Scalable Computing: Practice and Experience, ISSN 1895-1767, http://www.scpe.org
© 2025 SCPE. Volume 26, Issues 1, pp. 361–370, DOI 10.12694/scpe.v26i1.3533

GENOME SEQUENCE ANALYSIS OF SEVERE ACUTE RESPIRATORY SYNDROME USING GENOANALYTICA MODEL

SHIVENDRA DUBEY*, DINESH KUMAR VERMA†, AND MAHESH KUMAR‡

Abstract. We proposed a GenoAnalytica model for examining the SARS's genomics sequences. The technologies make proper data extraction from genomics sequences of viruses. We use the GenoAnalytica model, i.e. GenoCompute, and IGMiner Algorithm; to classify the range of genomics sequences, including recognizing the sequence variation from the datasets. The projected algorithm computes the nucleotide patterns and represents the nucleotide genome sequence of SARS (airborne virus) by IGMiner technique and works out on the GenoCompute to calculate computation time with minimum count in second. Along with this, we proposed a UMRA algorithm to compute the mutation rate of the genome sequence with minimum count in seconds as compared to traditional method. Furthermore, we work out the different datasets (China and Algeria datasets) and determine the whole variation at the index level inside the all genome sequence. This learning also signifies the performance evaluation on altering minsup using IGMiner and Aprori-based SPM. Also, we calculate the mutation rate of the genome sequence of airborne virus using Unique Mutation Rate Analysis algorithm. The severe acute respiratory syndrome coronavirus 2 has been responsible for the deadly COVID-19 pandemic. It has ruined limitless individuals all over the globe, and along with this, it continues to harm well-being and people's health. Healthcare specialists and Researchers can obtain insight into COVID-19's inherited variation or SAR-CoV-2 through cutting-edge Artificial Intelligence and genome sequence analysis tools.

Key words: IGMiner, UMRA, Genome Sequence, COVID-19, GenoCompute

1. Introduction. Corona, also called COVID-19, represents a severe respiratory disease brought about by a new corona virus named SARS-CoV-2. The infectious disease was initially discovered in December 2019 in the Chinese city of Wuhan; it since then has spread around the world, causing a global epidemic. When a person with the infection of sneezes, coughs, breathes or speaks loudly, COVID-19 typically spreads by droplets from their lungs [1]. The virus may also be transferred by contacting infected surfaces or items, especially the mouth, face, eyes, or nose. The moderate to severe COVID-19 symptoms include coughing, fever, exhaustion, shortness of breath, muscular or body pains, loss of smell or taste, sore throats, and digestive and headaches problems. In extreme circumstances, it can result in mortality, organ failure, ARDS (acute respiratory distress syndrome), and pneumonia, particularly among older persons and people with underlying medical issues [2, 3]. It is crucial to practise excellent hygiene to stop the spreading of COVID-19. Examples include: washing your hands frequently with water and soap for at least twenty seconds; using a hand sanitizer with a minimum alcohol content of 60%; wearing masks in public; engaging in social distancing; avoiding big gatherings. In several nations, vaccines have been created and are authorized for use during situations of crisis to help prevent COVID-19. To guarantee these vaccinations' effectiveness and safety, they undergo extensive testing. The SARS-CoV-2 genome is an RNA genome, meaning RNA (ribonucleic acid) is used to carry biological data [4, 5]. The virus's RNA with one strand genome generates several proteins required for interaction and replication with its host cells. Scientists have pinpointed the precise genetic makeup of SARS-CoV-2 thanks to the sequencing of the virus' genome. Researchers can follow the development and transmission of the virus's genome and better understand how it is evolving and adapting as time passes by comparing several virus genomes from various historical periods and places. Additionally essential to the creation of COVID-19 tests for diagnosis is genome sequencing. Scientists may create tests to identify the existence of the pathogenic virus in patient specimens

*Department of Computer Science and Engineering, Jaypee University of Engineering and Technology, Guna, Madhya Pradesh, India, 473226 (shivendrashivay@gmail.com).
†Department of Computer Science and Engineering, Jaypee University of Engineering and Technology, Guna, Madhya Pradesh, India, 473226 (dinesh.hpp@gmail.com).
‡Department of Computer Science and Engineering, Jaypee University of Engineering and Technology, Guna, Madhya Pradesh, India, 473226 (mahesh.chahar@gmail.com).

by focusing on some regions of the genome that contain the virus, especially its genes that generate the protein known as the spike, which is essential for viral entrance into host cells [6].

There some key contributions of this research as:

1. This article proposes the GenoAnalytica method that offers an innovative technique for sequence rule mining through the use of the effective data structure and the classes of equivalence.
2. This article performs in-depth tests on actual datasets to verify the efficiency enhancements of Geno-Analytica, illustrating its usefulness for massive sequence rule mining jobs.
3. This study recognizes a modest rise in memory usage as a cost of improved speed, offering important details on the approach's resource needs.
4. The enhanced efficiency of ERMiner is anticipated to be advantageous to a broad variety of uses which depend on successive rule mining, including bio informatics, website click stream analysis, and market basket analysis.

2. Related Work. A novel optical biosensor [10] developed and integrates thermal and optical properties to identify the coronavirus. Mostly gold nanoislands connected on an optical substrate make up the sensor. Coronavirus RNA sequences corresponding to synthetic DNA receptors have been discovered on nanoislands. The sensor's receptors work in conjunction with the visible virus. Localized Surface Plasmon Resonance is the name of this technique (Swiss Federal Laboratories for Materials Science and Technique). A brand-new biosensor built around specific cells with altered mammalian membranes was proposed [11]. By including 10% foetal bovine serum, the author examined the circumstances of Green Monkey's renal culture of cells.

Additionally, Vero/Anti-S1 membrane-engineered cells were used to fabricate sensors. This cutting-edge biosensor proved that it could be used for COVID-19 antigen surface large-scale screening in about 3 minutes and produced a remarkable outcome. Utilizing two approaches, including the ANFIS (adaptive network-based fuzzy inference system) and multi-layered perception modelling, suggested a comparative study between the soft computing and machine learning models for predicting a global epidemic disease, highlighting the possibility of using machine learning as a tool for solutions in healthcare [8]. Technological devices offer enormous potential for interacting with mental health. Wristbands, Cellphones, and Smartwatches are examples of wearable devices having integrated sensors that can communicate through Wi-Fi or Bluetooth. Gyroscopes and accelerometers are examples of sensors that sense inertia. The sense of human body heavily relies on biological sensors that measure heart rate and environmental sensors that measure temperature. To address the issues of identity, trust, and privacy IJPCC developed [12] a ubiquitous system for computing that utilizes the trust model. The framework used naive Bayes (NB) and Apriori models to extract behaviour patterns during the decision-making process. For a summary of the advancements and research on wearing biomarker structures, particularly for monitoring one's health, wearable biosensors demonstrated [13]. Using smartphones presented a method to identify human behaviour that includes health-related behaviours, including physical activity for sleeping and fitness activity. The recommended concept uses sensors contained within the Arduino GNO as input to power systems that run computers. The COVID-19 illness is forecast-ed and categorized for subsequent therapy using artificial intelligence technology and a model based on mathematics [14].

3. Material and Methods.

3.1. Nucleotide Sequences. The exact sequence of the nucleotide in a molecule of RNA or DNA is referred to as a nucleotide sequence. These molecules' fundamental components, known as nucleotides, are made from three primary parts: a phosphate group, a nitrogenous base, ribose in RNA and deoxyribose in DNA(a sugar molecule). Adenine (A), cytosine (C), guanine (G), and thymine (T) are each of the four distinct bases composed of nitrogen that may be discovered in DNA. Uracil (U) takes the role of thymine in RNA. The nucleotide sequence comprises the arrangement and order of these bases across the sugar-phosphate backbone [7].

The steps required to create proteins and other valuable compounds are included in the arrangement of nucleotide sequences, which also house the biological information of a living thing. It gives biological data to identify an individual's traits and characteristics [8].

3.2. Genome sequencing. Discovering the whole sequence of DNA of the genome of a living thing is a procedure called genome sequencing. It entails defining the arrangement of the bases that make up nucleotide

(adenine, cytosine, guanine, and thymine) within the DNA molecular structure, which transmits a living thing's genetic information. Genome sequencing may be done using various technologies and methodologies, with newer, more efficient ones made possible by technological breakthroughs [9]. The WGS (whole-genome sequencing) and TS (targeted sequencing) were the two main methods for sequencing the human genome.

3.3. Datasets. We have implemented this methodology on hp laptop with i5 processor; and the genome sequence dataset of SARS-CoV–2 of USA and Algeria collected from the NCBI repository. The genome sequence files has been use in common file extensions, like FASTA or GenBank, to allow for simple analysis and exchange across investigators. An investigator has examined patterns over a period of time by using the information set's potential inclusion of information on the time and dates each collecting the samples. The data sets have details about the site and geographic coordinates where the specimen took place. Using this, one may monitor the virus's international spread. It is often possible to identify and get particular sequences from a dataset by using the accession number, which is a code that is usually given to that particular sequence. The sequence information is usually structured within a certain pattern and is frequently expressed employing the common nucleotide coding (A, T, C, and G).

3.4. Proposed Method. The genome sequence's computing time is computed using Algorithm 1's and provided information. We use m = 3 and k = 9 to calculate the minimizes to comprehend the GenoCompute Algorithm better. These values were chosen based on experience with the conventional validation set technique. The frequency distribution vector-based representations are produced using the same methodology as the minimizes for specific nucleotide sequencing. We refer to this technique as Minimize Vector for convenience. A sequence database (SQDB) and the minconf and minsup thresholds are inputs to IGMiner. It first does a single scan of the database to create all equivalence classes of rules of Size 1*1, or including a single item in a single entity and antecedent in the consequent. The LMS (left search) function is then invoked to execute left merges across all left equivalence categories to find more extensive rules.

Similarly, the RMS (right search) function performs right merges for appropriate equivalence classes. Although left merges are permitted after right merges, it should be noted that the RMS (right search) technique may produce some additional left-equivalence classes. Those equivalence classes are kept in the left store structure. The processing of these classes of equivalence is done in a separate loop. The IGMiner Algorithm then returns the collection of discovered rules.

Algorithm 1 GenoCompute Algorithm

Input:
minconf: Minimum Confidence Threshold;
SQDB: A sequence database
Output:
Set of Valid Sequential Rules
1. Let Store = Empty set
2. Let rules = Empty set
3. Scan SQDB once to calculate EQ, the set of all uniformity classes of rules of size 1*1
4. For each left uniformity class C1 in EQ, do
5. LMS(C1, rules)
6. End
7. For each right uniformity class C2 in EQ, do
8. RMS(C2, rules, Store)
9. End
10. For each left uniformity class C3 in Store, do
11. RMS(C3)
12. End
13. Return rules

Each site in the COVID-19 sequences is compared to its appropriate place in the standard genome as part of the UMRA method, which determines the rate at which mutations occur. The total variety of modifications

Algorithm 2 IGMiner Algorithm

```
Input: Sequence s and integer k and m
Output: Set of Minimizes
minimizes = 0
def find_minimizes(sequence, k_length, m_length):
minimizes = set()
queue = []
current_min_index = 0
for i in range(len(sequence) - k_length + 1):
kmer = sequence[i:i+k_length]
queue.append(min(kmer[j:j+m_length] for j in range(k_length - m_length + 1)))
if len(queue) > m_length:
queue.pop(0)
if queue[-1] < queue[current_min_index]:
current_min_index = len(queue) - 1
if i >= k_length - m_length:
minimizes.add(queue[current_min_index])
return minimizes
```

and the overall frequency of aligned locations are counted. After that, the mutation rate is determined by multiplying the alteration score by the number of aligned positions. This approach can be enhanced and modified depending on the needs and demands for further study.

Algorithm 3 UMRA Algorithm (Unique Mutation Rate Analysis)

```
def calculate_mutation_rate(reference_sequence, covid_sequences):
mutation_count = 0
aligned_positions = 0
for i in range(len(reference_sequence)):
reference_nucleotide = reference_sequence[i]
for sequence in covid_sequences:
if i < len(sequence):
aligned_positions += 1
if sequence[i] != reference_nucleotide:
mutation_count += 1
mutation_rate = mutation_count / aligned_positions if aligned_positions > 0 else 0
return mutation_rate
# Example usage
reference_genome = "ATCGATCGATCG..."
covid_samples = ["ATCGATCGATCG...", "ATCGGTCGATCG...", "ATCGATCGCTCG..."]
rate = calculate_mutation_rate(reference_genome, covid_samples)
print("Mutation Rate:", rate)
```

4. Results. The total length of the genomics order, the computing power available, and the level of difficulty of the statistical methods are some variables that can affect how long it takes to analyze the COVID-19 genome sequence. The explanation of the variables that might impact calculation time is as follows:

Genome Sequence Length. SARS-CoV-2, the virus that causes COVID-19, has a genome sequence of around 30,000 base pair combinations long. The sequence length will affect how long it takes to process and analyze the data. Sequence comparison and alignment algorithms could take longer as the Size rises.

Analysis Algorithms. The decision regarding the analysis algorithms significantly impacts computation time. There are many levels of computational complexity for various analytical methods, including variant

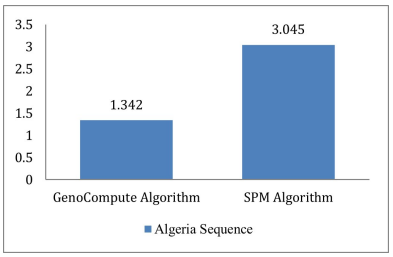

Fig. 4.1: Computation time of genetic sequences

calling, sequence alignment, and mutation detection. Several algorithms' implementations have been optimized to handle enormous datasets more quickly, cutting down on calculation time.

Computational Resources. Computation time may be significantly influenced by the available computational resources, such as the computer's processing speed and algorithm design. Analysis activities can be completed more quickly by dividing the workload across several processors or nodes into outstanding-performance cloud-based platforms or computing systems with concurrent processing capabilities. Here, we analyse the computation time of genetic sequences by two different algorithms which is shown in Figure 4.1.

Optimizing and paralleling. The effectiveness of the analytical methods and how they are implemented can impact computing time. Processing time can be decreased by optimized algorithms that minimize pointless operations or make use of effective data structures. The execution of analytical tasks concurrently can be made possible via parallelization techniques like distributed computing and multi-threading, which further reduces processing time.

Complexity and Size of the dataset. It might affect how quickly calculations are performed. The computational burden and processing time can be increased by analyzing excessive COVID-19 genomics sequences or incorporating more metadata, like clinical data or sample information. We worked on the Algeria sequences with GenoCompute algorithm and SPM algorithm (based method) to calculate the computation time, we get our proposed GenoCompute algorithm is much faster than SPM algorithm (see figure 4.1).

These variables significantly affect the calculation time, and it is challenging to give a specific time without considering the hardware resources, particular analytic pipelines, and dataset peculiarities. However, developments in parallel computing, computational biology, and algorithmic advances continue to increase computation time, the capacity for quicker analysis, and the effectiveness of genome sequence analysis of COVID-19 genome-wide information.

Knowing the COVID-19 dataset's genome sequence variation percentage offers essential information on the genetic variations of the virus's genome within a community or among many samples as shown in figure 2. Several factors make determining this proportion crucial, including the following:

Understanding Genetic Variability. The SARS-CoV-2 virus, the source of COVID-19, can demonstrate genetic changes or variations as it multiplies and spreads. The degree of biological variety within a dataset may be determined by computing the proportion of genome sequence diversity. This knowledge is essential for following the virus's development, spotting new variations, and comprehending their possible effects on pathogenic, transmission, or responsiveness to medications or vaccinations.

Monitoring the Spread and Transmission of Viruses. Analysis of Genome Sequence Variation (AGSV) enabled the tracking of viral transmission routes and the identification of virus clusters or lineages. We can evaluate the similarity of virus strains and monitor the spread of certain variations by comparing genomes from various places or periods. This data can help with contact tracing, public health actions, and tracking the success of control measures.

Genome sequence variation analysis supports epidemiological research and outbreak investigation. It assists researchers in comprehending the virus's genesis and dynamics of transmission, locating super-spreader events or clusters, and evaluating the effectiveness of treatments or containment measures. Sequences from various geographic or historical eras might be compared to recreate the virus' evolutionary history and learn more

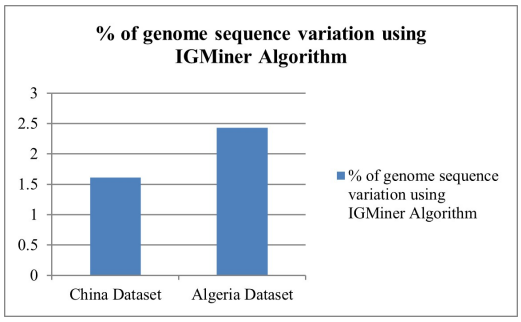

Fig. 4.2: Variation of genome sequence

Table 4.1: Variation in genome

Line	Location	Position	Change
1	23	23	T -> C
33	46	1966	T -> C
90	52	5392	G -> A
133	51	7971	A -> C
244	23	14603	T -> C
247	47	30305	A -> C

about how widely it has travelled.

We can investigate genetic variety, follow the transmission of the virus, evaluate the effectiveness of vaccines, and guide public health measures by measuring the proportion of the genome's sequence variations within COVID-19 databases. It is essential to expand our knowledge of this virus and enable decisions based on evidence in emergency response initiatives. Base method is not calculating percentage of genome sequence variation but our proposed IGMiner algorithm calculates this with high effectiveness (see figure 4.2).

There are several uses for determining the indexing of missing sequences in a COVID-19 dataset. Here are a few cases where deciding how to index missing sequences is crucial:

Evaluation of Completeness. A dataset's completeness can be inferred from the existence or absence of certain sequences. By determining missing sequences, researchers can evaluate whether particular geographic areas or viral strains are underrepresented or absent. Understanding the dataset's limits and any biases in the research is vital.

Comparative Analysis. In COVID-19 research, comparing the genetic sequences of several samples or areas is customary. The quality and thoroughness of comparison studies may suffer if specific sequences are absent from the dataset. Researchers can find deficiencies in the dataset and confidently decide whether the data are reliable and appropriate for particular analysis by computing the indices of missing sequences.

Identification of variations. Genetic variations or mutations can be identified by the absence or presence of specific sequences. Researchers can locate locations with novel or distinctive genetic variants by detecting missing sequences. Table 4.1 demonstrates the knowledge which can help with the detection and characterization of novel viral strains or variations in the population.

The integrity of Data and Control of Quality. One quality control tool is to determine how to index missing sequences. By eliminating any mistakes in data input or collecting, it helps to guarantee that the set of data is correct and complete. Researchers can preserve data integrity and improve the dataset's quality by checking for predicted sequences and locating missing ones.

Figure 4.3 Determining the order of indexing of missing sequences within the COVID-19 dataset enables researchers to judge the completeness of the dataset, ease comparison studies, pinpoint genetic variations, and guarantee quality control and data integrity. It allows for a more thorough and trustworthy comprehension of

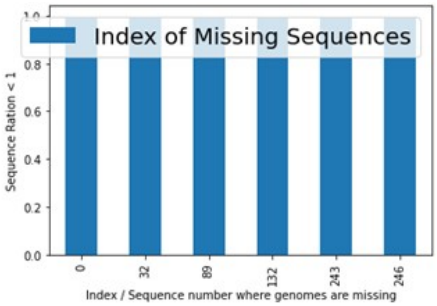

Fig. 4.3: Missing sequences

Table 4.2: Comparisons of Mutation with SPM-Point and GenoAnalytica Model

Dataset	**SPM-Point Mutation (sec)**	**UMRA Model (sec)**
China Dataset	3.96	2.73
Algeria Dataset	4.031	2.84

the virus and encourages reasoned judgement in public health and research activities.

We advise looking for information regarding the mutational study of the COVID-19 datasets in Algeria in reputable sources such as academic research papers, books from health organizations, or official materials provided by the Algerian Ministry of Healthcare or other authorities. In addition to any found variations or modifications in the COVID-19 virus in Algeria, these sites would offer current and reliable information regarding the mutation analysis, as shown in Figure 4.4.

It's crucial to understand that mutation analysis entails looking at the genetic structure of a virus and locating any alterations or variations in its molecular structure. This study aids in comprehending the genetic variety of the virus, monitoring the appearance and dissemination of various variations represented in Table 4.2, and evaluating the possible consequences for disease severity, transmission, therapies, diagnostics, and vaccinations. Here our proposed method worked on three data sets (USA, China and Algeria) to compute the mutations and we get our proposed method is so faster as compared to base method.

Figure 4.5 shows the comparison of support on various sequences with base method and IGMiner method. The time needed to identify linkages or correlations between objects in a dataset is called the "association time." The precise comparison relies on the data set's dimensions, complexity, chosen algorithm, and processing power. Direct comparisons of association periods are challenging without accurate information. However, the time required for association mining might vary greatly depending on the dataset's properties and methods. The particular dataset, including the amino acid sequences and their accompanying support values would typically be required to compare the support for sequencing found in the Algeria COVID-19 dataset. A simple comparison of support values for the genes in the Algeria COVID-19 dataset cannot be made with any accessibility to the actual dataset. We can, however, describe support in general and its importance in sequence analysis. Support in sequence analysis is the frequency or recurrence of an individual sequence pattern throughout a dataset. It reveals the relative frequency with which a specific design or sequence shows up in the dataset concerning the number of sequences. Support is essential for several data mining activities, such as sequential pattern mining and association rule mining. It assists in identifying frequent or significant sequences that are more likely to be noteworthy or instructive.

5. Conclusion. In this study, particular approaches for examining and investigating SARS-CoV-2 genomics sequences have been presented. The most important method, the pattern mining algorithm, recognizes frequent nucleotide bases inside the sequences and furthers their sequential and regular patterns. Numerous sequence prediction algorithms were experienced on genomics sequences. Furthermore, the outcomes point out

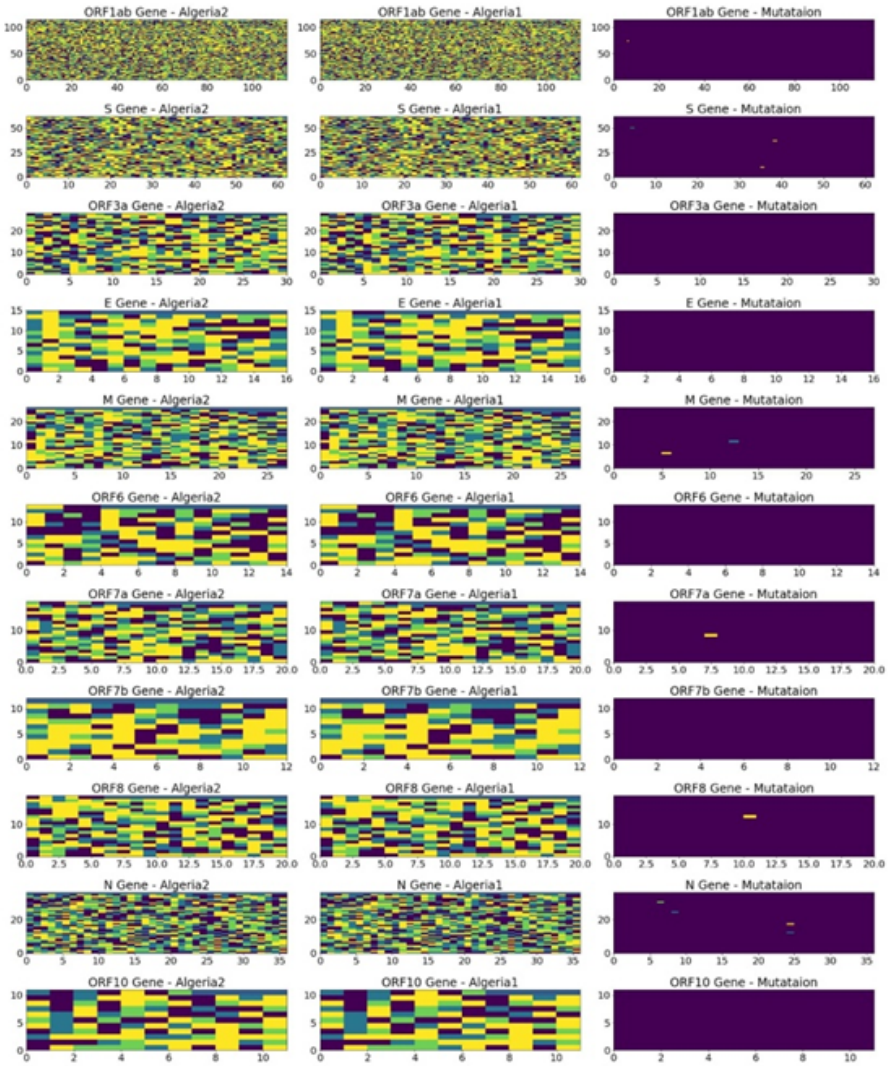

Fig. 4.4: Mutations

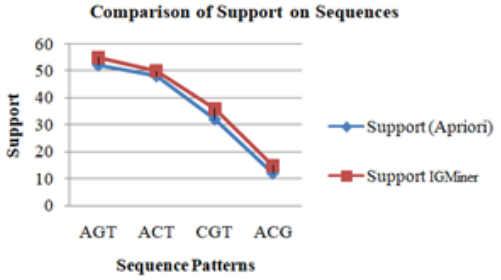

Fig. 4.5: Support sequence comparison

that IGMiner methods execute much better than supplementary methods considered to be state-of-the-art. The SARS-CoV-2 virus is not a simple virus that might be treated by the methods illustrated in this investigation. They may also be used to investigate other human virus analyses. It's probable that in the not-too-distant future, we may expand the span of our attempt to comprise the categorization examination of supplementary genome sequences. When it comes to genomics sequence categorization and analysis, investigators are powerfully urged to build alternative methods based on AI. We have worked on two Algeria data sets and the China data sets collected from the NCBI repository. Initially, we worked out the computation time by GenoCompute along with the comparison with the SPM method, which is evaluated in seconds; it is measured within 1.342 seconds and 3.045 seconds correspondingly. After that, we survey the fraction variation modification in genome sequence for particular datasets; also examine fundamental alteration at index level in genome sequence for different datasets; and investigation of frequent sets creation through IGMiner vs Apriori is also done for other patterns (A, C, G, and T). In this work, we have also calculated the mutation rate of the genome sequence of COVID-19 using the UMRA method with less time than the SPM method on China and Algeria datasets. In conclusion, GenoAnalytica, a unique sequential rule mining method, was proposed. It makes use of the idea of equivalence classes to facilitate the finding of rules and incorporates the ground-breaking effective data structure to facilitate the pruning of the search area. Multiple evaluations using real-world datasets showed that GenoAnalytica beats the most recent method, giving faster results at the cost of somewhat higher memory usage. Such findings highlight GenoAnalytica's potential as an asset for scaled sequence rule mining activities. In future, we can work on any other infectious disease using the analysis of genomes and their mutation. Mycobacterium tuberculosis is the bacterium that causes tuberculosis (TB), a communicable illness. Although it can affect other parts of the body, it mostly affects the lungs. Constant weight loss, coughing, a high temperature, sweats at night are some of the symptoms. When someone who is infected sneezes or coughs, the disease spreads by the breath of others. Limited access and antibiotic resistance are the main reasons why tuberculosis (TB) persists as a serious global medical concern even though it is preventable and curable.

REFERENCES

[1] Nawaz MS, Fournier-Viger P, Shojaee A, Fujita H., *Using artificial intelligence techniques for COVID-19 genome analysis*, in 1. Applied Intelligence. 2021 May;51:3086-103.

[2] J. Ahmed I, Jeon G, *Enabling artificial intelligence for genome sequence analysis of COVID-19 and alike viruses*, Interdisciplinary sciences: computational life sciences. 2022 Jun;14(2):504-19.

[3] Márquez S, Prado-Vivar B, Guadalupe JJ, Gutierrez B, Jibaja M, Tobar M, Mora F, Gaviria J, García M, Espinosa F, Ligña E, *Genome sequencing of the first SARS-CoV-2 reported from patients with COVID-19*, in Ecuador. MedRxiv. 2020 Jun 14 .

[4] Laamarti M, Alouane T, Kartti S, Chemao-Elfihri MW, Hakmi M, Essabbar A, Laamarti M, Hlali H, Bendani H, Boumajdi N, Benhrif O, *Large scale genomic analysis of 3067 SARS-CoV-2 genomes reveals a clonal geo-distribution and a rich genetic variations of hotspots mutations*, 1. PloS one. 2020 Nov 10;15(11):e0240345.

[5] Mousavizadeh L, Ghasemi S, . *Genotype and phenotype of COVID-19: Their roles in pathogenesis. Journal of Microbiology, Immunology and Infection*, 2021 Apr 1;54(2):159-63.

[6] Lu R, Zhao X, Li J, Niu P, Yang B, Wu H, Wang W, Song H, Huang B, Zhu N, Bi Y, *Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding*, The lancet. 2020 Feb 22;395(10224):565-74 .

[7] Ray M, Sable MN, Sarkar S, Hallur V, *Essential interpretations of bioinformatics in COVID-19 pandemic*, Meta Gene. 2021 Feb 1;27:100844 .

[8] Quazi S, *Artificial intelligence and machine learning in precision and genomic medicine.* , 1. Medical Oncology. 2022 Jun 15;39(8):120.

[9] Ahmed I, Ahmad M, Jeon G, Piccialli F, *A framework for pandemic prediction using big data analytics.*, Big Data Research. 2021 Jul 15;25:100190 .

[10] Dubeya S, Kumar M, Verma DK, *Machine Learning Approaches in Deal with the COVID-19: Comprehensive Study*, ECS Transactions. 2022 Apr 24;107(1):17815 .

[11] Tripathi A, Chourasia U, Dubey S, Arjariya A, Dixit P, *A Survey: Optimization Algorithms In Deep Learning.*, InProceedings of the International Conference on Innovative Computing & Communications (ICICC) 2020 Mar 31

[12] Dubey S, Verma D, Kumar M, *Severe acute respiratory syndrome Coronavirus-2 GenoAnalyzer and mutagenic anomaly detector using FCMFI and NSCE.*, International Journal of Biological Macromolecules, 2024, 258, 129051. .

[13] Ullah K, Ahmed I, Ahmad M, Rahman AU, Nawaz M, Adnan, *A. Rotation invariant person tracker using top view*, Journal of Ambient Intelligence and Humanized Computing. 2019 Oct 4:1-7.

[14] Ahmed I, Ahmad M, Khan FA, Asif M. , *Comparison of deep-learning-based segmentation models: Using top view person images*, IEEE Access. 2020 Jul 23;8:136361-73 .

[15] Ahmed I, Ahmad M, Ahmad A, Jeon G, *IoT-based crowd monitoring system: Using SSD with transfer learning*, Computers & Electrical Engineering. 2021 Jul 1;93:107226.

Edited by: Manish Gupta
Special issue on: Recent Advancements in Machine Intelligence and Smart Systems
Received: Mar 16, 2024
Accepted: Apr 25, 2024