SinkrOn
JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Issue 4, October 2025
DOI : https://doi.org/10.33395/sinkron.v9i4.15323
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Comparative Analysis of DNA Sequence Alignment Algorithms in SARS-CoV-2

Edi[1)*], Robet[2)], Nurhayati Harahap[3)]
[1)] Department of Information System STMIK TIME, Medan, Indonesia
[2)] Department of Informatics STMIK TIME, Medan, Indonesia
[3)] Department of Midwifery Universitas Bunda Thamrin, Medan, Indonesia
[1)]edi.foe84@gmail.com, [2)]robet@stmik-time.ac.id, [3)]yanihrhp14@gmail.com

Submitted :Sep 8, 2025 | **Accepted** : Oct 18, 2025 | **Published** : Oct 22, 2025

Abstract: Sequence alignment is fundamental in bioinformatics, with Smith-Waterman (local) and Needleman-Wunsch (global) algorithms widely applied. However, comparative analyses on highly similar viral genomes such as SARS-CoV-2 remain scarce. This study systematically evaluated both algorithms using the first 5,000 nucleotides of two SARS-CoV-2 genomes (29,903 and 29,684 nt) under four parameter configurations: standard, low gap penalty, high gap penalty, and high match reward. Performance was assessed through alignment score, sequence identity, gap distribution, execution time, and parameter sensitivity. Both algorithms produced identical sequence identity (97.80%), with 4,943 matches out of 5,054 positions. Smith-Waterman consistently yielded higher alignment scores (12.6-112 points advantage), while Needleman-Wunsch was substantially faster (0.7752 vs 3.9014 s), showing 5.03 times greater computational efficiency. These findings indicate that both methods are reliable for highly similar viral sequences, with a trade-off between scoring precision and computational speed. This study provides the first parameter-sensitive comparison for full SARS-CoV02 genomes, emphasizing how parameter tuning can influence performance outcomes. A key limitation is that the analysis was restricted to the first 5,000 nucleotides, which may not capture variability across the complete genome.

Keywords: Sequence alignment; Needleman-Wunsch; Smith-Waterman; SARS-CoV-2; Bioinformatics

INTRODUCTION

Sequence alignment is a foundational and indispensable task in bioinformatics, serving as a primary method for arranging sequences of DNA, RNA, or protein to identify regions of similarity (Cho et al., 2020; Saloom & Khafaji, 2023). These similarities often imply functional, structural, or evolutionary relationships between the sequences (Khodja et al., 2024; Lall & Tallur, 2023; Lee et al., 2020), making alignment a critical step in understanding genetic variation, gene function, and molecular evolution. The advent of Next-Generation Sequencing (NGS) has revolutionized genomics by producing massive volumes of sequence data, which has intensified the demand for accurate and efficient alignment algorithms (Kim et al., 2020). In the context of global health, the SARS-CoV-2 pandemic has underscored the urgent need for robust genomic analysis tools. Aligning SARS-CoV-2 genome sequences is crucial for tracking viral mutations, developing effective diagnostics, and understanding the molecular mechanisms of the virus (Heng et al., 2021;Banjarnahor et al., 2022).

At the core of sequence comparison are two seminal algorithms based on dynamic programming: the Needleman-Wunsch (NW) algorithm for global alignment and the Smith-Waterman (SW) algorithm for local alignment (Kim et al., 2020; Lee et al., 2020; Saada et al., 2024). The Needleman-Wunsch algorithm, proposed in 1970, finds the optimal alignment across the entire length of two sequences, making it most suitable for comparing sequences that are closely related and of similar length (Khodja et al., 2024; Kim et al., 2020; Parvez et al., 2020). In contrast, the Smith-Waterman algorithm, introduced in 1981, identifies the most highly conserved sub-regions between two sequences (Penaloza et al., 2021). This local alignment approach is particularly effective for analyzing divergent sequences, sequences of different lengths, or for discovering conserved motifs within longer, otherwise dissimilar sequences (Cho et al., 2020; Dhar et al., 2024; Khodja et al., 2024). Both algorithms are celebrated for their mathematical rigor in guaranteeing an optimal solution (Lee et al., 2020; Rashed et al., 2021) but are also known for their quadratic time and space complexity, O(MN), which presents significant

*name of corresponding author

computational challenges when aligning long genomes (Kalemati et al., 2025; Naghibzadeh et al., 2021; Xu et al., 2020).

The computational demands of these algorithms have spurred extensive research into performance optimization (Dineshdarsi et al., 2023; Kim et al., 2020; Lee et al., 2020). Numerous studies have focused on accelerating execution through hardware platforms like Graphics Processing Units (GPUs) and Field-Programmable Gate Arrays (FPGAs) (Kyal et al., 2020; Lee et al., 2020; Rashed et al., 2021), or by developing faster heuristic alternatives such as BLAST and FASTA (Kim et al., 2020; Lee et al., 2020; Rashed et al., 2021). Several studies have applied NW and SW algorithms to analyze SARS-CoV-2, successfully identifying nucleotide misalignments indicative of mutations (Heng et al., 2021; Khodja et al., 2024). Other research has compared the suitability of these algorithms in different contexts, such as the analysis of industrial alarm floods, concluding that NW is better for sequences of similar length while SW is more appropriate for sequences of disparate lengths (Parvez et al., 2020). However, while these algorithms are widely used, there remains a research gap in the systematic comparative analysis of how their fundamental architectural differences global versus local and the fine-tuning of scoring parameters specifically impact the alignment results for the complete SARS-CoV-2 genome. Previous applications to SARS-CoV-2 have often been preliminary or focused on specific open reading frames (Heng et al., 2021) or regions such as the spike protein (Khodja et al., 2024), rather than a comprehensive, parameter-sensitive comparison.

This study aims to fill this gap by providing a direct and detailed comparative analysis of the Needleman-Wunsch and Smith-Waterman algorithms for the alignment of SARS-CoV-2 DNA sequences. The primary objectives of this research are to: (1) investigate how the fundamental differences between local (Smith-Waterman) and global (Needleman-Wunsch) alignment methodologies affect the identification of similarities and variations in SARS-CoV-2 DNA sequences, (2) analyze the influence of key scoring parameters specifically match scores, mismatch penalties, and gap penalties on the alignment results and computational performance of both algorithms when applied to a SARS-CoV-2 genome dataset, (3) determine the specific analytical contexts in which the Smith-Waterman algorithm is superior to the Needleman-Wunsch algorithm, or vice versa, for SARS-CoV-2 genomic analysis, thereby offering clear guidance for researchers in the field.

The novelty of this research lies in providing the first parameter-sensitive comparative evaluation of the Needleman-Wunsch and Smith-Waterman algorithms. This study systematically explores how variations in scoring parameters affect alignment accuracy, execution time, and sensitivity across different algorithmic frameworks. The findings are expected to offer practical guidance for researchers in selecting appropriate alignment strategies based on analytical goals, computational constraints, and genomic similarity levels, thereby advancing methodological precision in viral genome analysis.

METHOD

This section outlines the research methodology employed in conducting the comparative analysis of the Needleman-Wunsch and Smith-Waterman algorithms. The overall workflow is illustrated in figure 1.

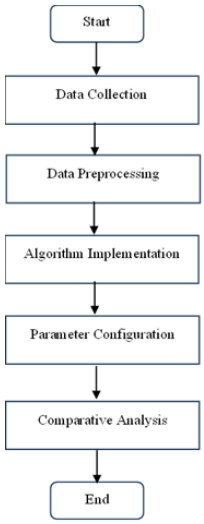

Fig. 1 Research methodology

Environment Setup

Table 1 present the hardware and software specifications utilized in this study.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Issue 4, October 2025
DOI : https://doi.org/10.33395/sinkron.v9i4.15323
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Table 1. Environment setup

Hardware/Software	Description
Visual Studio Code	Integrated development environment (IDE) for Python coding
Python 3.13.7	Sequence alignment implementation
NumPy 1.26 and Pandas 2.1	Numerical computation and data analysis
Matplotlib 3.8	Visualization
Biopython 1.83	Bioinformatics toolkit for sequence alignment

Data Collection

The reference sequence selection process begins with identifying the most appropriate genomic standard for SARS-CoV-2 comparative studies. NC_045512.2, representing the severe acute respiratory syndrome coronavirus 2 isolate Wuhan-Hu-1, serves as the primary reference genome due to its historical significance as the original sequenced SARS-CoV-2 genome and its widespread adoption in genomic research.

The comparative sequence selection focuses on temporal and geographic diversity to capture meaningful viral evolution. OL672836.1, designated as severe acute respiratory syndrome coronavirus 2 isolate SARS-CoV-2/human/BEL/rega-20174/2021, represents a Belgian isolate collected in 2021, providing both temporal separation (approximately two years) and geographic distance from the original Wuhan isolate. This sequence, containing 29,684 nucleotides, was obtained from the NCBI GenBank database and represents real-world viral evolution under selective pressure.

Data Preprocessing

To balance computational feasibility with biological relevance, both SARS-CoV-2 sequences (NC_045512.2 and OL672836.1) were truncated to 5,000 nucleotides from the natural 5′ terminus. This region encompasses the 5′ untranslated region (UTR), leader sequence, and the beginning of the ORF1ab polyprotein gene, which collectively represent the most functionally constrained and evolutionarily informative portion of the coronavirus genome.

The 5′ UTR contains critical regulatory elements including stem-loop structures, translation initiation signals, and transcription regulatory sequences (TRS) that govern viral replication, transcription, and translation efficiency. Any mutations in these regions are subject to strong purifying selection due to their direct impact on viral fitness. The adjacent ORF1ab region encodes early non-structural proteins (nsp1-nsp4 within the truncated segment) essential for viral replication machinery, including host response antagonists, viral proteases, and components of the replication-transcription complex. These early-expressed genes are functionally constrained and less prone to adaptive mutations compared to structural genes, making them ideal targets for detecting biologically meaningful sequence differences.

This region selection strategy minimizes noise from highly variable structural genes (such as spike protein) that are subject to immune selection pressure, while maximizing signal from functionally critical domains where sequence differences are more likely to reflect fundamental biological changes rather than neutral drift. The 5,000-nucleotide truncation provides sufficient sequence information for statistically robust alignment analysis while maintaining computational efficiency for dynamic programming algorithms.

Preliminary alignment of the truncated sequences revealed two consistent gap regions of 54 base pairs each at the termini (positions 0 and 5,000). These gap insertions increased the final processed sequence length to 5,054 nucleotides, ensuring equal-length alignment targets while preserving the complete regulatory and early coding sequences necessary for meaningful comparative genomic analysis.

Algorithm Implementation

Needleman-Wunsch Implementation

The Needleman-Wunsch algorithm is implemented to perform optimal global sequence alignment using dynamic programming. This global alignment approach seeks to find the best alignment across the entire length of both sequences, incorporating penalties for gaps at the sequence termini and throughout the alignment. The algorithm initializes a dynamic programming matrix where the first row and column are populated with cumulative gap penalties, reflecting the cost of aligning sequence prefixes with gaps. The matrix filling process evaluates three possible operations at each position: match/mismatch from the diagonal, insertion from the left cell, and deletion from the upper cell. The scoring system considers match rewards, mismatch penalties, gap opening costs, and gap extension penalties according to the specified parameter set. The algorithm guarantees finding the optimal global alignment by exhaustively evaluating all possible alignment paths and selecting the one with maximum cumulative score. The traceback procedure reconstructs the optimal alignment by following the path that produced the maximum score, generating the final aligned sequences with appropriate gap insertions. This implementation

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Issue 4, October 2025
DOI : https://doi.org/10.33395/sinkron.v9i4.15323
e-ISSN : 2541-2019
p-ISSN : 2541-044X

ensures reproducible results and provides comprehensive alignment coverage from the first to the last nucleotide of both sequences.

Smith-Waterman Implementation

The Smith-Waterman algorithm is implemented to identify optimal local sequence alignments within a larger genomic context. Unlike global alignment, the local approach seeks the highest-scoring subsequence alignment without requiring full-length coverage of both sequences. The algorithm initializes the dynamic programming matrix with zeros in the first row and column, allowing alignments to begin at any position within either sequence. A critical feature of the Smith-Waterman algorithm is the reset mechanism, where negative scores are automatically set to zero, preventing the extension of poor-quality alignment regions and allowing new high-scoring alignments to initiate at any matrix position. The scoring evaluation at each matrix cell considers the same three operations as Needleman-Wunsch (match/mismatch, insertion, deletion) but adds a fourth option of resetting to zero when all other options yield negative scores. This design enables the algorithm to identify and report only the most significant alignment regions while ignoring low-similarity flanking sequences. The traceback process begins from the highest-scoring matrix position and continues until reaching a zero-score cell, producing local alignments that represent the most conserved or similar regions between the two sequences. This implementation is particularly valuable for identifying functional domains or conserved motifs within divergent genomic sequences.

Parameter Configuration

The parameter set design encompasses four distinct scoring configurations that represent different biological scenarios and alignment stringencies commonly encountered in viral genomics. The parameter set design encompasses four distinct scoring configurations that represent different biological scenarios and alignment stringencies commonly encountered in viral genomics. The Standard parameter set (Match=+2, Mismatch=-1, Gap Open=-2, Gap Extend=-0.5) provides balanced scoring suitable for general genomic analysis, reflecting moderate penalties for mismatches and gaps while rewarding exact matches. The Low Gap Penalty configuration (Match=+1, Mismatch=-1, Gap Open=-1, Gap Extend=-0.1) creates a permissive environment for gap formation, which is appropriate for analyzing highly divergent sequences or regions with frequent insertions and deletions. The High Gap Penalty set (Match=+3, Mismatch=-2, Gap Open=-5, Gap Extend=-1) imposes stringent restrictions on gap formation, favoring alignments with minimal insertions or deletions, which is suitable for highly conserved genomic regions. The High Match Reward configuration (Match=+5, Mismatch=-4, Gap Open=-3, Gap Extend=-1) emphasizes exact sequence matching while heavily penalizing mismatches, creating conditions that favor high-identity alignments and are sensitive to sequence conservation. This systematic parameter design enables comprehensive evaluation of algorithm behavior across the spectrum of biological scenarios encountered in comparative viral genomics.

Comparative Analysis

Performance metrics evaluation encompasses four critical dimensions that capture both algorithmic efficiency and biological relevance of the alignment results. Alignment score analysis compares the optimal scores achieved by each algorithm under identical parameter conditions, providing direct quantitative assessment of algorithmic performance differences. Sequence identity calculation determines the percentage of exactly matching nucleotides in the final alignments, offering insights into the biological similarity detected by each approach. Execution time measurement captures computational performance in seconds, enabling assessment of algorithmic efficiency and scalability considerations for larger genomic datasets. Alignment length evaluation determines the total number of positions in the final aligned sequences, which differs between local and global approaches and affects interpretation of coverage and sensitivity.

RESULT

Table 2 presents the alignment score for both Needleman-Wunsch and Smith-Waterman algorithms under four different parameter configurations. Across all parameter sets, Smith-Waterman consistently produced slightly higher scores compared to Needleman-Wunsch.

Table 2. Alignment Score

Parameter Set	Algorithm	Score
Standard	Needleman-Wunsch	9826.0
	Smith-Waterman	9883.0
Low Gap Penalty	Needleman-Wunsch	4927.4

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Issue 4, October 2025
DOI : https://doi.org/10.33395/sinkron.v9i4.15323
e-ISSN : 2541-2019
p-ISSN : 2541-044X

	Smith-Waterman	4940.0
High Gap Penalty	Needleman-Wunsch	14707.0
	Smith-Waterman	14823.0
High Match Reward	Needleman-Wunsch	24591.0
	Smith-Waterman	24703.0

Figure 2 illustrates the comparative performance of both algorithms across the four parameter sets. The visualization highlights the score differences, showing that Smith-Waterman algorithm achieved higher alignment scores in all scenarios.

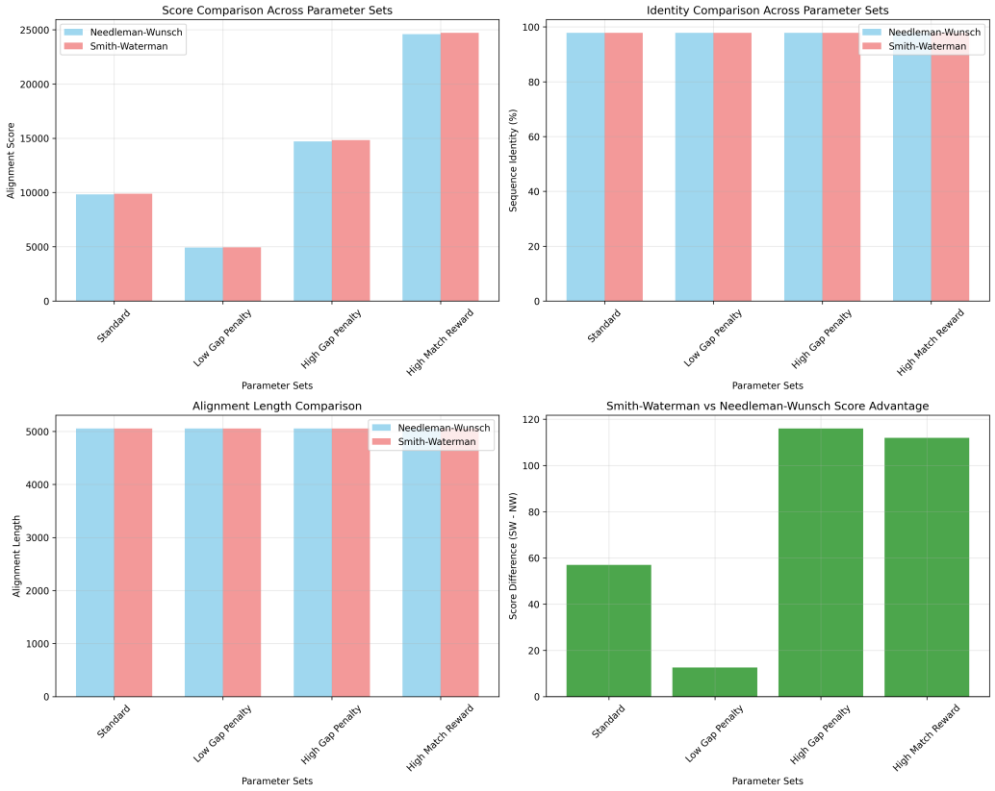

Fig. 2 Algorithm Performance Comparison Across Parameter Sets

Execution times for both algorithms are shown in figure 3. Needleman-Wunsch demonstrated faster runtime performance, while Smith-Waterman require substantially longer computation time under identical parameter conditions.

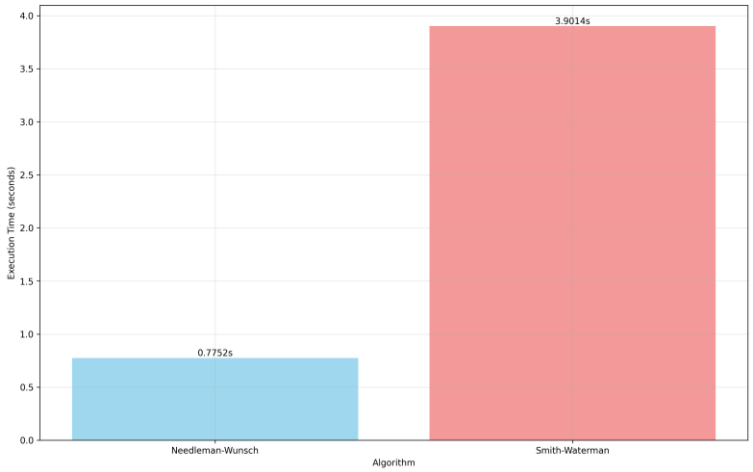

Fig. 3 Execution Time Performance Analysis

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Issue 4, October 2025
DOI : https://doi.org/10.33395/sinkron.v9i4.15323
e-ISSN : 2541-2019
p-ISSN : 2541-044X

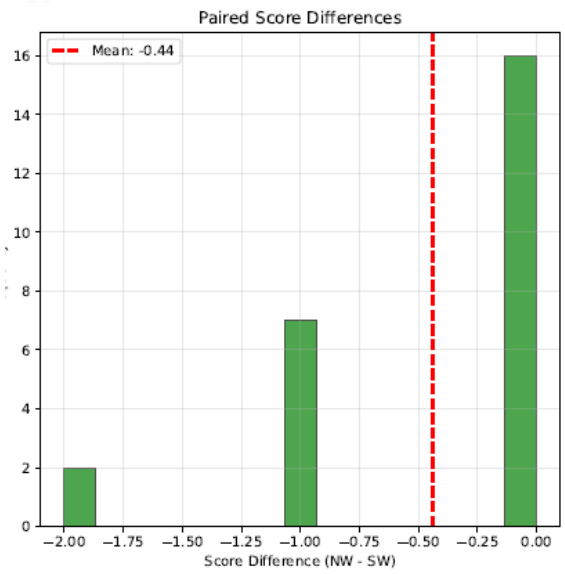

Fig. 4 Paired t-test between NW and SW

Paired t-test between NW and SW score are showed in figure 4. The results showed a statistically significant difference ($p = 0.002469$) with Smith-Waterman scoring slightly higher on average, though the effect size was very small (Cohen's $d = -0.0166$), indicating that while the difference is statistically detectable, it may not be practically meaningful for most applications.

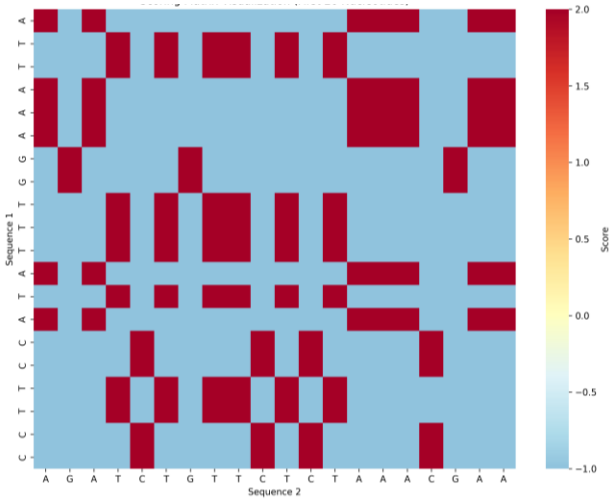

Fig. 5 Scoring matrix visualization (first 20 nucleotides)

Figure 5 provides a visualization of the scoring matrix for the first 20 nucleotides, showing how matches, mismatches, and gaps were scored during the alignment process. This matrix highlights the initial alignment regions used by both algorithms.

DISCUSSIONS

The comparative analysis of Smith-Waterman and Needleman-Wunsch algorithms on SARS-CoV-2 genome sequences highlights both their strengths and trade-offs. Both methods produced identical sequence identity (97.80%) and consistent gap patterns, showing that they capture the same biologically relevant features. However, Smith-Waterman consistently achieved slightly higher alignment scores across all parameter configurations, with advantages ranging from 12.6 to 116 points. These differences were most pronounced under stringent scoring conditions, suggesting that local alignment is particularly effective when penalties are high.

Despite its scoring advantage, Smith-Waterman required about five times longer to run compared to Needleman-Wunsch. This efficiency gap remained stable regardless of parameter settings, making Needleman-Wunsch a faster choice for large-scale analyses. In contrast, Smith-Waterman may be more useful when maximum

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Issue 4, October 2025
DOI : https://doi.org/10.33395/sinkron.v9i4.15323
e-ISSN : 2541-2019
p-ISSN : 2541-044X

alignment sensitivity is needed, especially for detecting conserved regions or optimizing under restrictive scoring schemes.

The findings are consistent with Heng et al. (2021), who reported that both algorithms yield highly similar alignment results for SARS-CoV-2 variants (Wuhan, Illinois, and India), but differ in computational demands. Our results also align with Khodja et al. (2024), who demonstrated that local alignment of the spike region using Biopython provides better viral family classification than global alignment. Together, these studies confirm that local alignment (Smith-Waterman) is superior for detecting fine-grained variations in critical regions, while global alignment (Needleman-Wunsch) remains efficient for full-length genome comparisons.

From a biological standpoint, the slightly higher local alignment scores observed for Smith-Waterman imply a stronger sensitivity in identifying subtle nucleotide substitutions within conserved functional domains particularly in the spike protein region, which governs host-cell receptor binding and immune response evasion. This suggests that when analyzing mutation hotspots like the spike gene, local alignment approaches may provide more information insights into variant evolution and transmissibility.

However, this study is limited to the first 5,000 nucleotides of closely related SARS-CoV-2 genomes rather than full-length (~30 kb) sequences. As shown by Khodja et al. (2024), local alignment accuracy and classification performance can vary when larger genomic contexts are included, especially in regions outside the spike domain. Expanding the analysis to entire genomes and more divergent variants would provide a more comprehensive understanding of algorithmic robustness across genomic scales.

CONCLUSION

This study shows that Smith-Waterman and Needleman-Wunsch produce identical biological accuracy for SARS-CoV-2 genome alignments, but differ in performance. Smith-Waterman consistently yields slightly higher alignment scores, while Needleman-Wunsch is over five times faster, making it more suitable for large-scale analyses. These results provide clear benchmarks to guide algorithm selection: Smith-Waterman for sensitivity, Needleman-Wunsch for efficiency. This study provides a benchmark framework for parameter-sensitive alignment evaluation.

While our analysis was limited to the first 5,000 nucleotides of closely related SARS-CoV-2 genomes and a few parameter sets, the findings highlight stable performance patterns across conditions. Future research should extend to full genome and integrate GPU-based acceleration, as well as test divergent sequences, other pathogens, and explore hybrid or optimized approaches that combine speed with sensitivity.

ACKNOWLEDGMENT

This research was funded by the Ministry of Higher Education, Science, and Technology.

REFERENCES

Banjarnahor, E., Bustamam, A., Siswantining, T., & Tampubolon, P. (2022). *Analyzing Kinship in Severe Acute Respiratory Syndrome Coronavirus 2 DNA Sequences Based on Hierarchical and K-Means Clustering Methods Using Multiple Encoding Vector*. *12*(6). https://www.gisaid.org.

Cho, H., Lim, S., Belenko, V., Kalinin, M., Zegzhda, D., & Nuralieva, E. (2020). Application and improvement of sequence alignment algorithms for intrusion detection in the Internet of Things. *Proceedings - 2020 IEEE Conference on Industrial Cyberphysical Systems, ICPS 2020*, 93–97. https://doi.org/10.1109/ICPS48405.2020.9274752

Dhar, A., Chauhan, I., Azad, H. K., Bamzai, A., & Yadav, T. (2024). Parallelization of the Smith-Waterman Algorithm to Accelerate DNA Sequence Alignment. *2024 3rd International Conference on Artificial Intelligence for Internet of Things, AIIoT 2024*. https://doi.org/10.1109/AIIoT58432.2024.10574609

Dineshdarsi, Rajesh, S., Krishna, P. J. S., & Sushma. (2023). Pairwise Sequence Alignment in Biological Sequences using Machine Learning. *2023 2nd International Conference on Advances in Computational Intelligence and Communication, ICACIC 2023*. https://doi.org/10.1109/ICACIC59454.2023.10435046

Heng, J. W., Juwono, F. H., & Reine, R. (2021a, July 7). Using Optimal Sequencing Algorithms for COVID-19 Case Study. *2021 International Conference on Green Energy, Computing and Sustainable Technology, GECOST 2021*. https://doi.org/10.1109/GECOST52368.2021.9538762

Heng, J. W., Juwono, F. H., & Reine, R. (2021b, July 7). Using Optimal Sequencing Algorithms for COVID-19 Case Study. *2021 International Conference on Green Energy, Computing and Sustainable Technology, GECOST 2021*. https://doi.org/10.1109/GECOST52368.2021.9538762

Kalemati, M., Nayeri, A. D., & Koohi, S. (2025). Pip-SW: Pipeline Architectures for Accelerating Smith-Waterman Algorithm on FPGA Platforms. *IEEE Transactions on Emerging Topics in Computing*, *13*(3), 628–639. https://doi.org/10.1109/TETC.2024.3472649

*name of corresponding author

Khodja, M., Melouki, A., Nabil, B., & Abdelouahab, D. (2024). Genomic analysis of SARS-COV2 using Biopython and Simplot comparison. *Journal of Bioscience and Applied Research*, *10*(6), 42–53. https://doi.org/10.21608/jbaar.2024.332656.1098

Kim, J., Ji, M., & Yi, G. (2020). A review on sequence alignment algorithms for short reads based on next-generation sequencing. *IEEE Access*, *8*, 189811–189822. https://doi.org/10.1109/ACCESS.2020.3031159

Kyal, C., Kumar, R., & Zamal, A. (2020, December 10). Performance-Based Analogising of Needleman Wunsch Algorithm to Align DNA Sequences Using GPU and FPGA. *2020 IEEE 17th India Council International Conference, INDICON 2020*. https://doi.org/10.1109/INDICON49873.2020.9342078

Lall, A., & Tallur, S. (2023). Deep reinforcement learning-based pairwise DNA sequence alignment method compatible with embedded edge devices. *Scientific Reports*, *13*(1). https://doi.org/10.1038/s41598-023-29277-6

Lee, Y. S., Kim, Y. S., & Uy, R. L. (2020). Serial and parallel implementation of Needleman-Wunsch algorithm. *International Journal of Advances in Intelligent Informatics*, *6*(1), 97–108. https://doi.org/10.26555/ijain.v6i1.361

Naghibzadeh, M., Babaei, S., Behkmal, B., & Hatami, M. (2021). Divide and Conquer Approach to Long Genomic Sequence Alignment. *ICCKE 2021 - 11th International Conference on Computer Engineering and Knowledge*, 399–405. https://doi.org/10.1109/ICCKE54056.2021.9721501

Parvez, M. R., Hu, W., & Chen, T. (2020, November 9). Comparison of the smith-waterman and needleman-wunsch algorithms for online similarity analysis of industrial alarm floods. *2020 IEEE Electric Power and Energy Conference, EPEC 2020*. https://doi.org/10.1109/EPEC48502.2020.9320080

Penaloza, R. M., Salgado, G. R., & Salazar, A. M. (2021). Optimization of a classical algorithm for the alignment of genomic sequences with artificial bee colony. *Proceedings - 2021 International Conference on Computational Science and Computational Intelligence, CSCI 2021*, 355–360. https://doi.org/10.1109/CSCI54926.2021.00131

Rashed, A. E. E. D., Amer, H. M., El-Seddek, M., & Moustafa, H. E. D. (2021). Sequence Alignment Using Machine Learning-Based Needleman-Wunsch Algorithm. *IEEE Access*, *9*, 109522–109535. https://doi.org/10.1109/ACCESS.2021.3100408

Saada, B., Zhang, T., Siga, E., Zhang, J., & Magalhães Muniz, M. M. (2024). Whole-Genome Alignment: Methods, Challenges, and Future Directions. In *Applied Sciences (Switzerland)* (Vol. 14, Issue 11). Multidisciplinary Digital Publishing Institute (MDPI). https://doi.org/10.3390/app14114837

Saloom, R. H., & Khafaji, H. K. (2023). Developing New Pairwise Sequence Alignment Method Based on Needleman-Wunsch Algorithm. *International Journal of Intelligent Engineering and Systems*, *16*(2), 580–590. https://doi.org/10.22266/ijies2023.0430.48

Xu, X., Chan, Y., Xu, K., Zhang, J., Wang, X., Yin, Z., & Liu, W. (2020). SLPal: Accelerating Long Sequence Alignment on Many-Core and Multi-Core Architectures. *Proceedings - 2020 IEEE International Conference on Bioinformatics and Biomedicine, BIBM 2020*, 2242–2249. https://doi.org/10.1109/BIBM49941.2020.9313429

*name of corresponding author