1.21665458	-1.64000406
(1, 8)	0.86837949	1.05762849	-0.62339642	-1.30261156
(1, 9)	0.91724544	1.088702	-0.72382697	-1.28212048

Fig. 12 shows the percentage change in nucleotide % concentration values for Genome 1 with all nine genome sequences. Fig. 12 also shows the absolute (difference) values. These values are used in deciding the evolutionary nature of virus genomes. Evolution of virus genomes is an involved study and it is not in the scope of this research document. Most of the virus genomes are active, that is they do not change their structures, in the sense that their nucleotide % concentration will not change drastically. However, their functional behavior depends on the distributions of their nucleotides. If the nucleotides % concentrations and their distribution patterns remain unchanged, then the viruses are passive in nature and an antidote or a vaccine would indeed deactivate them. On the other hand, if the nucleotide distribution patterns change in spite of their nucleotide % concentrations remaining same, they are classified as active, in the sense that they are evolutions from earlier ones. If the evolution pattern obeys a rule, then it is easy to predict future evolutions, where as if the evolution pattern is random, then it is difficult to predict future evolutions. If the random evolution is natural then it is almost impossible to control the pandemic nature of the virus. If the random evolution is empirical, that is the virus is modeled in a laboratory (God forbid such things should not happen), it is easy to predict and decrypt the evolutions by treating the evolutions as a random process and one can create in the laboratory an antidote, also termed as antiviral, which would act as a standard genetic vaccine. This antiviral would also change its evolutions, couple with any type of evolutionary rogue virus and deactivate it jointly. In order to achieve this, one has to carry out involved research with the help of theoretical scientists and genetic experts who work in wet laboratories.

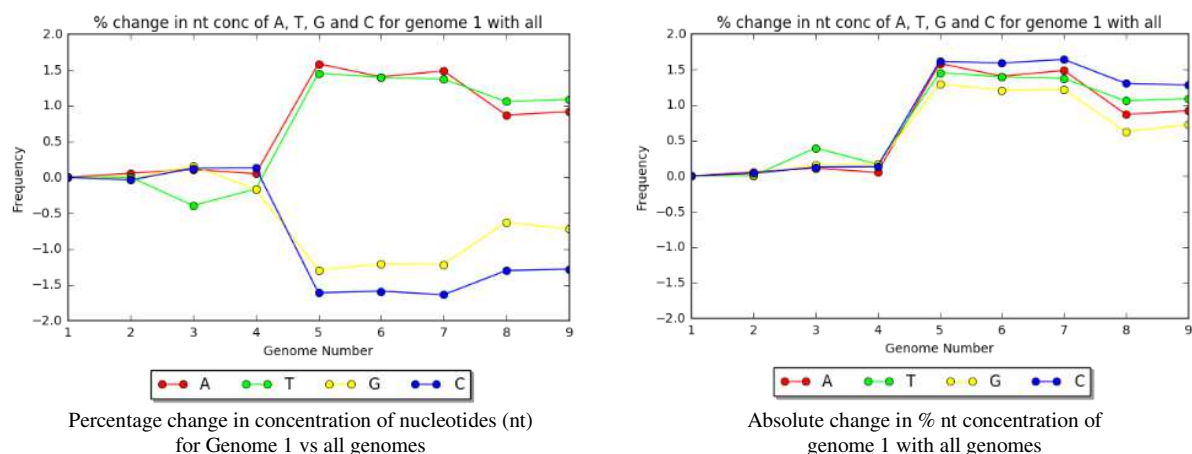


Fig. 12: Percentage change in concentration of nucleotides (nt) for Genome 1 vs all genomes and their differences

This exercise is repeated for all genome sequences and results presented below.

Differences in % nucleotide concentrations with respect to Genome 2

Genome (X, Y)	$p(A_x) - p(A_y)$	$p(T_x) - p(T_y)$	$p(G_x) - p(G_y)$	$p(C_x) - p(C_y)$
(2, 1)	-0.05702	-0.00269	0.021953	0.037761
(2, 2)	0	0	0	0
(2, 3)	0.053495	-0.39776	0.180788	0.163476
(2, 4)	-0.00697	-0.16721	-0.14641	0.169429
(2, 5)	1.527456	1.451436	-1.27267	-1.57496
(2, 6)	1.345555	1.392369	-1.18637	-1.55156
(2, 7)	1.427951	1.368994	-1.1947	-1.60224
(2, 8)	0.811358	1.054936	-0.60144	-1.26485
(2, 9)	0.860224	1.08601	-0.70187	-1.24436

Fig. 13 shows the percentage change in nucleotide % concentration values for Genome 2 with all nine genome sequences. Fig. 13 also shows the absolute (difference) values.

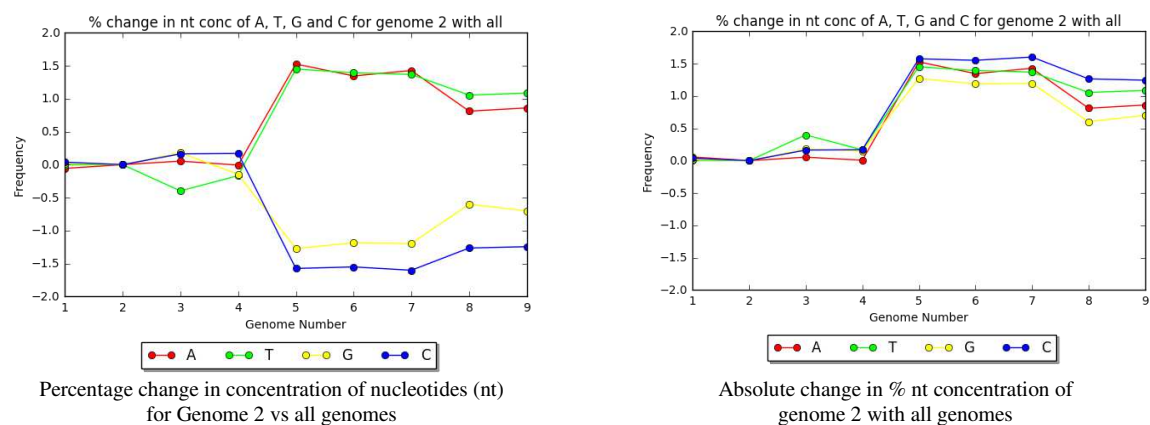


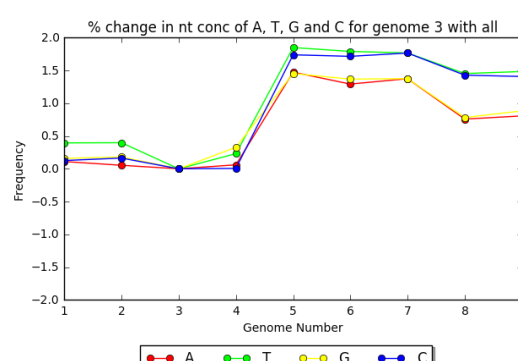
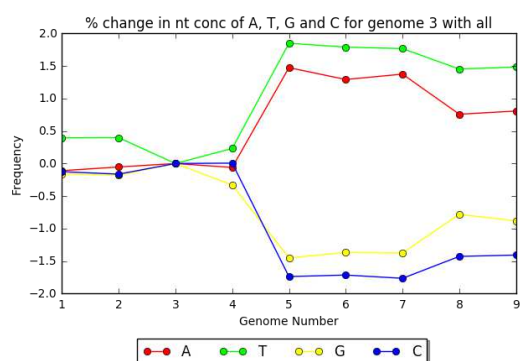
Fig. 13: Percentage change in concentration of nucleotides (nt) for Genome 2 vs all genomes and their differences

Genome sequences 3, 5, 6 and 7 show small differences from those of Genome 1.

Differences in % nucleotide concentrations with respect to Genome 3

Genome (X, Y)	$p(A_x) - p(A_y)$	$p(T_x) - p(T_y)$	$p(G_x) - p(G_y)$	$p(C_x) - p(C_y)$
(3, 1)	-0.11052	0.395066	-0.15883	-0.12571
(3, 2)	-0.05349	0.397758	-0.18079	-0.16348
(3, 3)	0	0	0	0
(3, 4)	-0.06046	0.230546	-0.3272	0.005953
(3, 5)	1.473961	1.849194	-1.45345	-1.73844
(3, 6)	1.29206	1.790128	-1.36716	-1.71503
(3, 7)	1.374456	1.766752	-1.37549	-1.76572
(3, 8)	0.757863	1.452694	-0.78223	-1.42833
(3, 9)	0.806729	1.483768	-0.88266	-1.40784

Fig. 14 shows the percentage change in nucleotide % concentration values for Genome 3 with all nine genome sequences. Fig. 14 also shows the absolute (difference) values.



Percentage change in concentration of nucleotides (nt) for Genome 3 vs all genomes

Absolute change in % nt concentration of genome 3 with all genomes

Fig. 14: Percentage change in concentration of nucleotides (nt) for Genome 3 vs all genomes and their differences

Genome sequence 3 shows considerable differences from all other sequences.

Differences in % nucleotide concentrations with respect to Genome 4

Genome (X, Y)	$p(A_x) - p(A_y)$	$p(T_x) - p(T_y)$	$p(G_x) - p(G_y)$	$p(C_x) - p(C_y)$
(4, 1)	-0.05005	0.16452	0.168361	-0.13167
(4, 2)	0.006967	0.167212	0.146408	-0.16943
(4, 3)	0.060462	-0.23055	0.327196	-0.00595
(4, 4)	0	0	0	0
(4, 5)	1.534423	1.618649	-1.12626	-1.74439
(4, 6)	1.352522	1.559582	-1.03996	-1.72099
(4, 7)	1.434918	1.536206	-1.04829	-1.77167
(4, 8)	0.818325	1.222149	-0.45504	-1.43428
(4, 9)	0.867191	1.253222	-0.55547	-1.41379

Fig. 15 shows the percentage change in nucleotide % concentration values for Genome 4 with all nine genome sequences. Fig. 15 also shows the absolute (difference) values.

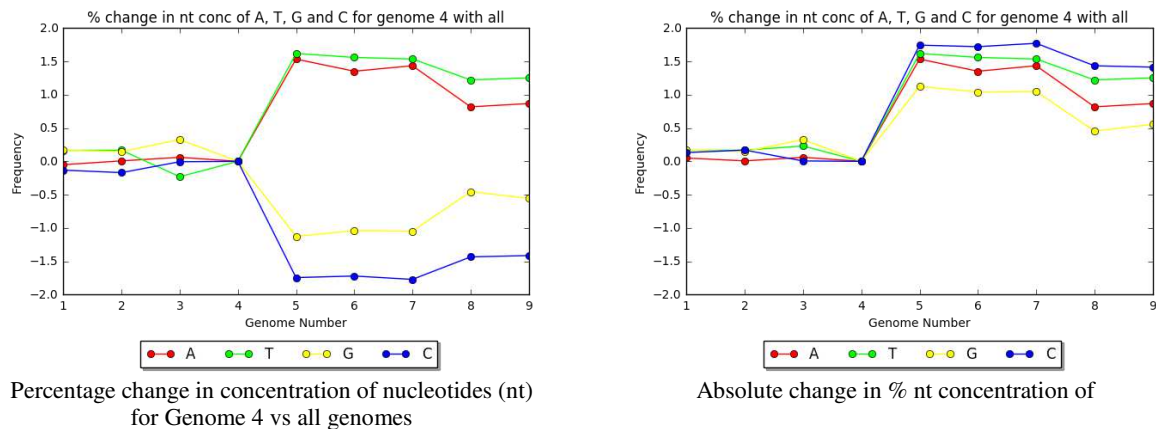


Fig. 15: Percentage change in concentration of nucleotides (nt) for Genome 4 vs all genomes and their differences. Genome sequence 4 shows considerable differences from all other sequences.

Differences in % nucleotide concentrations with respect to Genome 5

Genome (X, Y)	$p(A_x) - p(A_y)$	$p(T_x) - p(T_y)$	$p(G_x) - p(G_y)$	$p(C_x) - p(C_y)$
(5, 1)	-1.58448	-1.45413	1.29462	1.612722
(5, 2)	-1.52746	-1.45144	1.272667	1.574962
(5, 3)	-1.47396	-1.84919	1.453455	1.738437
(5, 4)	-1.53442	-1.61865	1.126259	1.744391
(5, 5)	0	0	0	0
(5, 6)	-0.1819	-0.05907	0.0863	0.023405
(5, 7)	-0.0995	-0.08244	0.077965	-0.02728
(5, 8)	-0.7161	-0.3965	0.671223	0.310111
(5, 9)	-0.66723	-0.36543	0.570793	0.330602

Fig. 16 shows the percentage change in nucleotide % concentration values for Genome 5 with all nine genome sequences. Fig. 16 also shows the absolute (difference) values.

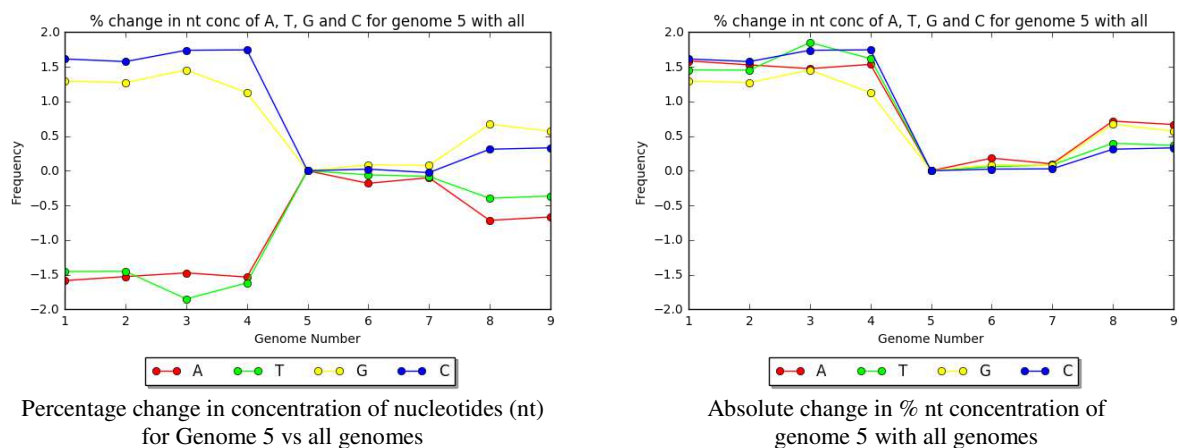


Fig. 16: Percentage change in concentration of nucleotides (nt) for Genome 5 vs all genomes and their differences. Genome sequence 5 shows almost a reverse characteristics when compared to all other Genome sequences 1, 2, 3, 4.

Differences in % nucleotide concentrations with respect to Genome 6

Genome (X, Y)	$p(A_x) - p(A_y)$	$p(T_x) - p(T_y)$	$p(G_x) - p(G_y)$	$p(C_x) - p(C_y)$
(6, 1)	-1.40258	-1.39506	1.20832	1.589318
(6, 2)	-1.34556	-1.39237	1.186367	1.551557
(6, 3)	-1.29206	-1.79013	1.367155	1.715033
(6, 4)	-1.35252	-1.55958	1.039959	1.720986
(6, 5)	0.181901	0.059067	-0.0863	-0.0234
(6, 6)	0	0	0	0
(6, 7)	0.082396	-0.02338	-0.00833	-0.05069
(6, 8)	-0.5342	-0.33743	0.584924	0.286706
(6, 9)	-0.48533	-0.30636	0.484493	0.307197

Fig. 17 shows the percentage change in nucleotide % concentration values for Genome 6 with all nine genome sequences. Fig. 17 also shows the absolute (difference) values.

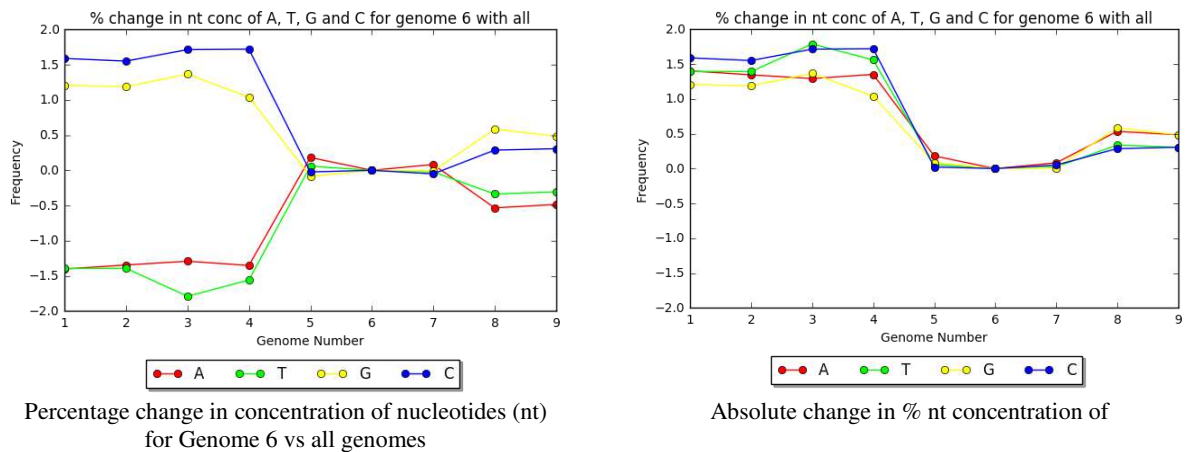


Fig. 17: Percentage change in concentration of nucleotides (nt) for Genome 6 vs all genomes and their differences

Like Genome sequence 5, Genome sequence 6 also shows almost a reverse characteristics when compared to other Genome sequences 1, 2, 3, 4.

Differences in % nucleotide concentrations with respect to Genome 7

Genome (X, Y)	$p(A_x) - p(A_y)$	$p(T_x) - p(T_y)$	$p(G_x) - p(G_y)$	$p(C_x) - p(C_y)$
(7, 1)	-1.48497	-1.37169	1.216655	1.640004
(7, 2)	-1.42795	-1.36899	1.194702	1.602243
(7, 3)	-1.37446	-1.76675	1.375489	1.765719
(7, 4)	-1.43492	-1.53621	1.048293	1.771672
(7, 5)	0.099505	0.082442	-0.07797	0.027282
(7, 6)	-0.0824	0.023376	0.008334	0.050686
(7, 7)	0	0	0	0
(7, 8)	-0.61659	-0.31406	0.593258	0.337392
(7, 9)	-0.56773	-0.28298	0.492828	0.357884

Fig. 18 shows the percentage change in nucleotide % concentration values for Genome 7 with all nine genome sequences. Fig. 18 also shows the absolute (difference) values.

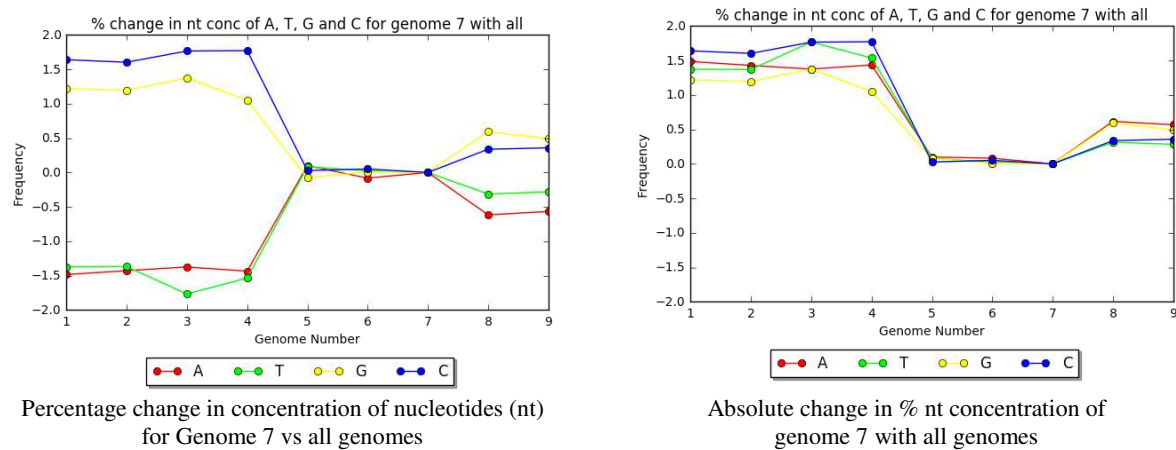


Fig. 18: Percentage change in concentration of nucleotides (nt) for Genome 7 vs all genomes and their differences

Like Genome sequences 5 and 6 Genome sequence 7 also shows almost a reverse characteristics when compared to other Genome sequences 1, 2, 3, 4.

Differences in % nucleotide concentrations with respect to Genome 8

Genome (X, Y)	$p(A_x) - p(A_y)$	$p(T_x) - p(T_y)$	$p(G_x) - p(G_y)$	$p(C_x) - p(C_y)$
(8, 1)	-0.86838	-1.05763	0.623396	1.302612
(8, 2)	-0.81136	-1.05494	0.601444	1.264851
(8, 3)	-0.75786	-1.45269	0.782231	1.428326
(8, 4)	-0.81833	-1.22215	0.455035	1.43428
(8, 5)	0.716098	0.3965	-0.67122	-0.31011
(8, 6)	0.534197	0.337433	-0.58492	-0.28671
(8, 7)	0.616593	0.314058	-0.59326	-0.33739
(8, 8)	0	0	0	0
(8, 9)	0.048866	0.031074	-0.10043	0.020491

Fig. 19 shows the percentage change in nucleotide % concentration values for Genome 8 with all nine genome sequences. Fig. 19 also shows the absolute (difference) values.

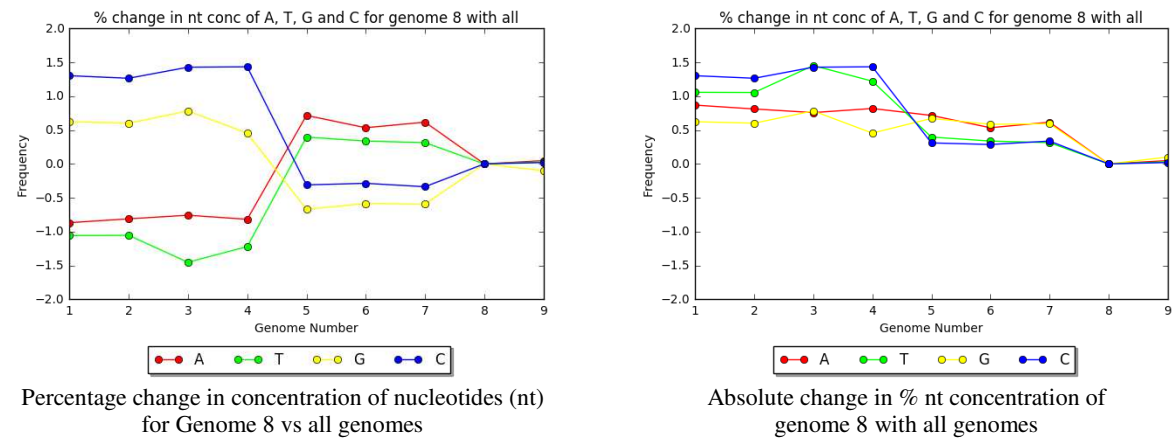


Fig. 19: Percentage change in concentration of nucleotides (nt) for Genome 8 vs all genomes and their differences

Like Genome sequences 5, 6 and 7 Genome sequence 8 also shows almost a reverse characteristics when compared to other Genome sequences 1, 2, 3, 4.

Difference in % nucleotide concentrations with respect to Genome 9

Genome (X, Y)	$p(A_x) - p(A_y)$	$p(T_x) - p(T_y)$	$p(G_x) - p(G_y)$	$p(C_x) - p(C_y)$
(9, 1)	-0.91725	-1.0887	0.723827	1.28212
(9, 2)	-0.86022	-1.08601	0.701874	1.24436
(9, 3)	-0.80673	-1.48377	0.882662	1.407835
(9, 4)	-0.86719	-1.25322	0.555466	1.413789
(9, 5)	0.667232	0.365426	-0.57079	-0.3306
(9, 6)	0.485331	0.30636	-0.48449	-0.3072
(9, 7)	0.567727	0.282984	-0.49283	-0.35788
(9, 8)	-0.04887	-0.03107	0.100431	-0.02049
(9, 9)	0	0	0	0

Fig. 20 shows the percentage change in nucleotide % concentration values for Genome 9 with all nine genome sequences. Fig. 20 also shows the absolute (difference) values.

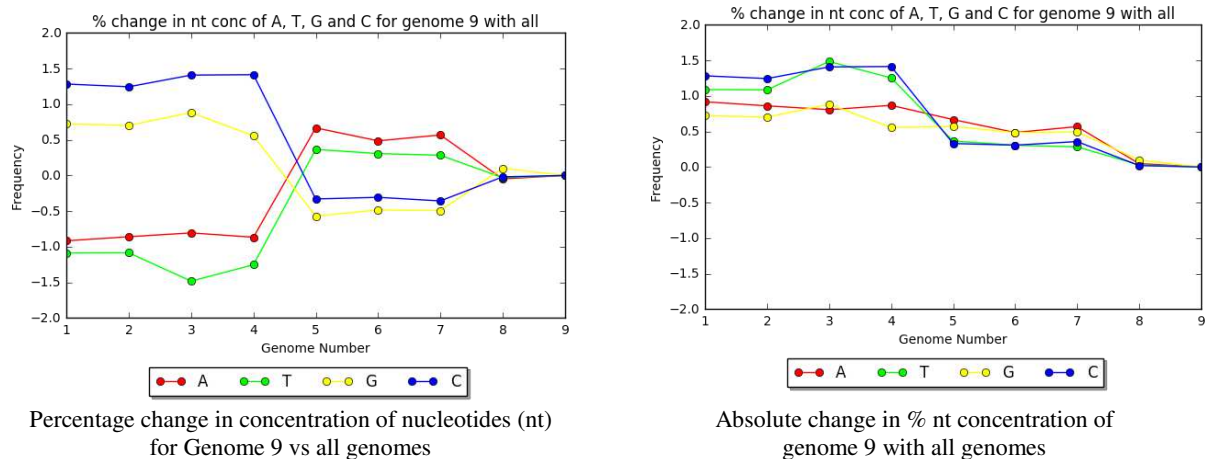


Fig. 20: Percentage change in concentration of nucleotides (nt) for Genome 9 vs all genomes and their differences

Like the Genome sequences 5, 6, 7 and 8 Genome sequence 9 also shows almost a reverse characteristics when compared to other Genome sequences 1, 2, 3, 4. **Thus, genomes 1, 2, 3 and 4 are of one type and genomes 5, 6, 7, 8 and 9 are of another type.**

IV. RECTANGULAR SUBSAMPLING OF ARRAYS OF ADJOINTS OF GENOME SEQUENCES AND CORRELATIONS

It is a usual practice in signal processing to subsample an analog signal, quantize the subsamples, convert them to digital bit sequences, and apply algorithms developed in the framework of discrete algebra and logic with the idea of processing the digital signals in order to extract hidden information and carry out pattern recognition. In the case of genomic signal processing also this kind of approach could be tried, mainly to extract hidden genetic information and carry out pattern recognition. The first step towards this is to convert a characteristic genome sequence into a number sequence by a coding morphism, for instance, A (adenine) is replaced by the number '1', T (thymine) is represented by the number '2', G (guanine) is represented by the number '3' and C (cytosine) is represented by the number '4'. Thus, a genomic characteristic sequence becomes a number sequence consisting of 1's, 2's, 3's and 4's. Next, adjoints of nucleotides have to be obtained from the original genome sequences (Ref Figs. 2 to 10). Now the adjoint arrays have to be subsampled using the following schema:

Rectangular checker board subsampling of an adjoint array for getting even and odd arrays

For example, let us consider an array pattern given in Fig. 21. The even and odd arrays are obtained as shown in Fig. 21.

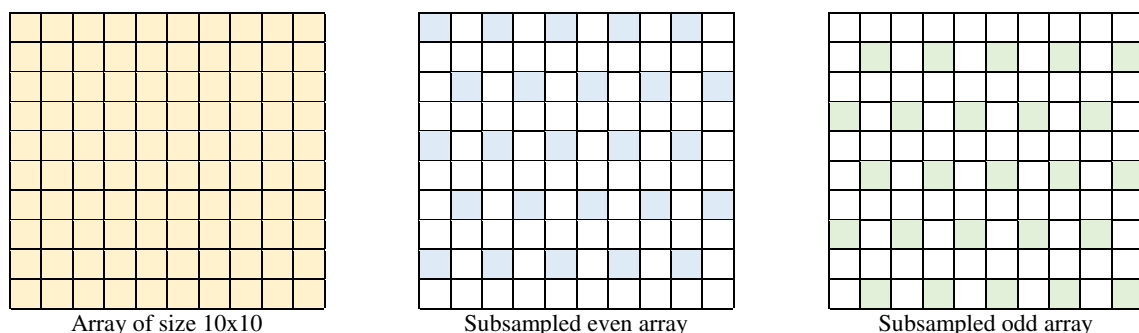
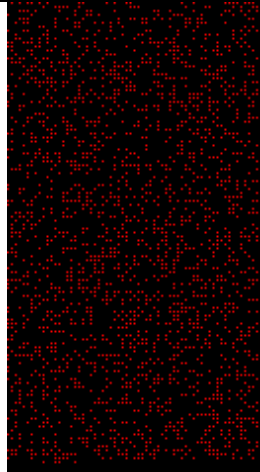
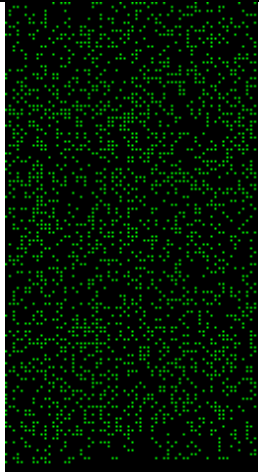
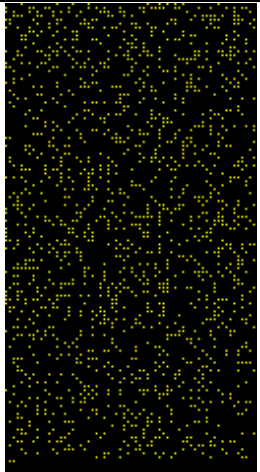
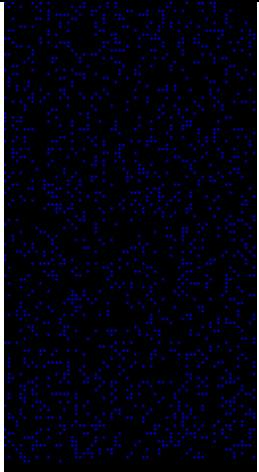
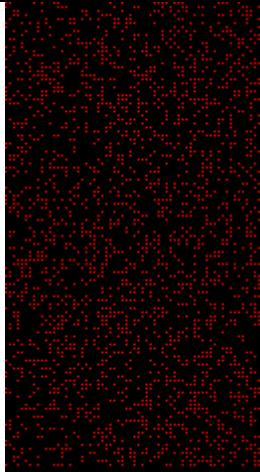
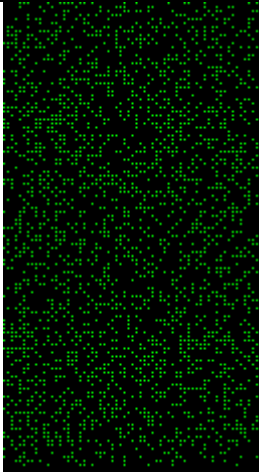
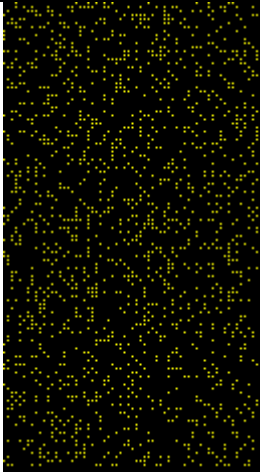
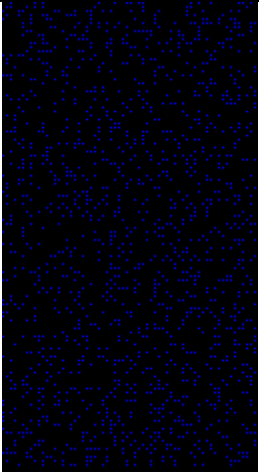
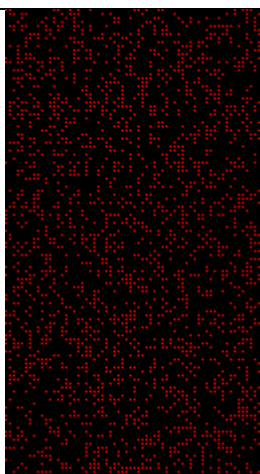
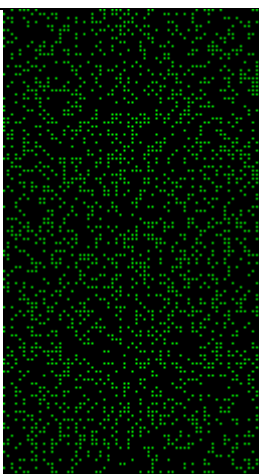
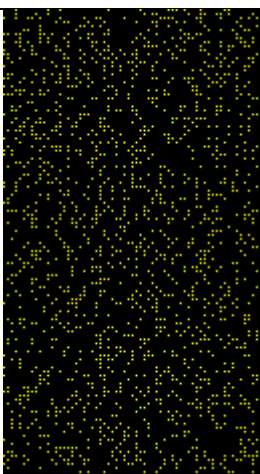
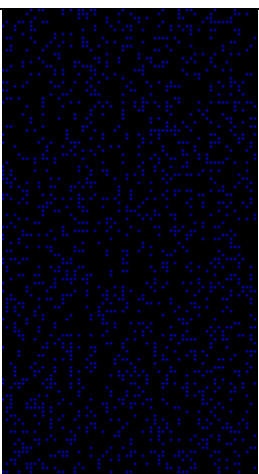
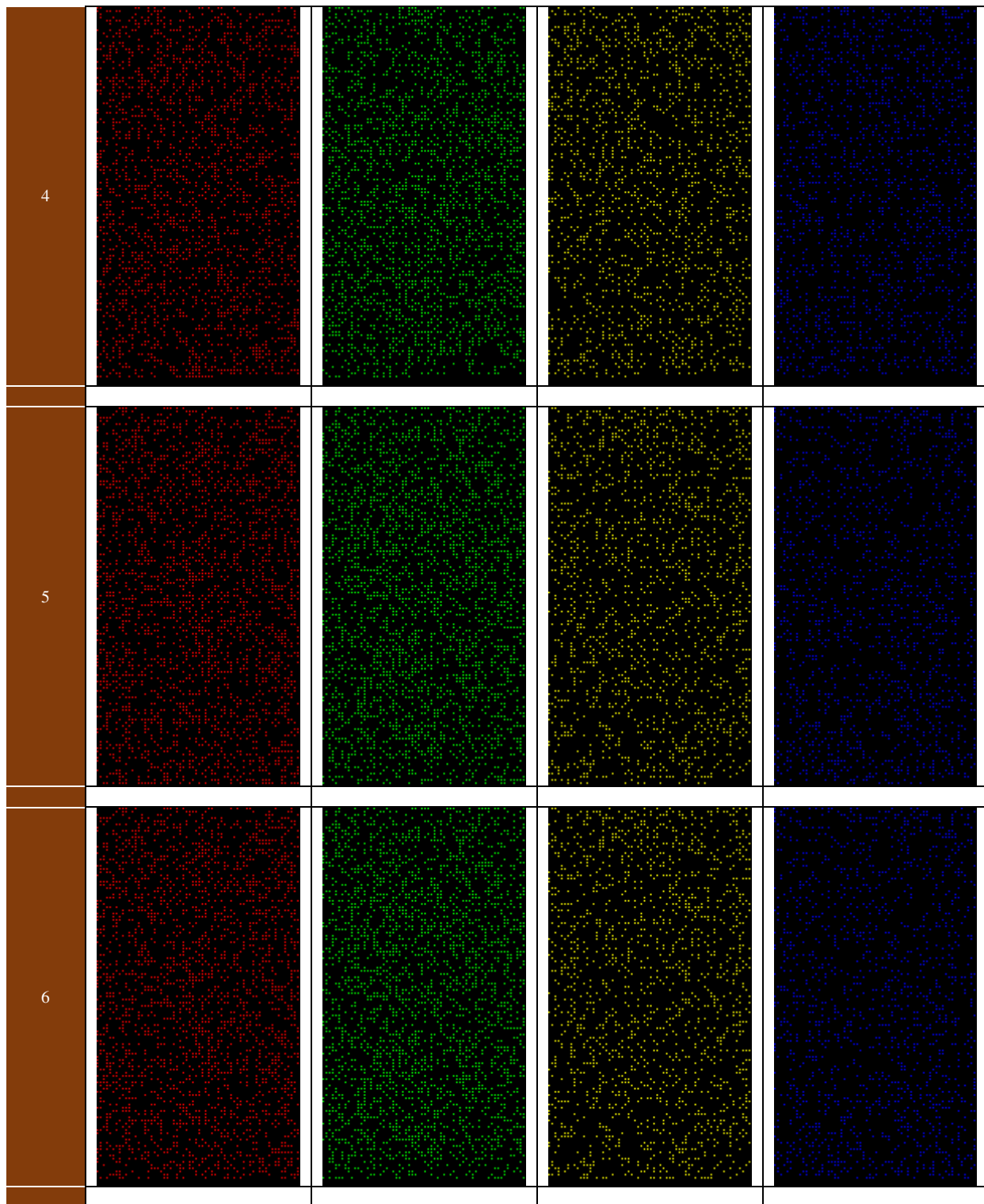


Fig. 21: Checker board and subsampled even and odd arrays

Adjoints arrays of nucleotides of all nine genomes are subsampled as shown above in Fig. 21 and the subsampled arrays are given below. Fig. 22 shows all even subsampled adjoints arrays. Fig. 24 shows all odd subsampled adjoints arrays.

Even arrays of adjoints of Genome 1 (AY304488.1), Genome 2 AY395003, Genome 3 DQ412043, Genome 4 KY417144.1, Genome 5 MN908947.1, Genome 6 MN908947.2, Genome 7 MN908947.3, Genome 8 Mg772933, Genome 9 Mg772934

Even subsampled adjoints arrays				
Genome	Adj(A)	Adj(T)	Adj(G)	Adj(C)
1				
2				
3				



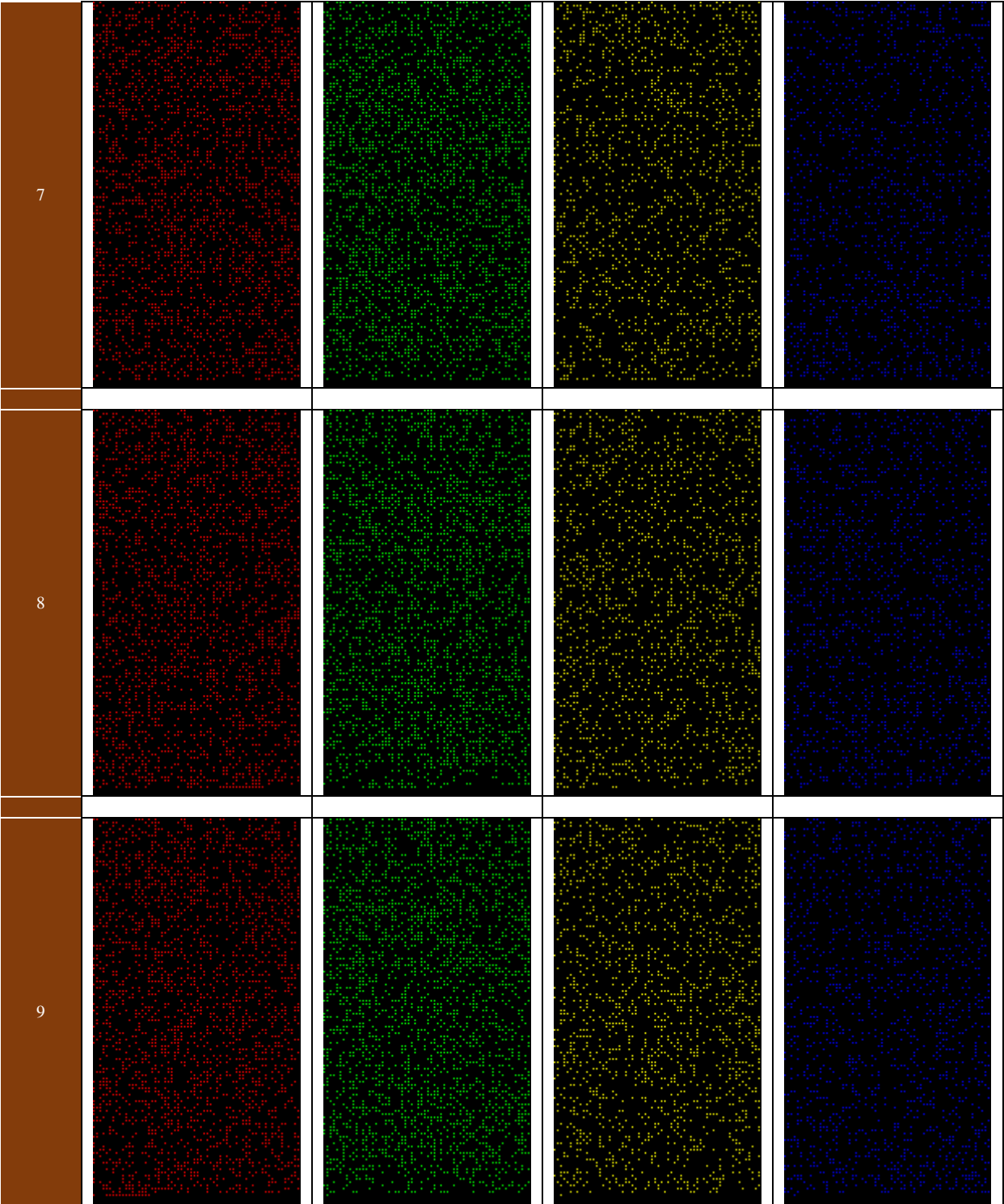


Fig. 22: Even subsampled adjoints arrays

The number of nucleotides present in the even subsampled adjoints arrays of all nine genomes are given below and in Fig. 23.

2142	1448	1558	2294
Genome 1 Adenine even subsampled	Genome 1 Cytosine even subsampled	Genome 1 Guanine even subsampled	Genome 1 Thymine even subsampled
2154	1454	1535	2281
Genome 2 Adenine even subsampled	Genome 2 Cytosine even subsampled	Genome 2 Guanine even subsampled	Genome 2 Thymine even subsampled
2096	1568	1591	2196
Genome 3 Adenine even subsampled	Genome 3 Cytosine even subsampled	Genome 3 Guanine even subsampled	Genome 3 Thymine even subsampled
2107	1553	1557	2244
Genome 4 Adenine even subsampled	Genome 4 Cytosine even subsampled	Genome 4 Guanine even subsampled	Genome 4 Thymine even subsampled
2297	1392	1502	2425
Genome 5 Adenine even subsampled	Genome 5 Cytosine even subsampled	Genome 5 Guanine even subsampled	Genome 5 Thymine even subsampled
2245	1371	1485	2387
Genome 6 Adenine even subsampled	Genome 6 Cytosine even subsampled	Genome 6 Guanine even subsampled	Genome 6 Thymine even subsampled
2254	1339	1476	2419
Genome 7 Adenine even subsampled	Genome 7 Cytosine even subsampled	Genome 7 Guanine even subsampled	Genome 7 Thymine even subsampled
2189	1419	1531	2338
Genome 8 Adenine even subsampled	Genome 8 Cytosine even subsampled	Genome 8 Guanine even subsampled	Genome 8 Thymine even subsampled
2179	1412	1522	2329
Genome 9 Adenine even subsampled	Genome 9 Cytosine even subsampled	Genome 9 Guanine even subsampled	Genome 9 Thymine even subsampled

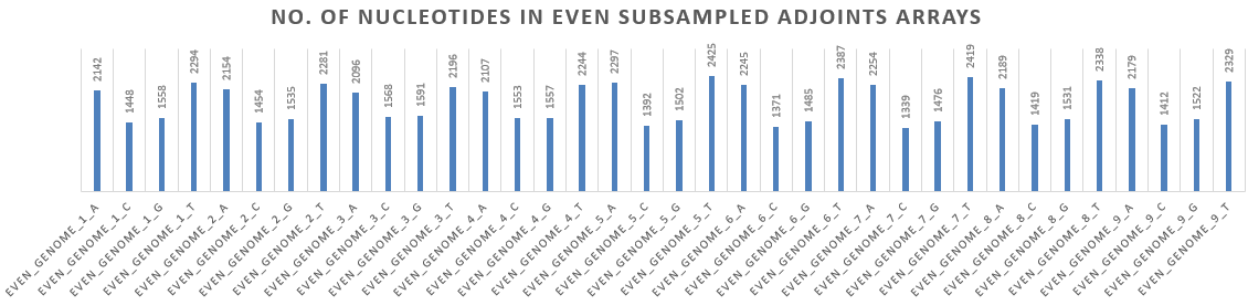
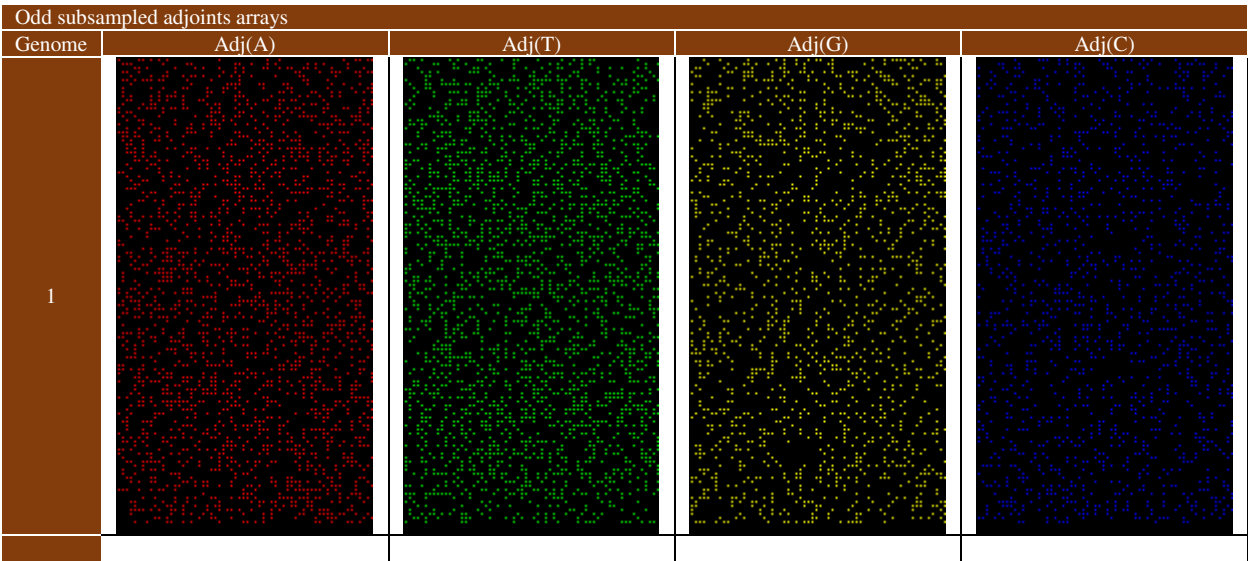
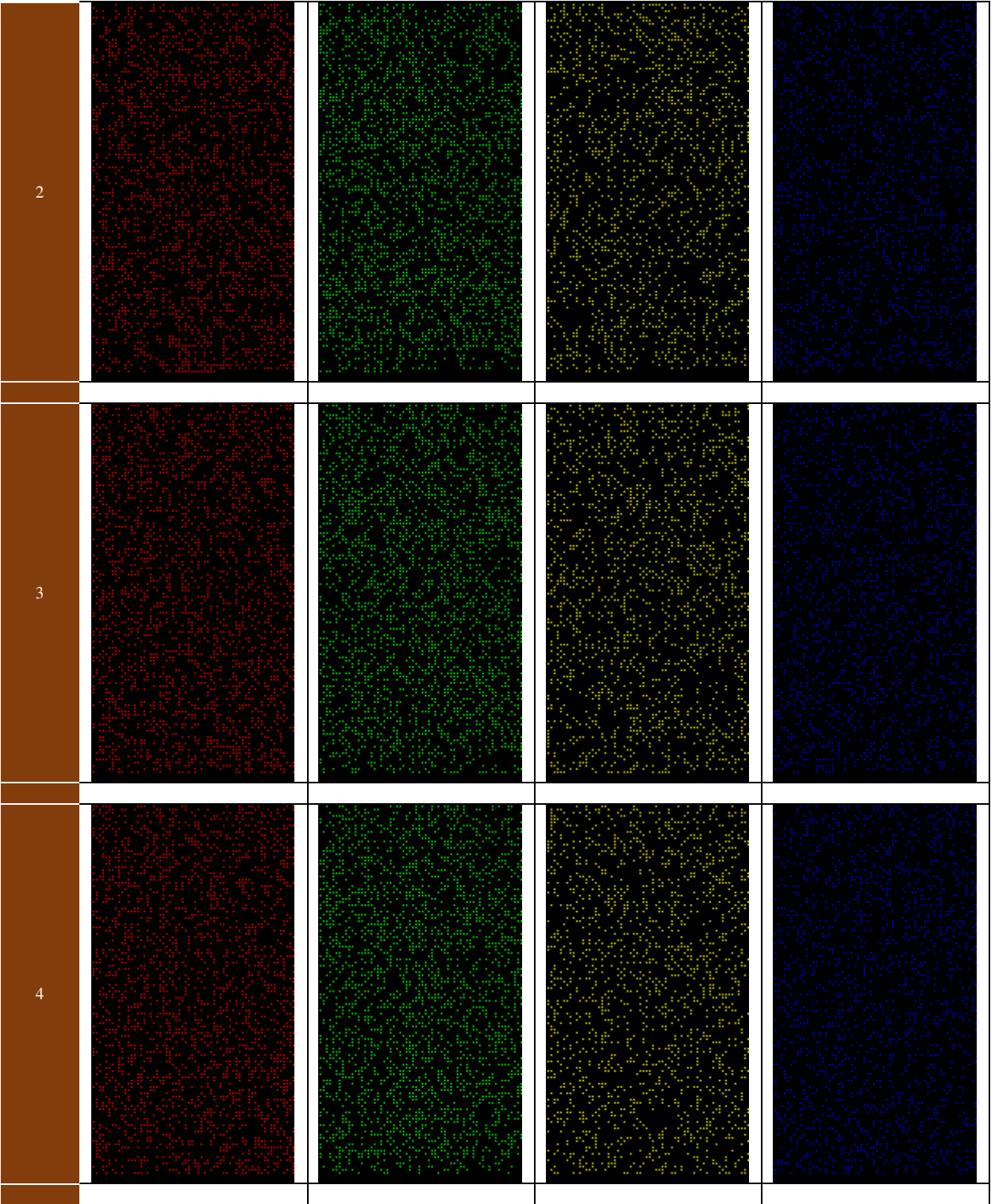
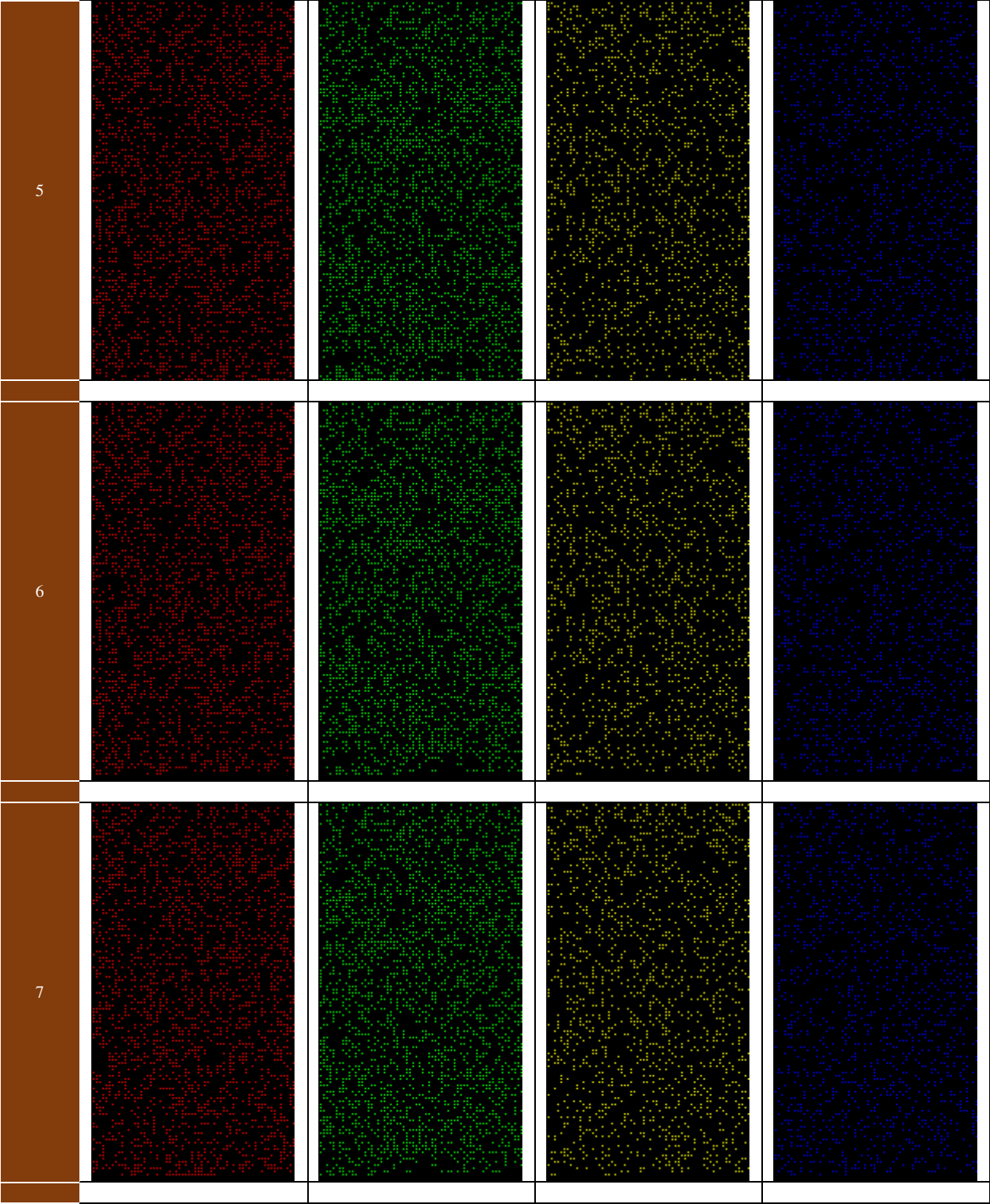


Fig. 23: Number of nucleotides present in even subsampled adjoints arrays of nine genomes

Odd arrays of adjoints of Genome 1 (AY304488.1), Genome 2 AY395003, Genome 3 DQ412043, Genome 4 KY417144.1, Genome 5 MN908947.1, Genome 6 MN908947.2, Genome 7 MN908947.3, Genome 8 Mg772933, Genome 9 Mg772934







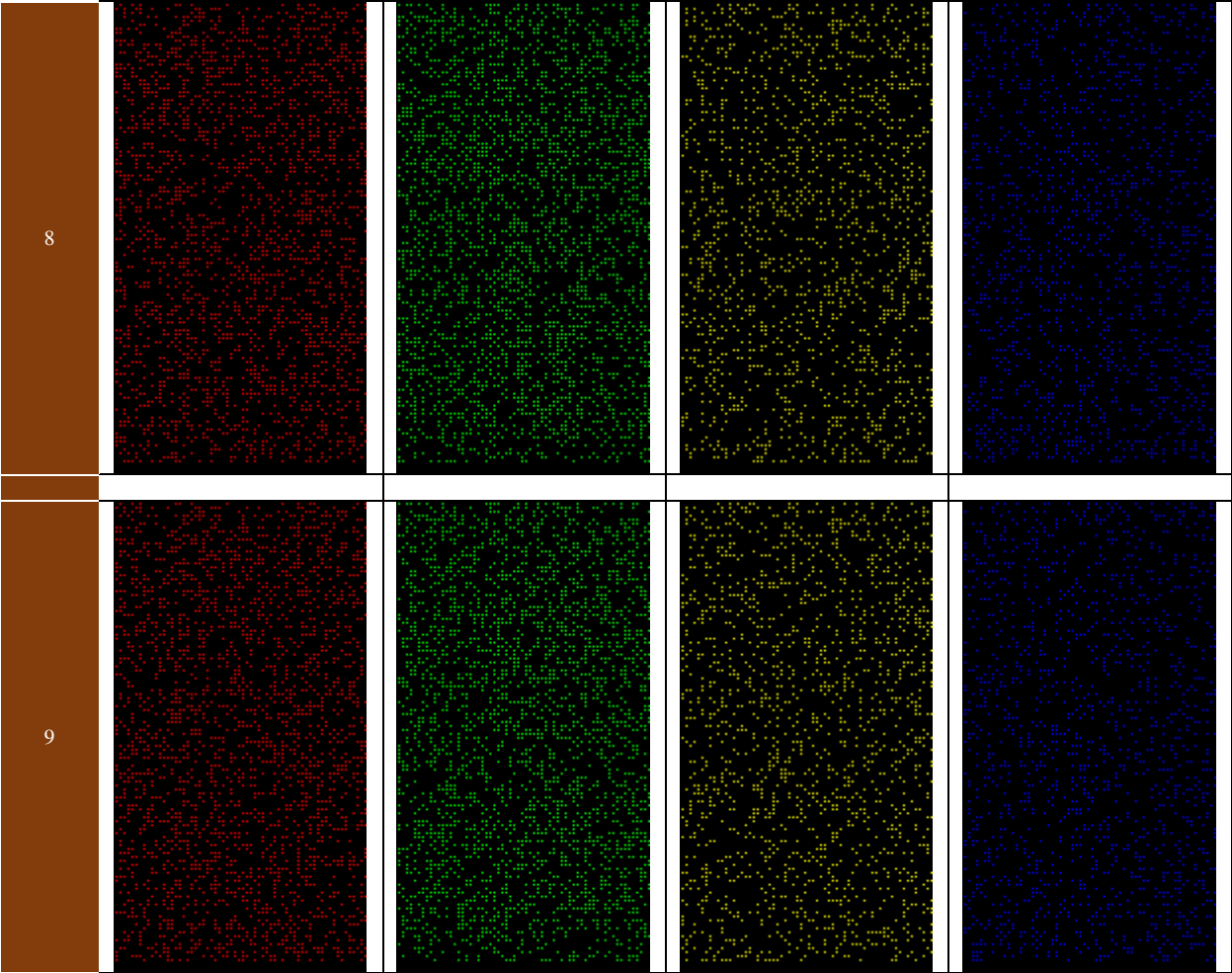


Fig. 24: Odd subsampled adjoints arrays

The number of nucleotides present in the odd subsampled adjoints arrays of all nine genomes are given below and in Fig. 25.

2135	1445	1574	2270
Genome 1 Adenine odd subsampled	Genome 1 Cytosine odd subsampled	Genome 1 Guanine odd subsampled	Genome 1 Thymine odd subsampled
2126	1448	1565	2260
Genome 2 Adenine odd subsampled	Genome 2 Cytosine odd subsampled	Genome 2 Guanine odd subsampled	Genome 2 Thymine odd subsampled
2090	1506	1574	2254
Genome 3 Adenine odd subsampled	Genome 3 Cytosine odd subsampled	Genome 3 Guanine odd subsampled	Genome 3 Thymine odd subsampled
2106	1511	1541	2266
Genome 4 Adenine odd subsampled	Genome 4 Cytosine odd subsampled	Genome 4 Guanine odd subsampled	Genome 4 Thymine odd subsampled
2272	1429	1504	2411
Genome 5 Adenine odd subsampled	Genome 5 Cytosine odd subsampled	Genome 5 Guanine odd subsampled	Genome 5 Thymine odd subsampled
2218	1402	1480	2349
Genome 6 Adenine odd subsampled	Genome 6 Cytosine odd subsampled	Genome 6 Guanine odd subsampled	Genome 6 Thymine odd subsampled
2227	1408	1464	2364
Genome 7 Adenine odd subsampled	Genome 7 Cytosine odd subsampled	Genome 7 Guanine odd subsampled	Genome 7 Thymine odd subsampled
2144	1429	1469	2382
Genome 8 Adenine odd subsampled	Genome 8 Cytosine odd subsampled	Genome 8 Guanine odd subsampled	Genome 8 Thymine odd subsampled
2198	1371	1491	2364
Genome 9 Adenine odd subsampled	Genome 9 Cytosine odd subsampled	Genome 9 Guanine odd subsampled	Genome 9 Thymine odd subsampled

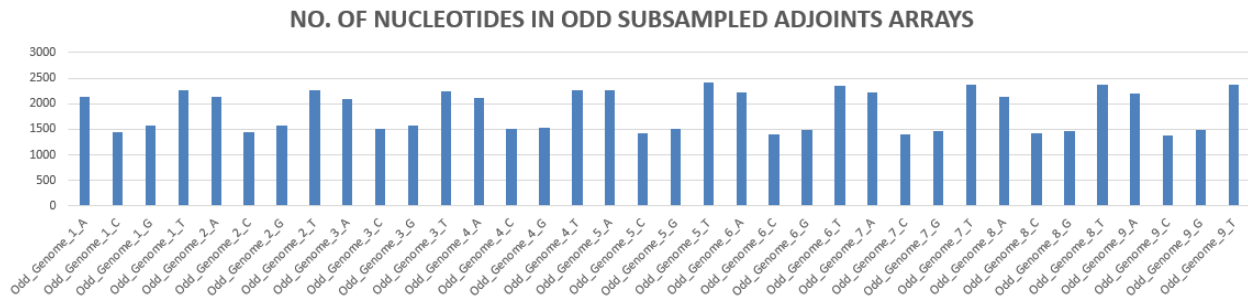


Fig. 25: Number of nucleotides present in odd subsampled adjoints arrays of nine genomes

V. PAIRWISE SPATIAL CORRELATION OF SUBSAMPLED ADJOINTS ARRAYS

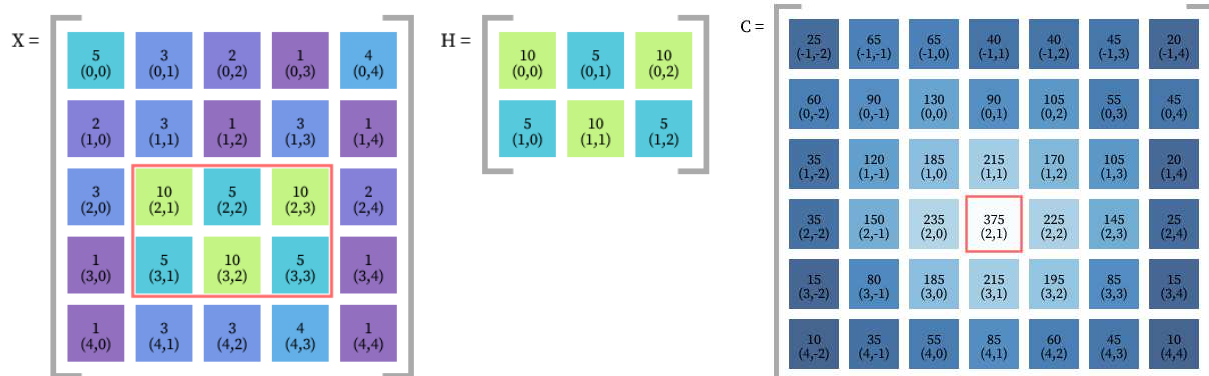
Usually two dimensional cross correlation of two matrices is evaluated using the equation:

$$C(k, l) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} X(m, n) \overline{H(m-k, n-l)}$$

where matrix X is of size $M \times N$ and matrix H is of size $P \times Q$. Then the resulting matrix C would be of size $(M+P-1) \times (N+Q-1)$ such that $-(P-1) \leq k \leq (M-1)$ and $-(Q-1) \leq l \leq (N-1)$. The bar over H denotes its complex conjugate. For example, consider the following matrices, X and H , where X is a 5-by-5 matrix ($M=5$, $N=5$) and H is a 2-by-3 matrix ($P=2$, $Q=3$):

$$X = \begin{bmatrix} 5 & 3 & 2 & 1 & 4 \\ 2 & 3 & 1 & 3 & 1 \\ 3 & 10 & 5 & 10 & 2 \\ 1 & 5 & 10 & 5 & 1 \\ 1 & 3 & 3 & 4 & 1 \end{bmatrix}$$

$$H = \begin{bmatrix} 10 & 5 & 10 \\ 5 & 10 & 5 \end{bmatrix}$$



The resulting cross correlated matrix C of size 6x7 is given above.

Courtesy: <https://observablehq.com/@lemonnish/cross-correlation-of-2-matrices>

Now, two matrices of the same size $M \times N$ would yield the cross correlation matrix of size $(2M-1) \times (2N-1)$. As a special case of two-dimensional Cross Correlation (CC), one can think of 'Dyadic Cross Correlation (DCC)' as *pointwise correlation* of two equal size matrices. This is evaluated using 'logical AND' instead of 2D convolution operations 'shift, multiply and add'. In such a case, the size of the DCC matrix will remain $M \times N$. For example, consider two matrices X and Y both of size 9x9 as shown in Fig. 26.

0	1	0	1	1	0	0	1	1
1	1	1	0	1	0	1	1	0
0	0	0	1	1	1	0	0	1
1	1	0	0	0	1	1	0	0
0	0	1	1	1	0	0	1	1
1	0	1	0	0	1	1	0	1
0	1	1	0	1	0	1	1	1
0	0	1	1	1	1	0	0	0
1	0	0	1	0	1	1	0	1

Matrix - X

0	1	0	0	1	0	0	1	0
1	0	1	0	1	1	1	0	0
1	1	0	1	1	0	0	0	1
1	0	1	0	1	1	0	1	0
0	1	0	1	1	1	0	0	1
1	0	1	0	0	0	1	0	0
0	1	0	0	1	0	1	0	1
0	0	1	0	1	0	0	0	0
1	0	1	1	0	1	0	0	0

Matrix - Y

0	1	0	0	1	0	0	1	0
1	0	1	0	1	0	1	0	0
1	1	0	1	1	0	0	0	1
1	0	0	0	0	1	0	0	0
0	0	0	1	1	0	0	0	1
1	0	1	0	0	0	1	0	0
0	1	0	0	1	0	1	0	1
0	0	1	0	1	0	0	0	0
1	0	0	1	0	1	0	0	0

Matrix -

Fig. 26: Sample images of matrices X, Y and the DCC marix C

One may observe that the DCC matrix is a member of Cross Correlation matrix space of two equisize matrices. For example, let us consider the above example given in Fig. 26. One can have total of 289 Cross Correlation (CC) components having common intersected cell positions, some of them being shown in Fig. 27.

0	1	0	0	1	0	0	1	0
1	0	1	0	1	1	1	0	0
1	1	0	1	1	0	0	0	1
1	0	1	0	1	1	0	1	0
0	1	0	1	1	1	0	0	1
1	0	1	0	0	0	1	0	0
0	1	0	0	1	0	1	0	1
0	0	1	0	1	0	0	0	0
1	0	1	1	0	1	0	0	0

Component 1

0	1	0	0	1	0	0	1	0
1	0	1	0	1	1	1	0	0
1	1	0	1	1	0	0	0	1
1	0	1	0	1	1	0	1	0
0	1	0	1	1	1	0	0	1
1	0	1	0	0	0	1	0	0
0	1	0	0	1	0	1	0	1
0	0	1	0	1	0	0	0	0
1	0	1	1	0	1	0	0	0

Component 9

0	1	0	0	1	0	0	1	0
1	0	1	0	1	1	1	0	0
1	1	0	1	1	0	0	0	1
1	0	1	0	1	1	0	1	0
0	1	0	1	1	1	0	0	1
1	0	1	0	0	0	1	0	0
0	1	0	0	1	0	1	0	1
0	0	1	0	1	0	0	0	0
1	0	1	1	0	1	0	0	0

Component 17

0	1	0	0	1	0	0	1	0
1	0	1	0	1	1	1	0	0
1	1	0	1	1	0	0	0	1
1	0	1	0	1	1	0	1	0
0	1	0	1	1	1	0	0	1
1	0	1	0	0	0	1	0	0
0	1	0	0	1	0	1	0	1
0	0	1	0	1	0	0	0	0
1	0	1	1	0	1	0	0	0

Component 18

0	1	0	0	1	0	0	1	0
1	0	1	0	1	1	1	0	0
1	1	0	1	1	0	0	0	1
1	0	1	0	1	1	0	1	0
0	1	0	1	1	1	0	0	1
1	0	1	0	0	0	1	0	0
0	1	0	0	1	0	1	0	1
0	0	1	0	1	0	0	0	0
1	0	1	1	0	1	0	0	0

Component 26

0	1	0	0	1	0	0	1	0
1	0	1	0	1	1	1	0	0
1	1	0	1	1	0	0	0	1
1	0	1	0	1	1	0	1	0
0	1	0	1	1	1	0	0	1
1	0	1	0	0	0	1	0	0
0	1	0	0	1	0	1	0	1
0	0	1	0	1	0	0	0	0
1	0	1	1	0	1	0	0	0

Component 34

0	1	0	0	1	0	0	1	0
1	0	1	0	1	1	1	0	0
1	1	0	1	1	0	0	0	1
1	0	1	0	1	1	0	1	0
0	1	0	1	1	1	0	0	1
1	0	1	0	0	0	1	0	0
0	1	0	0	1	0	1	0	1
0	0	1	0	1	0	0	0	0
1	0	1	1	0	1	0	0	0

Component 137

0	1	0	0	1	0	0	1	0
1	0	1	0	1	1	1	0	0
1	1	0	1	1	0	0	0	1
1	0	1	0	1	1	0	1	0
0	1	0	1	1	1	0	0	1
1	0	1	0	0	0	1	0	0
0	1	0	0	1	0	1	0	1
0	0	1	0	1	0	0	0	0
1	0	1	1	0	1	0	0	0

Component 145

0	1	0	0	1	0	0	1	0
1	0	1	0	1	1	1	0	0
1	1	0	1	1	0	0	0	1
1	0	1	0	1	1	0	1	0
0	1	0	1	1	1	0	0	1
1	0	1	0	0	0	1	0	0
0	1	0	0	1	0	1	0	1
0	0	1	0	1	0	0	0	0
1	0	1	1	0	1	0	0	0

Component 153

0	1	0	0	1	0	0	1	0
1	0	1	0	1	1	1	0	0
1	1	0	1	1	0	0	0	1
1	0	1	0	1	1	0	1	0
0	1	0	1	1	1	0	0	1
1	0	1	0	0	0	1	0	0
0	1	0	0	1	0	1	0	1
0	0	1	0	1	0	0	0	0
1	0	1	1	0	1	0	0	0

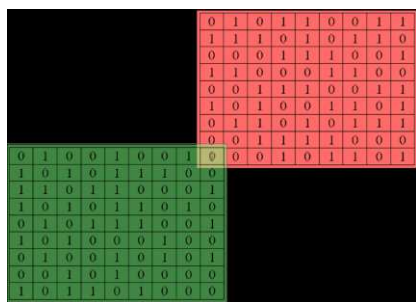
Component 256

0	1	0	0	1	0	0	1	0
1	0	1	0	1	1	1	0	0
1	1	0	1	1	0	0	0	1
1	0	1	0	1	1	0	1	0
0	1	0	1	1	1	0	0	1
1	0	1	0	0	0	1	0	0
0	1	0	0	1	0	1	0	1
0	0	1	0	1	0	0	0	0
1	0	1	1	0	1	0	0	0

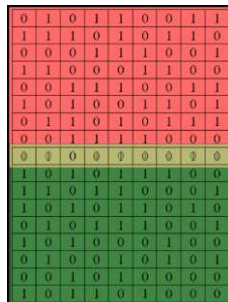
Component 264

0	1	0	0	1	0	0	1	0
1	0	1	0	1	1	1	0	0
1	1	0	1	1	0	0	0	1
1	0	1	0	1	1	0	1	0
0	1	0	1	1	1	0	0	1
1	0	1	0	0	0	1	0	0
0	1	0	0	1	0	1	0	1
0	0	1	0	1	0	0	0	0
1	0	1	1	0	1	0	0	0

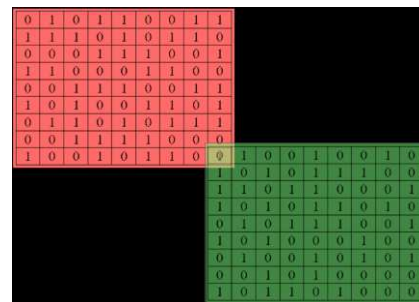
Component 272



Component 273



Component 281



Component

Fig. 27: Scan positions for example given in Fig. 26

All these 289 scan positions have common intersected cell positions that are given below.

Rows	Columns																	
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
	1	1	2	3	4	5	6	7	8	9	8	7	6	5	4	3	2	1
	2	2	4	6	8	10	12	14	16	18	16	14	12	10	8	6	4	2
	3	3	6	9	12	15	18	21	24	27	24	21	18	15	12	9	6	3
	4	4	8	12	16	20	24	28	32	36	32	28	24	20	16	12	8	4
	5	5	10	15	20	25	30	35	40	45	40	35	30	25	20	15	10	5
	6	6	12	18	24	30	36	42	48	54	48	42	36	30	24	18	12	6
	7	7	14	21	28	35	42	49	56	63	56	49	42	35	28	21	14	7
	8	8	16	24	32	40	48	56	64	72	64	56	48	40	32	24	16	8
	9	9	18	27	36	45	54	63	72	81	72	63	54	45	36	27	18	9
	10	8	16	24	32	40	48	56	64	72	64	56	48	40	32	24	16	8
	11	7	14	21	28	35	42	49	56	63	56	49	42	35	28	21	14	7
	12	6	12	18	24	30	36	42	48	54	48	42	36	30	24	18	12	6
	13	5	10	15	20	25	30	35	40	45	40	35	30	25	20	15	10	5
	14	4	8	12	16	20	24	28	32	36	32	28	24	20	16	12	8	4
	15	3	6	9	12	15	18	21	24	27	24	21	18	15	12	9	6	3
	16	2	4	6	8	10	12	14	16	18	16	14	12	10	8	6	4	2
	17	1	2	3	4	5	6	7	8	9	8	7	6	5	4	3	2	1

Correlation Component 145 has the possibility of having maximum number (81) of matches between two equized matrices. Components 1, 17, 273 and 289 have the possibilities of having minimum number (1) of matches between two equized matrices. Fig. 28 shows the cross correlation graph in 3D form.

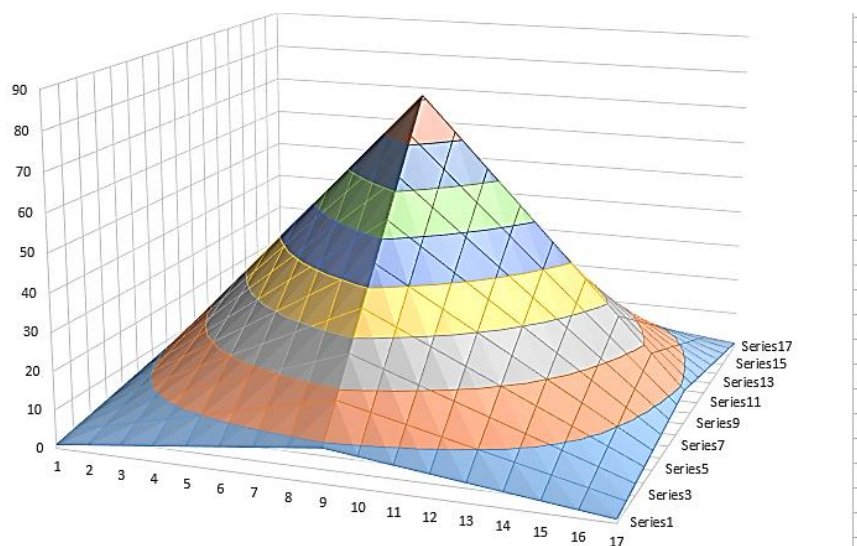


Fig. 28: 3D graph showing 2D cross correlation between sample arrays X and Y shown in Fig. 26

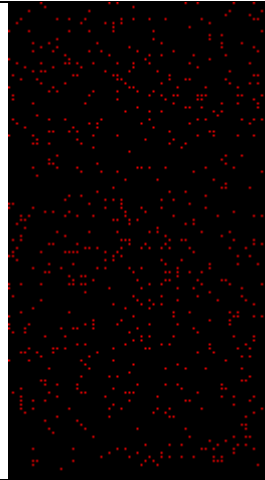
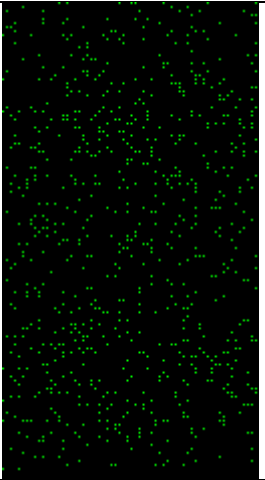
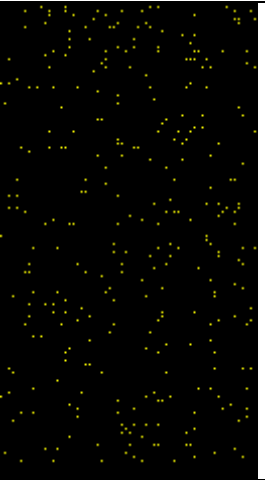
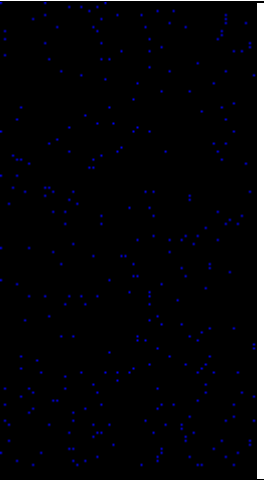
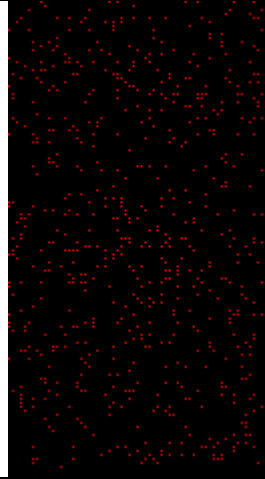
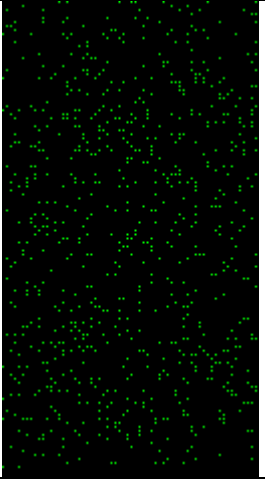
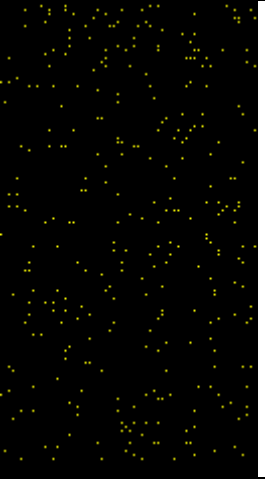
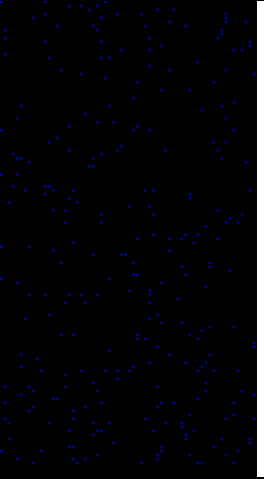
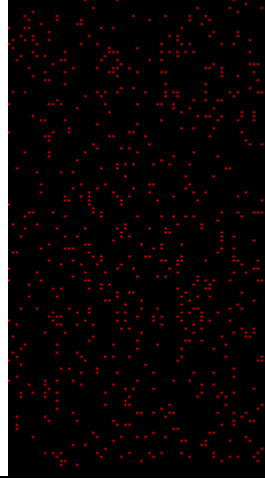
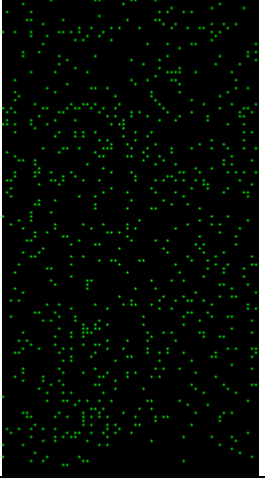
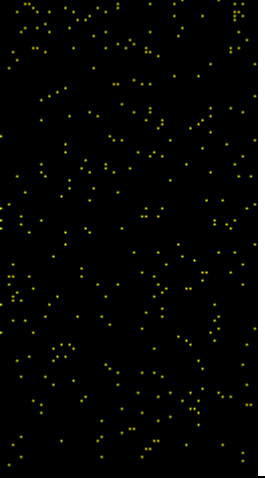
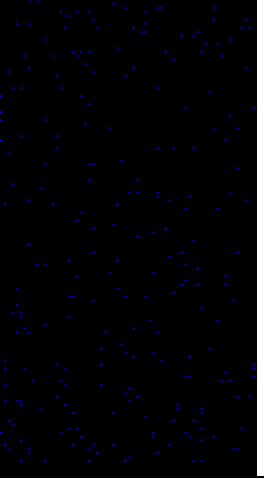
Pairwise spatial correlation (intersection) of even subsampled adjoints arrays

First, even subsampled adjoint arrays are examined for cross correlation, dyadically. Dyadic cross correlation is a logical AND operation that involves two different adjoint arrays. The purpose of doing this is to find out how many nucleotides, say for example, adenines are present in both the genome sequences at the same locations. The number of such adenines present in the same locations of both the genome sequences is called 'Dyadic Cross Correlation Coefficient (DCCC or DC^3)'. In other words, $DC^3(A)$, $DC^3(T)$, $DC^3(G)$ and $DC^3(C)$ for genome sequences X and Y yield an apparent spatial correlation between them. Let genome X represent the

even subsampled adjoint array of AY304488.1 and Genome Y represent the even subsampled adjoint array of AY395003. For one pair of even subsampled adjoints arrays, one can have **1,21,125 Cross Correlation (CC) components** and **60,563rd Cross Correlation (CC) component** will correspond to the **‘Dyadic Cross Correlation’ (DCC)**. In what follows, the dyadic cross correlations of even subsampled adjoints arrays are presented. **DC³** represents **‘Dyadic Cross Correlation Coefficient’** connecting two arrays. Figs. 29 to 36 show all possible DCCs and their coefficient values.

Dyadic Cross Correlations of Even Subsampled Adjoints Arrays

Spatial Dyadic Cross Correlations of even subsampled adjoints arrays				
Genome (X, Y)	Adenine	Thymine	Guanine	Cytosine
(1, 2)				
	DC ³ (A) = 644	DC ³ (T) = 735	DC ³ (G) = 376	DC ³ (C) = 282
(1, 3)				
	DC ³ (A) = 612	DC ³ (T) = 684	DC ³ (G) = 322	DC ³ (C) = 295
(1, 4)				
	DC ³ (A) = 606	DC ³ (T) = 751	DC ³ (G) = 353	DC ³ (C) = 320

(1, 5)				
	DC ³ (A) = 649	DC ³ (T) = 766	DC ³ (G) = 313	DC ³ (C) = 278
(1, 6)				
	DC ³ (A) = 649	DC ³ (T) = 766	DC ³ (G) = 313	DC ³ (C) = 278
(1, 7)				
	DC ³ (A) = 637	DC ³ (T) = 715	DC ³ (G) = 297	DC ³ (C) = 237

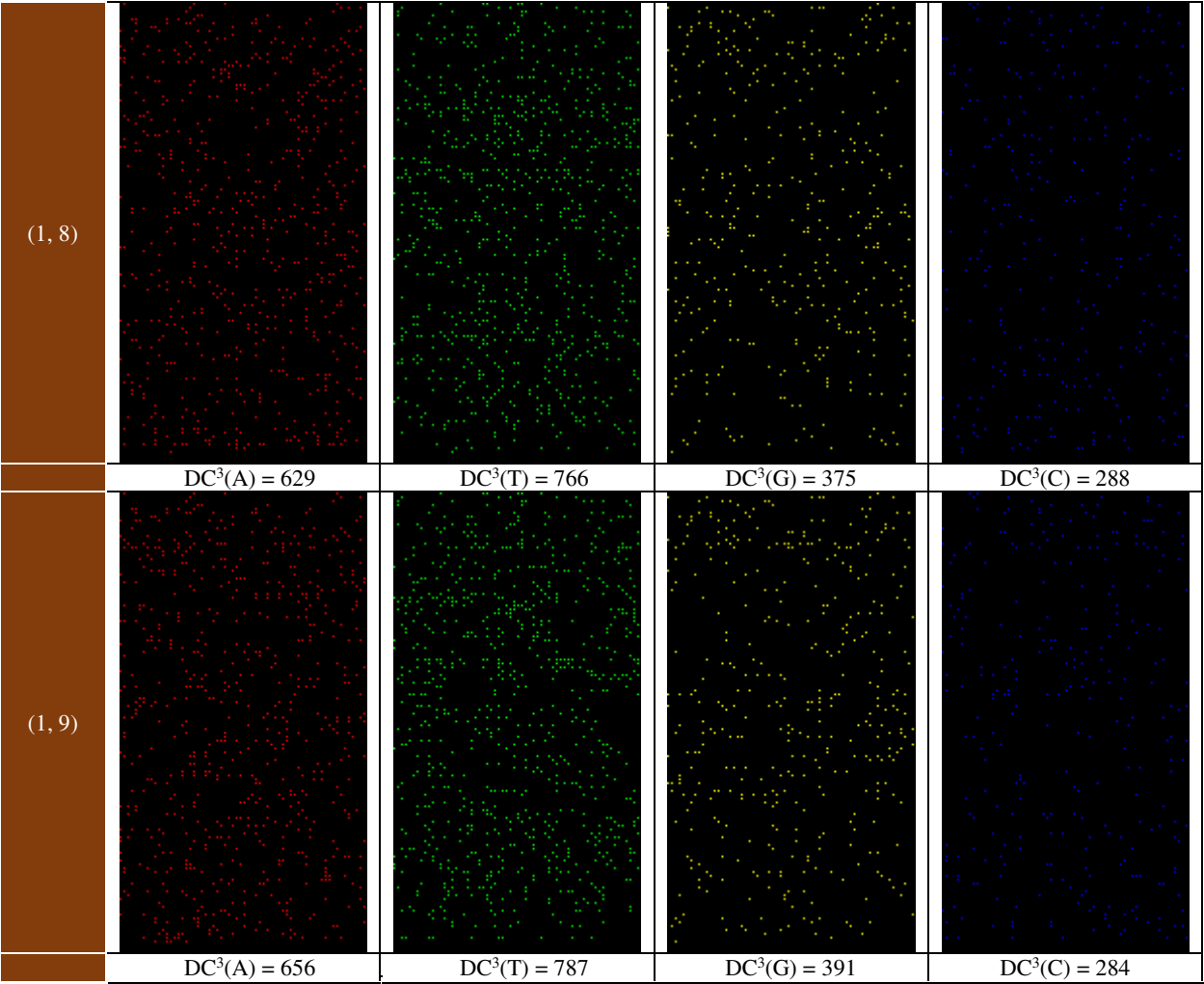
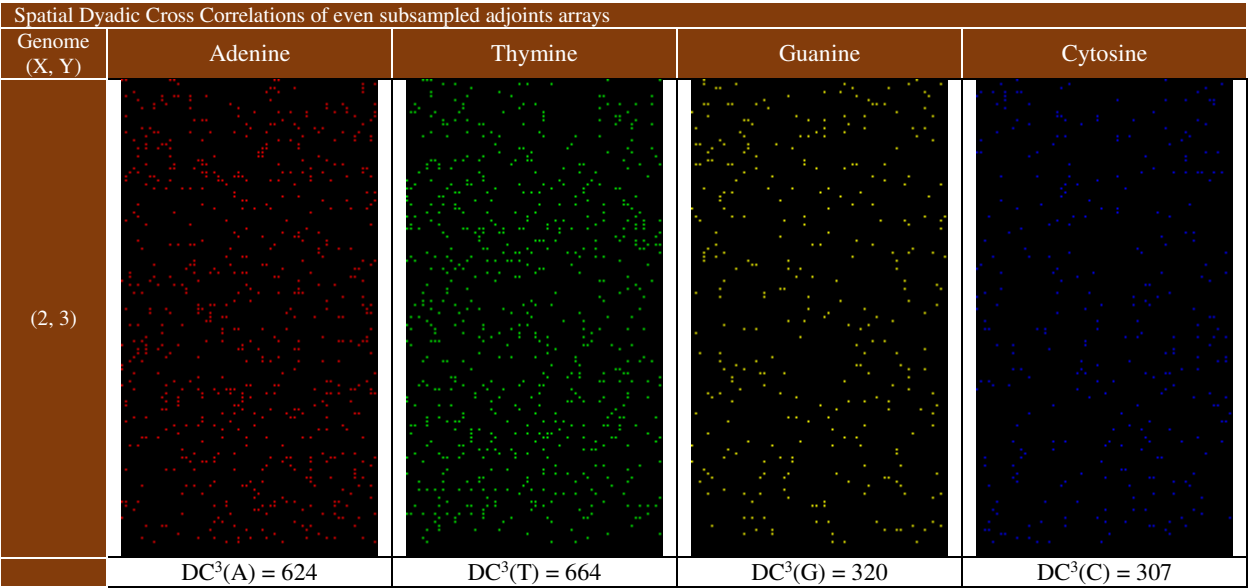
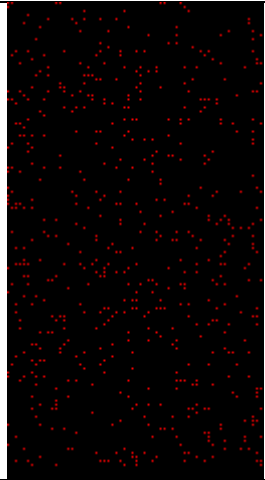
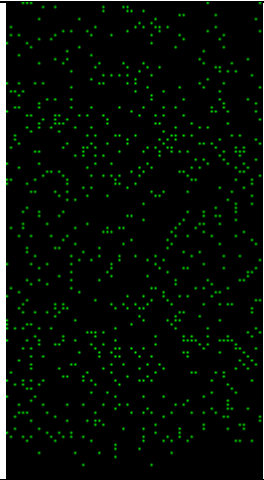
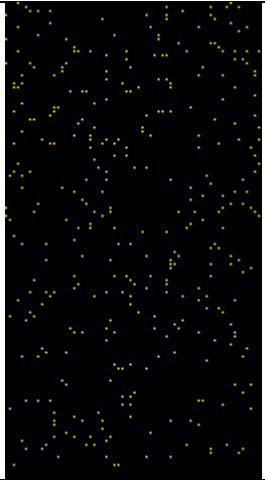
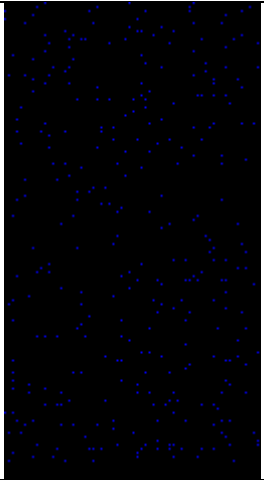
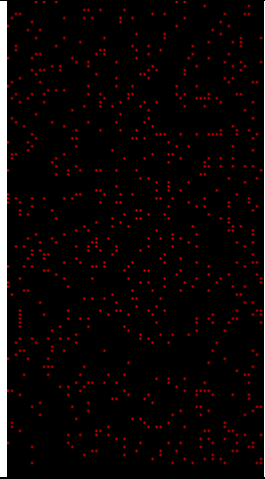
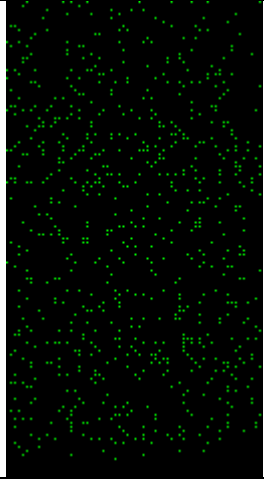
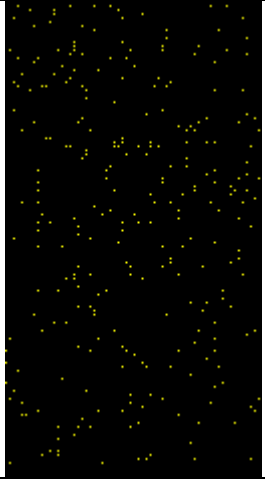
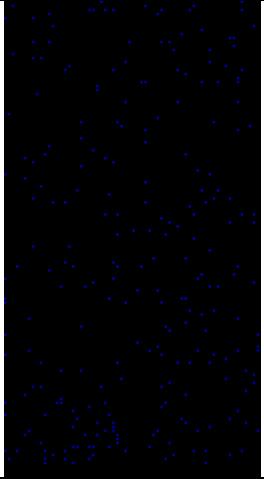
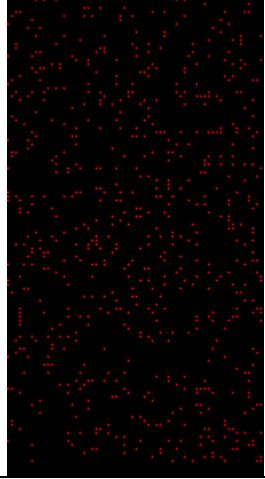
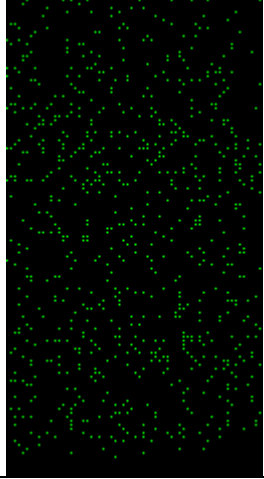
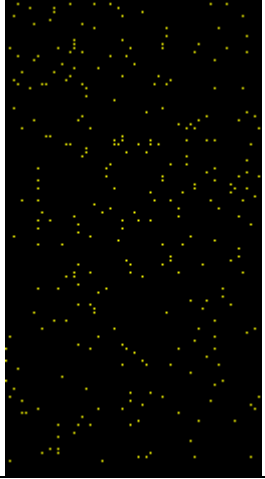
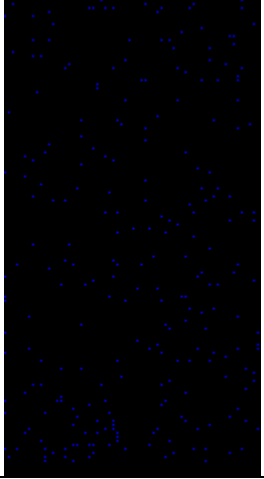


Fig. 29: Pairwise DC³s of genome 1 with other genomes



(2, 4)				
	DC ³ (A) = 595	DC ³ (T) = 736	DC ³ (G) = 349	DC ³ (C) = 321
(2, 5)				
	DC ³ (A) = 652	DC ³ (T) = 748	DC ³ (G) = 279	DC ³ (C) = 248
(2, 6)				
	DC ³ (A) = 652	DC ³ (T) = 748	DC ³ (G) = 279	DC ³ (C) = 248

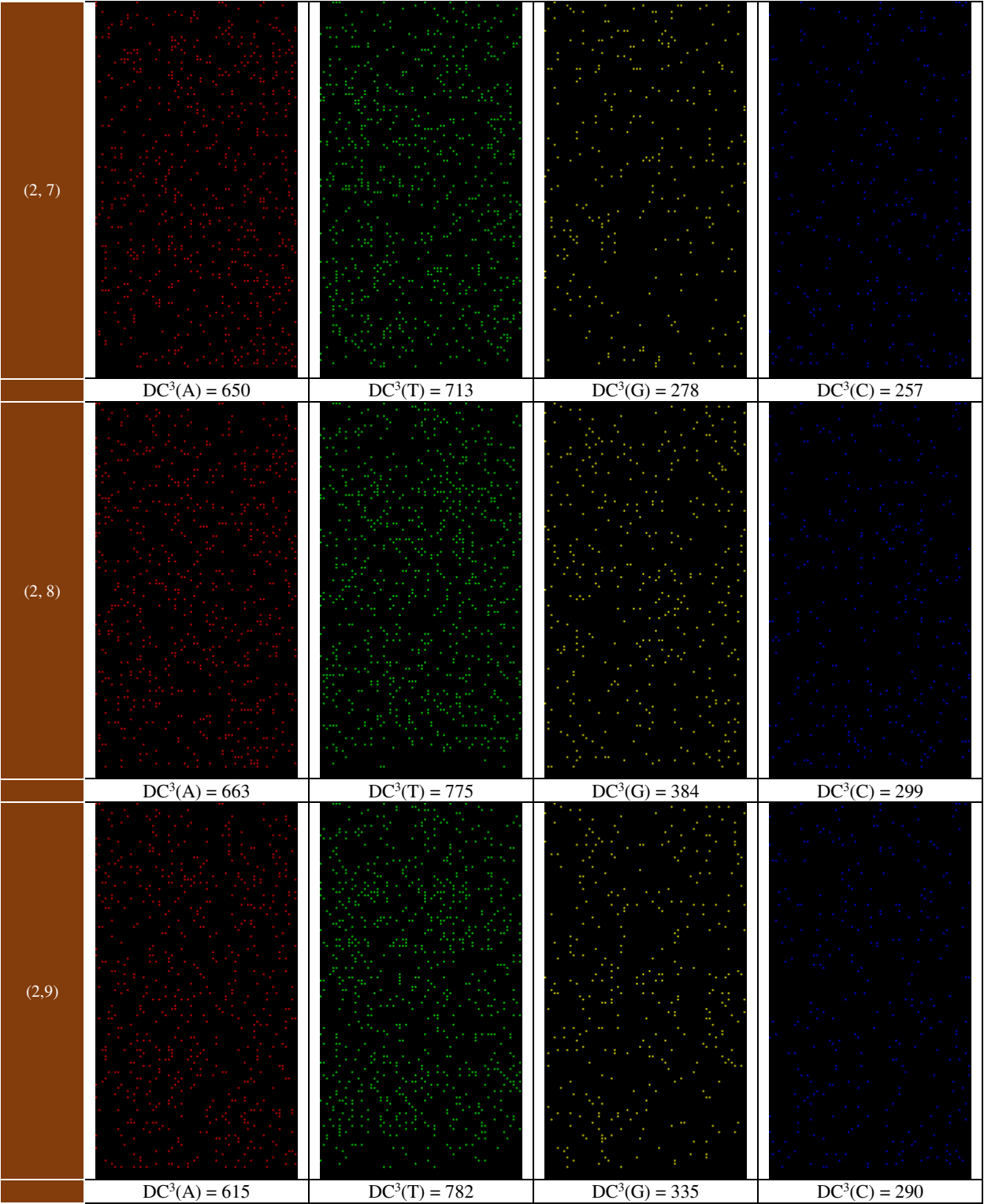
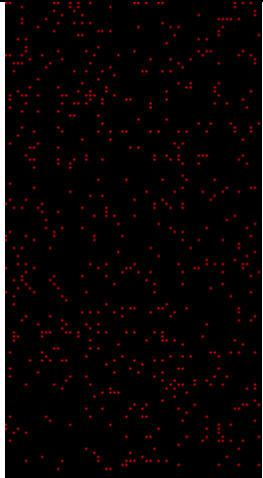
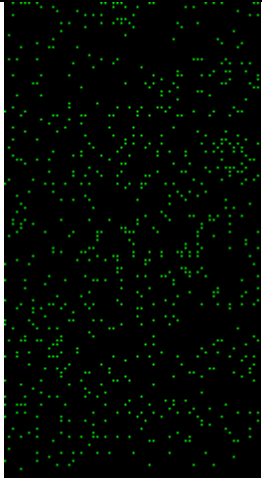
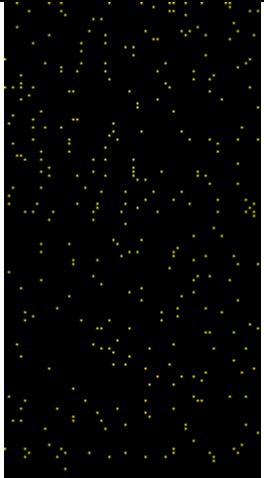
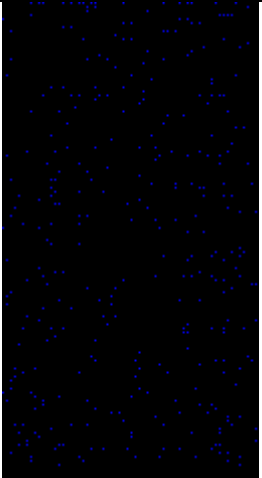
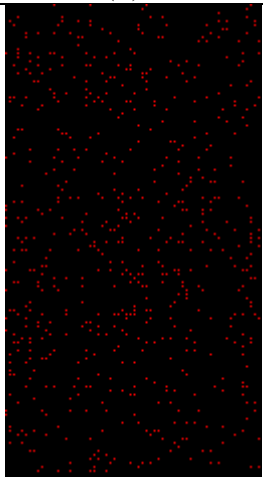
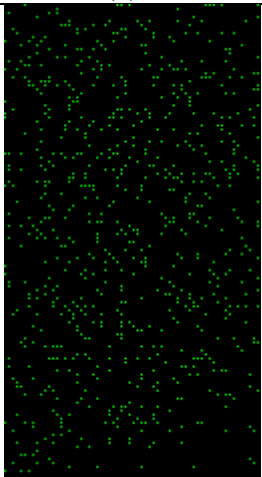
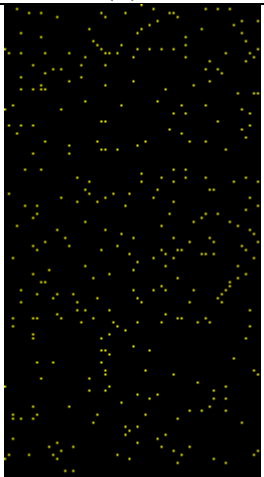
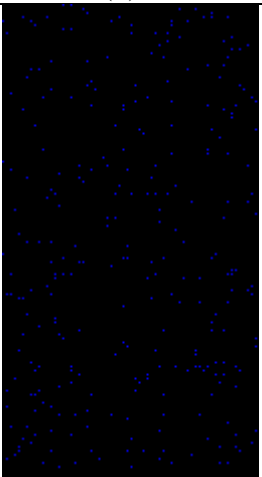
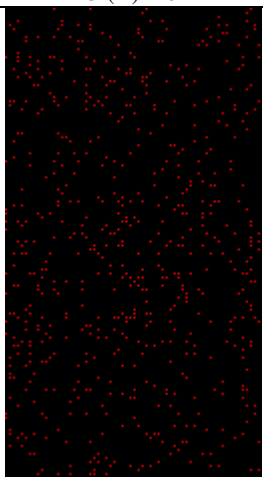
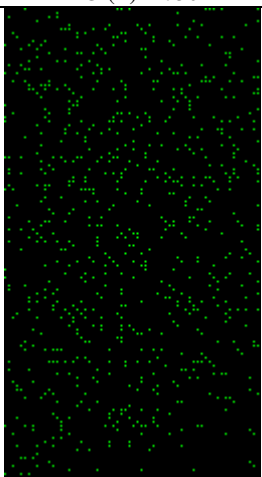
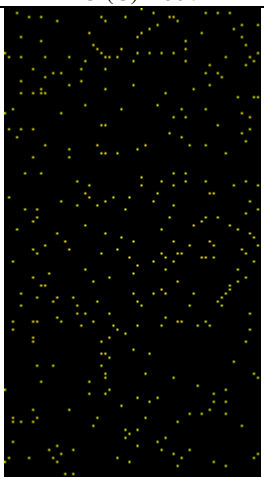
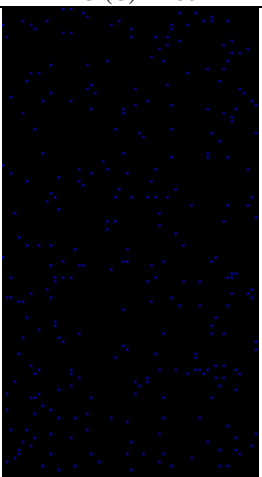


Fig. 30: Pairwise DC₃s of genome 2 with other genomes

Spatial Dyadic Cross Correlations of even subsampled adjoints arrays				
Genome (X, Y)	Adenine	Thymine	Guanine	Cytosine
(3, 4)				
	DC ³ (A) = 581	DC ³ (T) = 692	DC ³ (G) = 314	DC ³ (C) = 331
(3, 5)				
	DC ³ (A) = 621	DC ³ (T) = 750	DC ³ (G) = 337	DC ³ (C) = 280
(3, 6)				
	DC ³ (A) = 621	DC ³ (T) = 750	DC ³ (G) = 337	DC ³ (C) = 280

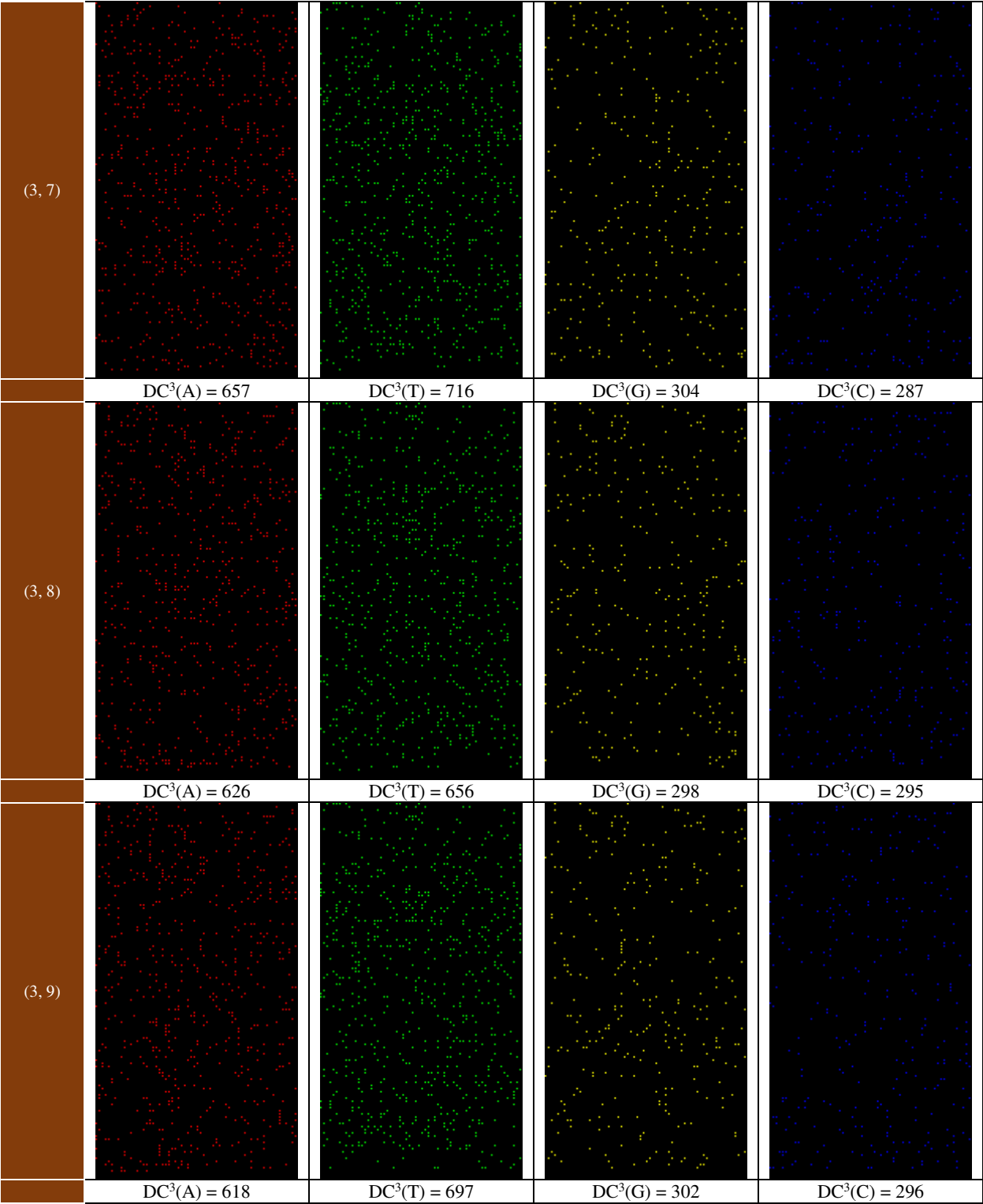
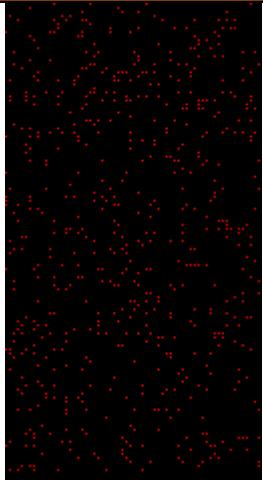
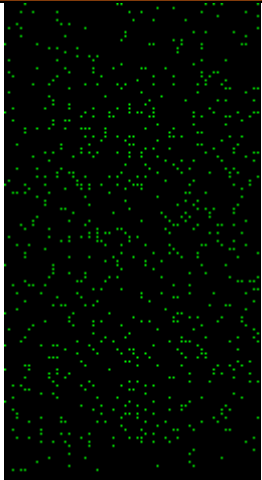
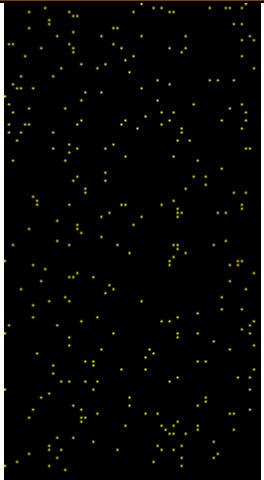
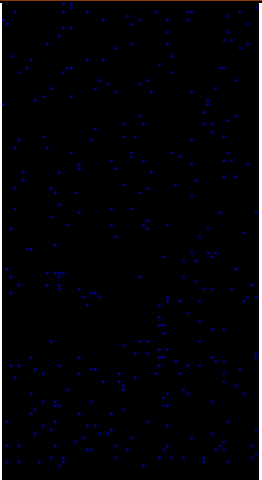
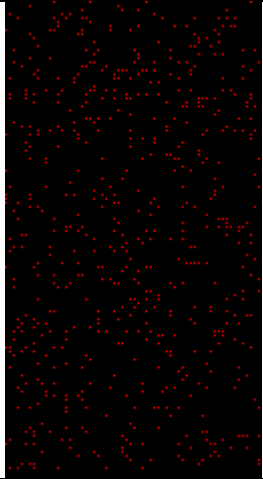
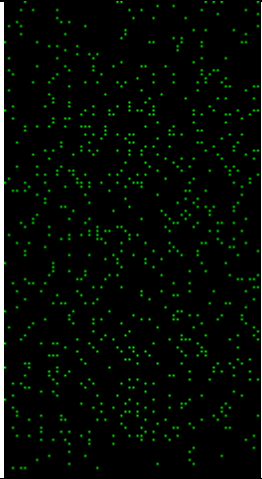
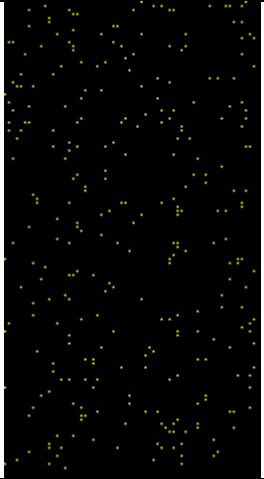
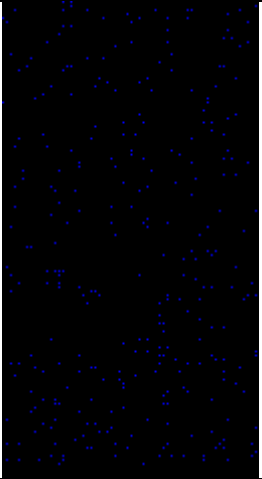
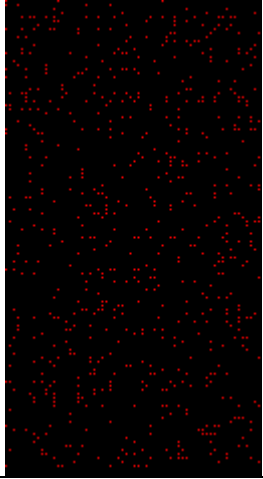
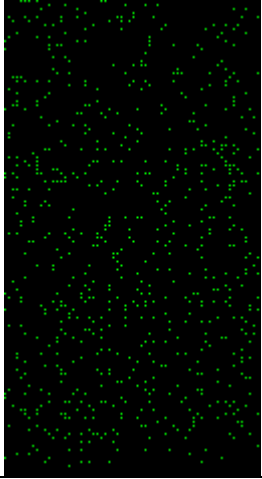
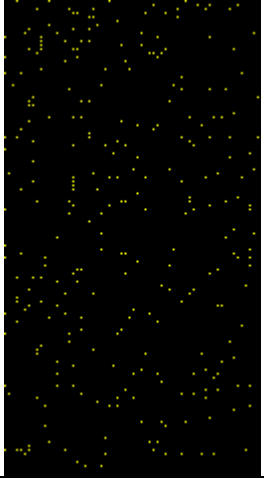
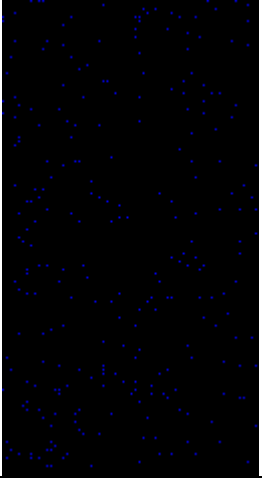


Fig. 31: Pairwise DC³s of genome 3 with other genomes

Spatial Dyadic Cross Correlations of even subsampled adjoints arrays				
Genome (X, Y)	Adenine	Thymine	Guanine	Cytosine
(4, 5)				
	DC ³ (A) = 608	DC ³ (T) = 726	DC ³ (G) = 297	DC ³ (C) = 299
(4, 6)				
	DC ³ (A) = 608	DC ³ (T) = 726	DC ³ (G) = 297	DC ³ (C) = 299
(4, 7)				
	DC ³ (A) = 669	DC ³ (T) = 709	DC ³ (G) = 289	DC ³ (C) = 259

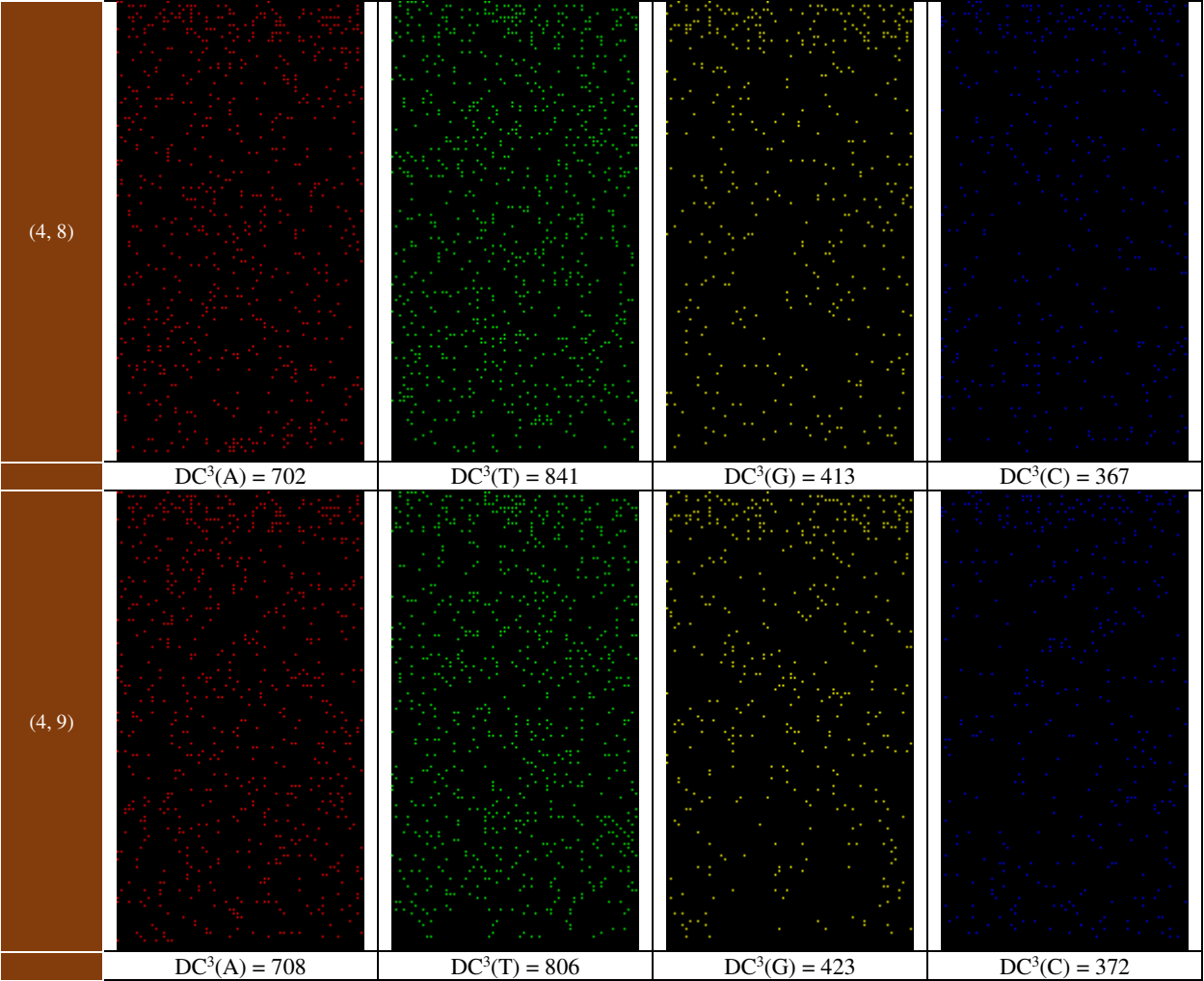
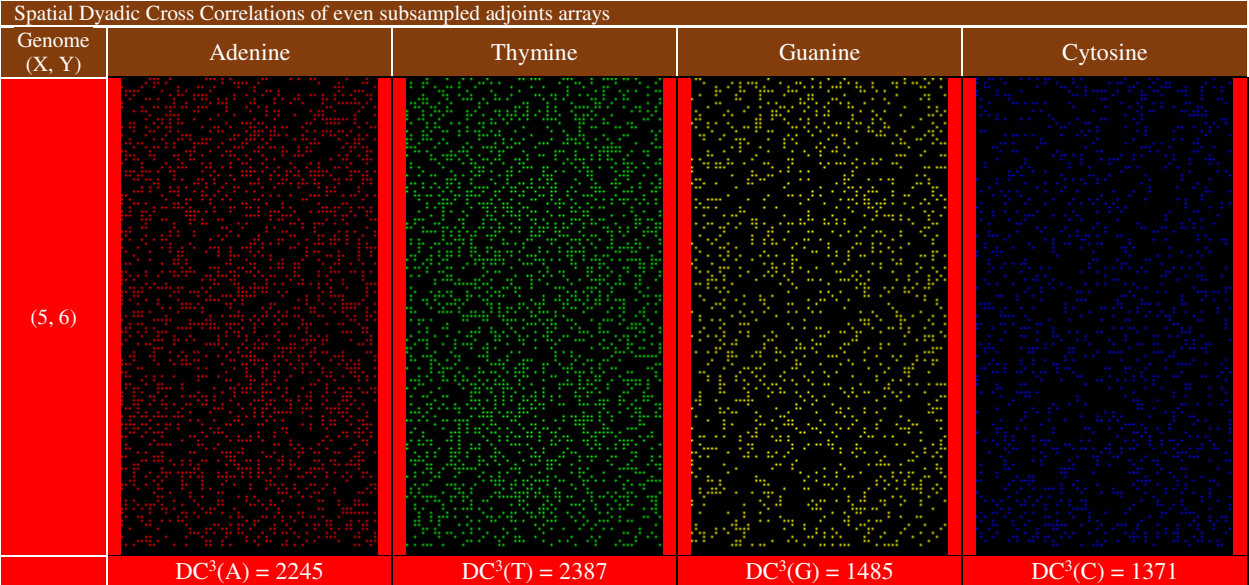


Fig. 32 : Pairwise DC³s of genome 4 with other genomes



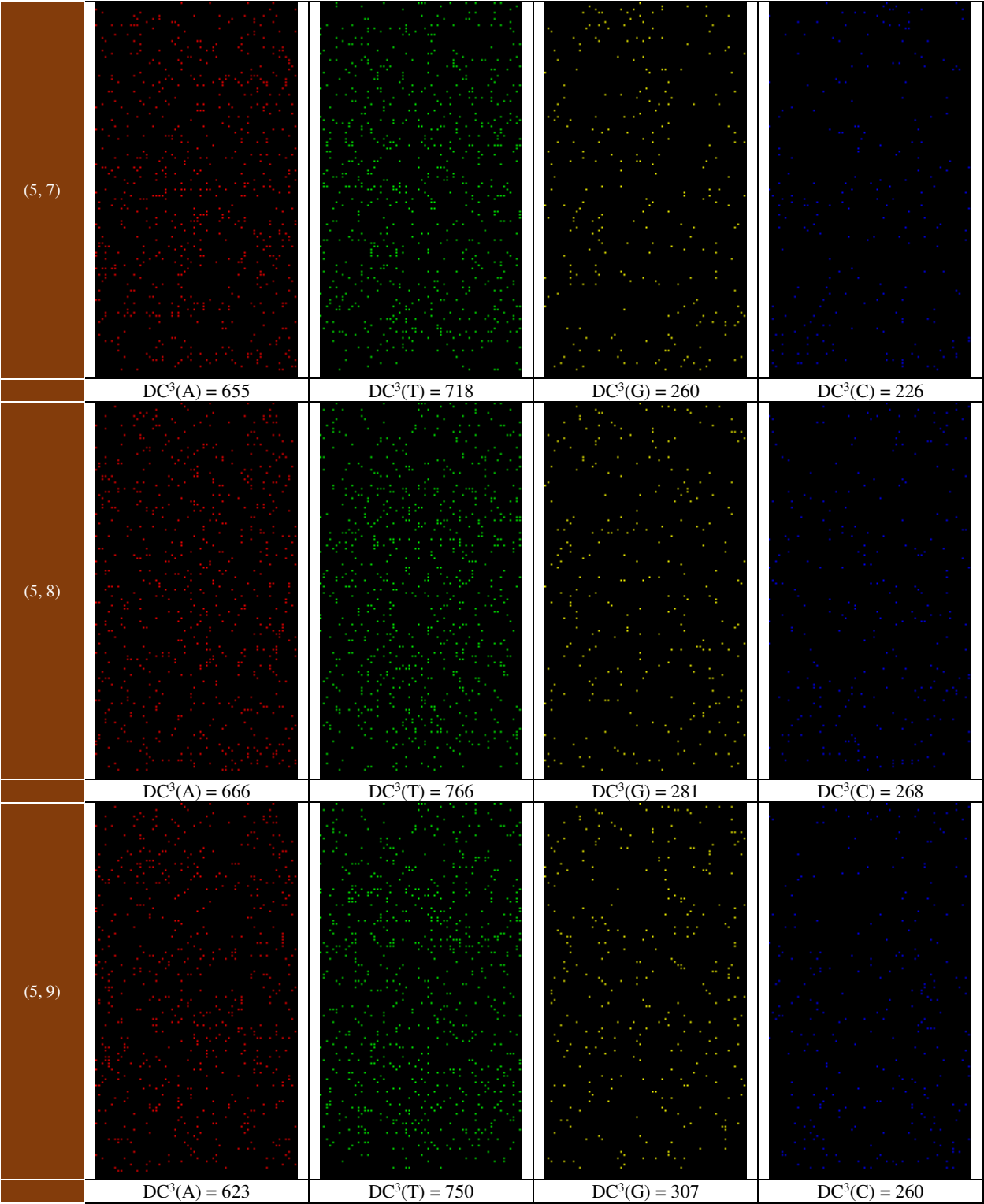


Fig. 33: Pairwise DC³s of genome 5 with other genomes

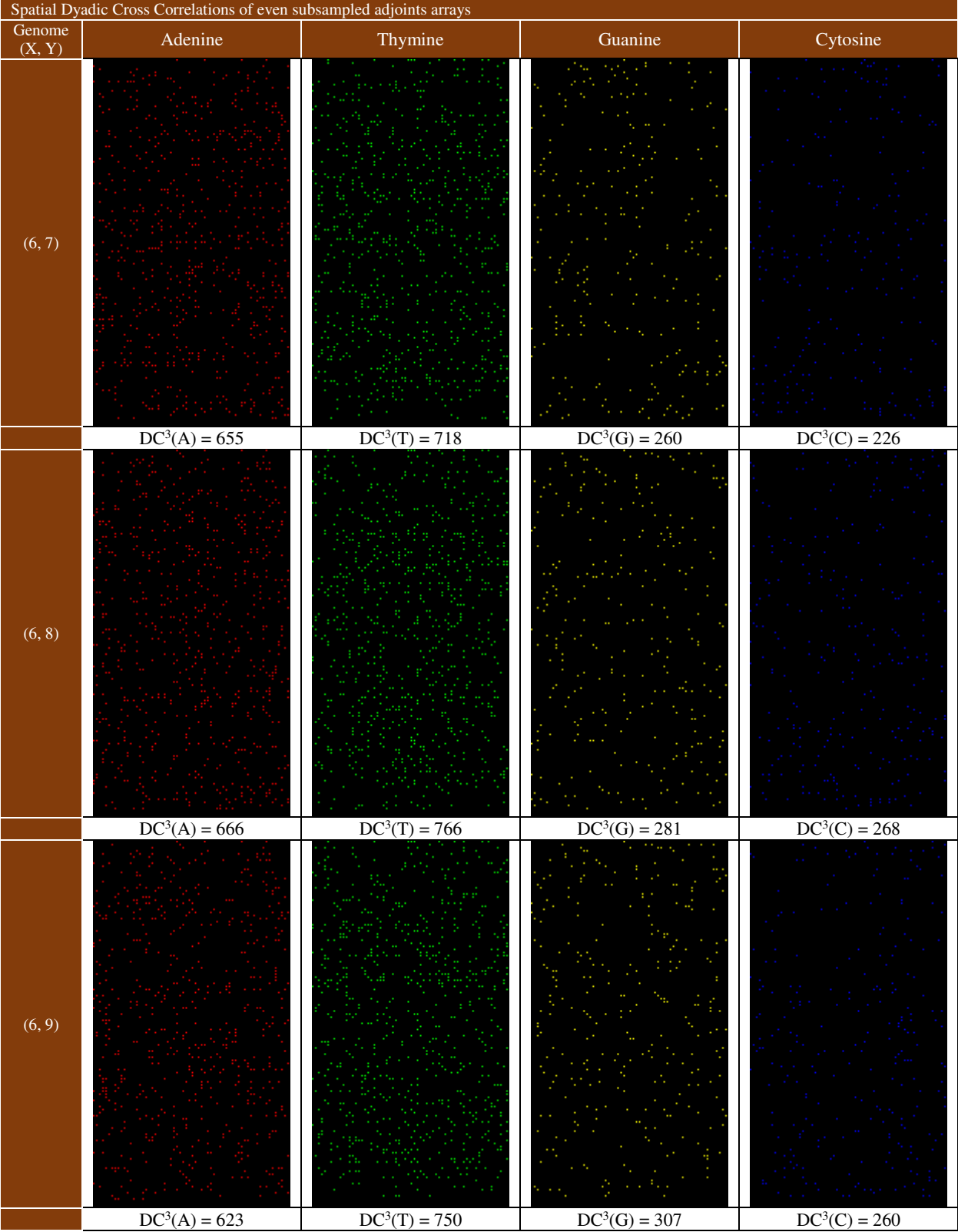


Fig34: Pairwise DC³s of genome 6 with other genomes

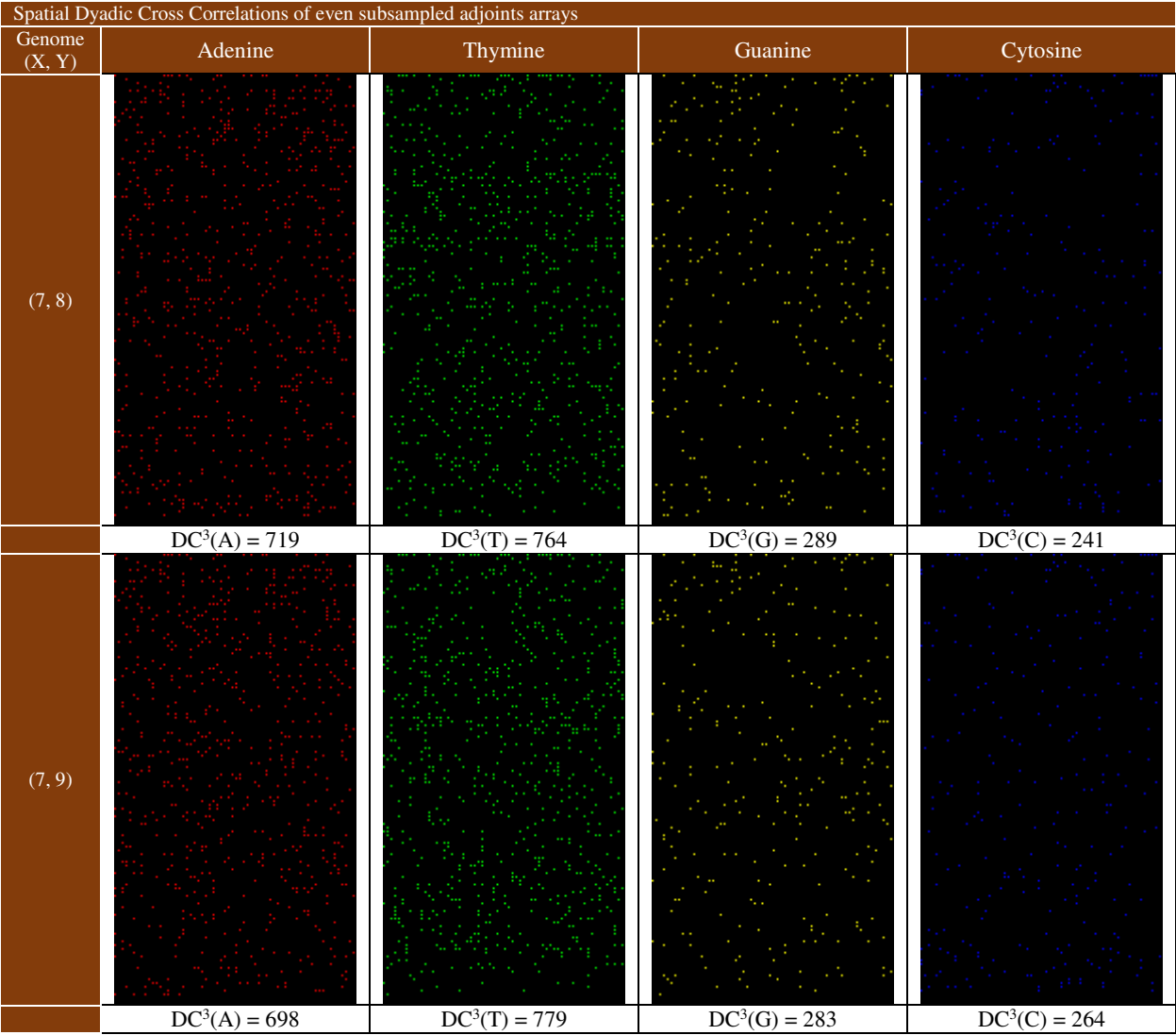


Fig. 35: Pairwise DC³s of genome 7 with other genomes

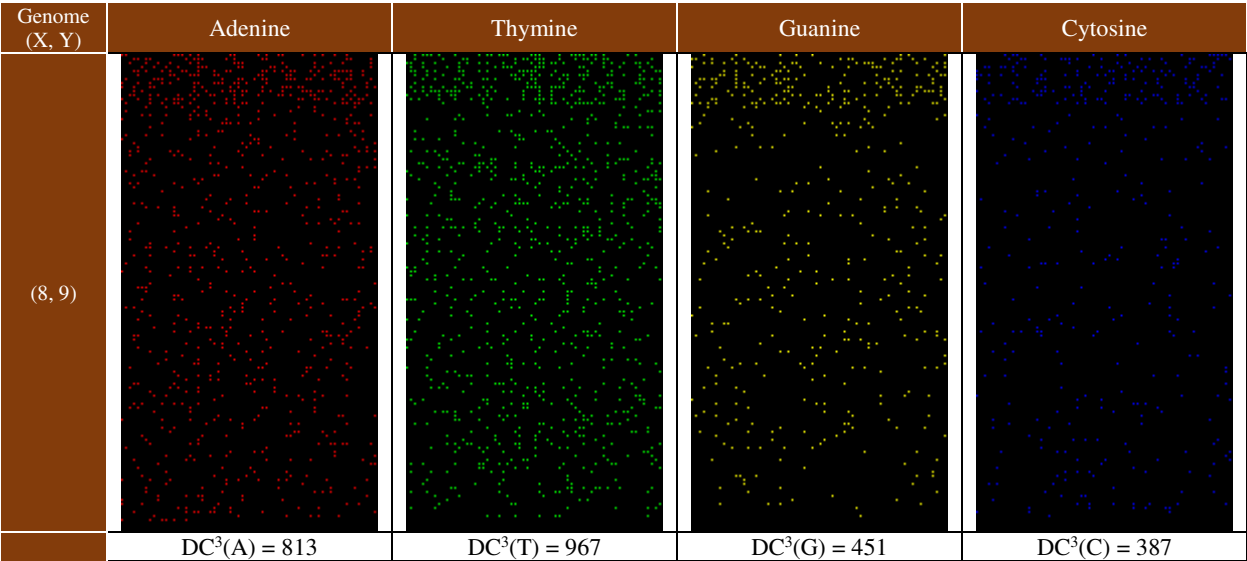


Fig. 36: Pairwise DC³s of genome 8 with genome 9

Fig. 37 shows the graph connecting various adjoint pairs and their DC^3 values.

Pairwise intersection property of even arrays [$DC^3(A)$, $DC^3(T)$, $DC^3(G)$ and $DC^3(C)$]

Sl. No.	Even subsampled Adjoints Genome Pair (X, Y)	$DC^3(A)$	$DC^3(T)$	$DC^3(G)$	$DC^3(C)$
1	(1, 2)	644	735	376	282
2	(1, 3)	612	684	322	295
3	(1, 4)	606	751	353	320
4	(1, 5)	649	766	313	278
5	(1, 6)	649	766	313	278
6	(1, 7)	637	715	297	237
7	(1, 8)	629	766	375	288
8	(1, 9)	656	787	391	284
9	(2, 3)	624	664	320	307
10	(2, 4)	595	736	349	321
11	(2, 5)	652	748	279	248
12	(2, 6)	652	748	279	248
13	(2, 7)	650	713	278	257
14	(2, 8)	663	775	384	299
15	(2, 9)	615	782	335	290
16	(3, 4)	581	692	314	331
17	(3, 5)	621	750	337	280
18	(3, 6)	621	750	337	280
19	(3, 7)	657	718	304	287
20	(3, 8)	626	656	298	295
21	(3, 9)	618	697	302	296
22	(4, 5)	608	726	297	299
23	(4, 6)	608	726	297	299
24	(4, 7)	669	709	289	259
25	(4, 8)	702	841	413	367
26	(4, 9)	708	806	423	372
27	(5, 6)	2245	2387	1485	1371
28	(5, 7)	655	718	260	226
29	(5, 8)	666	766	281	268
30	(5, 9)	623	750	307	260
31	(6, 7)	655	718	260	226
32	(6, 8)	666	766	281	268
33	(6, 9)	623	750	307	260
34	(7, 8)	719	764	289	241
35	(7, 9)	698	779	283	264
36	(8, 9)	813	967	451	387

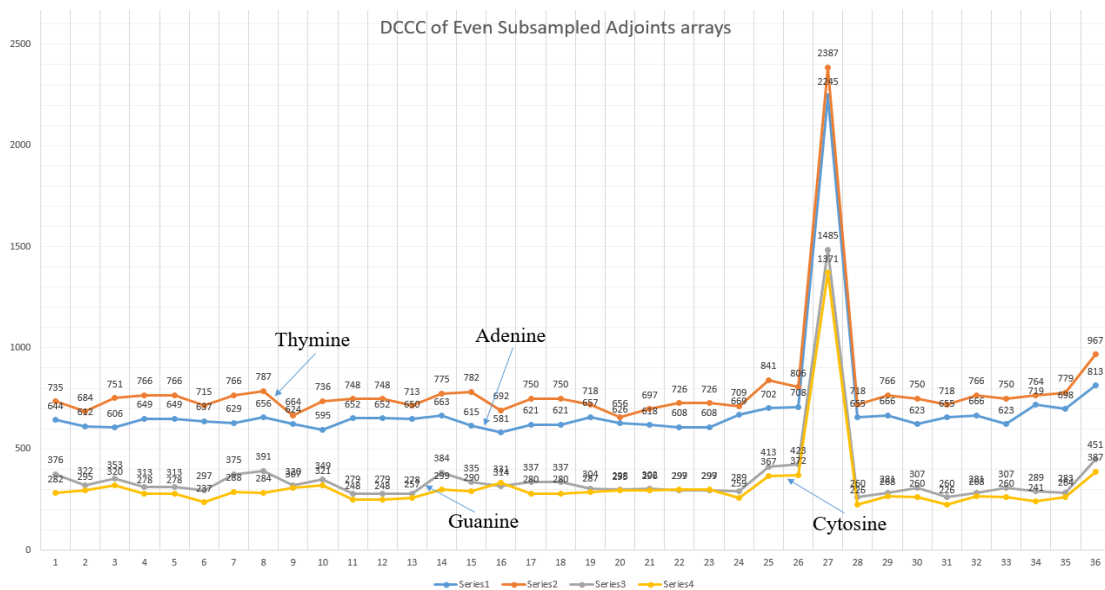


Fig. 37: Graph connecting adjoint pairs and DC^3 values

It is observed that the 27th unordered pair (Genome 5, Genome 6) shows up a **maximum number of matches of nucleotides** and this calls for special investigation of these two virus genomes. A visual presentation of adjoints of A, T, G and C of Genome 5 and Genome 6 along with Genome 6 superimposed on Genome 5 are shown in Figs. 38 - 41. Note that the length of Genome 5 (length = 30473) is longer than that of Genome 6 (length = 29875). It is to be noted that all images are of size 128x238 = 30,464.

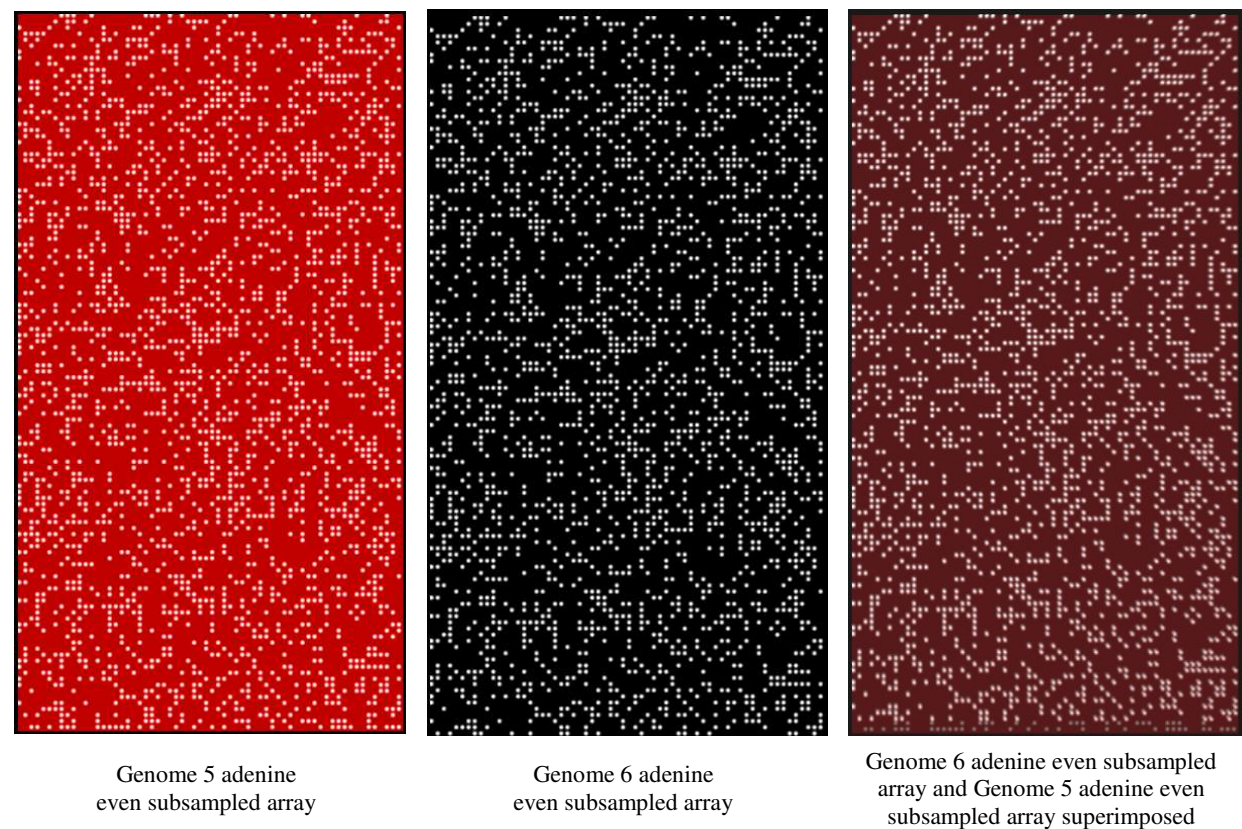


Fig. 38: Adenine even subsampled arrays of Genome 5 and 6 and their superimposed version

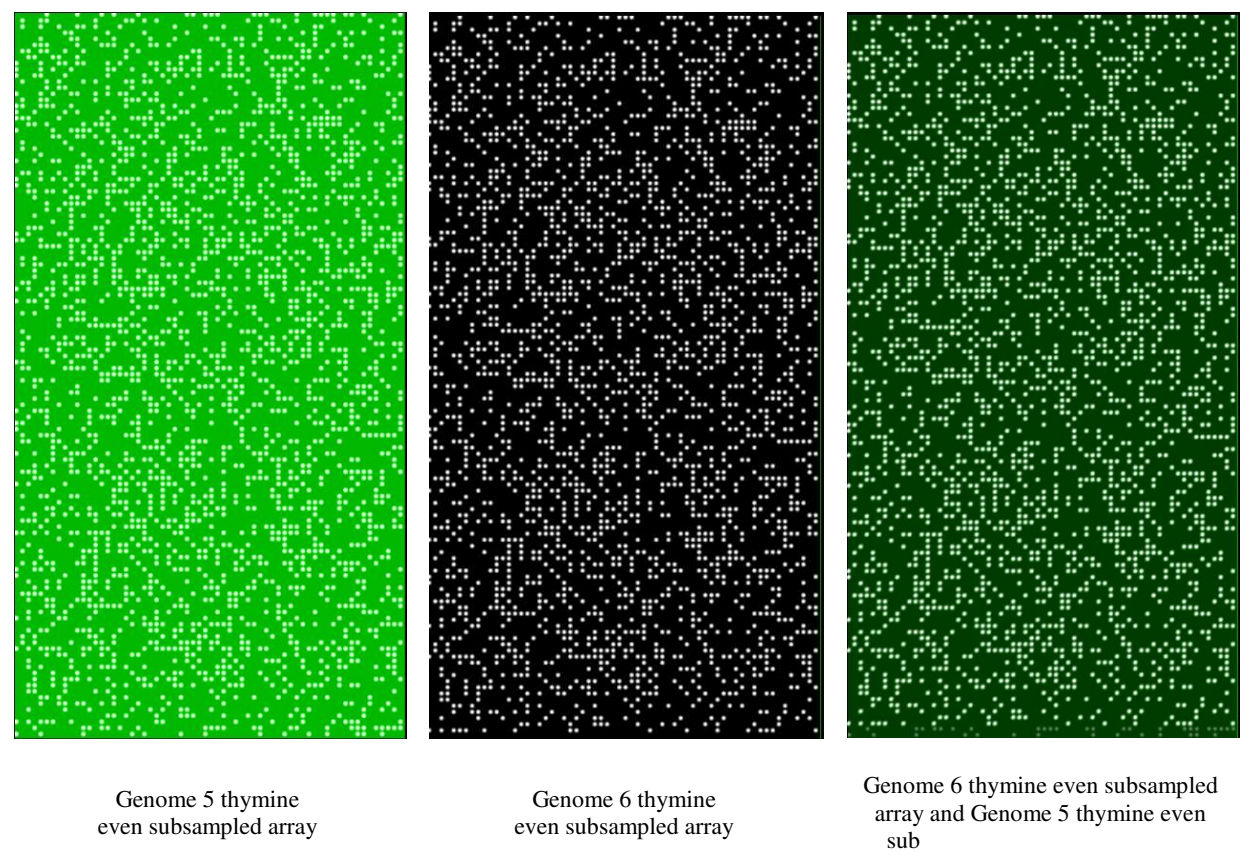


Fig. 39: Thymine even subsampled arrays of Genome 5 and 6 and their superimposed version

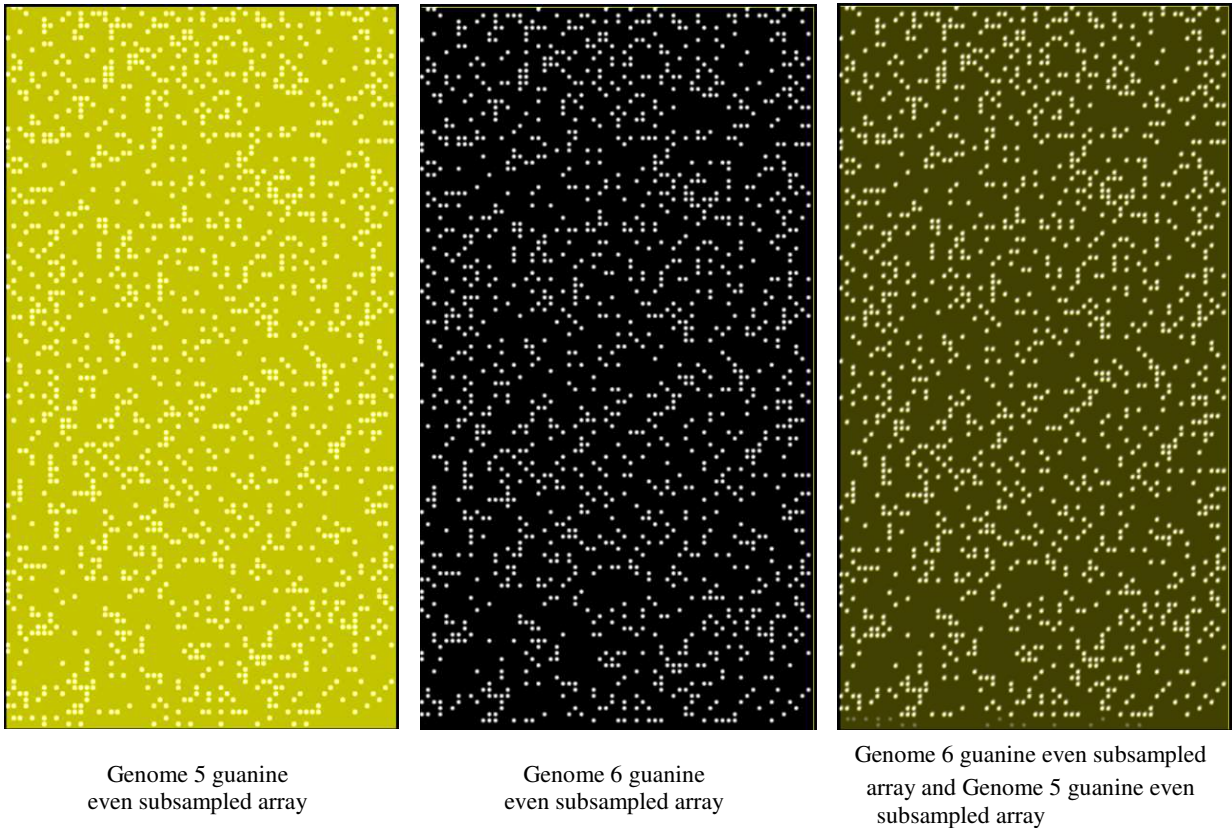


Fig. 40: Guanine even subsampled arrays of Genome 5 and 6 and their superimposed version

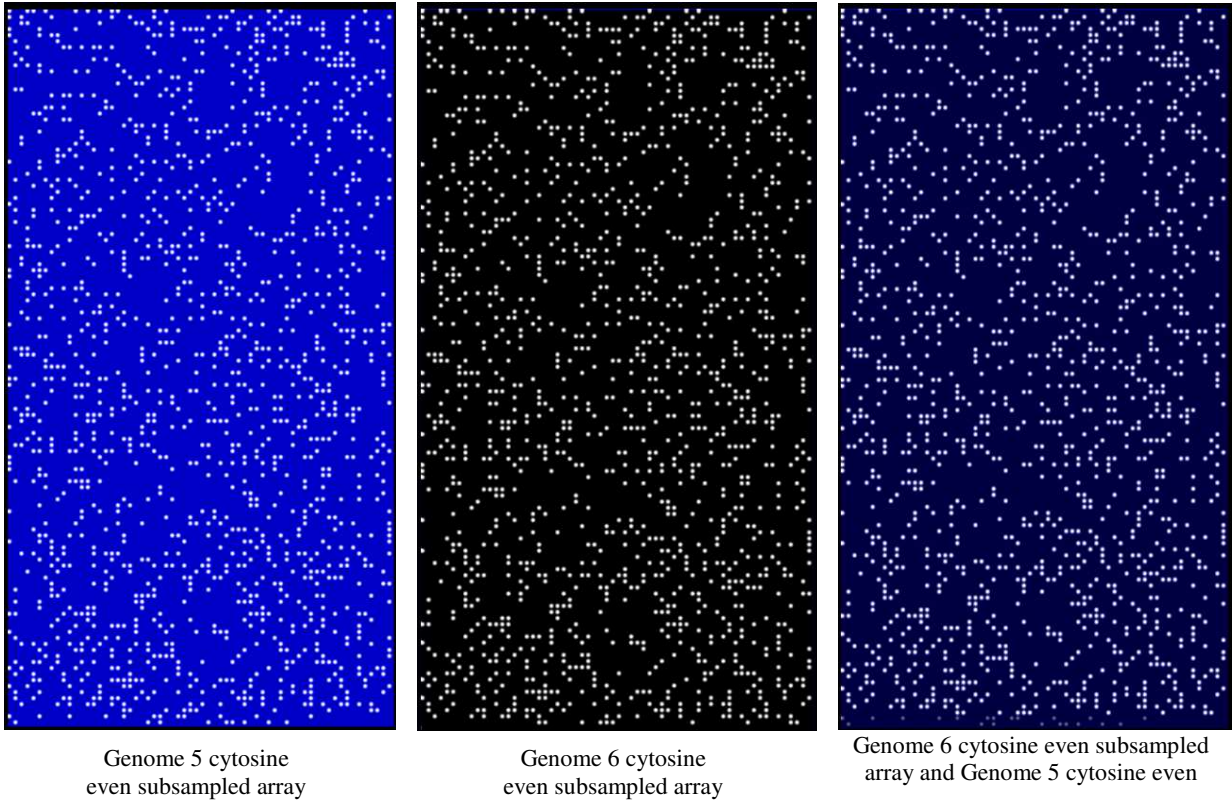


Fig. 41: Cytosine even subsampled arrays of Genome 5 and 6 and their superimposed version

For one pair of even subsampled adjoints arrays, one can have 121,125 (i.e., 255×475) Cross Correlation (CC) components and 60,563rd Cross Correlation (CC) component is the only CC component that corresponds to the ‘**Dyadic Cross Correlation**’ (DCC). One has to compare 4,84,500 (i.e., $121,125 \times 4$) Cross Correlation (CC) components just for one genome. For nine genomes, one has to compare 43,60,500 CC components. This argument is meant for comparing even subsampled adjoints arrays of nine genomes. So, one has to compare 87,21,000 CC components for both even and odd subsampled adjoints arrays. This number increases exponentially when the number of genomes to be compared is increased. There are 36 even subsampled adjoint array pairs for all nine genomes. Out of 36 pairs, one pair, that is, 27th pair exhibits maximum DC³. Let us consider the 36th pair (Genome 8 and Genome 9) and find out its DC³. Fig. 42 shows adenine even subsampled arrays of Genome 8 and 9 and superimposed version

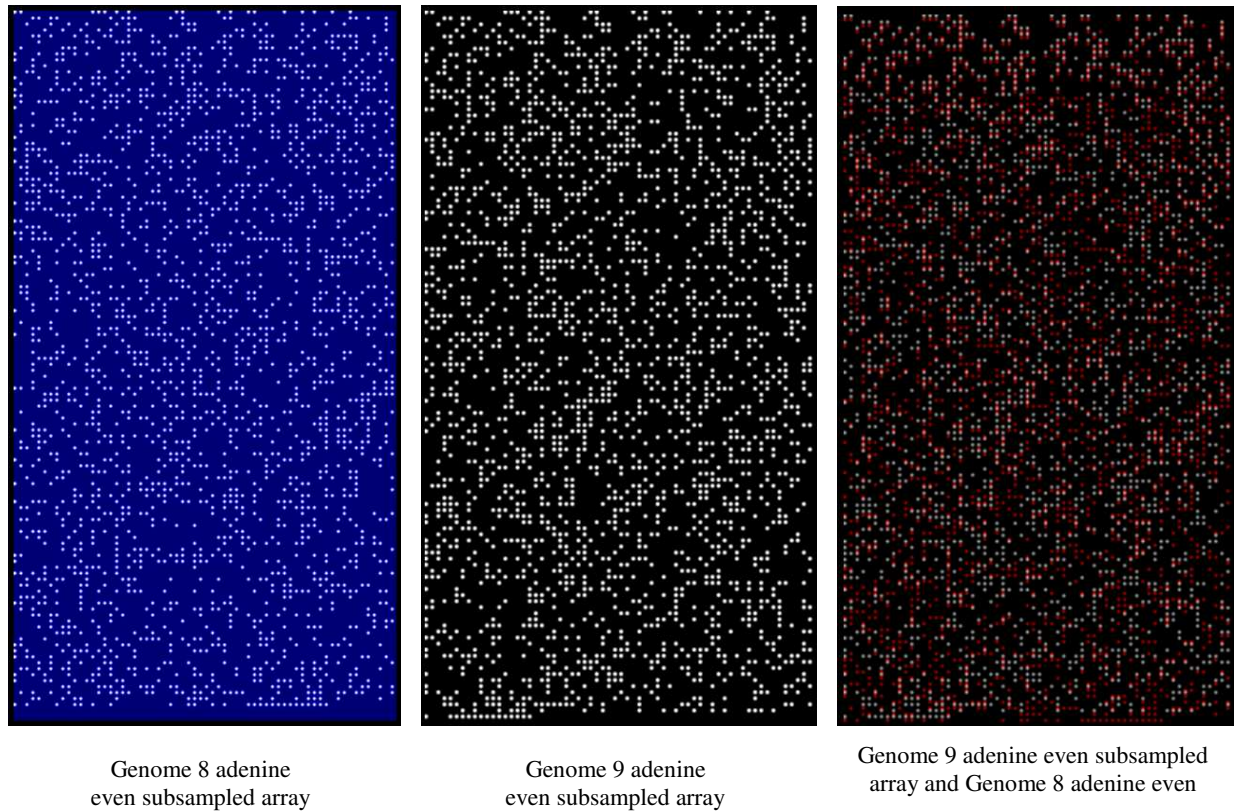


Fig. 42: Adenine even subsampled arrays of Genome 8 and 9 and their superimposed version

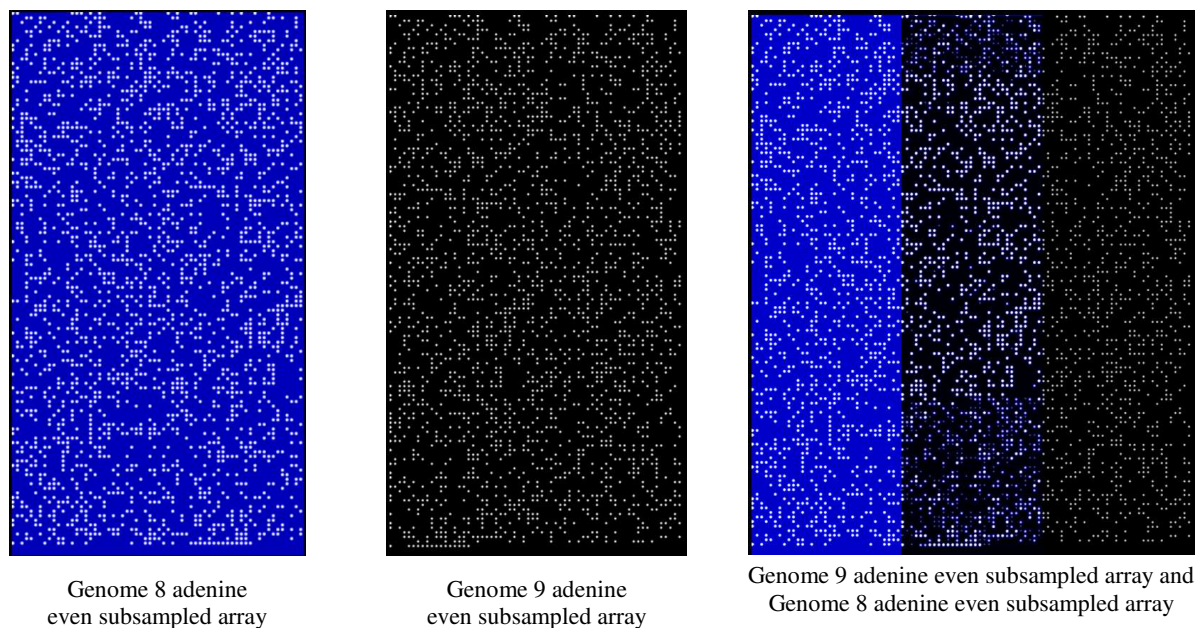


Fig. 43: Adenine even subsampled arrays of Genome 8 and 9 and their partially superimposed version

Fig. 43 shows adenine even subsampled arrays of Genome 8 and 9 and their partially superimposed version. Some significant observations made among Fig. 42 and Fig. 43 are given below.

1. The 27th pair (5, 6) among 36 pairs exhibits the highest DC^3 value, that is, 2245, and the next highest DC^3 value 813 is exhibited by the 36th pair (8, 9).
2. Though the dyadic comparison of the pair (8, 9) shows 813 adenine positions common to both arrays, they are seen to be randomly distributed, that is, sparsely, and such random distribution does not convey any meaningful interpretation or a common pattern as regards their spatial correlation is concerned (Ref. Fig. 42).
3. On the other hand, when Genome 9 adenine even subsampled array is horizontally shifted towards right with respect to Genome 8 adenine even subsampled array and partially superimposed, the middle portion of the intersected region shows a common pattern (Ref. Fig. 43).
4. This amounts to saying that one can shift one array in eight possible directions with respect to another array in a digital lattice and look for common digital patterns. One can have 4,84,496 partial superpositions of the arrays

Each of the 4,84,496 superpositions of the even subsampled adjoint arrays may yield a common digital pattern individually. These common digital patterns are the objects of interest for studying spatial correlations between various genome adjoint arrays. All these correlations are called non-dyadic cross correlations.

Non-Dyadic Cross Correlations

There are 1,21,124 non-dyadic cross correlations for a pair of even subsampled adjoint arrays. Every non-dyadic cross correlation has a Cross Correlation Coefficient (C^3). Now, every C^3 is not DC^3 , but every DC^3 is essentially a C^3 . Many non-dyadic cross correlations are observed to yield large common digital patterns. Given a pair of arrays, one can shift in eight directions with respect to another. For example, Fig. 44 shows two equisized arrays and eight shift

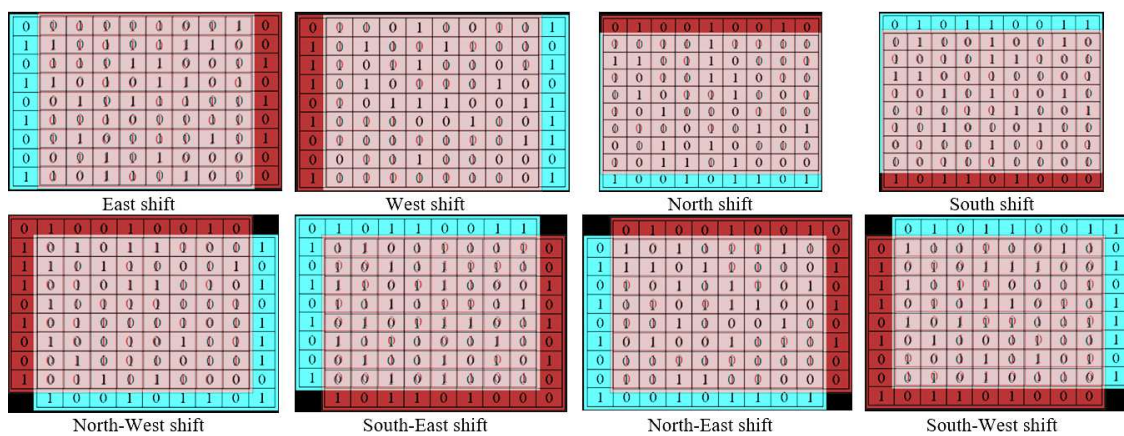


Fig. 44: Eight shifts of one equisized array with respect to the other

Some of the non-dyadic cross correlations are shown in Fig. 45, Fig. 46 and Fig. 47.

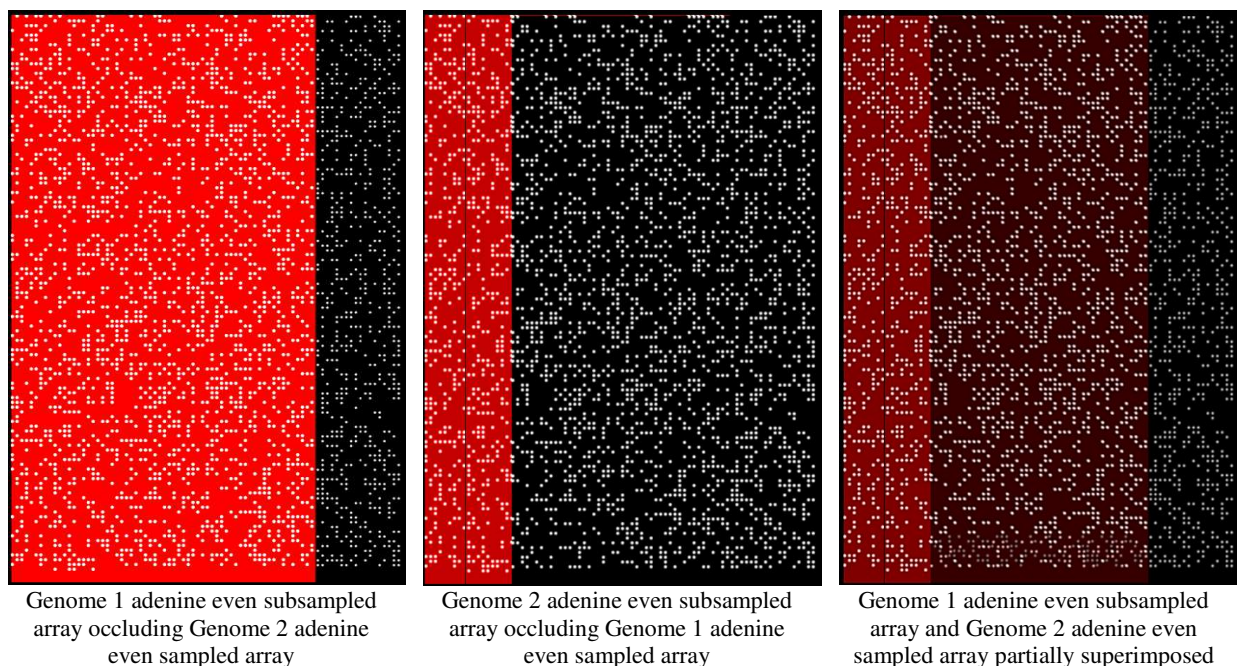


Fig. 45: Genome 1 adenine and Genome 2 adenine adjoint even subsampled arrays superimposed (maximum match found)

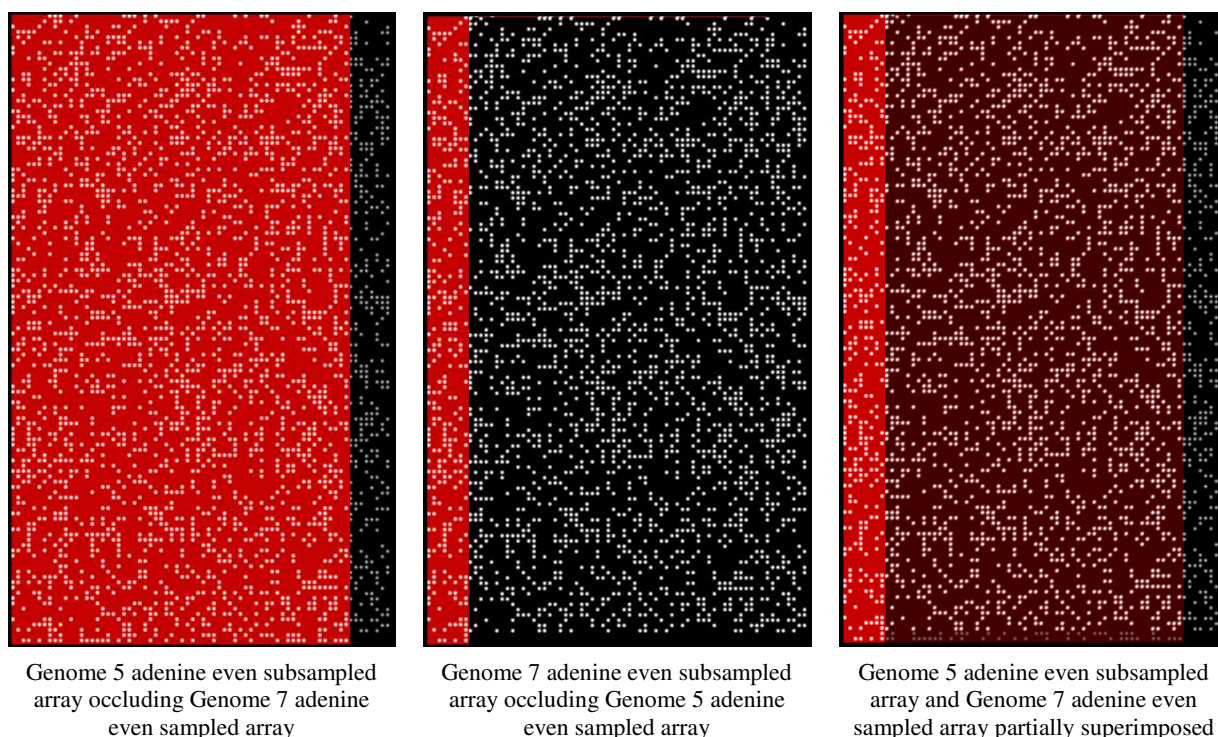


Fig. 46: Genome 5 adenine and Genome 7 adenine adjoints even subsampled arrays superimposed (maximum match found)

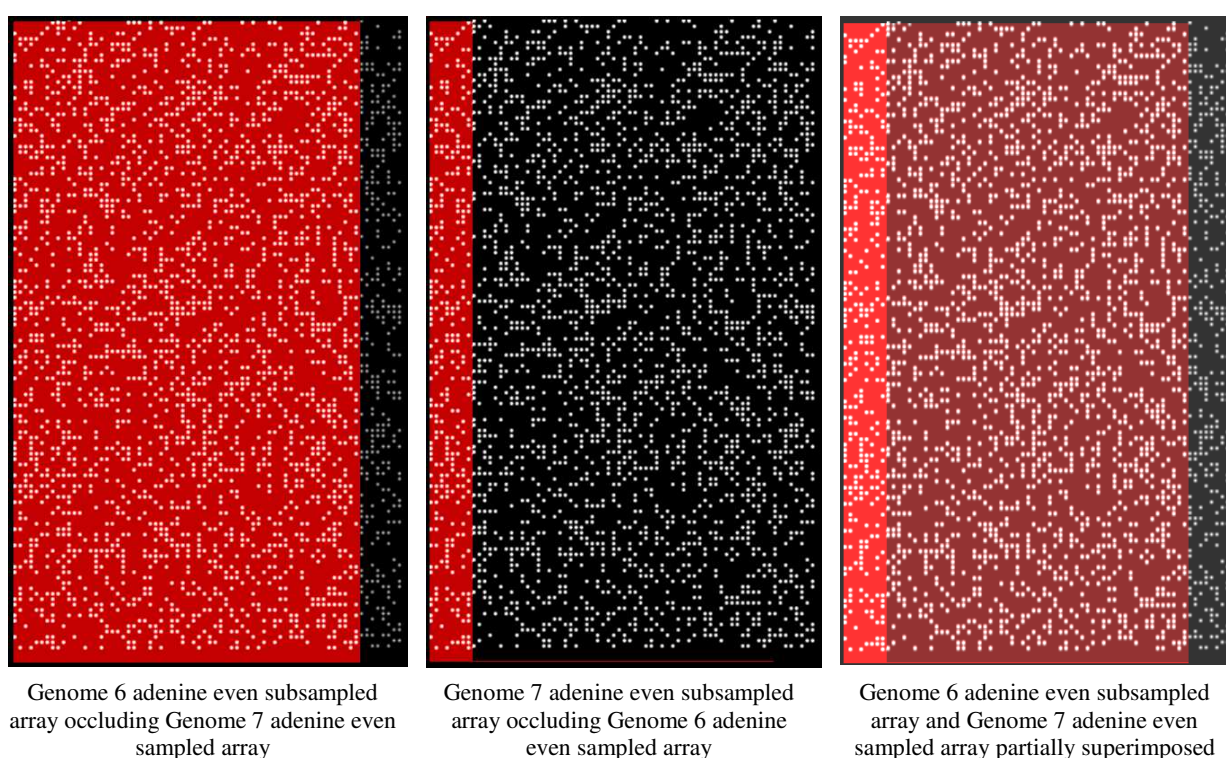


Fig. 47: Genome 6 adenine and Genome 7 adenine adjoints even subsampled arrays superimposed (maximum match found)

One can clearly visualize in Fig. 45, Fig. 46 and Fig. 47, common digital patterns of different widths but almost of the same height of those of the corresponding partially superimposed even subsampled adjoints arrays. This observation paves a constructive way to analyze the respective genes in a laboratory, especially in the light of virus genome evolution.

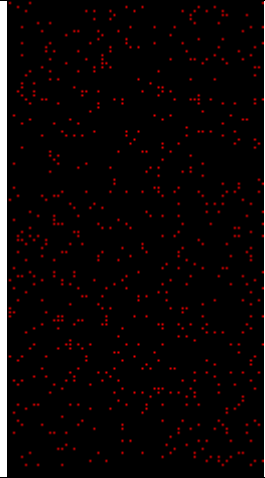
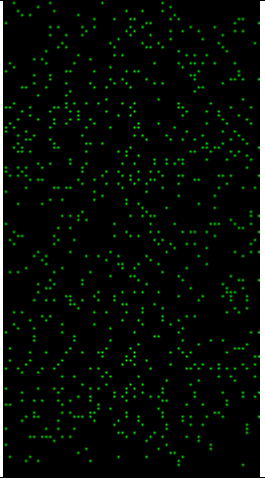
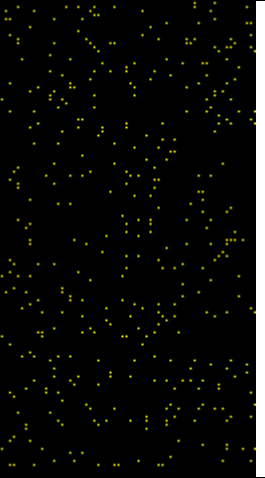
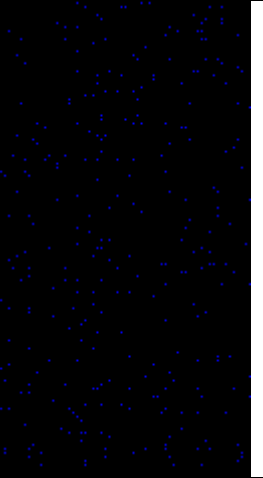
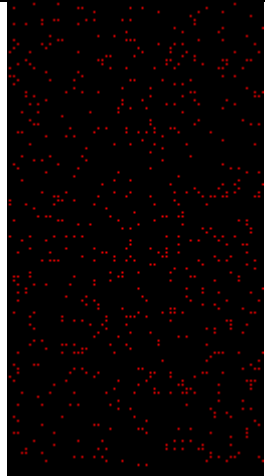
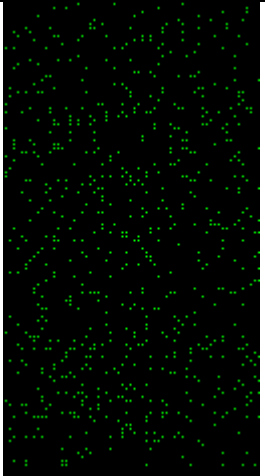
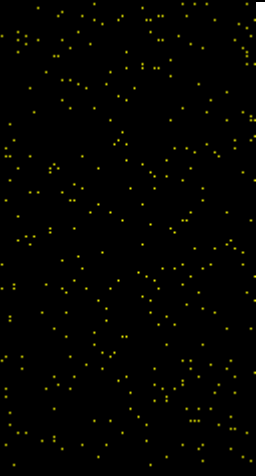
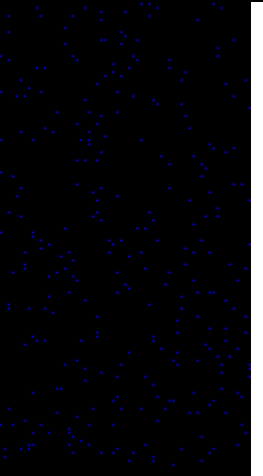
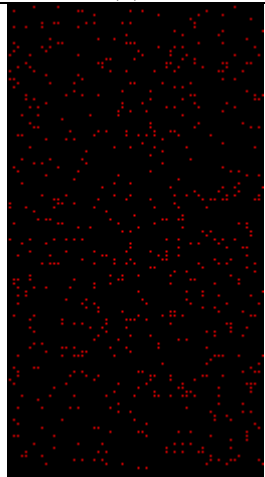
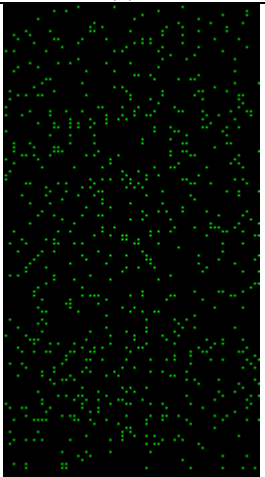
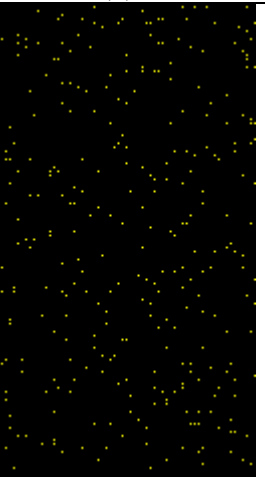
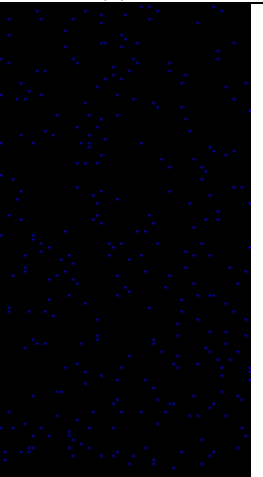
Having studied various aspects of cross correlations of even subsampled adjoints arrays of all nine virus genomes, let us turn our attention on a similar study of cross correlations of odd subsampled adjoints arrays of all nine virus genomes.

Pairwise spatial correlation (intersection) of odd subsampled adjoints arrays

As it was done in the case of even subsampled arrays, the odd subsampled adjoints arrays are examined for cross correlation, dyadically. Dyadic cross correlation is a logical AND operation that involves two different adjoint arrays. The purpose of doing this is to find out how many nucleotides, say for example, adenines are present in both the genome sequences at the same locations. The number of such adenines present in the same locations of both the genome sequences is called ‘Dyadic Cross Correlation Coefficient (DCCC or DC^3)’. In other words, $DC^3(A)$, $DC^3(T)$, $DC^3(G)$ and $DC^3(C)$ for genome sequences X and Y yield an apparent spatial correlation between them. For one pair of odd subsampled adjoints arrays, one can have **1,21,125 Cross Correlation (CC) components** and **60,563rd Cross Correlation (CC) component** will correspond to the ‘**Dyadic Cross Correlation**’ (DCC). In what follows, the dyadic cross correlations of odd subsampled adjoints arrays are presented. DC^3 represents ‘**Dyadic Cross Correlation Coefficient**’ connecting two arrays. Figs. 48 to xx show all possible DCCs and their coefficient values.

Dyadic Cross Correlations of Odd Subsampled Adjoints Arrays

Spatial Dyadic Cross Correlations of odd subsampled adjoints arrays				
Genome (X, Y)	Adenine	Thymine	Guanine	Cytosine
(1, 2)				
	$DC^3(A) = 607$	$DC^3(T) = 741$	$DC^3(G) = 374$	$DC^3(C) = 275$
(1, 3)				
	$DC^3(A) = 589$	$DC^3(T) = 661$	$DC^3(G) = 327$	$DC^3(C) = 286$

(1, 4)				
	DC ³ (A) = 625	DC ³ (T) = 750	DC ³ (G) = 388	DC ³ (C) = 295
(1, 5)				
	DC ³ (A) = 626	DC ³ (T) = 717	DC ³ (G) = 301	DC ³ (C) = 291
(1, 6)				
	DC ³ (A) = 626	DC ³ (T) = 717	DC ³ (G) = 301	DC ³ (C) = 291

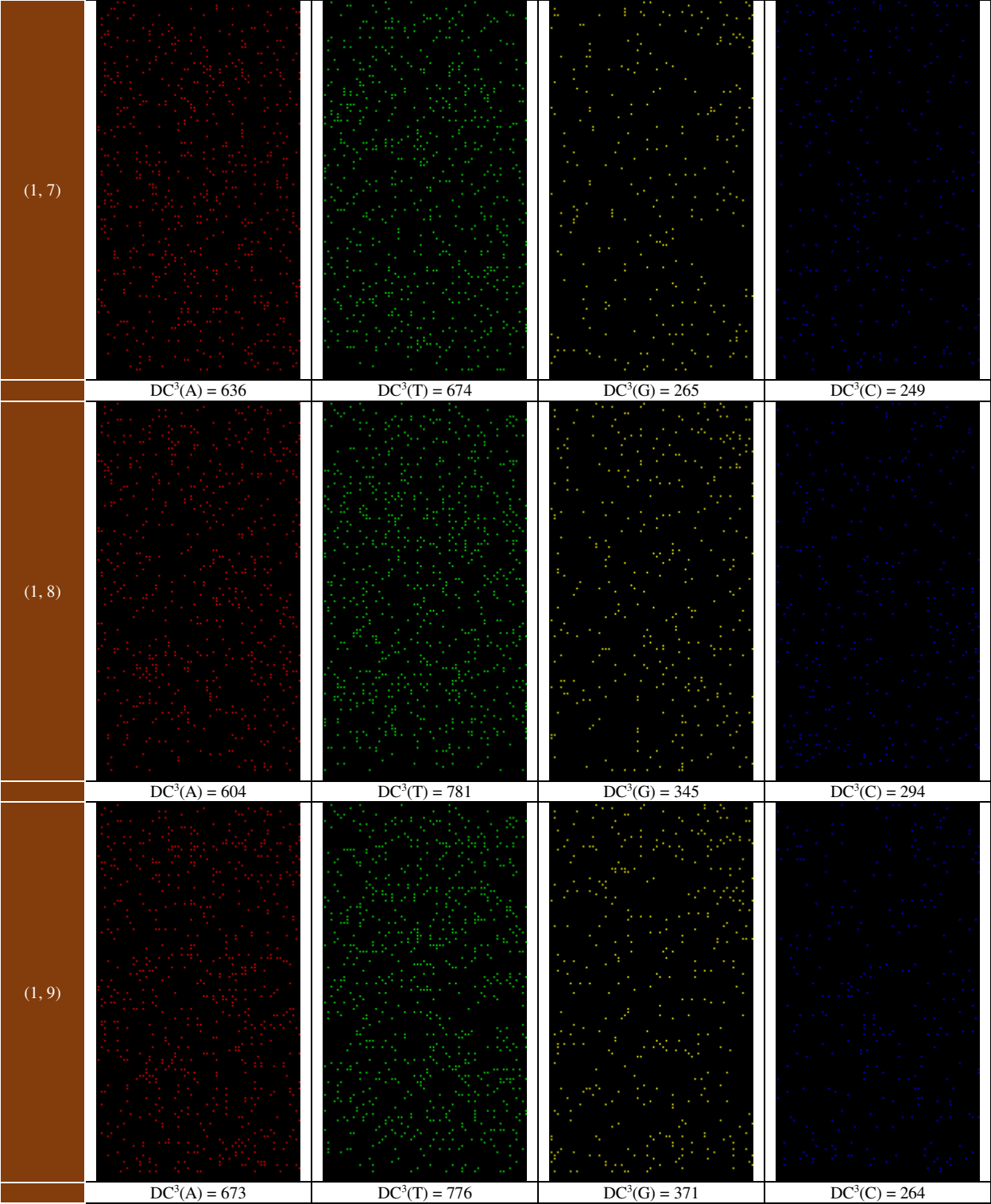
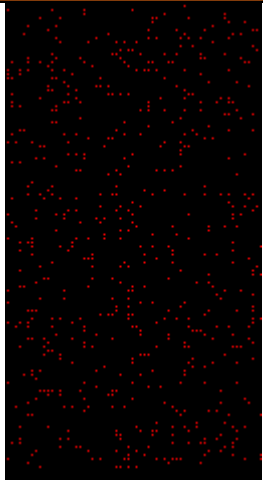
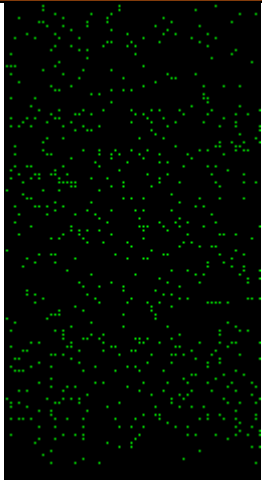
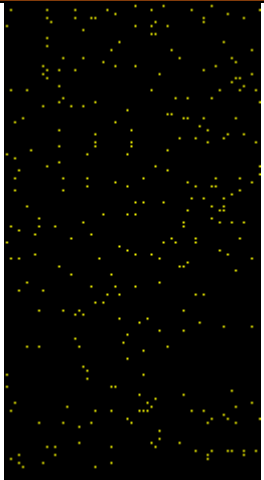
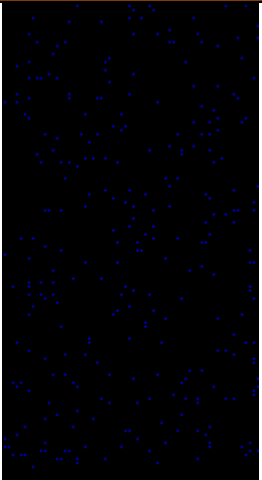
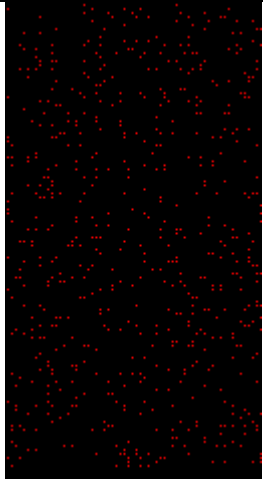
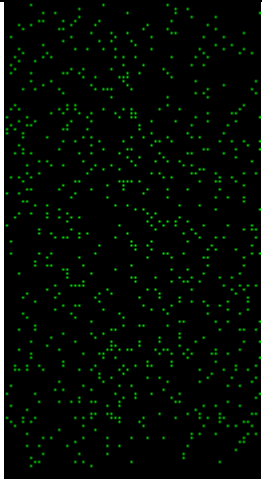
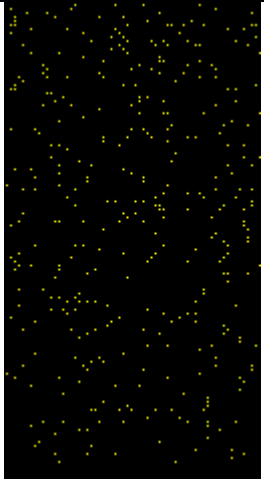
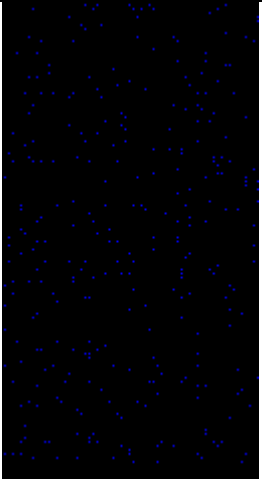
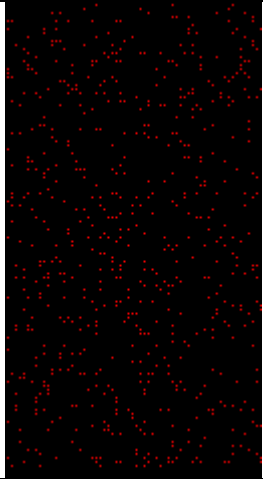
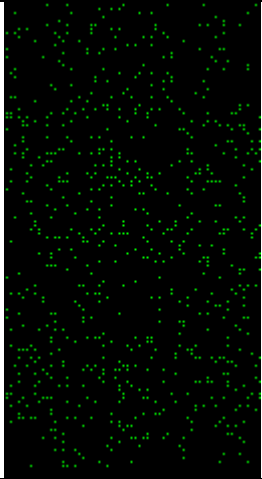
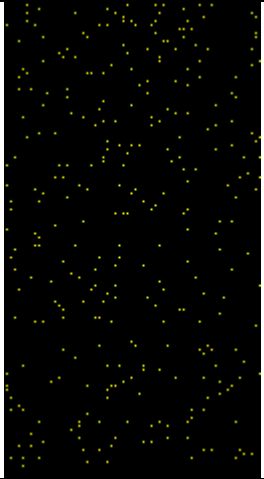
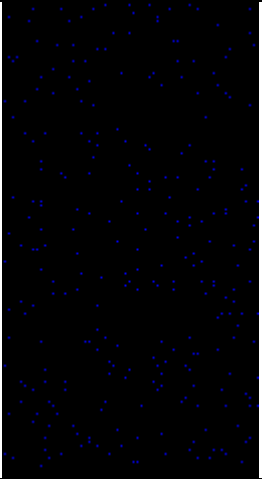
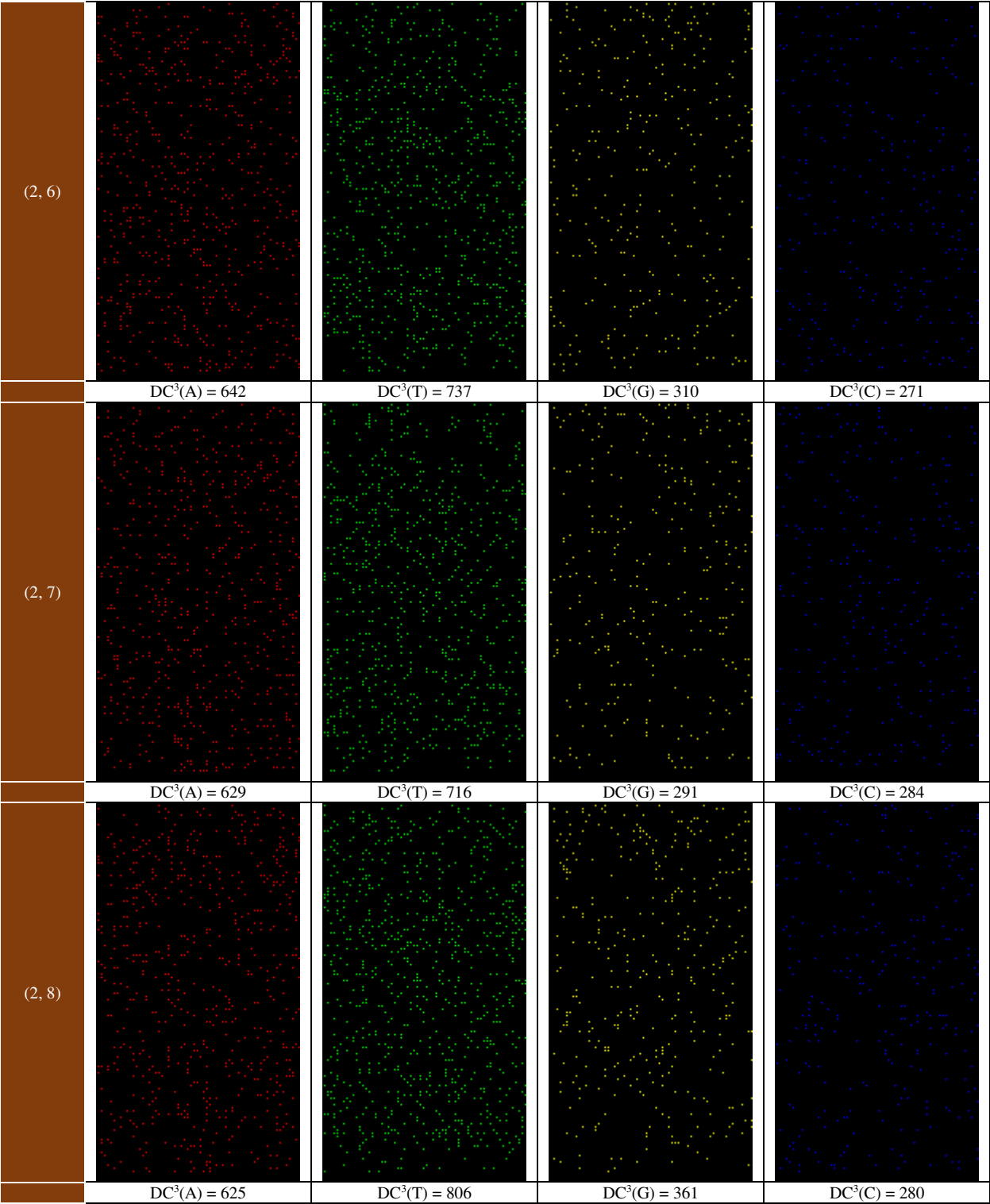


Fig. 48: Pairwise DC³s of genome 1 with other genomes

Spatial Dyadic Cross Correlations of odd subsampled adjoints arrays				
Genome (X, Y)	Adenine	Thymine	Guanine	Cytosine
(2, 3)				
	DC ³ (A) = 603	DC ³ (T) = 643	DC ³ (G) = 288	DC ³ (C) = 285
(2, 4)				
	DC ³ (A) = 608	DC ³ (T) = 726	DC ³ (G) = 365	DC ³ (C) = 301
(2, 5)				
	DC ³ (A) = 642	DC ³ (T) = 737	DC ³ (G) = 310	DC ³ (C) = 271



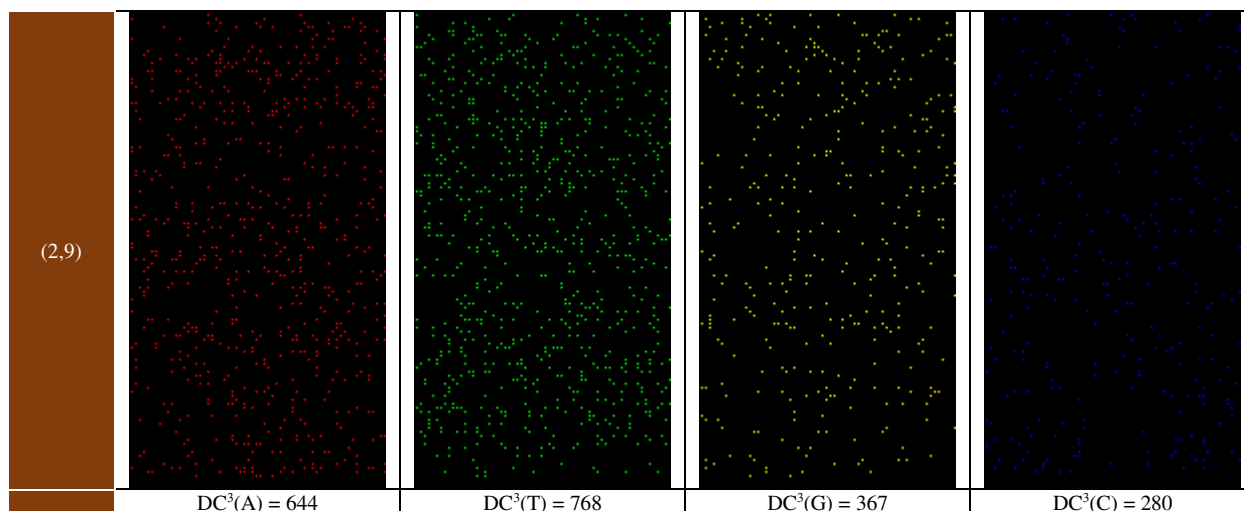
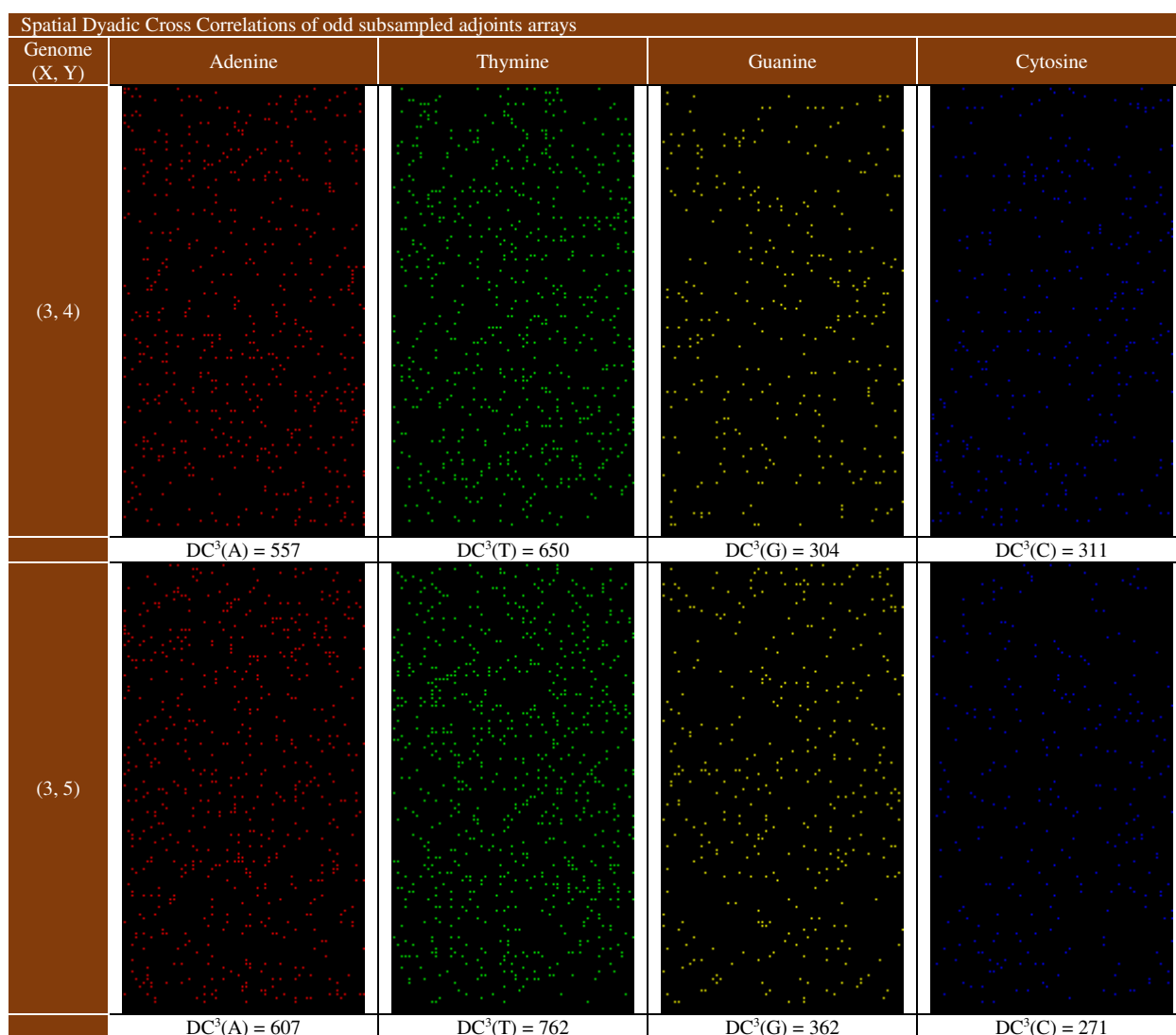
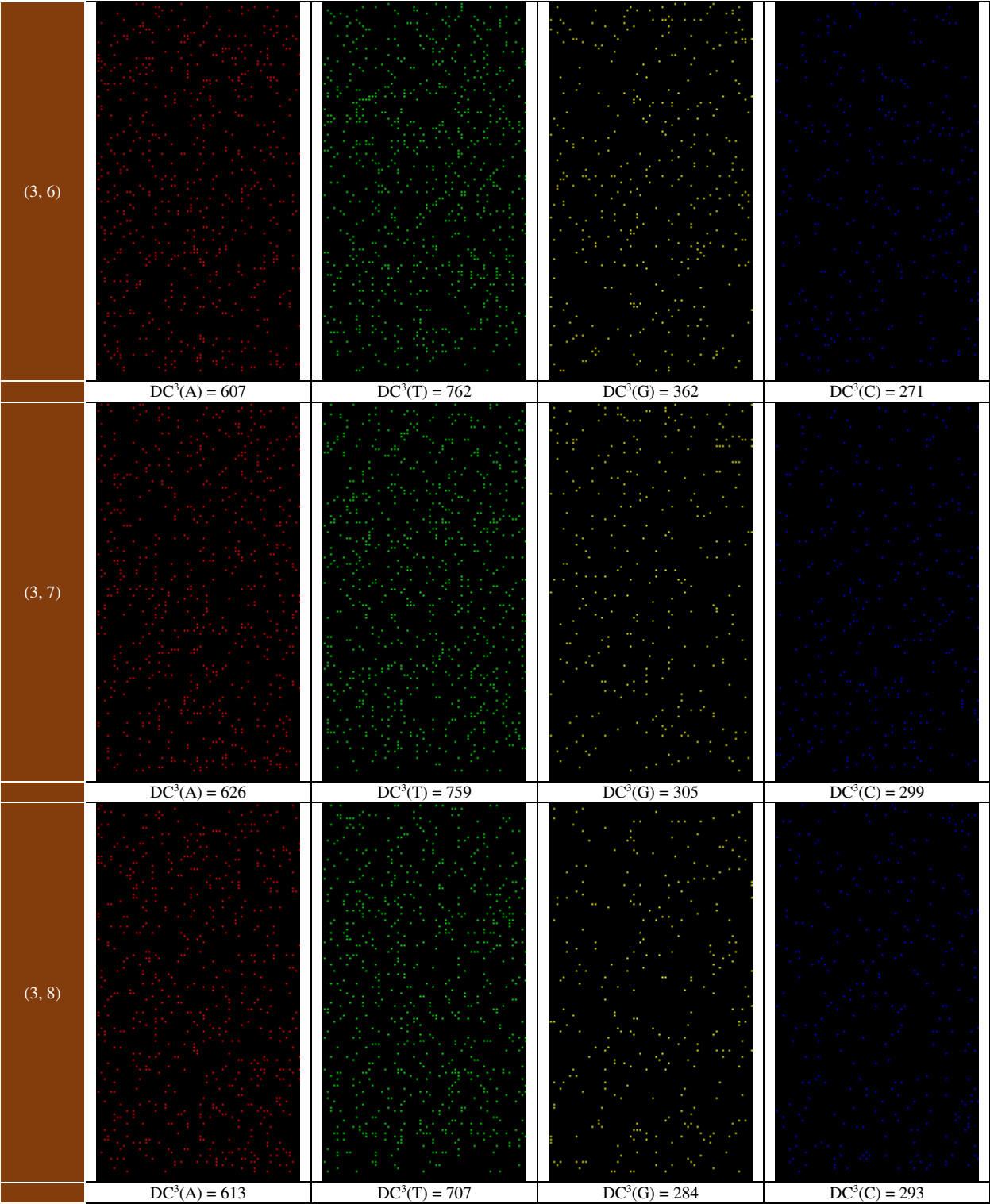


Fig. 49: Pairwise DC³s of genome 2 with other genomes





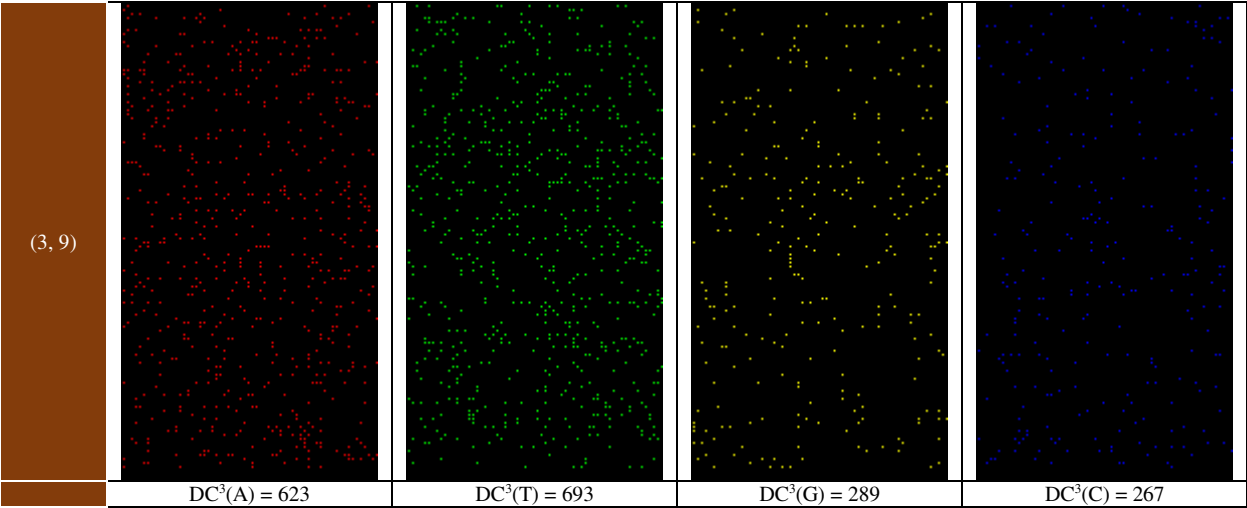
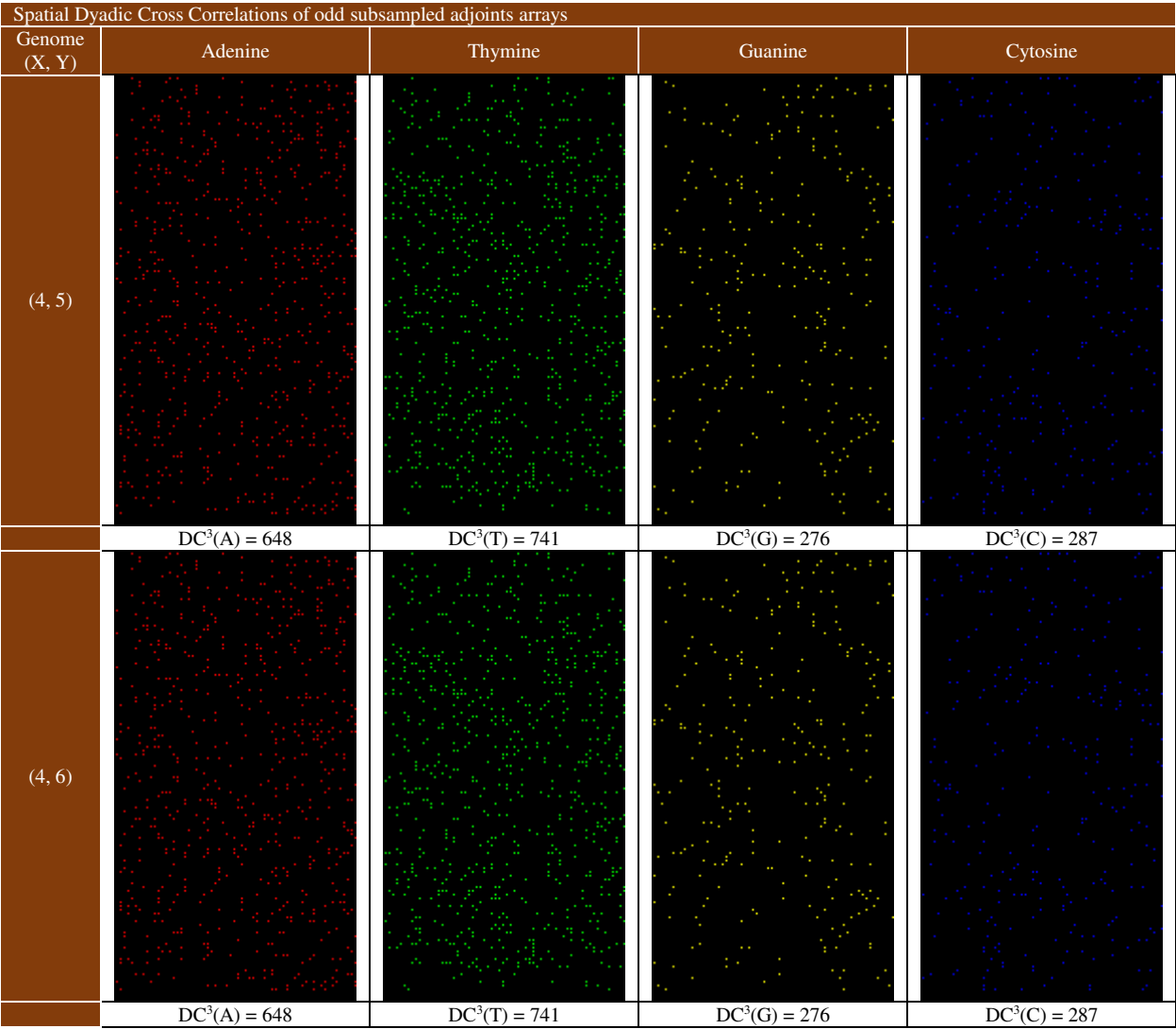


Fig. 50: Pairwise DC³s of genome 3 with other genomes



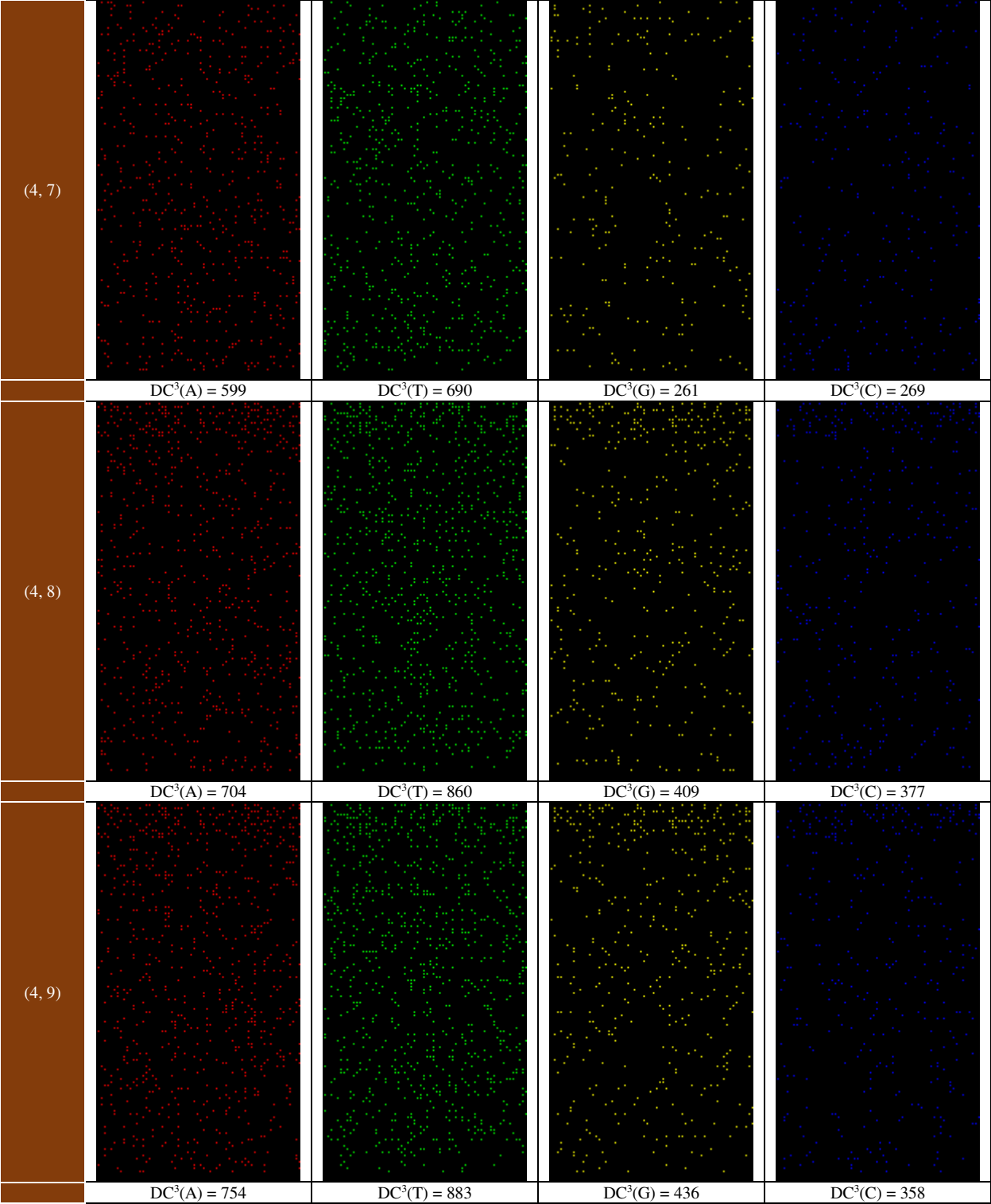
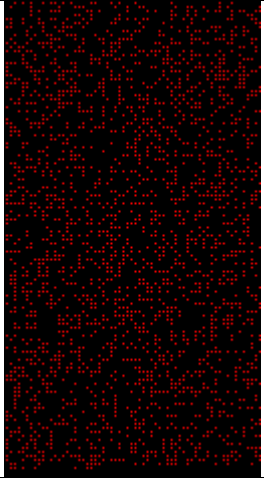
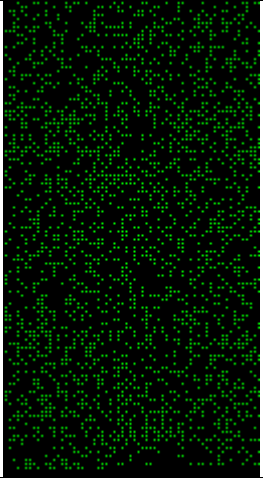
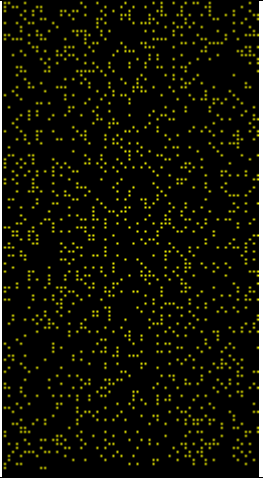
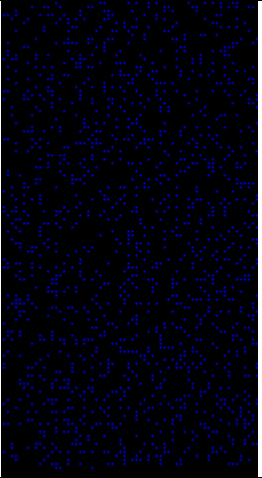
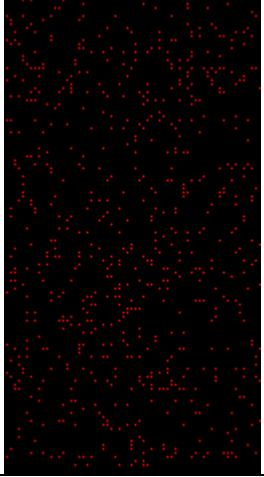
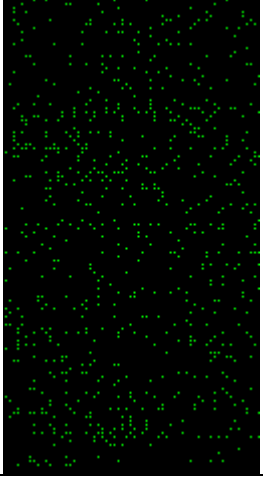
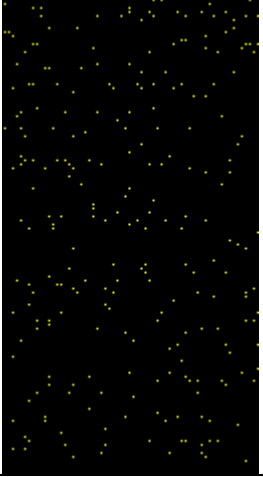
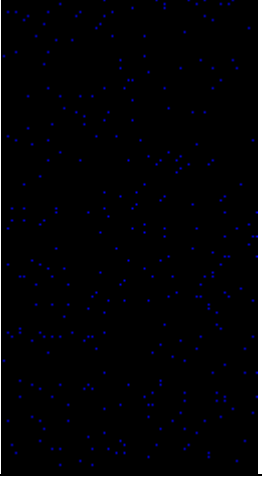
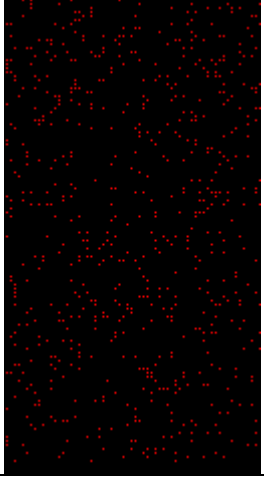
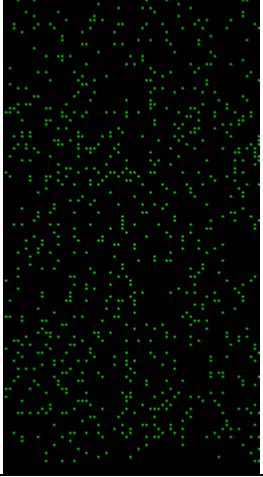
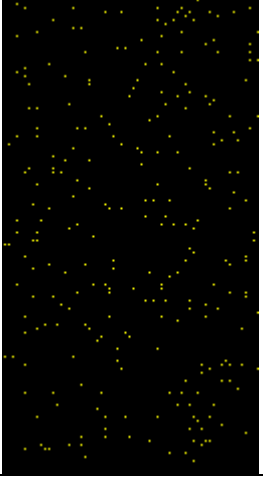
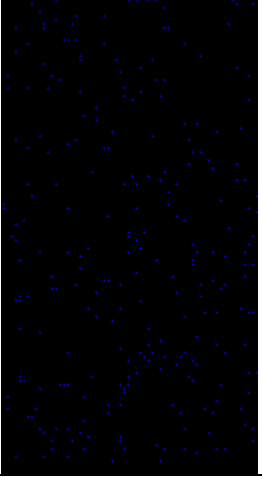


Fig .51: Pairwise DC³s of genome 4 with other genomes

Spatial Dyadic Cross Correlations of odd subsampled adjoints arrays				
Genome (X, Y)	Adenine	Thymine	Guanine	Cytosine
(5, 6)				
	DC ³ (A) = 2218	DC ³ (T) = 2349	DC ³ (G) = 1480	DC ³ (C) = 1402
(5, 7)				
	DC ³ (A) = 670	DC ³ (T) = 732	DC ³ (G) = 250	DC ³ (C) = 288
(5, 8)				
	DC ³ (A) = 648	DC ³ (T) = 774	DC ³ (G) = 264	DC ³ (C) = 252

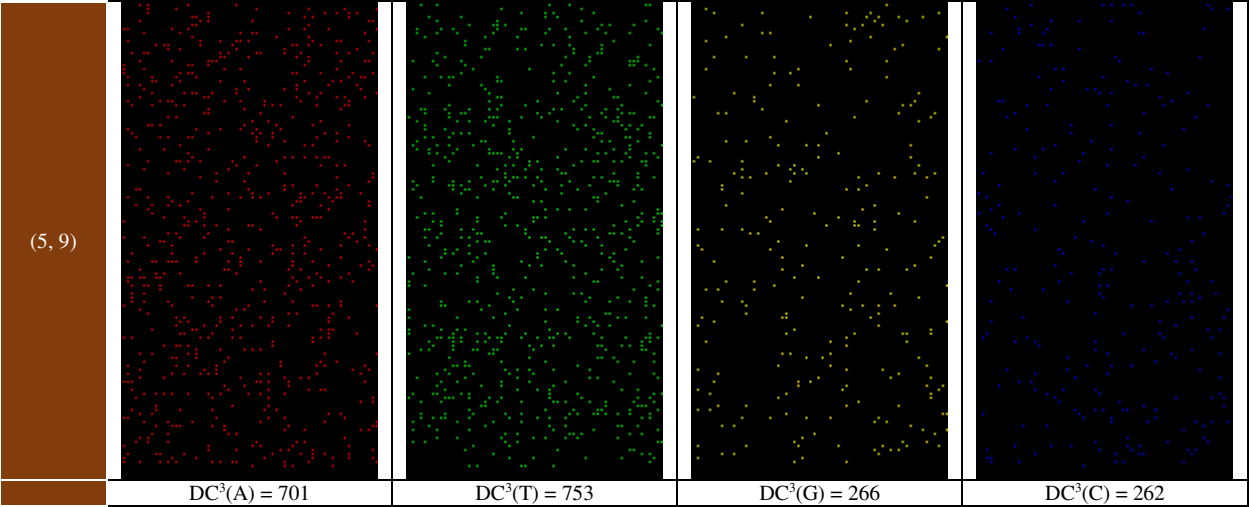
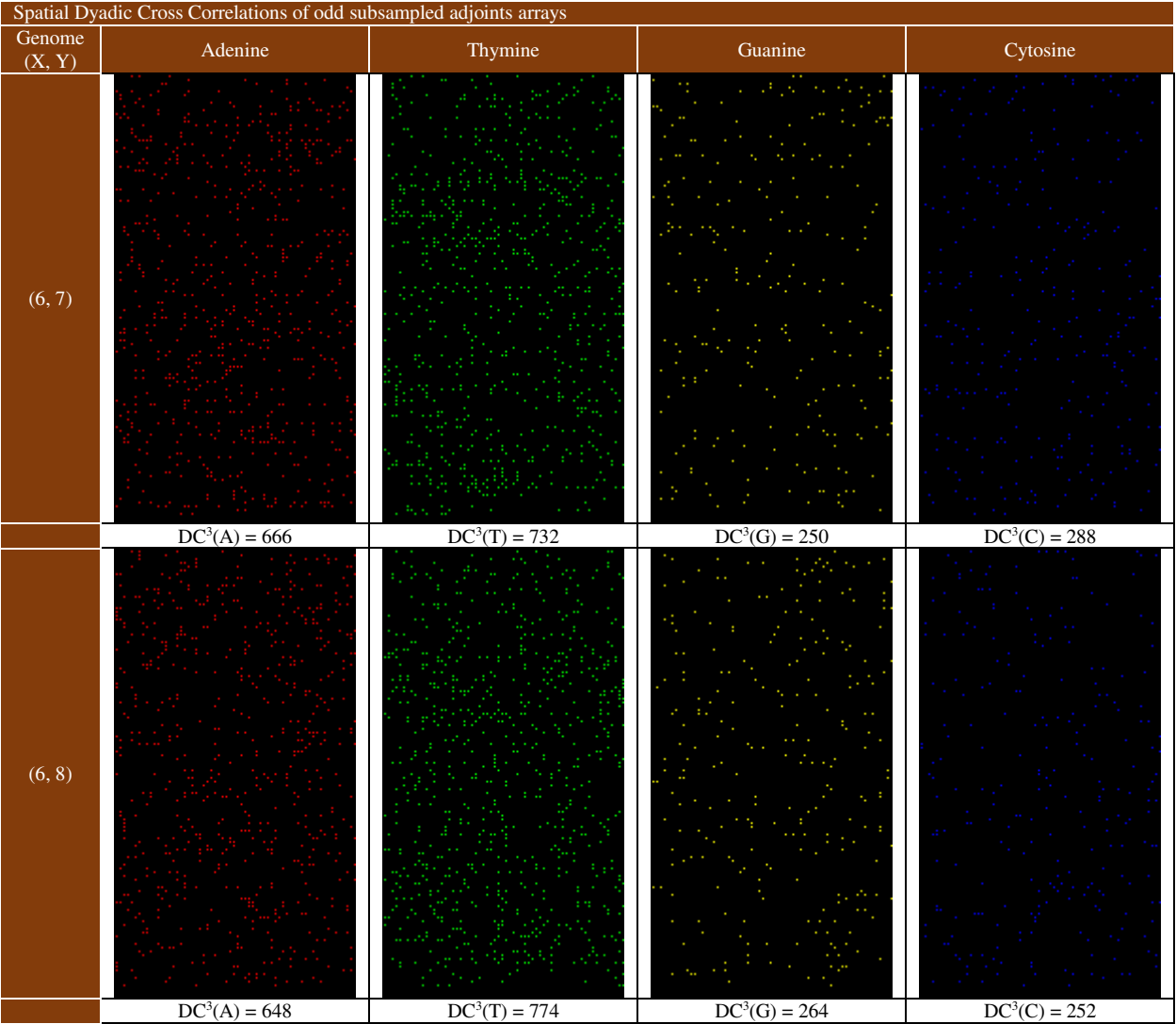


Fig. 52: Pairwise DC³s of genome 5 with other genomes



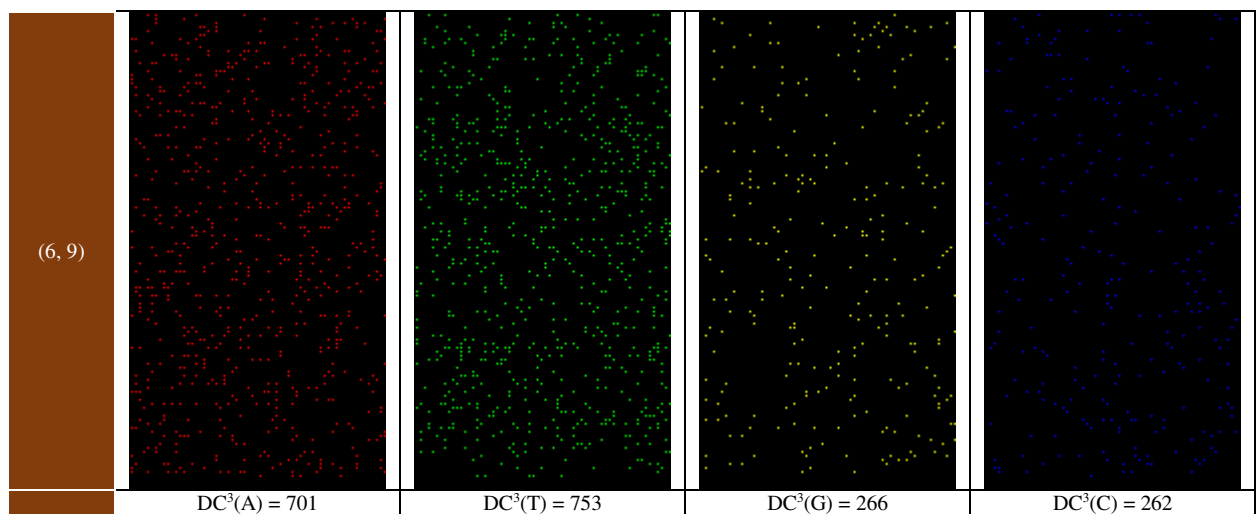


Fig. 53: Pairwise DC³s of genome 6 with other genomes

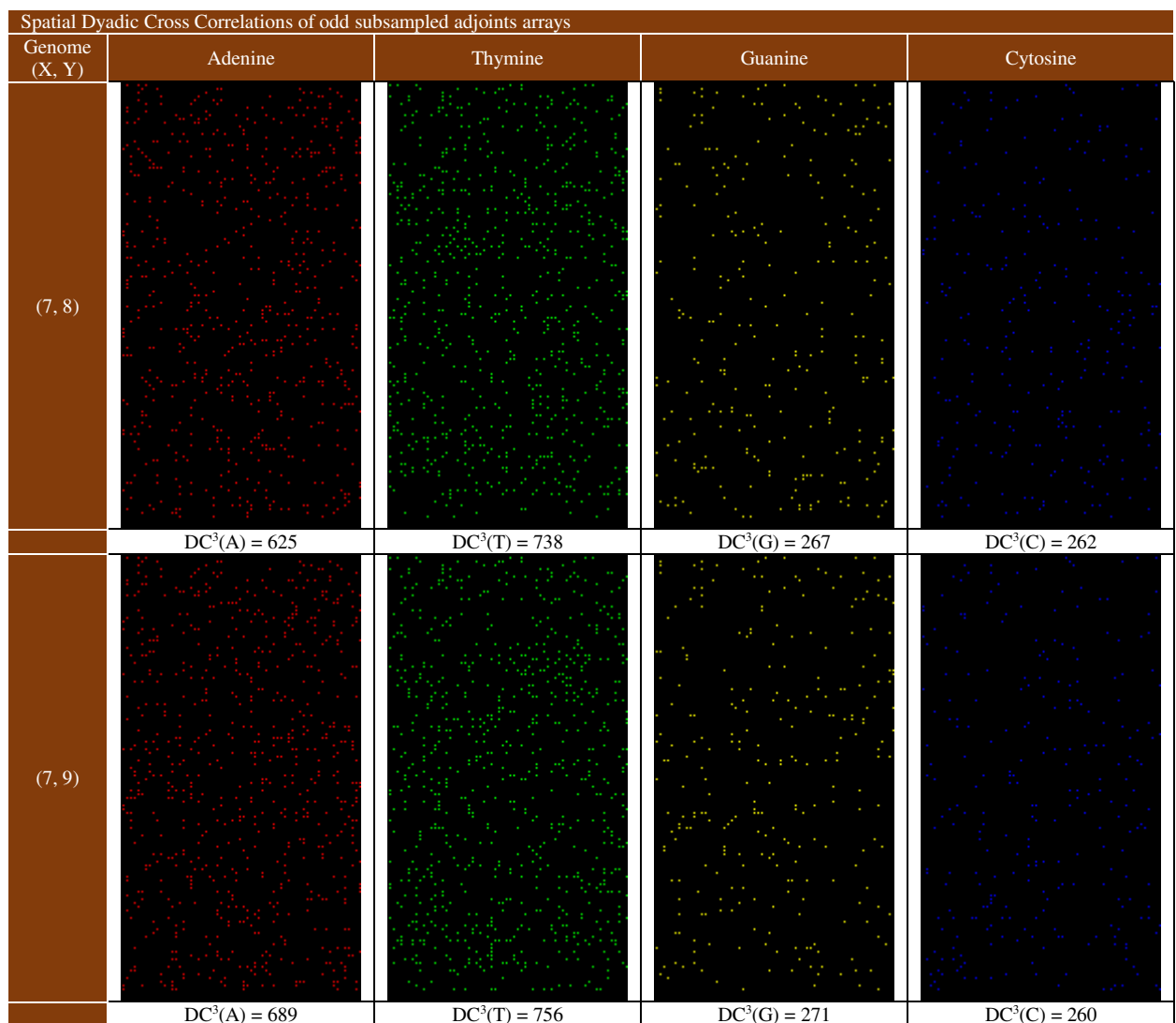


Fig. 54: Pairwise DC³s of genome 7 with other genomes

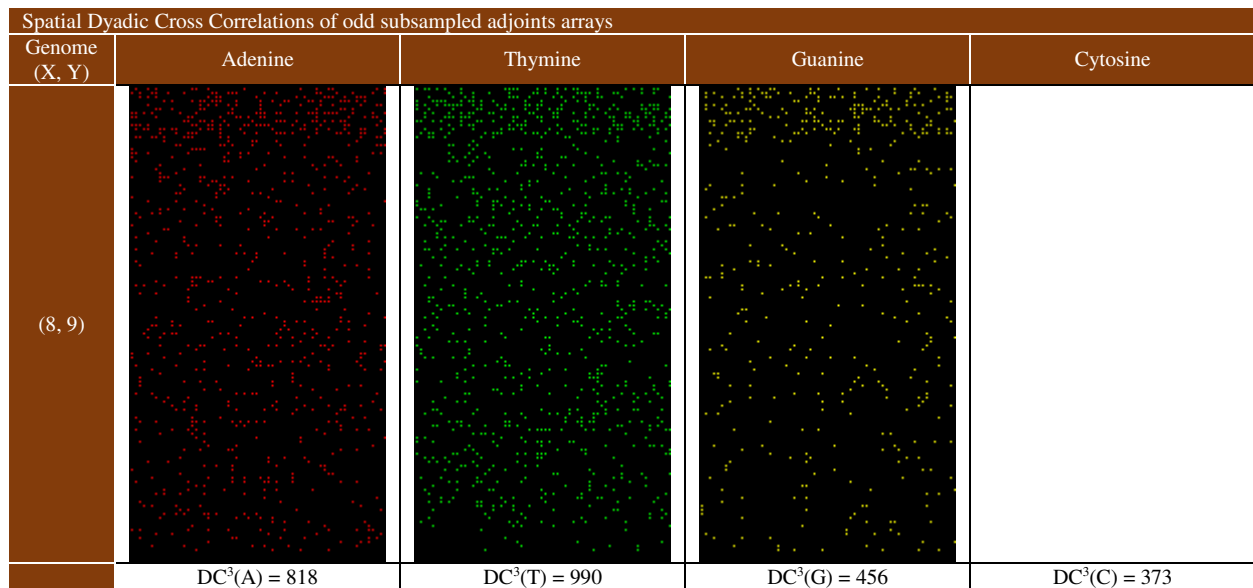


Fig. 55: Pairwise DC³s of genome 8 with genome 9

Fig. 56 shows the graph connecting various adjoint pairs and their DC³ values.

Pairwise intersection property of odd arrays [DC³(A), DC³(T), DC³(G) and DC³(C)]

Sl. No.	Odd subsampled Adjoints Genome Pair (X, Y)	DC ³ (A)	DC ³ (T)	DC ³ (G)	DC ³ (C)
1	(1, 2)	607	741	374	275
2	(1, 3)	589	661	327	286
3	(1, 4)	625	750	388	295
4	(1, 5)	626	717	301	291
5	(1, 6)	626	717	301	291
6	(1, 7)	636	674	265	249
7	(1, 8)	604	781	345	294
8	(1, 9)	673	776	371	264
9	(2, 3)	603	643	288	285
10	(2, 4)	608	726	365	301
11	(2, 5)	642	737	310	271
12	(2, 6)	642	737	310	271
13	(2, 7)	629	716	291	284
14	(2, 8)	625	806	361	280
15	(2, 9)	644	768	367	280
16	(3, 4)	557	650	304	311
17	(3, 5)	607	762	362	271
18	(3, 6)	607	762	362	271
19	(3, 7)	626	759	305	299
20	(3, 8)	613	707	284	293
21	(3, 9)	623	693	289	267
22	(4, 5)	648	741	276	287
23	(4, 6)	648	741	276	287
24	(4, 7)	599	690	261	269
25	(4, 8)	704	860	409	377
26	(4, 9)	754	883	436	358
27	(5, 6)	2218	2349	1480	1402
28	(5, 7)	670	732	250	288
29	(5, 8)	648	774	264	252
30	(5, 9)	701	753	266	262
31	(6, 7)	666	732	250	288
32	(6, 8)	648	774	264	252
33	(6, 9)	701	753	266	262
34	(7, 8)	625	738	267	262
35	(7, 9)	689	756	271	260
36	(8, 9)	818	990	456	373

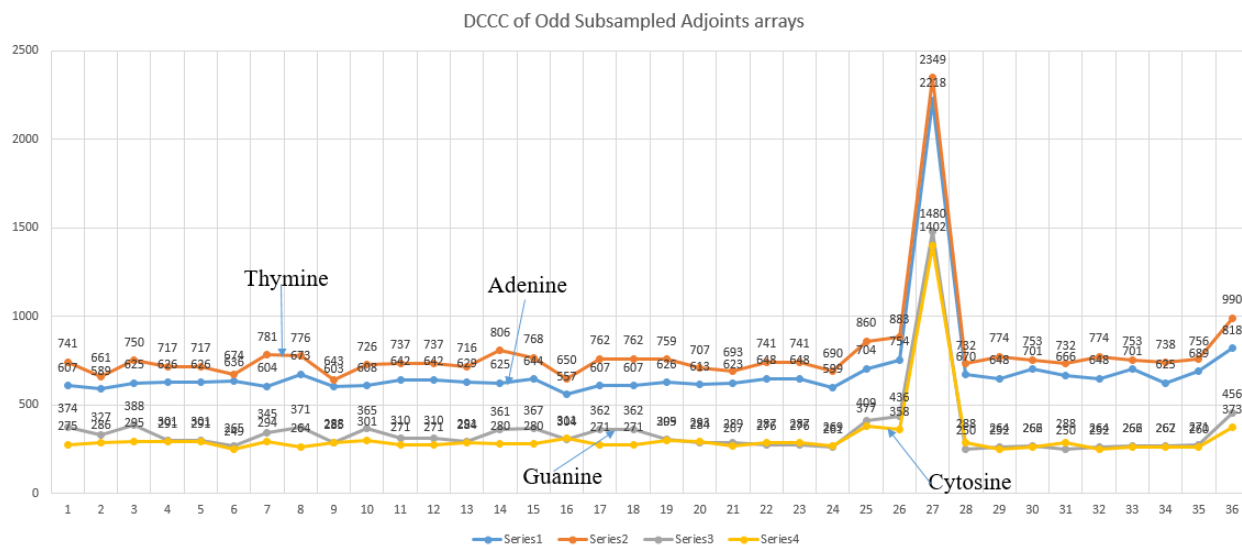
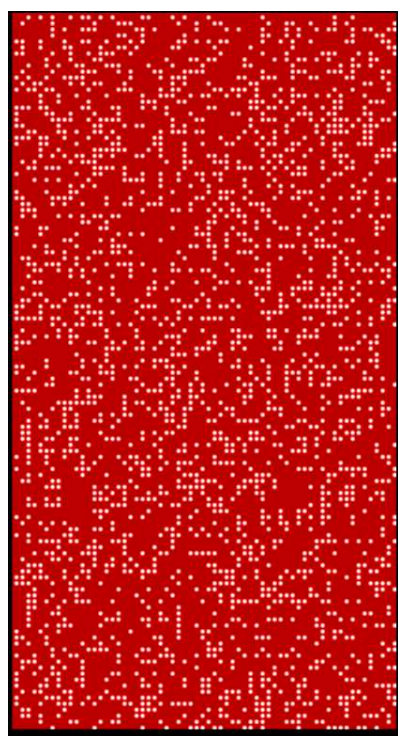


Fig. 56: Graph connecting adjoint pairs and DC^3 values

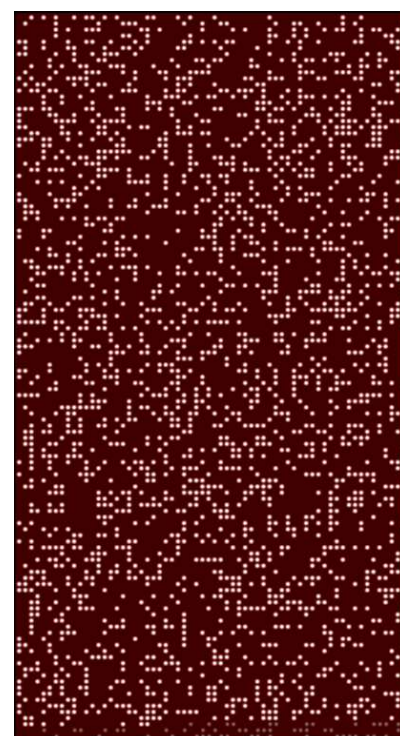
As it was observed in the case of even subsampled adjoints arrays, it is observed in the case of odd subsampled adjoints arrays also that the 27th unordered pair (Genome 5, Genome 6) shows up a **maximum number of matches of nucleotides** and indeed this calls for special investigation of these two virus genomes. A visual presentation of adjoints of A, T, G and C of Genome 5 and Genome 6 along with Genome 6 superimposed on Genome 5 are shown in Figs. 57 - 60. Note that the length of Genome 5 (length = 30473) is longer than that of Genome 6 (length = 29875). It is to be noted that all images are of size $128 \times 238 = 30,464$.



Genome 5 adenine
odd subsampled array



Genome 6 adenine
odd subsampled array



Genome 6 adenine odd subsampled
array and Genome 5 adenine odd

Fig. 57: Adenine odd subsampled arrays of Genome 5 and 6 and their superimposed version

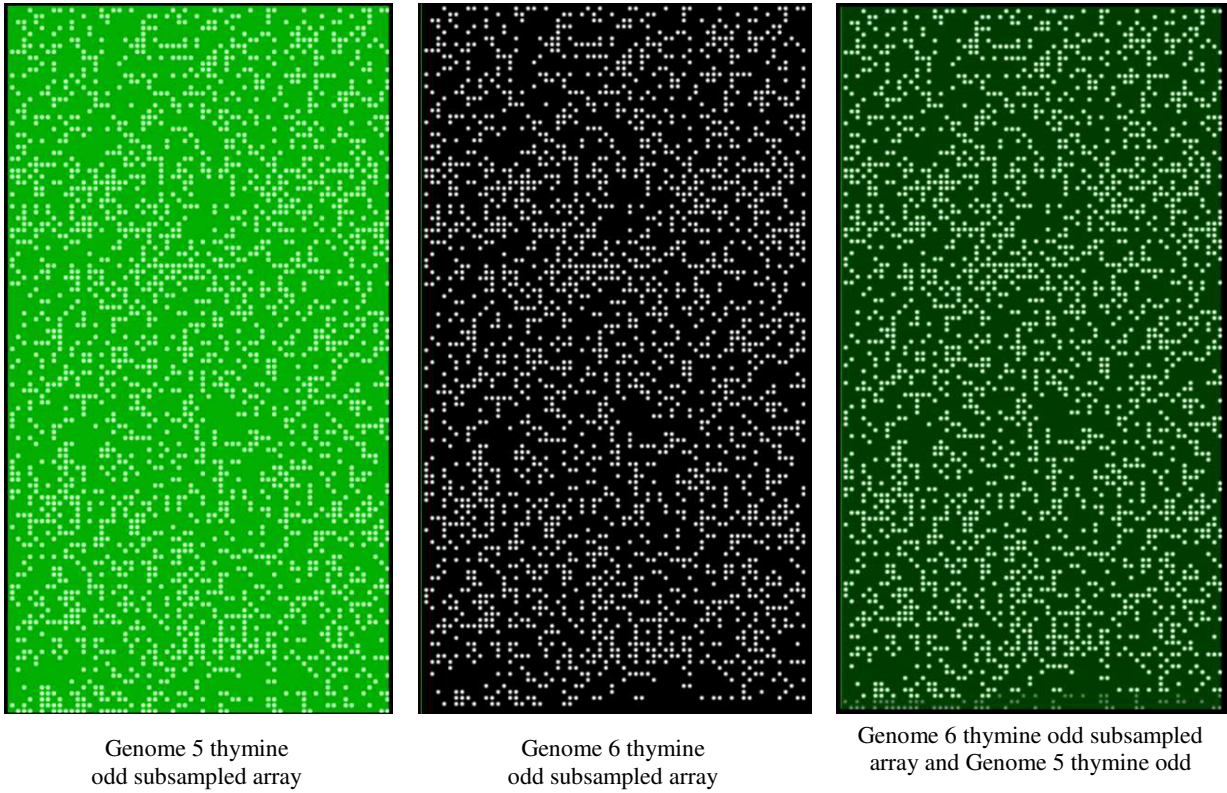


Fig. 58: Thymine odd subsampled arrays of Genome 5 and 6 and their superimposed version

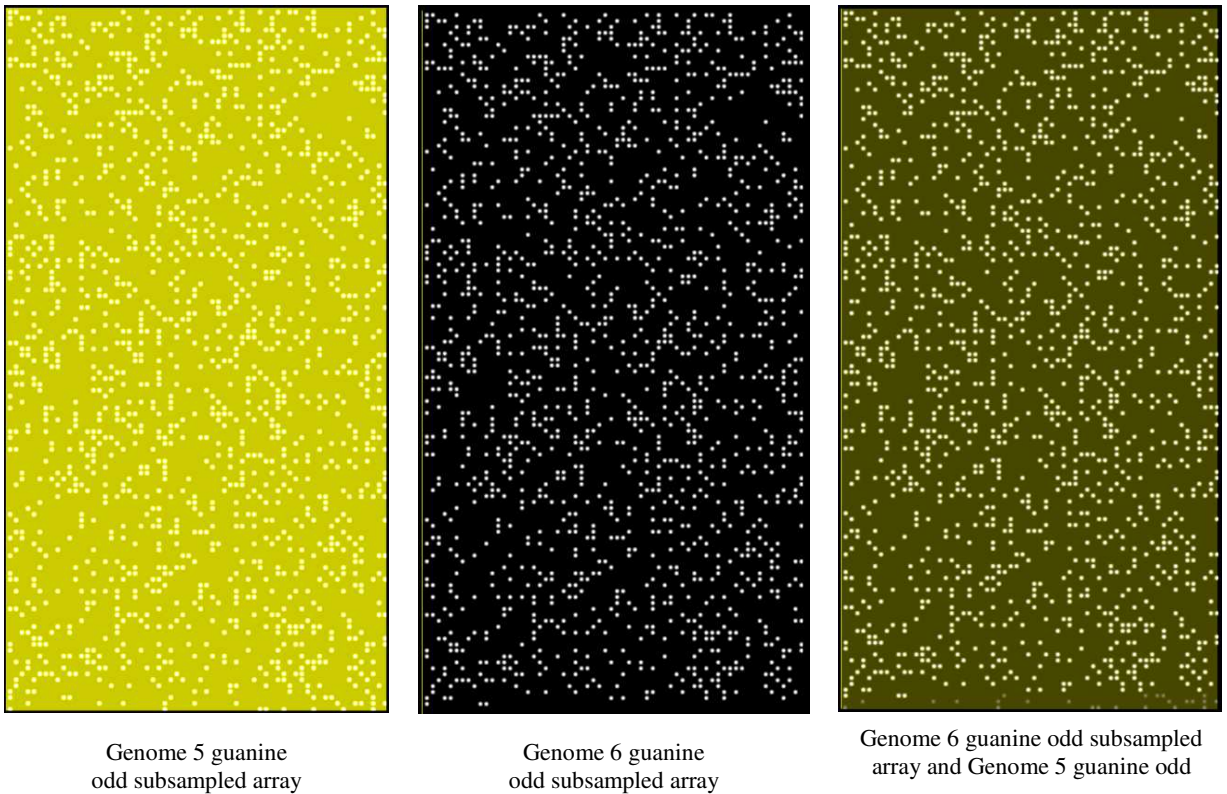


Fig. 59: Guanine odd subsampled arrays of Genome 5 and 6 and their superimposed version

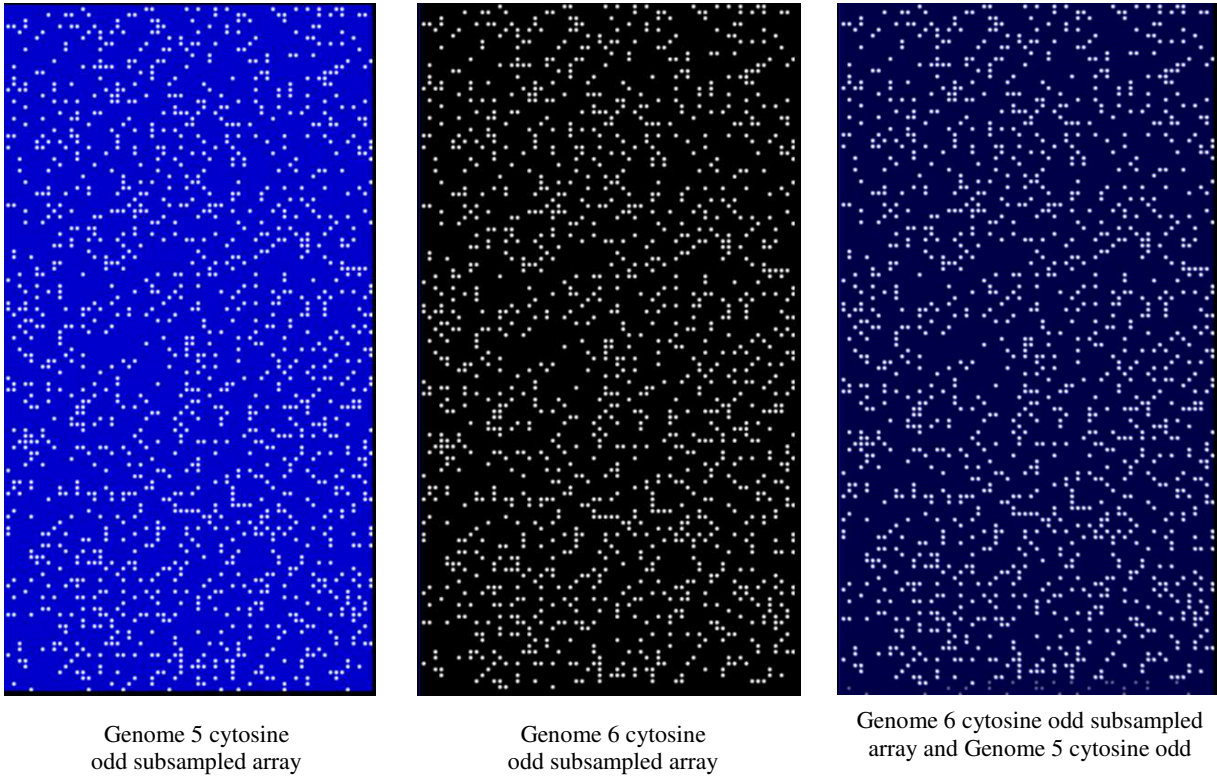


Fig. 60: Cytosine odd subsampled arrays of Genome 5 and 6 and their superimposed version

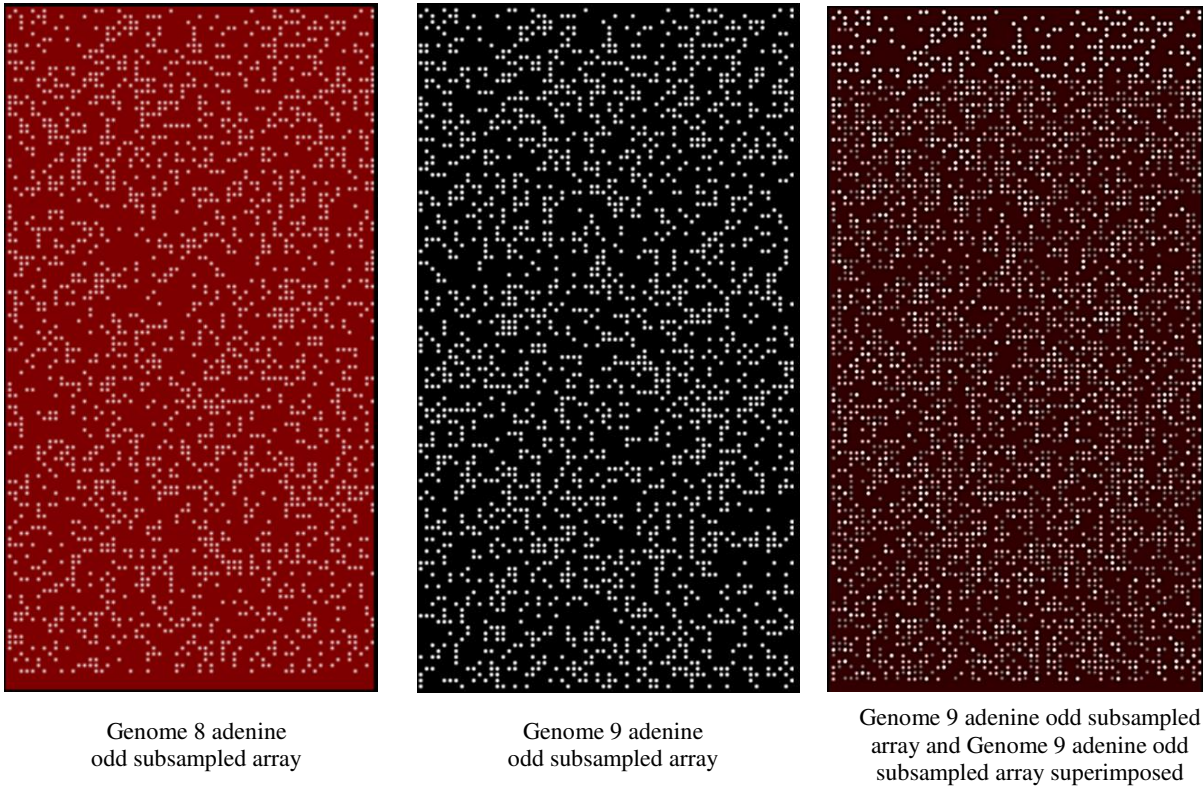


Fig. 61: Adenine odd subsampled arrays of Genome 8 and 9 and their superimposed version

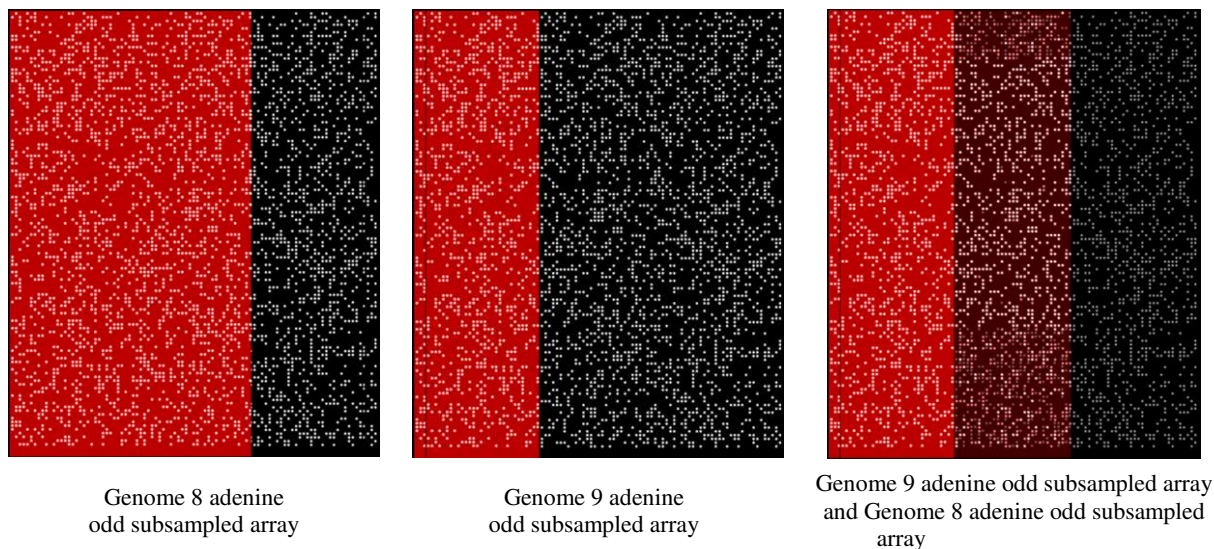


Fig. 62: Adenine odd subsampled arrays of Genome 8 and 9 partially superimposed

Fig. 62 shows adenine odd subsampled arrays of Genome 8 and 9 and their partially superimposed version. Some significant observations made among Fig. 61 and Fig. 62 are given below.

1. The 27th pair (5, 6) among 36 pairs exhibits the highest DC^3 value, that is, 2349 (thymine pair), and the next highest DC^3 value 990 (thymine pair) is exhibited by the 36th pair (8, 9).
2. Though the dyadic comparison of the pair (8, 9) shows 818 adenine positions common to both arrays, they are seen to be randomly distributed, that is, sparsely, and such random distribution does not convey any meaningful interpretation or a common pattern as regards their spatial correlation is concerned (Ref. Fig. 61).
3. On the other hand, when Genome 9 adenine odd subsampled array is horizontally shifted towards right with respect to Genome 8 adenine odd subsampled array and partially superimposed, the middle portion of the intersected region shows a common pattern (Ref. Fig. 62).
4. This amounts to saying that one can shift one array in eight possible directions with respect to another array in a digital lattice and look for common digital patterns. One can have 4,84,496 partial superpositions of the arrays.

Each of the 4,84,496 superpositions of the odd subsampled adjoint arrays may yield a common digital pattern individually. These common digital patterns are the objects of interest for studying spatial correlations between various genome adjoint arrays. All these correlations are called non-dyadic cross correlations.

Non-Dyadic Cross Correlations

There are 1,21,124 non-dyadic cross correlations for a pair of odd subsampled adjoint arrays. Every non-dyadic cross correlation has a Cross Correlation Coefficient (C^3). Now, every C^3 is not DC^3 , but every DC^3 is essentially a C^3 . Many non-dyadic cross correlations are observed to yield large common digital patterns. Given a pair of arrays, one can shift in eight directions with respect to another. Details of shifting and comparing array pairs for common digital patterns are already explained in previous sections. Hence, similar exercise is not repeated here.

Observation on genomes 5 and 6

Genome 5 and genome 6 are almost identical except the fact that the length of genome 5, that is MN908947.1, is 30473 and that of genome 6, that is MN908947.2, is 29875. Yet, we conducted experiments on both the genomes separately and found that they are one and the same but with different lengths. In what follows, the DC^3 values of all the nine pairs of even subsampled and odd subsampled adjoint arrays are given and to illustrate this point, the superimposed images of the pairs (1, 5) and (1, 6) are shown in Fig. 63. The same observation holds good for other pairs also, be it even or odd subsampled adjoint arrays.

Observation on even subsampled adjoint arrays					
Sl. No.	Even subsampled Adjoints Genome Pair (X, Y)	$DC^3(A)$	$DC^3(T)$	$DC^3(G)$	$DC^3(C)$
1	(1, 5)	649	766	313	278
2	(1, 6)	649	766	313	278
3	(2, 5)	652	748	279	248
4	(2, 6)	652	748	279	248
5	(3, 5)	621	750	337	280
6	(3, 6)	621	750	337	280
7	(4, 5)	608	726	297	299
8	(4, 6)	608	726	297	299
9	(5, 6)	2245	2387	1485	1371

The same observation holds good for odd subsampled adjoint arrays also.

Sl. No.	Odd subsampled Adjoints Genome Pair (X, Y)	DC ³ (A)	DC ³ (T)	DC ³ (G)	DC ³ (C)
1	(1, 5)	626	717	301	291
2	(1, 6)	626	717	301	291
3	(2, 5)	642	737	310	271
4	(2, 6)	642	737	310	271
5	(3, 5)	607	762	362	271
6	(3, 6)	607	762	362	271
7	(4, 5)	648	741	276	287
8	(4, 6)	648	741	276	287
9	(5, 6)	2218	2349	1480	1402

Fig. 63 shows highlighted versions of superpositions of even subsampled adenine adjoints of pairs (1, 5) and (1, 6) separately.

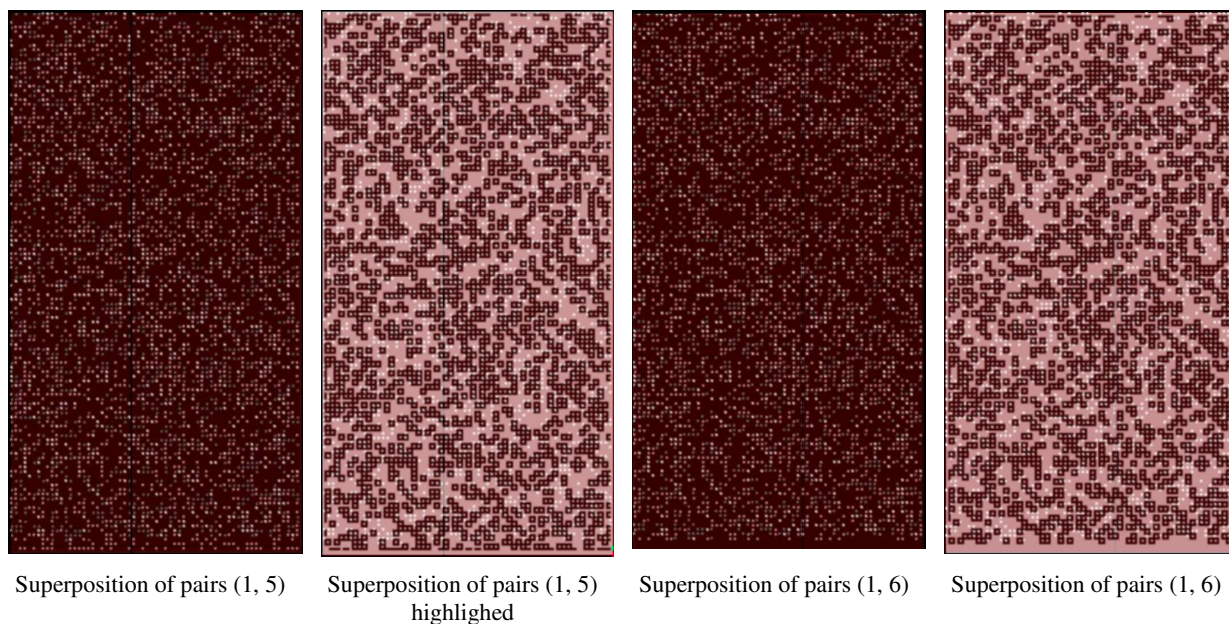


Fig. 63: Adenine even subsampled arrays of Genome 1 and 5 superimposed

Fig. 64 shows the results of Genome 6 even subsampled adjoints arrays subtracted from Genome 5 even subsampled adjoints arrays

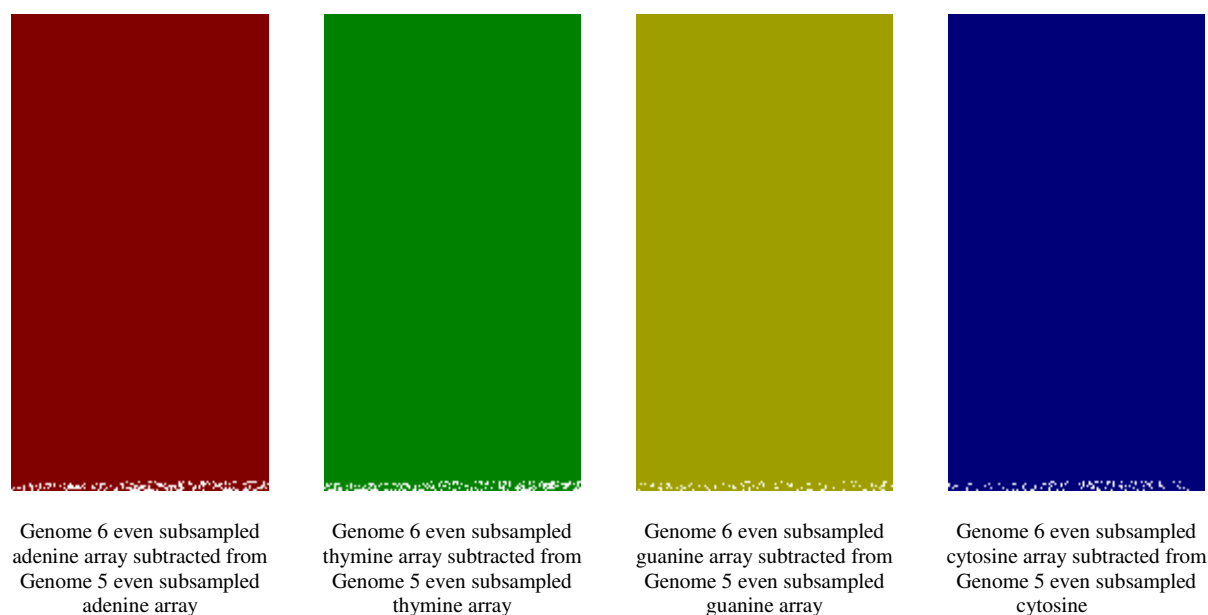


Fig. 64: Results of Genome 6 even subsampled adjoints arrays subtracted from Genome 5 even subsampled adjoints arrays

VI. CONSOLIDATED OBSERVATIONS AND CONCLUDING REMARKS

1. The nine genome sequences taken under the radar for this report show that the percentage nucleotide concentration is almost same for all of them. This observation highlights the point that all these viruses share the same foundation of life.
2. Changes in percentage nucleotide concentrations exhibit a relationship between viruses 1, 2, 3, 4, which form one class. Viruses 5, 6, and 7 form one class and viruses 8 and 9 form another class. Basically all are one and the same and they evolve from one generic form.
3. Dyadic cross correlation coefficients for all genomes when compared with one another fall within a similar range for all nucleotides indicating that there is some kind of a structural similarity between the genomes.
4. There is a significant structural relationship between genomes, which has been indicated in this report. Such relationship can be further understood when spectral domain characterization is carried out.

Such observations give rise to the possibility that the viruses with the same concentration as the others if can occur in nature can also be synthesized in a laboratory.

While the understanding of the genome sequences in spatial domain has opened up the possibilities of modeling the structural patterns as a random process, the viruses do not mutate indiscriminately, but they evolve systematically showing common digital patterns in their evolutions. In this regard, one can model evolutions of virus strains in the framework of stochastic cellular automata

However, spectral domain study of virus genome data may open up new dimensions in this research.

A crucial observation to make is the percentage nucleotide concentration. When the nucleotide concentration of a virus is known, a similar virus can be theoretically created. Belongingness of all these genomes to the same family and having the same concentration of nucleotides can extrapolate the possibility of creation of a virus under un-natural forces and influences.

Question

Is the present day Covid-20 pandemic 'natural or man-made' ? Further research in this direction presented in this paper may yield an answer to this question and show a way to counter such threats.

स्वस्ति प्रजाभ्यः परिपालयंतां न्यायेन मार्गेन महीम महीशाः ।
गोब्राह्मणेभ्य शुभमस्तु नित्यं लोका समस्ता सुखिनो भवन्तु ।।

May there be happiness for all people.

May the rulers righteously rule the earth.

May there be welfare for the animals and men of wisdom at all times.

May all beings be happy and prosperous.

An Activity of Digital Health Initiative Started by the Team



Scan to know paper details and
author's profile

Future Automobile Market and Role of Lithium ION Batteries in it

Niraj Sunil Kulkarni

ABSTRACT

Automobile sector is going through a vast nature of changes from the start of 2015. The new norms of BS6 in India and updated Euro norms have increased difficulties in the business of the automobile sector. Nearly 25% cost increase is seen to fulfill the requirements of an BS6 engine. Many of the automotive giants have stopped the manufacturing of diesel engines and many are planning to do so in future. While the crude oil market is not so preferable by government authorities, a new technology like Lithium Batteries supporting E-Vehicles has been seen in the last few years. The cost per kilometre of the vehicles have made a vital impact on the public. While the traditional vehicles are softly getting shifted over electric ones, the new technology also lacks at some points. Chemistry is playing an important role in the automobile sector. The use of different kinds of fuel has made an obstacle to traditional fuel companies. This paper discuss the design, advantage, disadvantage and future scope of electric vehicles.

Keywords: lithium ion battery, e-vehicles, design of lithium ION battery.

Classification: A.1

Language: English



LJP Copyright ID: 975843
Print ISSN: 2514-863X
Online ISSN: 2514-8648

London Journal of Research in Computer Science and Technology



Volume 20 | Issue 1 | Compilation 1.0

© 2020. Niraj Sunil Kulkarni. This is a research/review paper, distributed under the terms of the Creative Commons Attribution-Noncom-mercial 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.

Future Automobile Market and Role of Lithium ION Batteries in it

Niraj Sunil Kulkarni

ABSTRACT

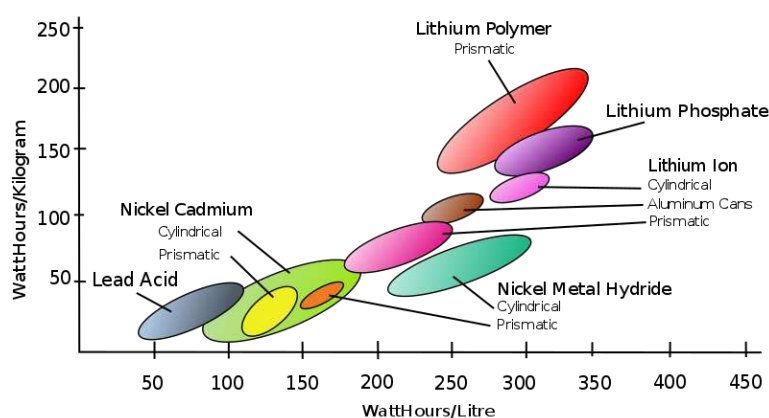
Automobile sector is going through a vast nature of changes from the start of 2015. The new norms of BS6 in India and updated Euro norms have increased difficulties in the business of the automobile sector. Nearly 25% cost increase is seen to fulfill the requirements of an BS6 engine. Many of the automotive giants have stopped the manufacturing of diesel engines and many are planning to do so in future. While the crude oil market is not so preferable by government authorities, a new technology like Lithium Batteries supporting E-Vehicles has been seen in the last few years. The cost per kilometre of the vehicles have made a vital impact on the public. While the traditional vehicles are softly getting shifted over electric ones, the new technology also lacks at some points. Chemistry is playing an important role in the automobile sector. The use of different kinds of fuel has made an obstacle to traditional fuel companies. This paper discuss the design, advantage, disadvantage and future scope of electric vehicles.

Keywords: lithium ion battery, e-vehicles, design of lithium ion battery.

Author: (Research & Development Engineer, Bachelor of Mechanical Engineering, MGM's Jawaharlal Nehru Engineering College, Aurangabad, Maharashtra, India.

I. LITHIUM ION BATTERIES

Primarily used in mobile phones, the Lithium Ion batteries are now getting place of traditional Lead acid batteries that used to drive the electric automobile world. Being nearly $1/4^{\text{th}}$ of lead acid batteries with same output and long lasting capacity, lead acid batteries are losing their market. Lithium batteries are lightweight and provide more energy per kwh than lead acid batteries. Primarily, lithium batteries were used to be in pouch type. But, for automobiles the prismatic type of lithium ion batteries are developed making it easier to mount and formulate the required voltage inside an battery pack. The tech giant Tesla company made its own choice and went through the another type of lithium ion battery type for its vehicles i.e. cylindrical type of lithium ion battery.



Being smaller, the 18650 and 21700 lithium ion cells are easily arranged in small plastic cases. This made the batteries of various sizes available according to the cabinet size provided inside an vehicle. Because of cylindrical lithium ion batteries, the manufacturing of various size cabinets and achieving different voltage and current levels is possible. The 48V, 60V, 96V batteries are popularly used in automobiles. Along

with their sizes, the batteries are classified into different chemistries like Lithium Ion Batteries (LFP), Lithium Nickel Iron Batteries (NMC), Lithium Ion Titanate Batteries (LTO), etc. The classification can also be made according to the energy output of the cell types used. Basically the 18650 cell has first 2 digits 18 equals the diameter of cell i.e. 18mm and later two digits 65 means the height of the cell i.e. 65mm.

II. ROLE OF LITHIUM ION BATTERIES IN AUTOMOBILE



Lithium batteries mainly used in automobiles are of 48V or 60V in general. Later 72V and 96V batteries are proposed to capture the market. The lithium ion batteries provide a better output to the vehicles consuming less space. These batteries are made as per the voltage and current requirement in electric vehicles.

III. DESIGN OF 48V18AH LITHIUM ION BATTERIES

Considering and LFP battery cell 18650 of 3.7V2.6Ah capacity for design of 48V18Ah battery pack, the series combination equals: $48/3.7$ i.e. 12.97 which equals 13 cells in series.

Parallel combination equals: $18/2.6$ i.e. 6.92 which equals 7 cells in parallel.

So, for an 48V18Ah battery pack: 13x7 numbers of cells are required i.e. 91 numbers of cells will be

required. The same numbers can be calculated for any other battery pack. The conducting and connecting material also plays an important in energy output of lithium ion batteries.

Advantages of Lithium Ion batteries

Higher energy density than lead acid battery.

- Lightweight .
- Can be easily fit into a small region.
- Rechargeable and reusable.
- Higher life cycle than any other battery type.
- Disadvantages of Lithium Ion batteries:
- Initial cost is higher.
- Technology is not yet proven.
- Import of lithium is required for many countries.
- Storage at certain 25°C temperature is required.

Future Scope:

- Lithium Ion batteries might drive aeroplanes and submarines.
- Hydrogen fuel cell technology might prove a challenge for lithium technology.
- Vehicle to vehicle wireless transmission of electric pulses to charge another battery of lower charge.
- Battery swapping technology taking over petrol pumps.

REFERENCES:

1. https://www.researchgate.net/publication/282791010_Li-ion_batteries_basics_progress_and_challenges/link/581fe1d908ae12715af927db/download
2. <https://www.sciencedirect.com/science/article/pii/S1369702114004118>

This page is intentionally left blank



Scan to know paper details and
author's profile

A Review on Face Recognition Technique using the Artificial Neural Network

Savita Bairagi, Hemant Gupta, Ashok Panwar & Bhavana Panwar

ABSTRACT

Face recognition is one of the classical domains of research and development. Significant contributions on face recognition technique are available in literature. There are various kinds of techniques and methodologies which can be divided in two major categories i.e. feature classification based approach and reconstruction based approaches. In this paper, we are first providing the survey on existing efficient and recent techniques of face recognition. Additionally on the basis of concluded consequences a new model for recognizing the partial face. In this context a new model using LDA based features and BPN (back propagation neural network) based is introduced. Additionally their functional aspects are discussed. Finally the conclusion and future working guidelines are provided.

Keywords: face recognition, partial faces, neural network, LDA, survey, machine learning.

Classification: I.5.1

Language: English



London
Journals Press

LJP Copyright ID: 975844
Print ISSN: 2514-863X
Online ISSN: 2514-8648

London Journal of Research in Computer Science and Technology



Volume 20 | Issue 1 | Compilation 1.0

© 2020. Savita Bairagi, Hemant Gupta, Ashok Panwar & Bhavana Panwar. 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.

A Review on Face Recognition Technique using the Artificial Neural Network

Savita Bairagi^α, Hemant Gupta^σ, Ashok Panwar^ρ & Bhavana Panwar^β

ABSTRACT

Face recognition is one of the classical domains of research and development. Significant contributions on face recognition technique are available in literature. There are various kinds of techniques and methodologies which can be divided in two major categories i.e. feature classification based approach and reconstruction based approaches. In this paper, we are first providing the survey on existing efficient and recent techniques of face recognition. Additionally on the basis of concluded consequences a new model for recognizing the partial face. In this context a new model using LDA based features and BPN (back propagation neural network) based is introduced. Additionally their functional aspects are discussed. Finally the conclusion and future working guidelines are provided.

Keywords: face recognition, partial faces, neural network, LDA, survey, machine learning.

Author α: Research Scholar at Lakshmi Narain College of Technology and Science (RIT), Indore.

σ: HOD in Computer Science & Engineering Department at LNCT, Indore.

ρ: Technical Officer (C) in ECIL (NPCIL), Tarapur, Mumbai.

β: Research Scholar as well as Faculty from SVBPHSS, Kasrawad, Khargone, MP.

I. INTRODUCTION

Now in these days the machine learning and its techniques are played essential role in various real world applications. These applications are used for pattern recognition, classification,

categorization, prediction, motion detection and various others [1]. Additionally the machine learning algorithms also works well with the images for motion detection, activity recognition and others. Among them the face recognition is one of the most essential applications. That is used in wide range of security, surveillance, monitoring and authentication [2].

Face is one of the most essential identities, which is almost different and unique for all the individuals. In a number of applications it is used as a biometric identity and utilized for authentication and access control [3]. But these applications are feasible if the recognition of the faces becomes accurate. The low rate of face recognition can degrade the performance of these applications. The basic issues of these systems are the pose variation, partial capturing of faces, small changes on faces, low resolution images, noisy images and others [4]. Due to these, improvements on the traditional face recognition systems are required. The proposed work is intended to design and develop an accurate face recognition technique that works on different variants of faces and can able to recognize any of the pose obtained by the user data [5].

Therefore in this paper the ORL dataset is used for designing and developing the face recognition system which contains different poses of faces of the same class. In addition of that the LDA (linear discriminant analysis) technique is used for computing face features. The multi-pose based features are used with the back-propagation neural network for training of faces.

After training of the neural network the system is able to accept any of the training face pose and can recognize accurately.

This section provides the basic overview of the proposed methodology and the next section provides the essential contributions placed for recognizing the faces. Further the literature summary is offered. Additionally on the basis of concluded facts a new face recognition model is proposed. Finally the conclusion and future directions of the work is provided.

II. LITERATURE SURVEY

This section provides the collection of recently available contributions and efforts that are promising for accurately recognizing the faces.

S. S. Farfade et al [6] consider the problem of multi-view face detection. Current state-of-the-art approaches for this task require annotation of facial landmarks, or annotation of poses. They also require training dozens of models to capture faces in all orientations. Authors propose Deep Dense Face Detector (DDFD), a method that does not require pose/landmark annotation and able to detect faces in a wide range of orientations using a single model based on deep CNN. The method has minimal complexity; and not requires additional components. They analyzed scores of the face detector for faces in different orientations and found that 1) method is able to detect faces from different angles, 2) there seems to be a correlation between distribution of positive examples in the training set and scores of the given face detector. The method's performance was improved using better sampling and sophisticated data. Evaluations on benchmark datasets show that single-model face detector has better performance compared to the previous methods.

Automatic affect recognition is a challenging task due to the various modalities emotions. Applications can be found in many areas i.e. multimedia retrieval and human computer

interaction. *P. Tzirakis et al [7]* propose an emotion recognition system using auditory and visual modalities. To capture the emotions for various styles of speaking, robust features need to be extracted. This purpose, authors utilize a CNN to extract features from the speech, for visual modality a deep residual network (ResNet). Additionally, a machine learning algorithm needs also to be insensitive to outliers while being able to model the context. To tackle this, Long Short-Term Memory networks are utilized. System is then trained in an end-to-end fashion where – by also taking advantage of the correlations of the each of the streams – author manage to outperform the traditional approaches based on auditory and visual handcrafted features.

Recently, CNN have demonstrated excellent performance on various tasks, including classification of two-dimensional images. *W. Hu et al [8]*, deep CNN are employed to classify hyper-spectral images. More specifically, the architecture of the classifier contains five layers with weights which are the input layer, the convolutional layer, the max pooling layer, the full connection layer, and the output layer. These layers are implemented on each spectral signature to discriminate against others. Results based on several data sets show that the method can achieve better performance than traditional methods, such as SVM and traditional deep learning- methods.

Z. Yu et al [9] report image based static facial expression recognition method for the Emotion Recognition. Authors focus on the sub-challenge of the SFEW 2.0 dataset, where seeks to classify a set of static images into 7 emotions. The method contains a face detection module based on ensemble of three face detectors, followed by a classification module with the ensemble of multiple deep CNN. Each CNN model is initialized randomly and pre-trained on a larger dataset. The pre-trained models are then fine-tuned on the training set. To combine CNN models, author present two schemes for learning:

by minimizing the log likelihood loss, and by minimizing the hinge loss. The method generates result on the FER dataset. It also achieves 55.96% and 61.29% respectively.

X. Yin et al [10] explores multi-task learning (MTL) for face recognition. They answer the questions of how and why MTL can improve. First, author propose a multi-task CNN for face recognition where identity classification is the main task and pose, illumination, and expression estimations are the side tasks. Second, they develop a dynamic-weighting scheme to automatically assign the loss weight to each side task. Third, they propose a pose-directed multi-task CNN by grouping different poses to learn pose-specific identity features. They propose an energy-based weight analysis method to explore how CNN-based MTL works. They observe that the side tasks serve as regularizations to disentangle the variations from the learnt identity features. Experiments demonstrate the effectiveness of the approach. The approach is also applicable for pose-invariant face recognition and better performance.

The problems facing by blind people fall in the category of navigating through indoor and outdoor consisting of various obstacles and recognition of person in front of them. Identification of objects or person only with perceptive and audio information is difficult. This helps blind people to navigate with the help of a Smartphone, global positioning system (GPS) and a system with ultrasonic sensors. Face recognition can be done using neural learning techniques. The images of friends, relatives are stored in the database. Whenever a person comes in front of the blind user, the application gives the voice aid to the user. Thus system can replace the regular imprecise use of guide dogs and white sticks to help the navigation and face recognition process. *P. M. Kumar et al [11]* have proposed a novel image recognition and navigation system that provides precise and quick messages in the form of audio to visually challenged people.

The complexity of the human face and the changes due to different effects make it more challenging to design and implement a system for face recognition. *M. A. Abuzneid et al [12]* presented an approach to improve face recognition using a BPNN and features extraction based on correlation. A key contribution is the generation of a new set called the T-Dataset from the training data set, which is used to train the BPNN. The correlated T-Dataset provides a high distinction layer between the training images, it helps BPNN to converge faster and achieve better accuracy. Additionally authors used LBP histogram to prove that there is potential improvement. They applied five distance measurement algorithms and combine them to obtain the T-Dataset. Authors achieved higher recognition accuracy with less computational cost. Furthermore, the author evaluates methods on one of the benchmark data set, Labeled Faces, where they produce a competitive face recognition performance.

Thermal face recognition has received increasing attention. The traditional methods for TFR concentrate on the hand-crafted features, which requires more efforts to select and extract features and usually has lower recognition rate. *Z. Wu et al [13]* present CNN architecture for TFR. It is a new type of ANN that can learn features from the raw data. Results on face database show that CNN achieves higher recognition rate.

S. Kumaar et al [14] present real-time deep neural network architecture for disguised face verification. The model consists of two ANN, first one being CNN that predicts 20 facial key-points in the image and the second ANN classifies the subject. The accuracies are 67.4% and 74.8% for prediction and classification respectively.

Identification of individual livestock has become a pressing issue as intensification practices continue to be adopted and precise objective measurements are required. Current best practice involves the use of RFID tags which are

time-consuming task. To overcome this, non-invasive biometrics is proposed by *M. F. Hansen et al [15]*. They test this in a farm, on 10 individual pigs using three techniques: Fisher-faces, the VGG-Face pre-trained face CNN model and own CNN model that train using an artificially augmented data. Results show that

accurate pig recognition is possible with accuracy rates of 96.7% on 1553 images.

III. LITERATURE SUMMARY

This section provides the summary of the collected literature. The table 3.1 includes the different aspects of the proposed study.

Table 3.1 : Review summary

Authors and publication	Contribution	Outcomes
2015 ACM, S. S. Farfade et al [6]	Consider problem of multi-view face detection and proposed Deep Dense Face Detector. That does not require pose/landmark annotation to detect faces of orientations based on deep CNN. They analyzed scores of face detector for faces in different orientations and found 1) able to detect faces of different angles, 2) a correlation between distribution of positive samples in the training set and scores of the given face detector.	The method's performance was improved using sampling and sophisticated data. Evaluations on datasets show that single-model face detector has better performance.
Journal of Latex Class Files 2015, P. Tzirakis et al [7]	An emotion recognition system using auditory and visual modalities. To capture emotions of speaking. A CNN to extract features from the speech, for visual modality a deep residual network proposed to be insensitive to outliers to model the context. So, Long Short-Term Memory networks are utilized. System trained in an end-to-end fashion and also taking advantage of the correlations.	It outperforms as compared to the traditional approaches based on auditory and visual handcrafted features.
2015 Hindawi, W. Hu et al [8]	Deep CNN are employed to classify hyper-spectral images. The architecture of classifier contains five layers with weights (input, convolutional, max pooling, full connection, and output layer.	Results based on several data sets show that the method can achieve better performance than traditional methods, such as SVM and traditional deep learning- methods.
2015 ACM, Z. Yu et al [9]	Focus on SFEW 2.0 dataset, seeks to classify a set of images into 7 emotions. It contains a face detection module, followed by classification module using ensemble of multiple deep CNN. It is initialized randomly and pre-trained.	First minimizing the log likelihood loss, by minimizing hinge loss. The method generates result on FER dataset. It also achieves 55.96% and 61.29% respectively.

arXiv 9 May 2017, X. Yin et al [10]	They explore multi-task learning for face recognition. A multi-task CNN for face recognition where identity classification is the main task. Second, develop a weighting scheme to assign the loss weight. Third, a pose-directed multi-task CNN by grouping different poses to learn features.	An energy-based weight analysis to show CNN- MTL work. They observe that the side tasks serve as regularizations to disentangle the variations from the learnt identity features. Experiments demonstrate the effectiveness of the approach. The approach is also applicable for pose-invariant face recognition and better performance.
Springer Nature 2017, P. M. Kumar et al [11]	The problems facing by blind people fall in the category of navigation through indoor and outdoor consisting of various obstacles and recognition of person in front of them.	Stored images of friends, relatives in the database to recognize a person in front of blind user, to gives voice aid to user. It is a face recognition and navigation system. It provides precise and quick messages.
IEEE 2018, M. A. et al [12]	That generates a dataset called the T-Dataset used to train the BPNN. Provides a high distinction layer between training images to converge faster and better accuracy and also used LBP histogram to prove that.	Achieved higher recognition accuracy with less cost. They produce a competitive face recognition performance.
2016 IEEE, Z. Wu et al [13]	Thermal face recognition depends on the hand-crafted features, requires efforts to select features and has lower recognition rate.	CNN architecture for TFR. A new type of ANN can learn features from raw data. Results on face database show CNN achieves higher recognition rate.
2018 IEEE, S. Kumaar et al [14]	It consists of two ANN, first CNN predicts 20 facial key-points in image and second ANN classifies subject.	The accuracies are 67.4% and 74.8% for prediction and classification respectively.
Elsevier 2018, M. F. Hansen et al [15]	Non-invasive biometrics, CNN model that train using an artificially augmented data.	Accurate pig recognition with rates of 96.7% on 1553 images.

III. PROPOSED WORK

The proposed work is motivated to study the multiple pose based face recognition system. A number of highly accurate models for recognizing faces in multiple poses are presented in recent years by different contributors, but most of the techniques either less accurate or computationally cost effective. Therefore the proposed work is focused on exploring the accurate pose based face recognition model. Additionally the proposed model is extended to offer the technique for partial face recognition.

Thus we established solution for two limitations of the current available approaches:

1. Recognition of partial faces
2. Recognition of face after small change in face images (i.e. smiling face, crying face)

Basically, even small changes in face can impact on the performance of the learning algorithm for face recognition. Therefore need to be developing a feature selection and classification method by which we achieve the required consequences. Therefore the required model is demonstrated in figure 3.1.

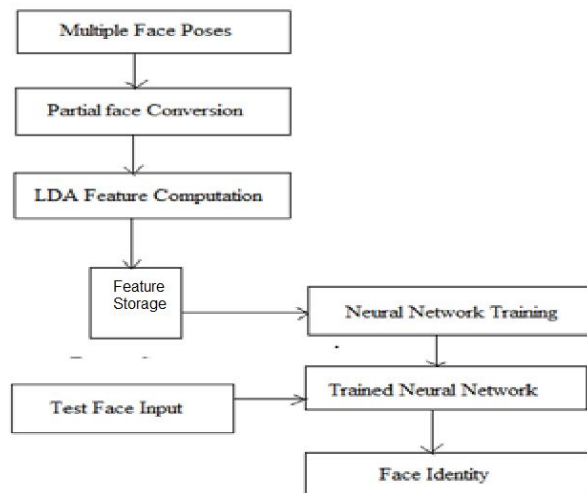


Figure 3.1: Proposed Model

The proposed model works in two modules. In first the training is provided to the model and then testing module is provided. Thus to demonstrate the required model ORL dataset is used. In this dataset we consider on 40 classes or identities and for each class we use 10 different poses. Means total of 400 images of 40 people is used. Among them the 70% of randomly selected images are used to train the model and 30% of data is used for testing. The system accepts the input sample images and for partial pose recognition the images are partitioned. These partitioned dataset is used for training and testing of the system.

Further the features from the face images are computed and stored in a database. In order to compute the features the LDA is used. The stored image features in database are learned by the artificial neural network (ANN). After learning the ANN is able to classify the data or faces by producing the test image samples. Finally based on the classification outcomes the performance of the model is computed. The proposed partial face recognition model is used in various other tasks also using multi-pose face recognition technique.

1. Biometric identity based authentication system

2. Identification of faces even some changes on face are found

This section provides the overview of the proposed model which is being developed in further.

V. CONCLUSION & FUTURE WORK

The face recognition is a classical domain of research and development. Therefore a review of existing tools and technology is presented and summarized first. According to the summary of face recognition models we concluded a number of different domains and applications. A significant developments are contributed in this domain i.e. authentication, navigation, and others. But a number of challenges are remaining to fix. In this work we consider the issue of changes in faces to recognize the multiple poses and also able to recognize the identity by partially captured face images. Therefore in this presented work an artificial neural network is employed for recognizing the human partial faces. Those concept usages the LDA based feature extraction to carry out this task. By using this we demonstrate the model for future development and implementation. The proposed model implemented in near future and their performance is measured and reported.

REFERENCES

1. D. Bhamare, P. Suryawanshi, "Review on Reliable Pattern Recognition with Machine Learning Techniques", *Fuzzy Information and Engineering*, 10:3, 362-377
2. F. P. Mahdia, Md. M. Habibb, Md. A. R. Ahadb, S. Mckeeverc, A. S. M. Moslehuddinb, P. Vasanta, "Face recognition-based real-time system for surveillance", *Intelligent Decision Technologies* 11 (2017) 79-92
3. A. K. Jain, S. Pankanti, S. Prabhakar, L. Hong, A. Ross, J. L. Wayman, "Biometrics: A Grand Challenge", *Proceedings of International Conference on Pattern Recognition*, Cambridge, UK, Aug. 2004

4. M. Hassaballah, S. Aly, "Face recognition: challenges, achievements and future directions", *IET Comput. Vis.*, 2015, Vol. 9, Iss. 4, pp. 614–626, 614 & The Institution of Engineering and Technology 2015
5. M. P. Beham, S. M. M. Roomi, "A Review of Face Recognition Methods", *International Journal of Pattern Recognition and Artificial Intelligence* Vol. 27, No. 4 (2013) 1356005 (35 pages)
6. S. S. Farfade, M. Saberian, L. J. Li, "Multi-view Face Detection Using Deep Convolutional Neural Networks", *ICMR'15*, June 23–26, 2015, Shanghai, China Copyright is held by the owner/author(s). ACM 978-1-4503-3274-3/15/06
7. P. Tzirakis, G. Trigeorgis, M. A. Nicolaou, B. Schuller, S. Zafeiriou, "End-to-End Multi-modal Emotion Recognition using Deep Neural Networks", *Journal of Latex Class Files*, Vol. 14, No. 8, Aug 2015
8. W. Hu, Y. Huang, L. Wei, F. Zhang, H. Li, "Deep Convolutional Neural Networks for Hyperspectral Image Classification", *Hindawi Publishing Corporation Journal of Sensors* Volume 2015, Article ID 258619, 12 pages
9. Z. Yu, C. Zhang, "Image based Static Facial Expression Recognition with Multiple Deep Network Learning", *ICMI 2015*, November 9–13, 2015, Seattle, WA, USA c 2015 ACM ISBN 978-1-4503-3912-4/15/11
10. X. Yin, X. Liu, "Multi-Task Convolutional Neural Network for Pose-Invariant Face Recognition", *arXiv:1702.04710v2 [cs.CV]* 9 May 2017
11. P. M. Kumar, U. Gandhi, R. Varatharajan, G. Manogaran, R. Jidhesh, T. Vadivel, "Intelligent face recognition and navigation system using neural learning for smart security in Internet of Things", *Springer Science+Business Media, LLC, part of Springer Nature* 2017
12. M. A. Abuzneid, A. Mahmood, "Enhanced Human Face Recognition Using LBPH Descriptor, Multi-KNN, and Back-Propagation Neural Network", 2169-3536, 2018 IEEE, Vol 6, 2018
13. Z. Wu, M. Peng, T. Chen, "Thermal Face Recognition Using Convolutional Neural Network", 2016 *International Conference on Optoelectronics and Image Processing*, 978-1-5090-0880-3/16/\$31.00 ©2016 IEEE
14. S. Kumaar, R. M. Vishwanath, S. N. Omkar, A. Majeedi, A. Dogra, "Disguised Facial Recognition Using Neural Networks", 2018 *IEEE 3rd International Conference on Signal and Image Processing*, 978-1-5386-6396-7/18/\$31.00 ©2018 IEEE
15. M. F. Hansen, M. L. Smith, L. N. Smith, M. G. Salter, E. M. Baxter, M. Farish, B. Grieved, "Towards on-farm pig face recognition using convolutional neural networks", *Computers in Industry*, 2018, 98, 145-152.