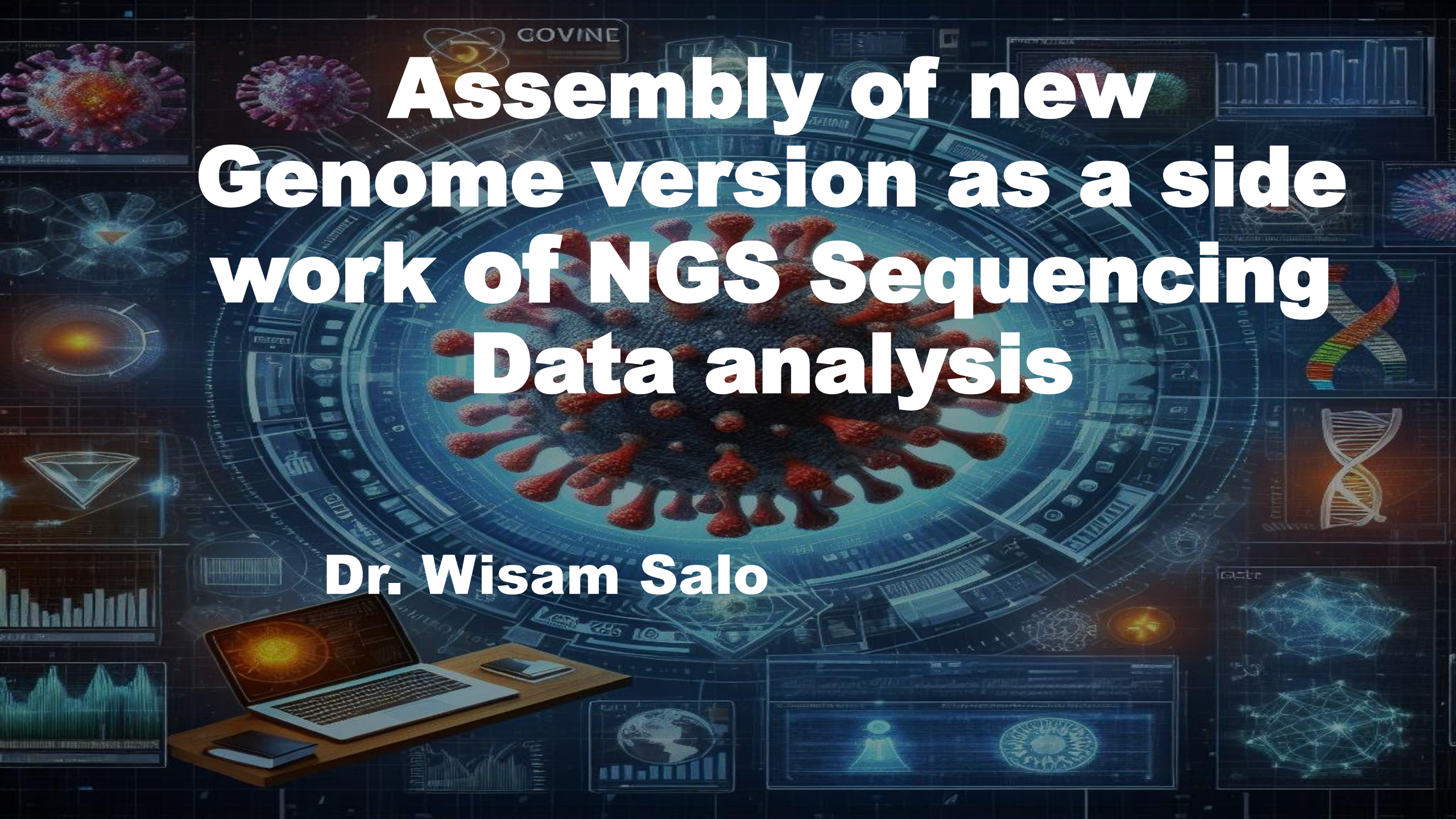
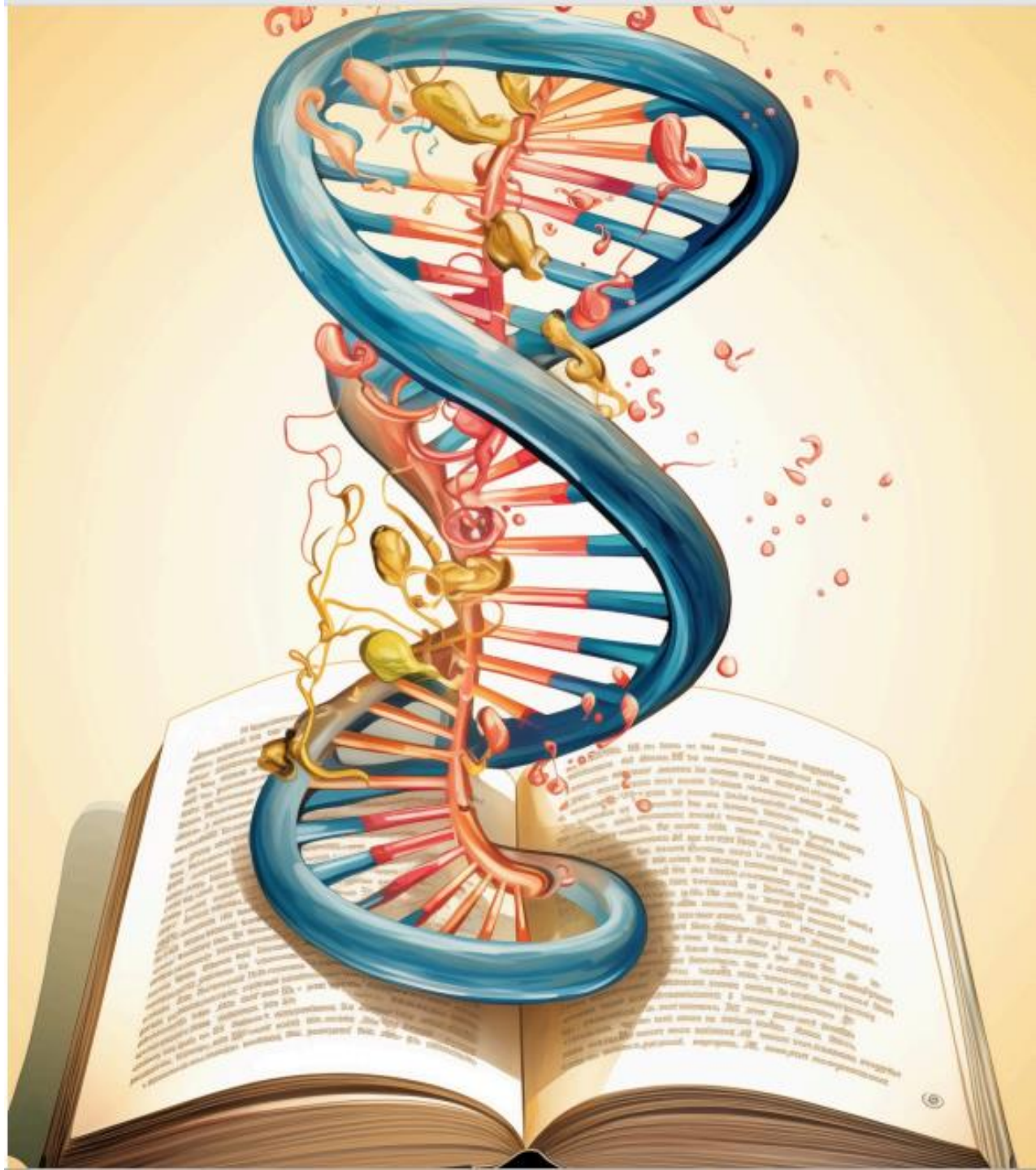


Assembly of new Genome version as a side work of NGS Sequencing Data analysis

Dr. Wisam Salo

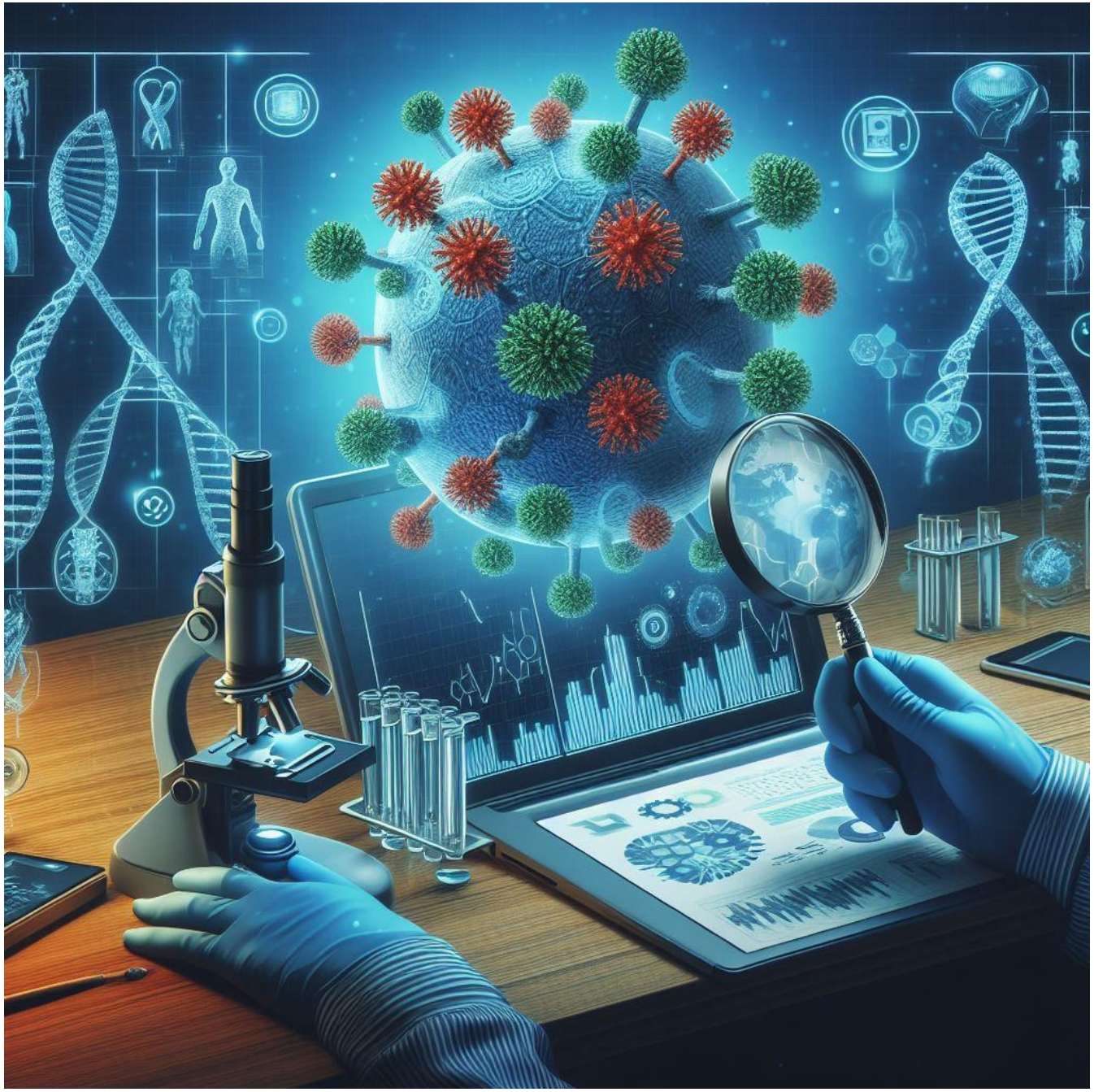
Introduction

- The genetic code of viruses plays a crucial role in understanding their pathogenicity and evolution.
- Viral genome analysis provides a window into the evolutionary history and genetic diversity of viruses.
- Understanding the genetic variability of viral strains informs the design of effective vaccines that target conserved regions of the viral genome.



NGS in viral genome

NGS has revolutionized the field of genomics, enabling high-throughput sequencing of DNA. This technology allows for the rapid and cost-effective sequencing of entire viral genomes, providing valuable insights into viral diversity and evolution.



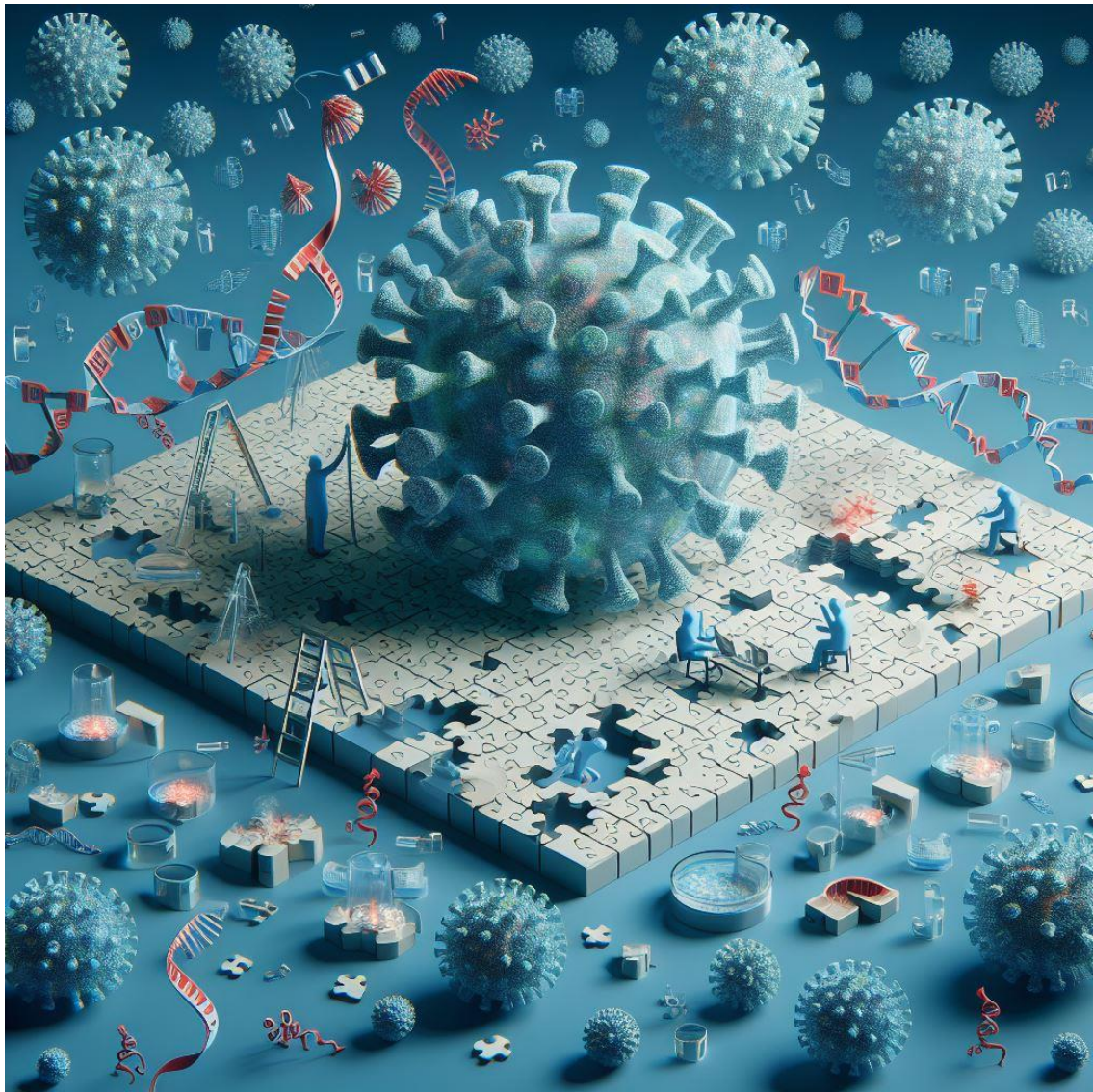
Viral Genome Assembly

The process of viral genome assembly involves reconstructing the complete viral genome from short sequencing reads. Advanced bioinformatics tools and algorithms are utilized to align and assemble the fragmented genetic information into a coherent genome.

Challenges in Viral Genome Assembly

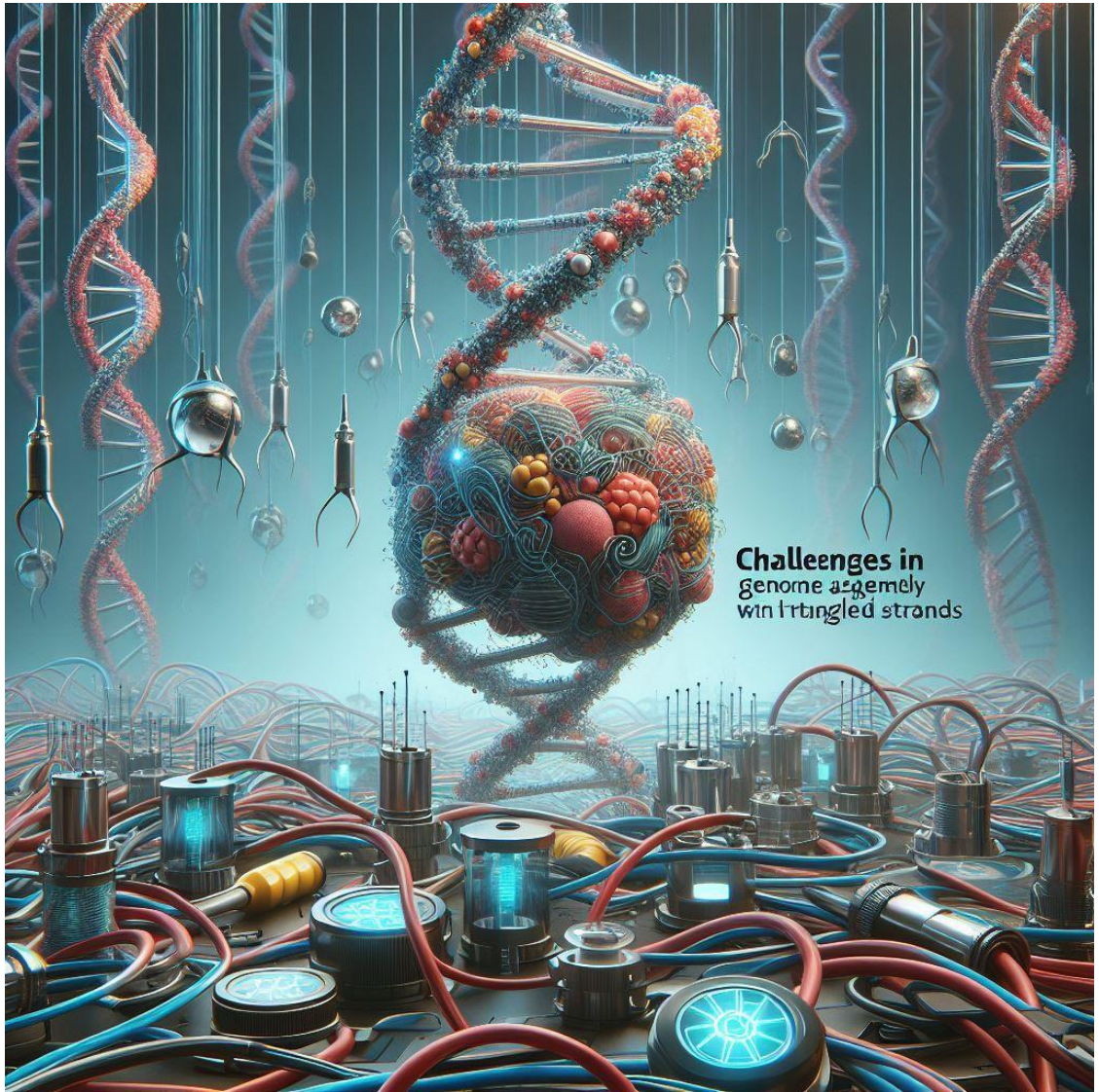
Genomic Heterogeneity: Viral populations often exhibit genetic diversity, including point mutations, insertions, deletions, and genomic rearrangements, which complicate sequence alignment and variant calling.

Sample Contamination and Bias: Contamination from host DNA, environmental microbes, or laboratory reagents can introduce bias and artifacts into sequencing data, leading to erroneous interpretations.



Challenges in Viral Genome Assembly

Data Interpretation and Integration:
Integrating genomic data with clinical, epidemiological, and experimental metadata requires sophisticated bioinformatics tools and interdisciplinary collaboration to extract meaningful insights from complex datasets.



Challenges in Viral Genome Assembly

- The host contamination is the major problem in viral genomic analysis, since Its almost impossible to get pure viral genome for sequencing

- This problem could be tackled if a
- new approach of genomics studies

scientific reports

3): 265-276
49 (online)
s reserved.

Check for updates

REVIEW

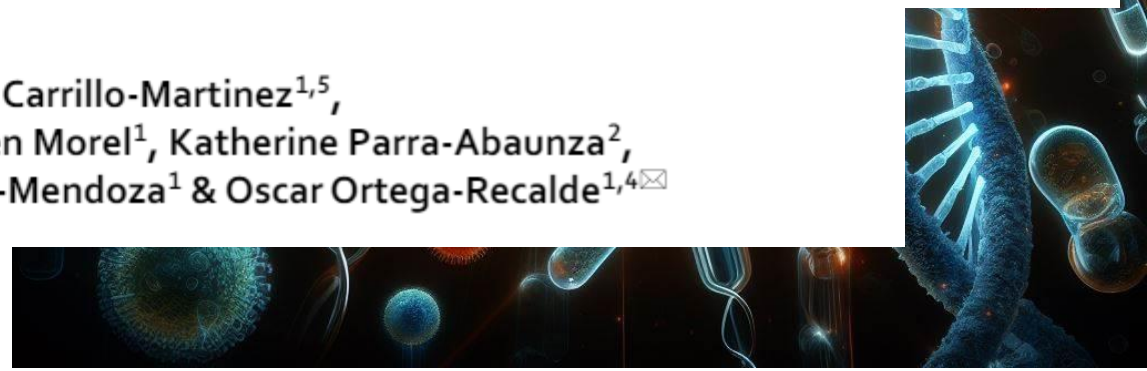
very

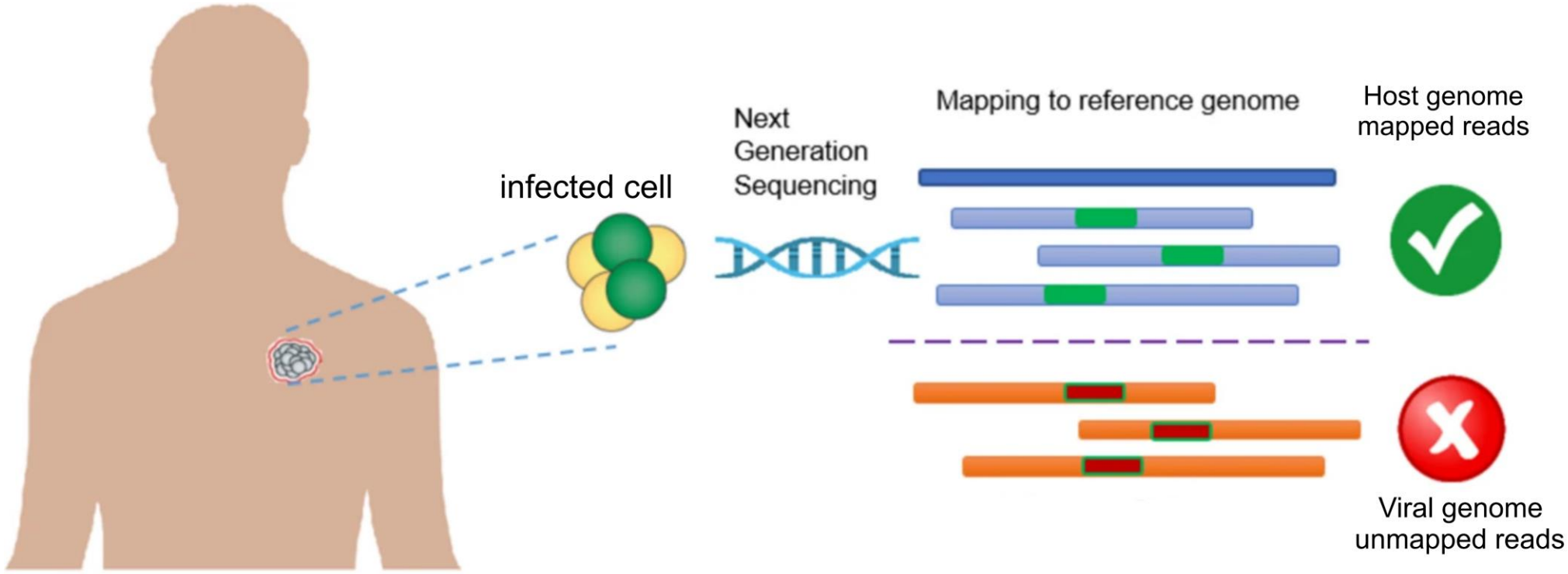
/eer

OPEN

Next-generation sequencing of host genetics risk factors associated with COVID-19 severity and long-COVID in Colombian population

Mariana Angulo-Aguado[1,5], Juan Camilo Carrillo-Martinez[1,5], Nora Constanza Contreras-Bravo[1], Adrien Morel[1], Katherine Parra-Abaunza[2], William Usaquén[3], Dora Janeth Fonseca-Mendoza[1] & Oscar Ortega-Recalde[1,4]✉

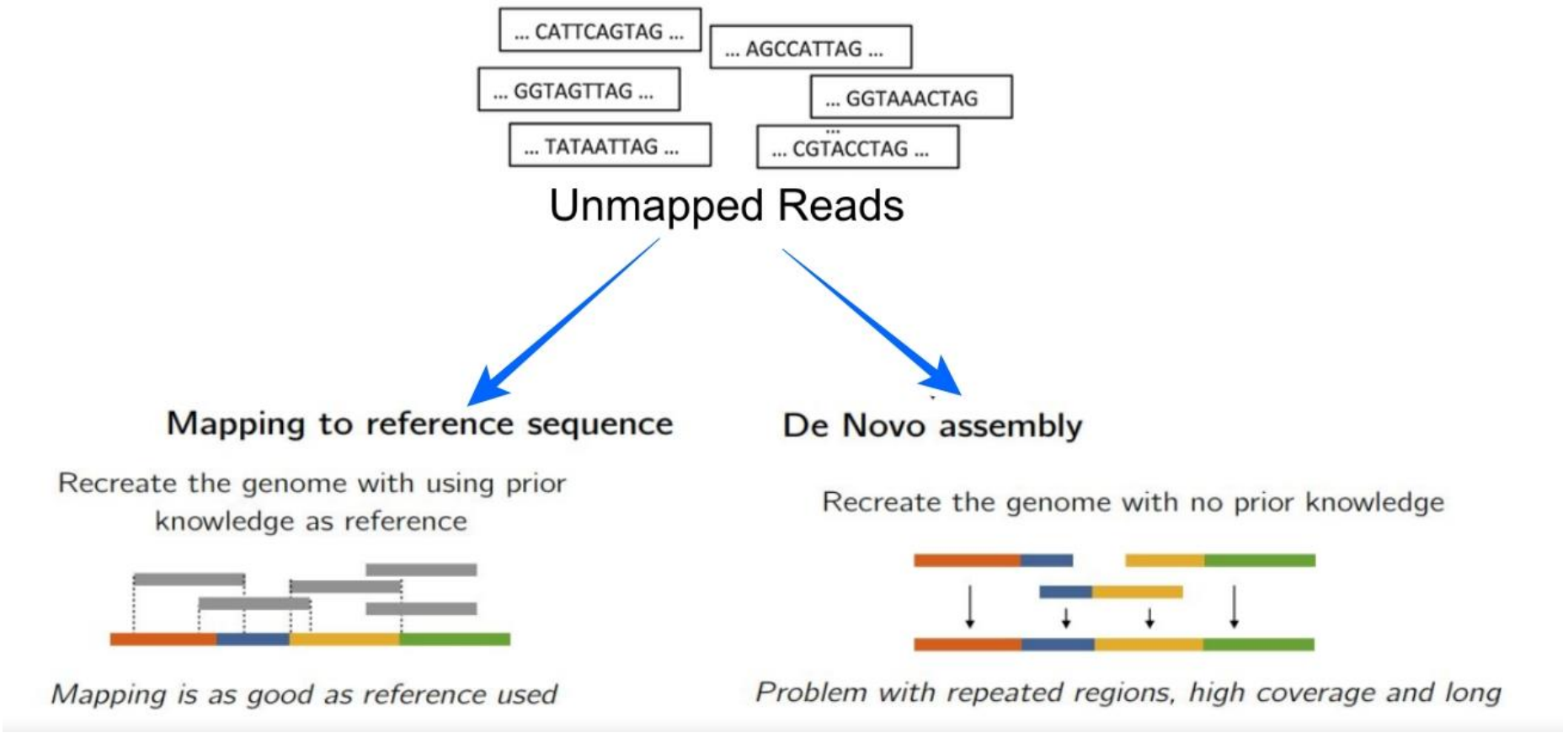




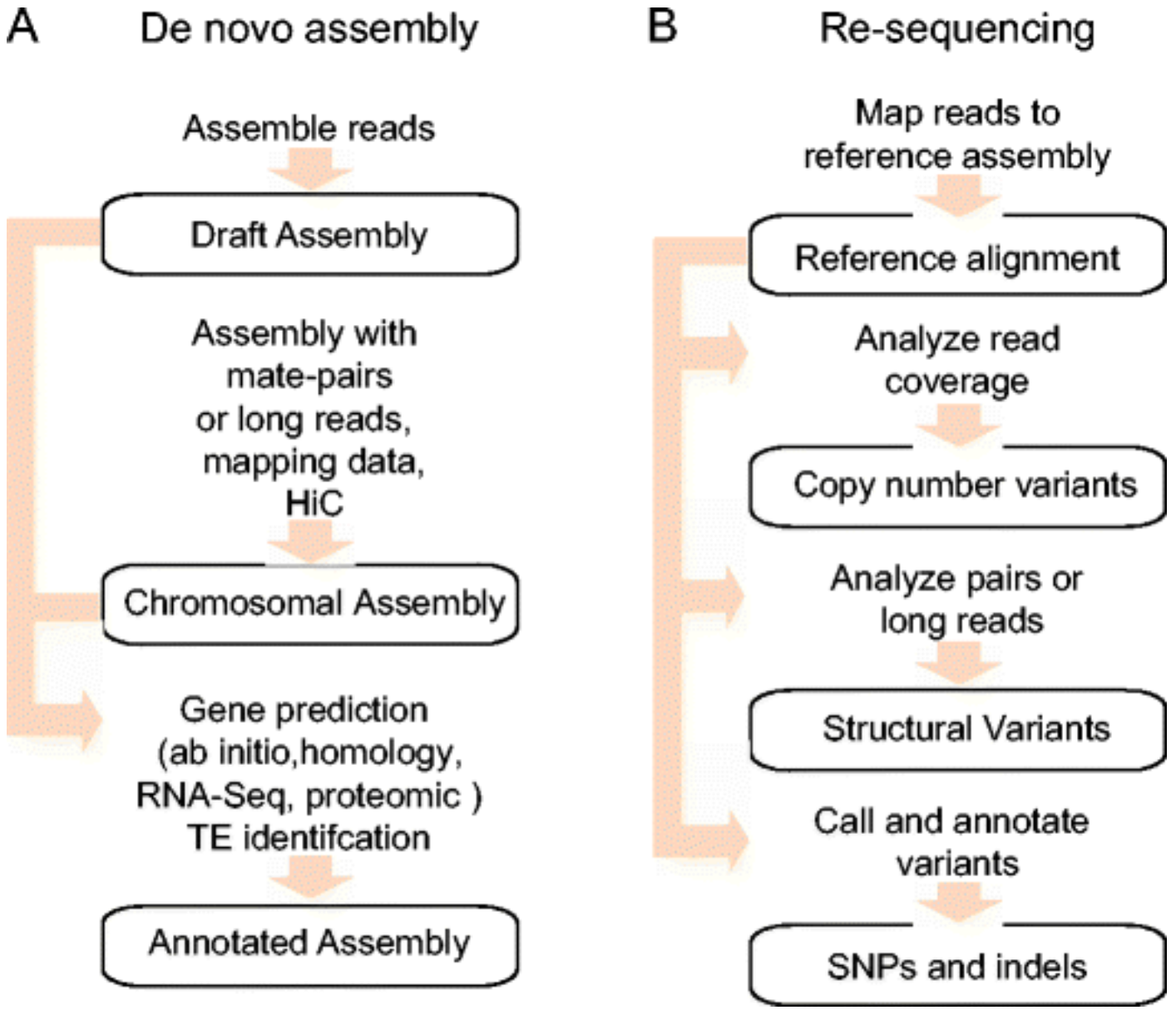
- Normally, infected samples has both the host genome and the viral genome
- Researchers, normally, mapping the host Genome reads to a reference genomes, usually the unmapped reads would be generated as a side data file

Unmapped reads Analysis

- There are two way for unmapped reads analysis with bioinformatics tool



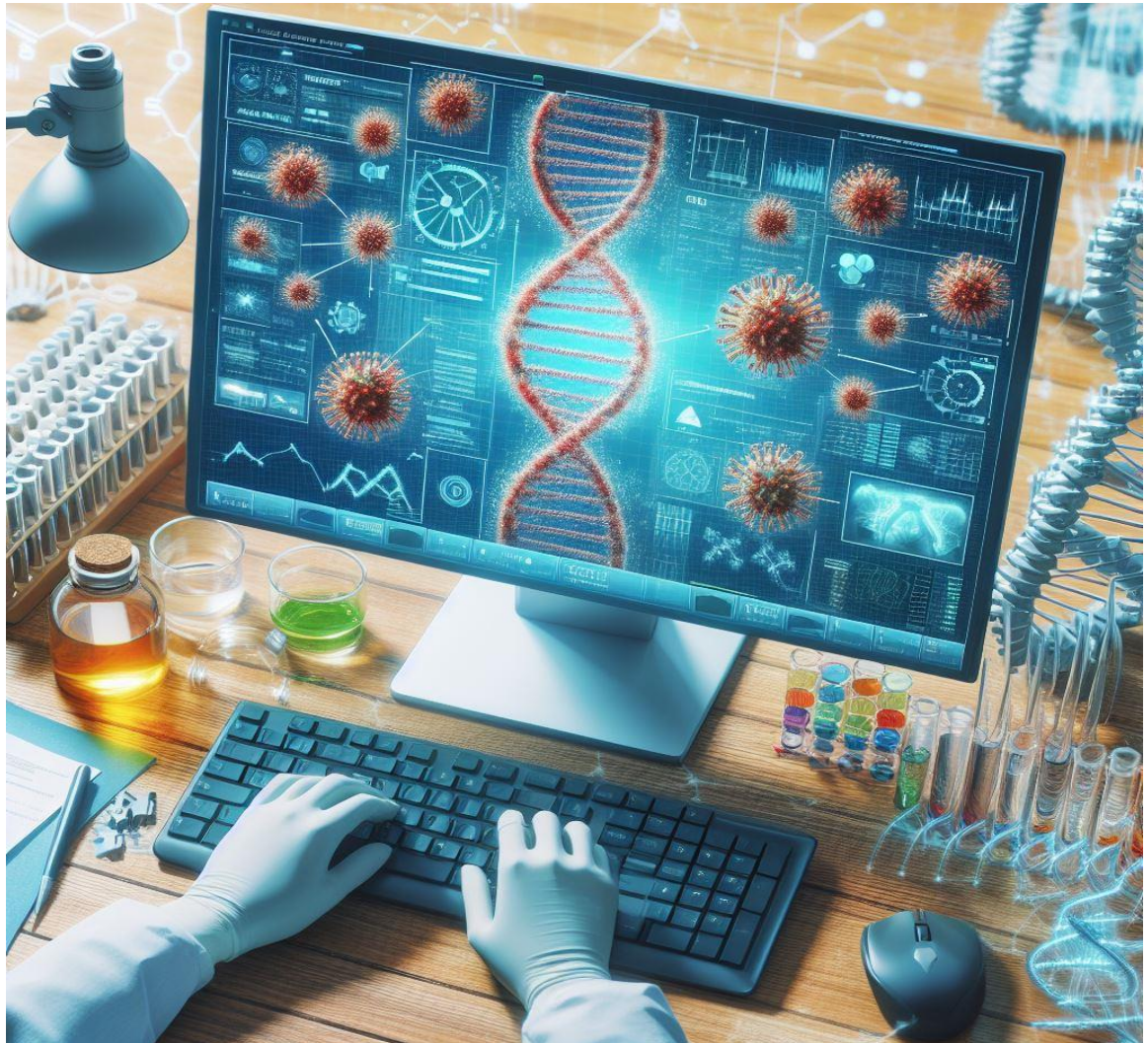
Unmapped reads analysis



Bioinformatics Tools and Resources:

A plethora of bioinformatics tools and software packages are available for viral genome

- Assembly such as SPAdes, Velvet, and IDBA.
- Alignment and mapping such as Bowtie, BWA, minimap2, those were the best aligner algorithms with short reads
- Variant calling, and annotation: GATK, and ViralSeq, as well as specialized pipelines and workflows developed by research groups and bioinformatics communities.



Viral Database

Databases and repositories such as GenBank, GISAID, and NCBI Virus provide access to annotated viral genomes, raw sequencing data, and metadata, facilitating data sharing, collaboration, and comparative analysis.

Quality Control

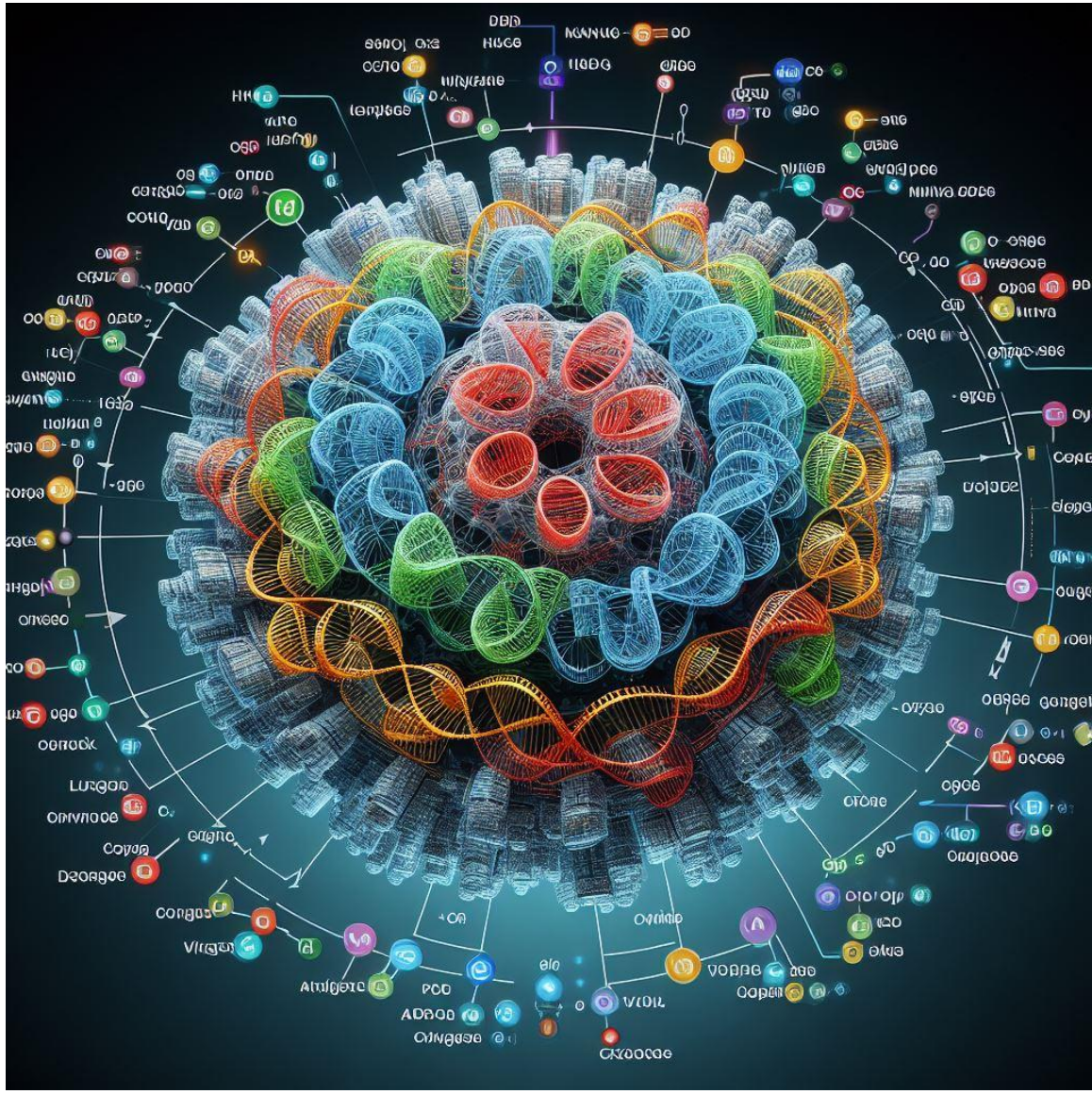
- Quality control measures involve assessing sequencing depth, coverage uniformity, and identifying potential sequencing errors or artifacts that may impact the accuracy of the assembled genome

Viral Genome Annotation

- Viral Genome Annotation

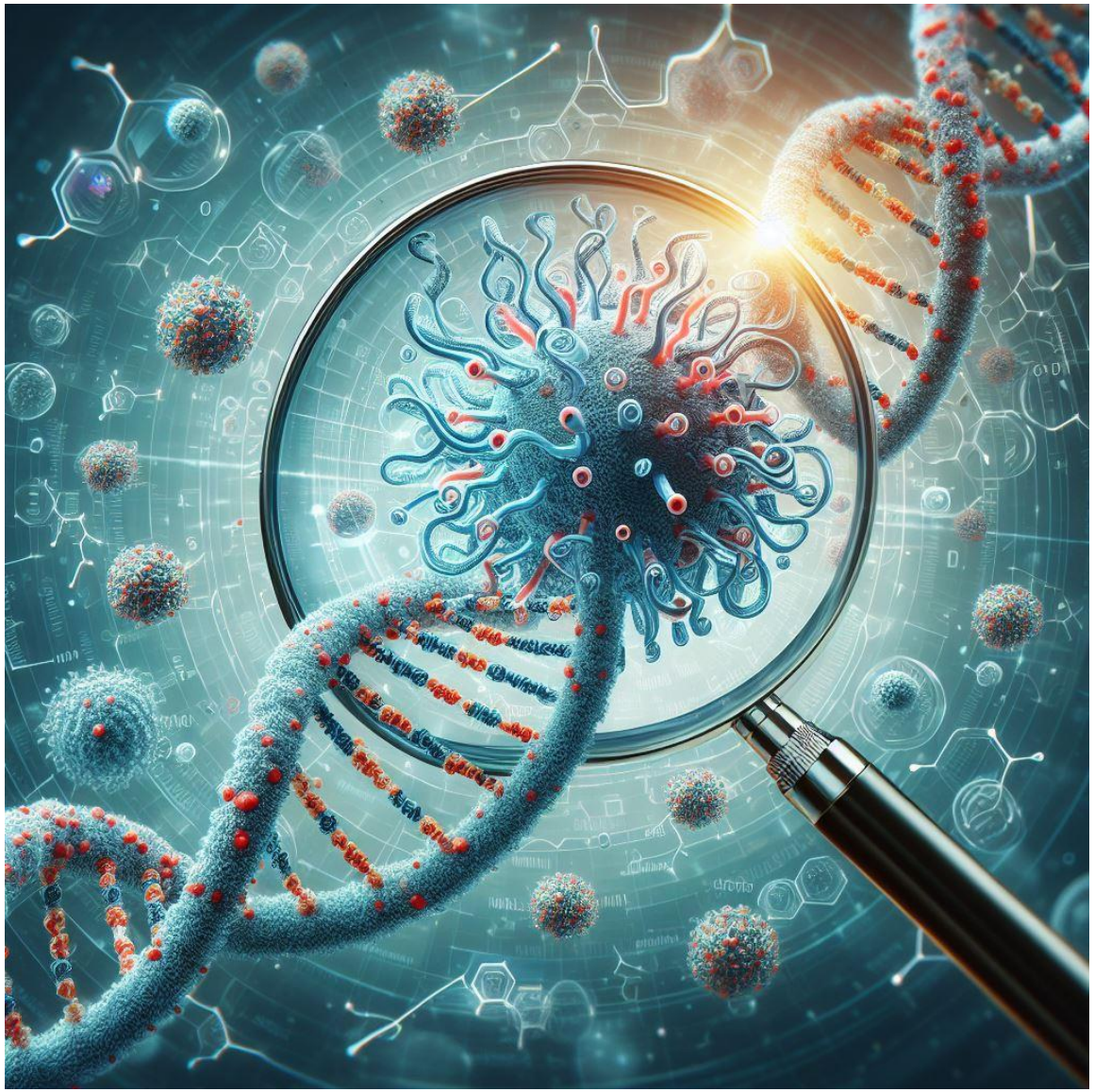
Genome annotation involves identifying genes, regulatory elements, and functional regions within the viral genome.

Accurate annotation provides valuable insights into viral gene expression, protein function, and evolutionary relationships.



Functional Analysis

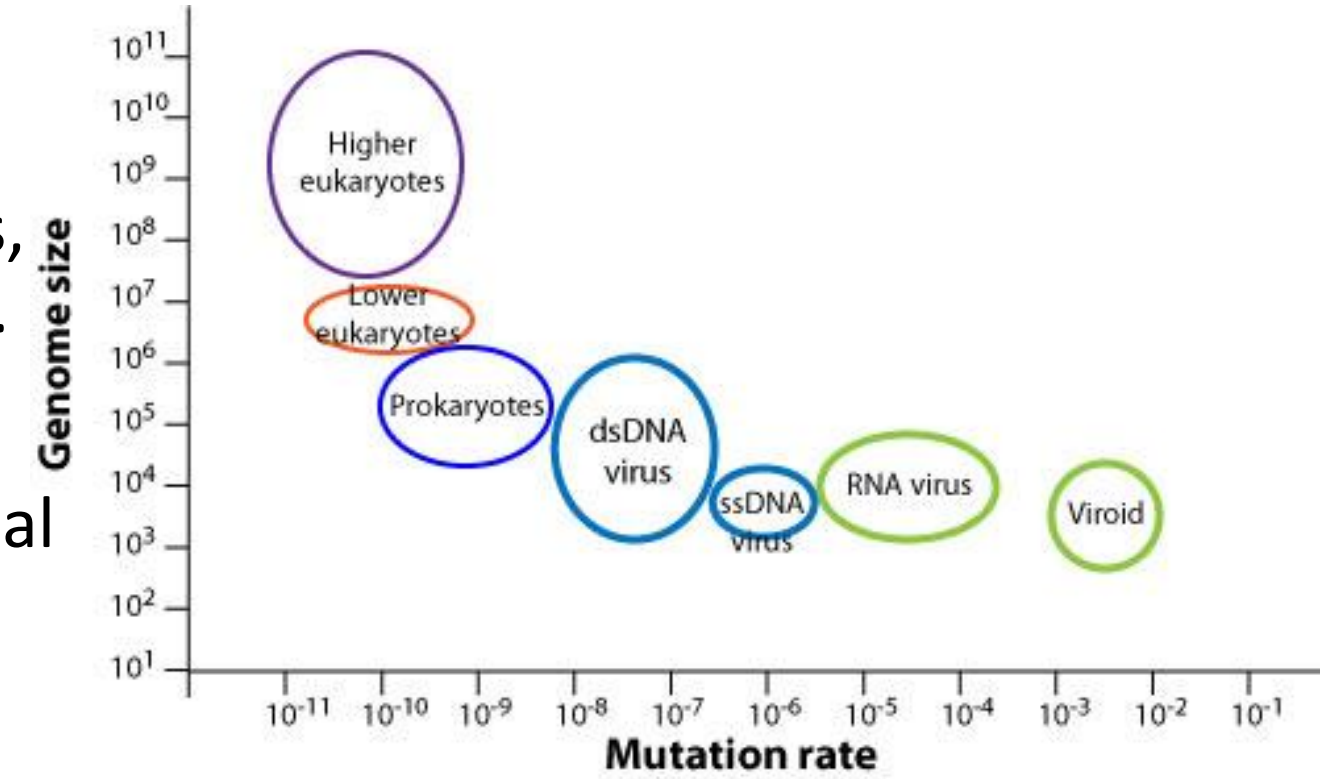
Functional analysis of viral genomes involves predicting the biological functions of encoded proteins and understanding the mechanisms underlying viral pathogenicity, host interaction, and immune evasion. This analysis aids in the development of antiviral strategies and vaccines.



Applications of Viral Genome Analysis:

- Studying viral genome sequences provides valuable insights into viral evolutionary dynamics, including mutation rates, recombination events, and adaptation to host environments.

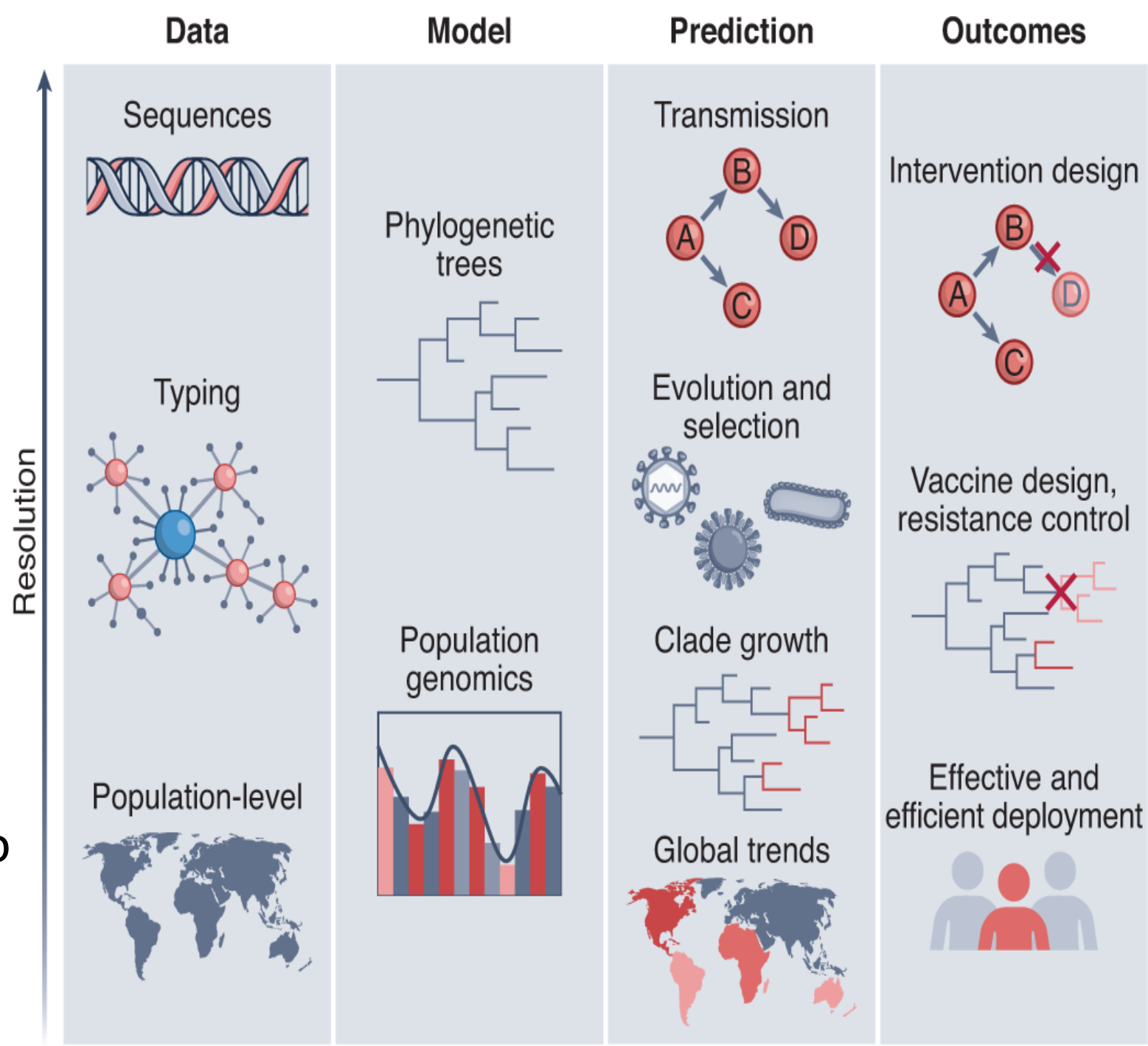
- Understanding viral evolution is crucial for managing viral outbreaks and designing effective control strategies.



viral genome analysis applications

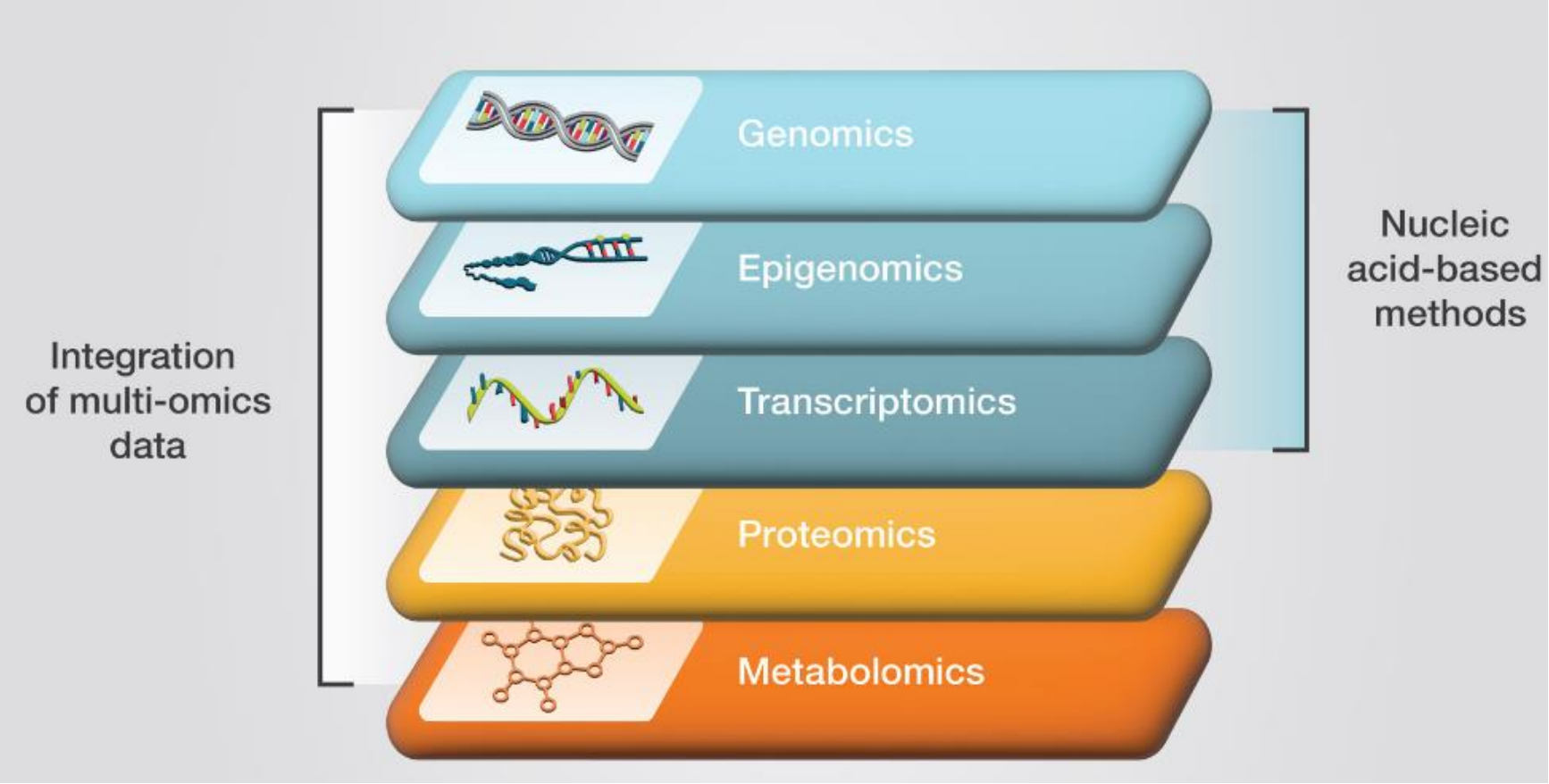
1- epidemiological surveillance,
2- outbreak investigation
3- vaccine development
4- understanding viral transmission dynamics.

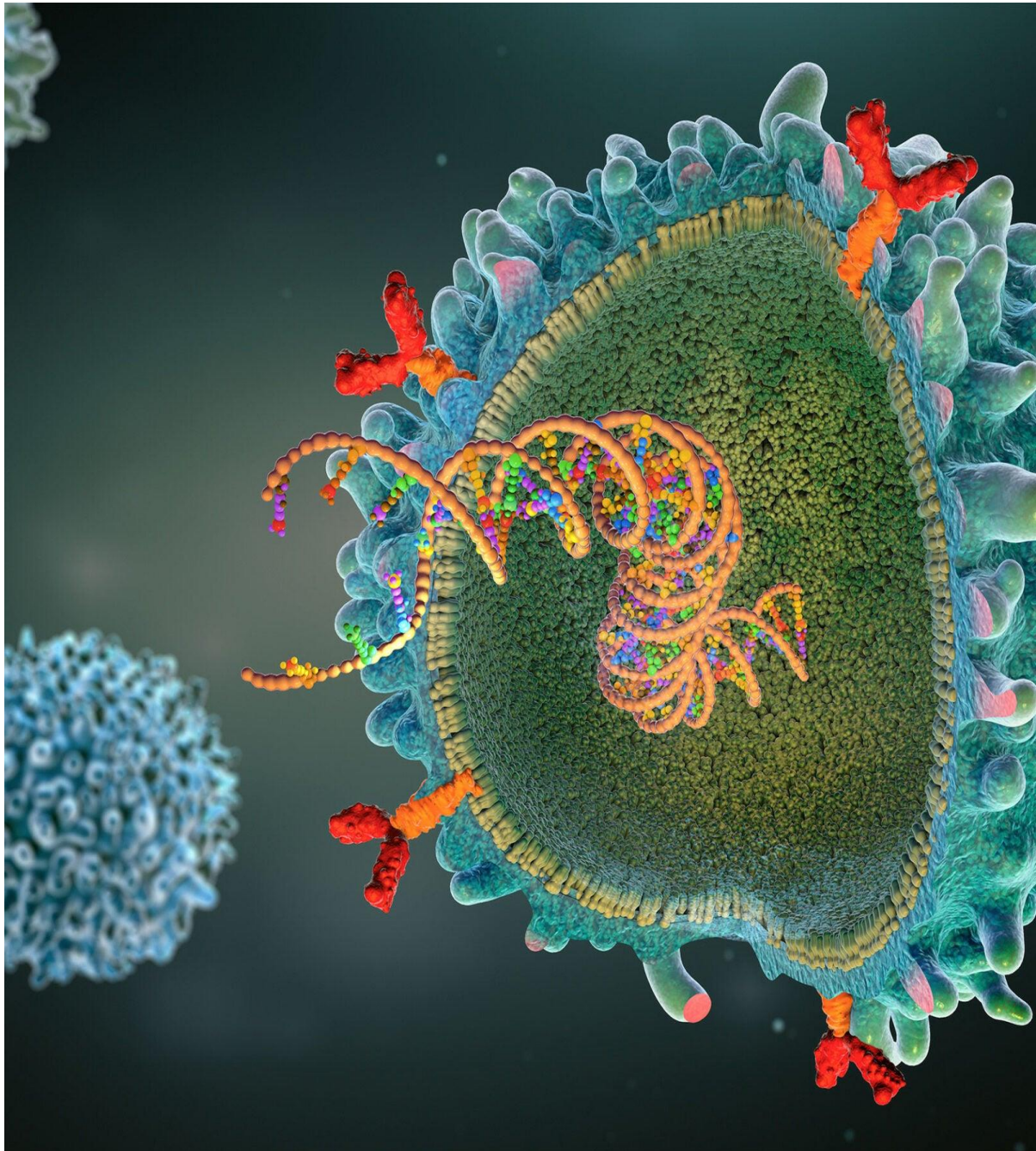
These applications contribute to global efforts in combating viral diseases.



Future Perspectives

- The integration of multiomics approaches and machine learning algorithms holds promise for deeper insights into viral biology and evolution.



Conclusion

In conclusion, viral genome assembly from NGS data provides a powerful means of uncovering the genetic blueprints of viruses, enabling a deeper understanding of viral biology, evolution, and pathogenesis. This knowledge is instrumental in developing effective strategies for combating viral diseases.