

Review of Bioinformatics Methods for Phylogenetic Analysis of RNA Viruses

Sooka Agapito*

Department of Health Sciences, Magna Graecia University, Catanzaro, Italy

Introduction

The study of RNA viruses is crucial due to their rapid evolution, high mutation rates, and significant impact on human and animal health. Understanding their evolutionary relationships helps in tracking virus origins, predicting outbreaks, and designing effective vaccines and treatments. Phylogenetic analysis, a key tool in studying the evolution of RNA viruses, involves reconstructing evolutionary trees based on genetic sequences. With advances in bioinformatics, numerous computational techniques have emerged to facilitate phylogenetic studies, allowing researchers to analyze large datasets efficiently and accurately. This report reviews bioinformatics methods for RNA virus phylogenetics, highlighting key approaches, tools, challenges, and future directions.

Description

Phylogenetic analysis of RNA viruses typically begins with the collection of viral genome sequences from publicly available databases such as GenBank, GISAID, and ViPR. These databases store thousands of viral sequences collected from different hosts, geographic locations, and time points. Once the sequences are retrieved, preprocessing steps such as sequence alignment and quality control are performed to ensure accurate analysis. Multiple Sequence Alignment (MSA) is a critical step in phylogenetic analysis, as it arranges sequences in a way that reflects evolutionary relationships. Common alignment tools include Clustal Omega, MUSCLE, and MAFFT, each offering different advantages in terms of speed and accuracy. After alignment, phylogenetic tree construction methods are applied to infer evolutionary relationships among RNA virus strains. Traditional methods include distance-based, maximum likelihood, and Bayesian approaches. Distance-based methods, such as the neighbor-joining algorithm, estimate evolutionary distances between sequences and build trees based on these distances. While computationally efficient, they may oversimplify complex evolutionary relationships. Maximum likelihood methods, implemented in software such as RAxML and PhyML, provide more accurate tree estimations by evaluating multiple tree topologies and selecting the one that best fits the data. Bayesian inference, as used in MrBayes and BEAST, incorporates probabilistic models to assess uncertainty in phylogenetic trees, making it particularly useful for analyzing viral evolution over time [1].

One of the major challenges in RNA virus phylogenetics is dealing with their high mutation rates and recombination events. RNA viruses, such as influenza, HIV, and coronaviruses, evolve rapidly, leading to genetic diversity that complicates tree reconstruction. Recombination, where different viral strains exchange genetic material, can result in misleading phylogenetic signals. To address this, specialized bioinformatics tools such as RDP, GARD, and SimPlot are used to detect and correct for recombination before tree inference. Additionally, selection pressure analysis using tools like HyPhy and

PAML helps identify regions of the viral genome that are undergoing positive selection, which may indicate adaptive evolution in response to host immune pressure. Another important aspect of RNA virus phylogenetics is molecular clock analysis, which estimates the timing of evolutionary events based on sequence divergence. Molecular clocks assume that mutations accumulate at a constant rate, allowing researchers to infer when specific viral strains emerged or when cross-species transmission occurred. Bayesian frameworks, such as BEAST, are widely used for molecular clock dating, providing insights into the origins and spread of viral outbreaks. This approach has been instrumental in tracing the emergence of major pandemics, including HIV, Ebola, and SARS-CoV-2 [2].

Phylogenetic analysis is also essential for epidemiological surveillance and outbreak tracking. Real-time phylogenetics, enabled by tools like Nextstrain and ViPR, allows researchers to monitor viral evolution as new sequences become available. These platforms integrate phylogenetic trees with geographic and temporal data, providing interactive visualizations that help identify emerging variants and transmission patterns. The ability to track RNA virus evolution in near real time has been particularly valuable during the COVID-19 pandemic, where phylogenetics played a crucial role in identifying variants of concern and assessing their potential impact on public health. Despite advancements in bioinformatics methods, challenges remain in RNA virus phylogenetics. One issue is the incomplete sampling of viral diversity, as sequencing efforts often focus on specific geographic regions or host populations, leading to biases in phylogenetic reconstructions. Additionally, computational limitations can arise when analyzing large datasets, as phylogenetic methods require significant processing power and memory. Parallel computing and cloud-based solutions, such as Google Colab and AWS, are increasingly being used to overcome these challenges by enabling large-scale phylogenetic analyses [3].

Another emerging area in RNA virus phylogenetics is the integration of machine learning techniques. Machine learning algorithms are being explored to predict viral evolution, classify strains based on genetic features, and improve phylogenetic tree accuracy. These approaches leverage large-scale sequence datasets and evolutionary models to identify patterns that may not be apparent through traditional phylogenetic methods. As artificial intelligence continues to advance, its integration with bioinformatics is expected to enhance our ability to study RNA virus evolution more efficiently. Future directions in RNA virus phylogenetics involve improving sequencing technologies, developing more accurate evolutionary models, and enhancing computational efficiency. Third-generation sequencing technologies, such as nanopore sequencing, are enabling real-time viral genome sequencing with high accuracy, providing new opportunities for rapid phylogenetic analysis. Additionally, refining evolutionary models to account for complex mutation patterns and host interactions will improve the accuracy of phylogenetic inferences. The development of user-friendly bioinformatics platforms will also make advanced phylogenetic tools more accessible to researchers worldwide, facilitating global collaboration in viral surveillance [4,5].

Conclusion

Overall, bioinformatics methods have revolutionized RNA virus phylogenetics by enabling rapid, large-scale analyses of viral genomes. From sequence alignment and tree reconstruction to recombination detection and molecular clock dating, these methods provide crucial insights into virus evolution, epidemiology, and outbreak dynamics. As new computational tools and sequencing technologies emerge, the field of RNA virus phylogenetics will continue to advance, improving our ability to monitor and control viral

*Address for Correspondence: Sooka Agapito, Department of Health Sciences, Magna Graecia University, Catanzaro, Italy, E-mail: agapitosoo@gmail.com
Copyright: © 2025 Agapito S. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Received: 02 January, 2025, Manuscript No. IJdr-25-163395; Editor Assigned: 04 January, 2025, Pre QC No. P-163395; Reviewed: 17 January, 2025, QC No. Q-163395; Revised: 23 January, 2025, Manuscript No. R-163395; Published: 31 January, 2025, DOI: 10.37421/2277-1506.2025.14.488

diseases. The integration of real-time phylogenetics, machine learning, and high-throughput sequencing will further enhance our understanding of RNA viruses, ultimately contributing to better public health strategies and pandemic preparedness.

Acknowledgement

None.

Conflict of Interest

None.

References

1. Garbuglia, Anna Rosa, Silvia Pauciullo, Verdiana Zulian and Paola Del Porto. "Update on Hepatitis C Vaccine: Results and Challenges." *Viruses* 16 (2024): 1337.
2. Feng, Chunyu, Yuting Liu, Guangqi Lyu and Songyang Shang, et al. "Adaptive Evolution of the Fox Coronavirus Based on Genome-Wide Sequence Analysis." *Biomed Res Int* 2022 (2022): 9627961.
3. Gorbalenya, Alexander E. and Chris Lauber. "Bioinformatics of virus taxonomy: Foundations and tools for developing sequence-based hierarchical classification." *Curr Opin Virol* 52 (2022): 48-56.
4. Hamim, Islam, Syun-ichi Urayama, Osamu Netsu and Akemi Tanaka, et al. "Discovery, genomic sequence characterization and phylogenetic analysis of novel RNA viruses in the turfgrass pathogenic *Colletotrichum* spp. in Japan." *Viruses* 14 (2022): 2572.
5. Charon, Justine, Jan P. Buchmann, Sabrina Sadiq and Edward C. Holmes. "RdRp-scan: A bioinformatic resource to identify and annotate divergent RNA viruses in metagenomic sequence data." *Virus Evol* 8 (2022): veac082.

How to cite this article: Agapito, Sooka. "Review of Bioinformatics Methods for Phylogenetic Analysis of RNA Viruses." *Int J Drug Res Tech* 14 (2025): 488.