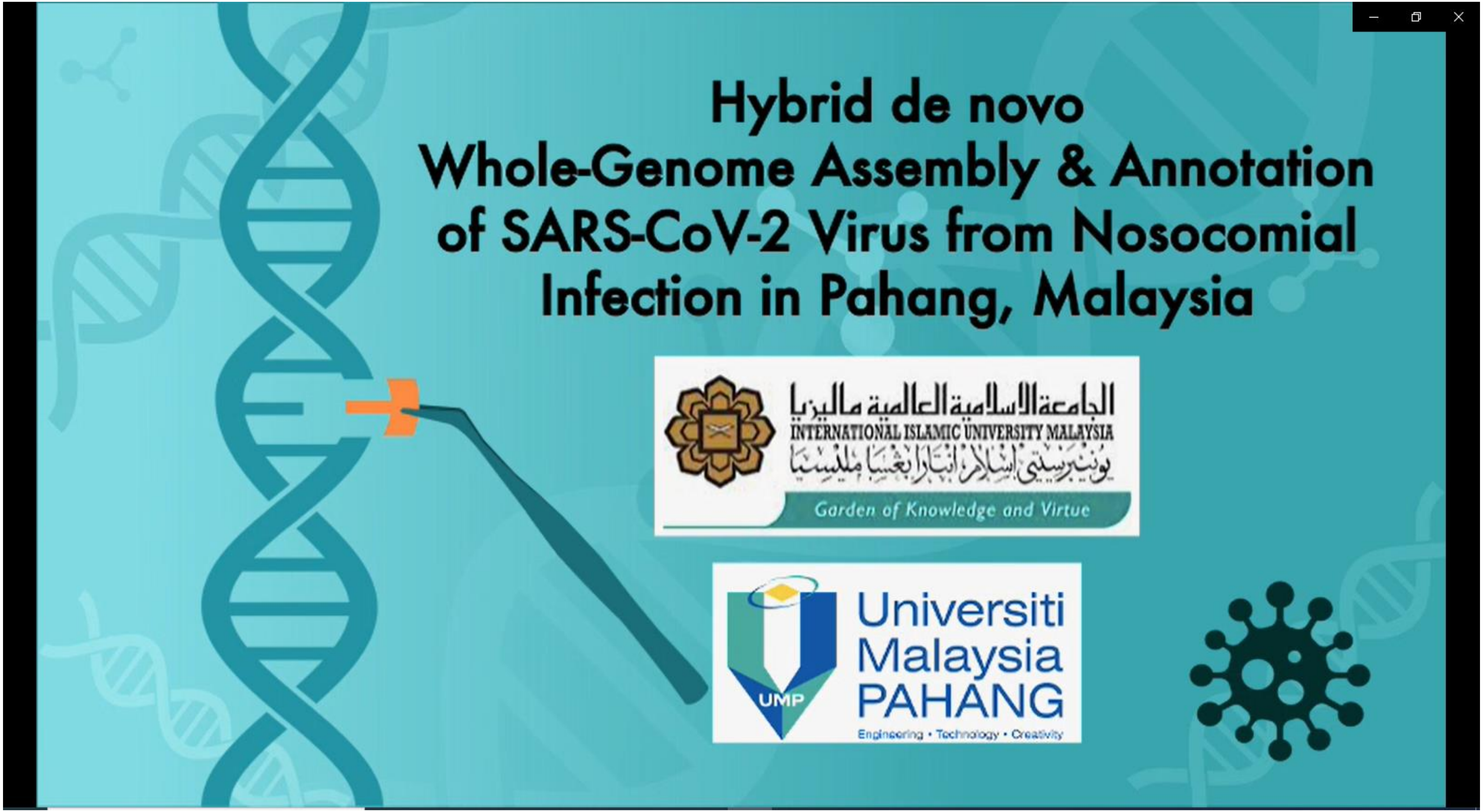


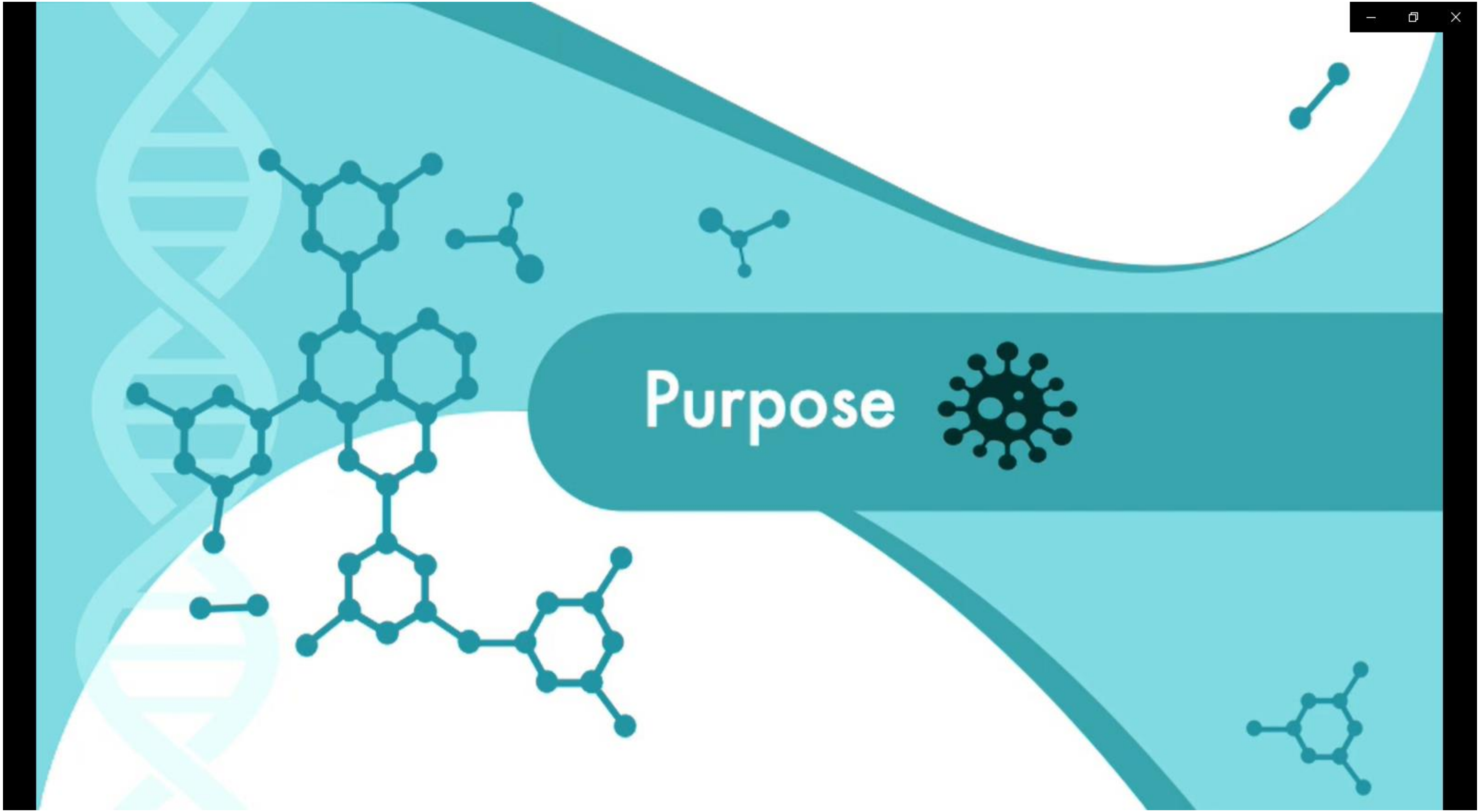
Hybrid de novo Whole-Genome Assembly & Annotation of SARS-CoV-2 Virus from Nosocomial Infection in Pahang, Malaysia

الجامعة الإسلامية العالمية ماليزيا
INTERNATIONAL ISLAMIC UNIVERSITY MALAYSIA
يونيۏرسيتي إسلام انتارابڠسا مليسيا
Garden of Knowledge and Virtue

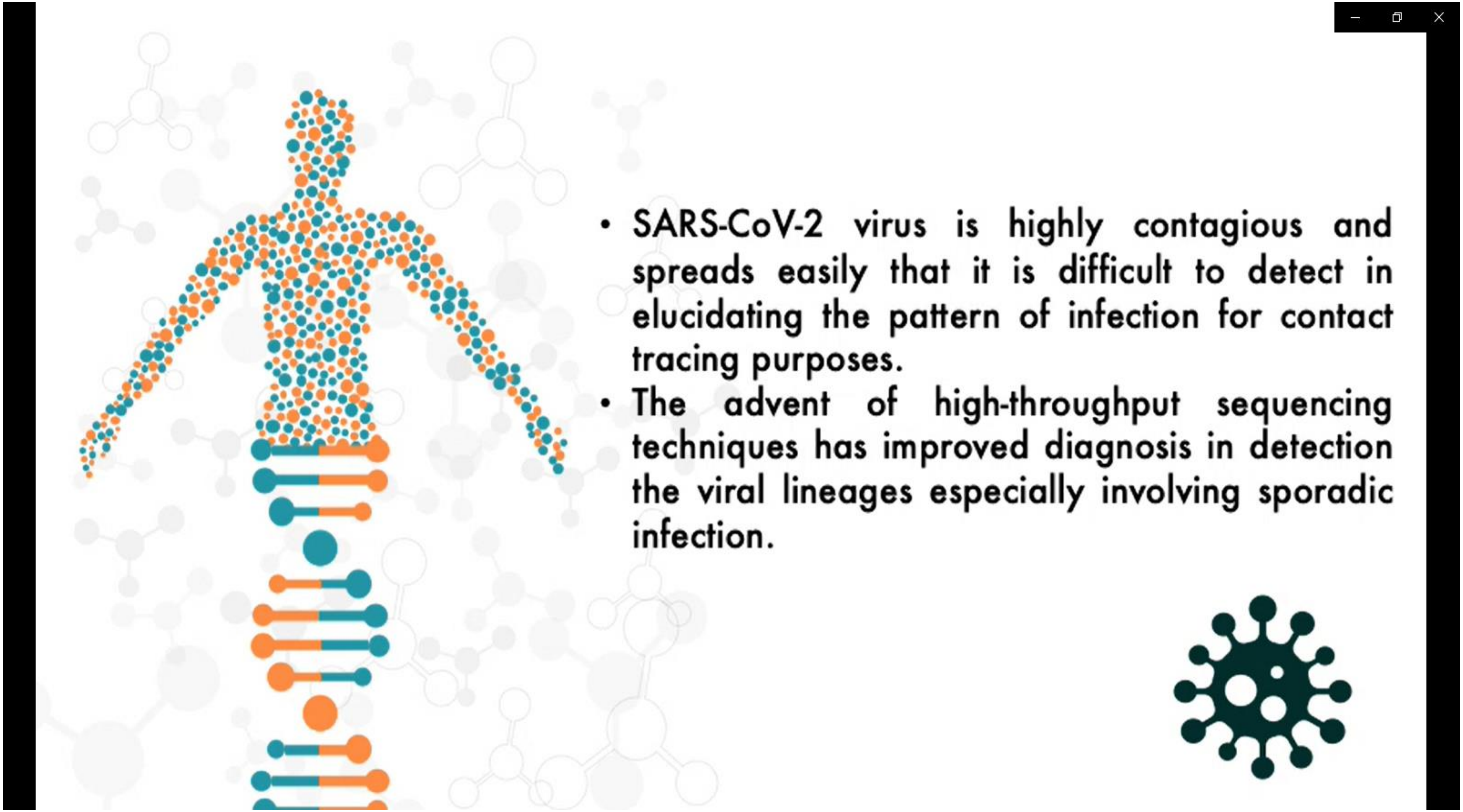
UMP
Universiti
Malaysia
PAHANG
Engineering • Technology • Creativity

Researchers

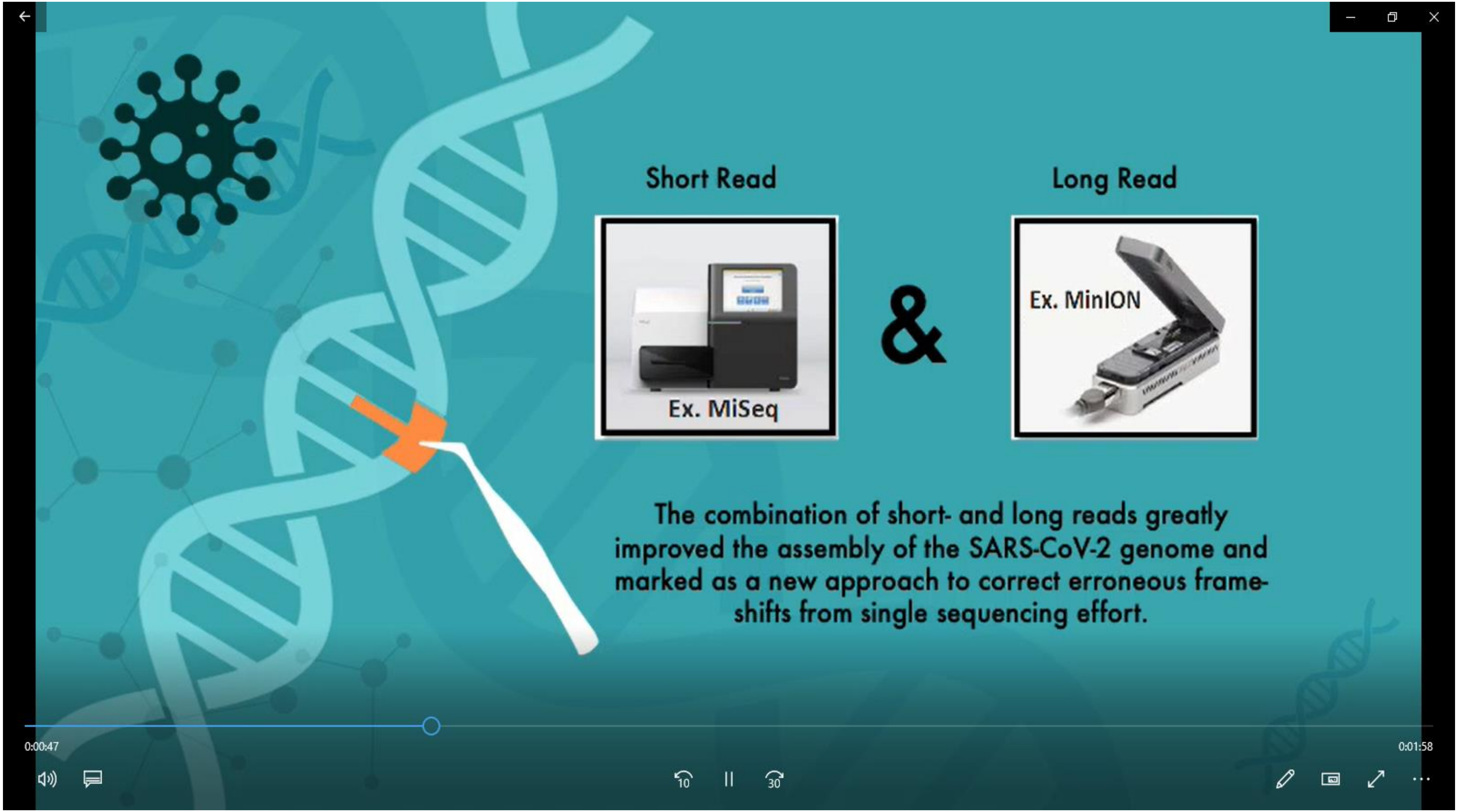
- Ummu Afeera Zainulabid
- Norhidayah Kamarudin
- Razman Mohd Rus
- Sharmeen Nellisa Soffian,
- Aini Syahida Mat Yassin
- How Soon Hin
- Hajar Fauzan Ahmad



Purpose



- SARS-CoV-2 virus is highly contagious and spreads easily that it is difficult to detect in elucidating the pattern of infection for contact tracing purposes.
- The advent of high-throughput sequencing techniques has improved diagnosis in detection the viral lineages especially involving sporadic infection.



Short Read

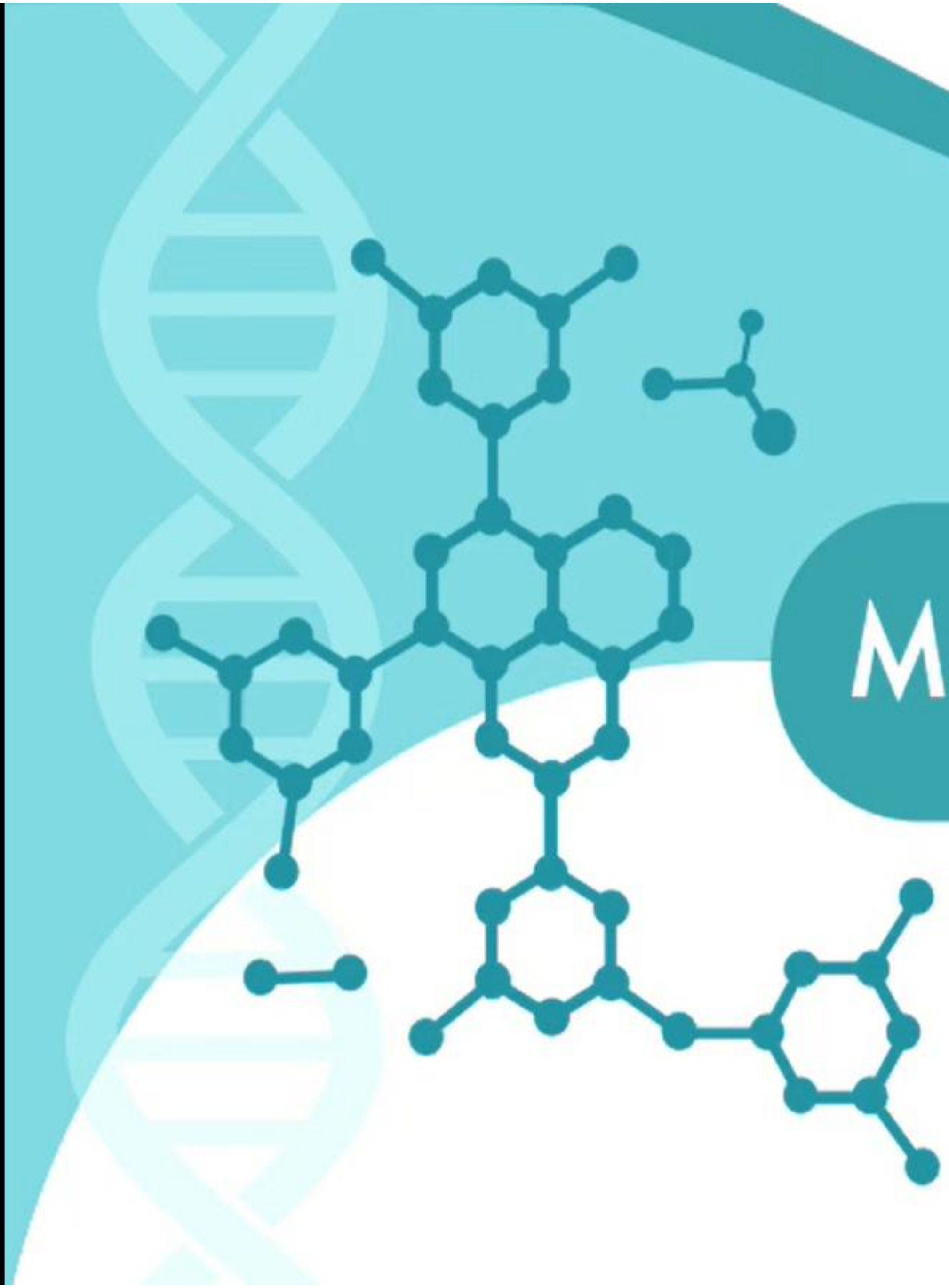
Ex. MiSeq

&

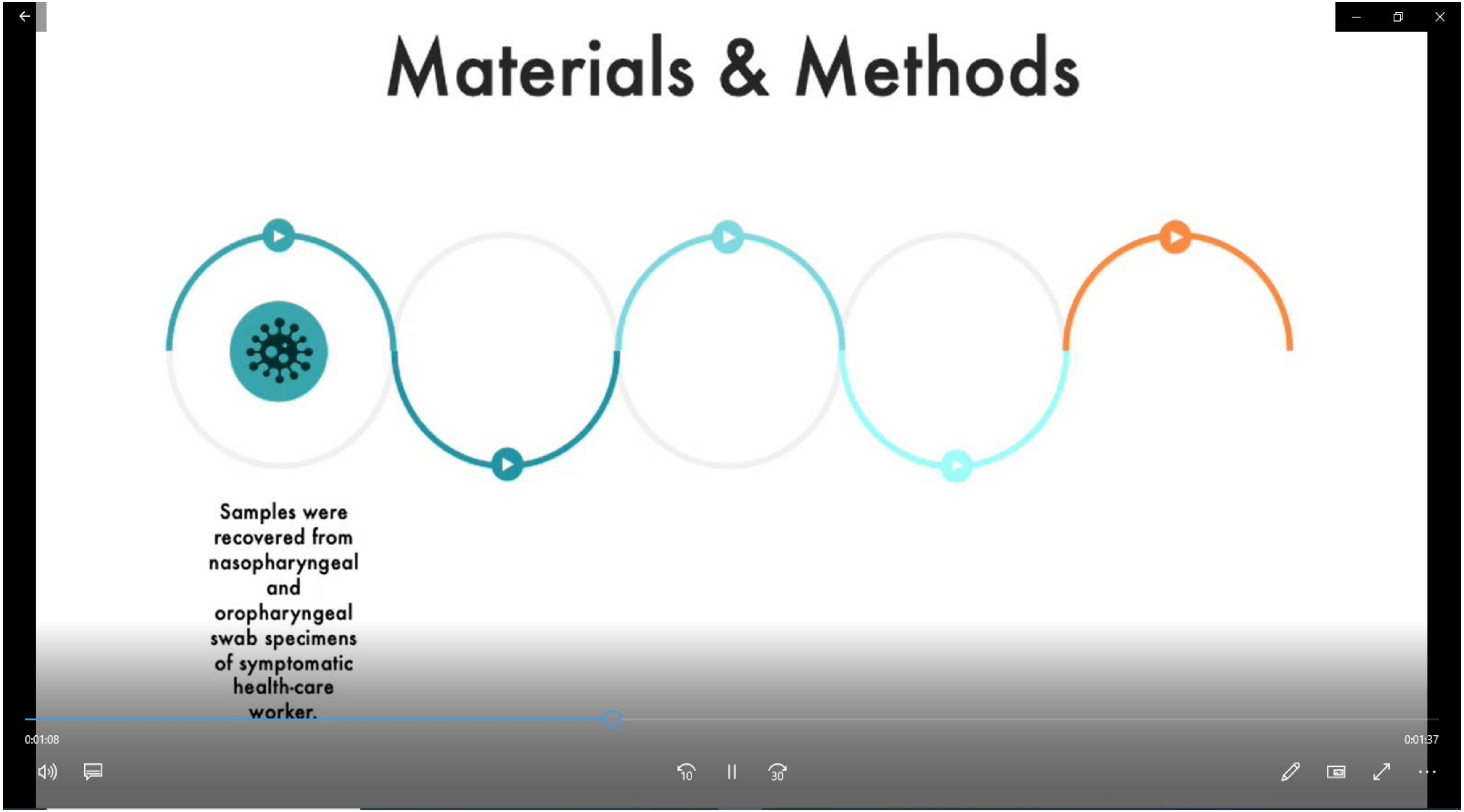
Long Read

Ex. MinION

The combination of short- and long reads greatly improved the assembly of the SARS-CoV-2 genome and marked as a new approach to correct erroneous frame-shifts from single sequencing effort.

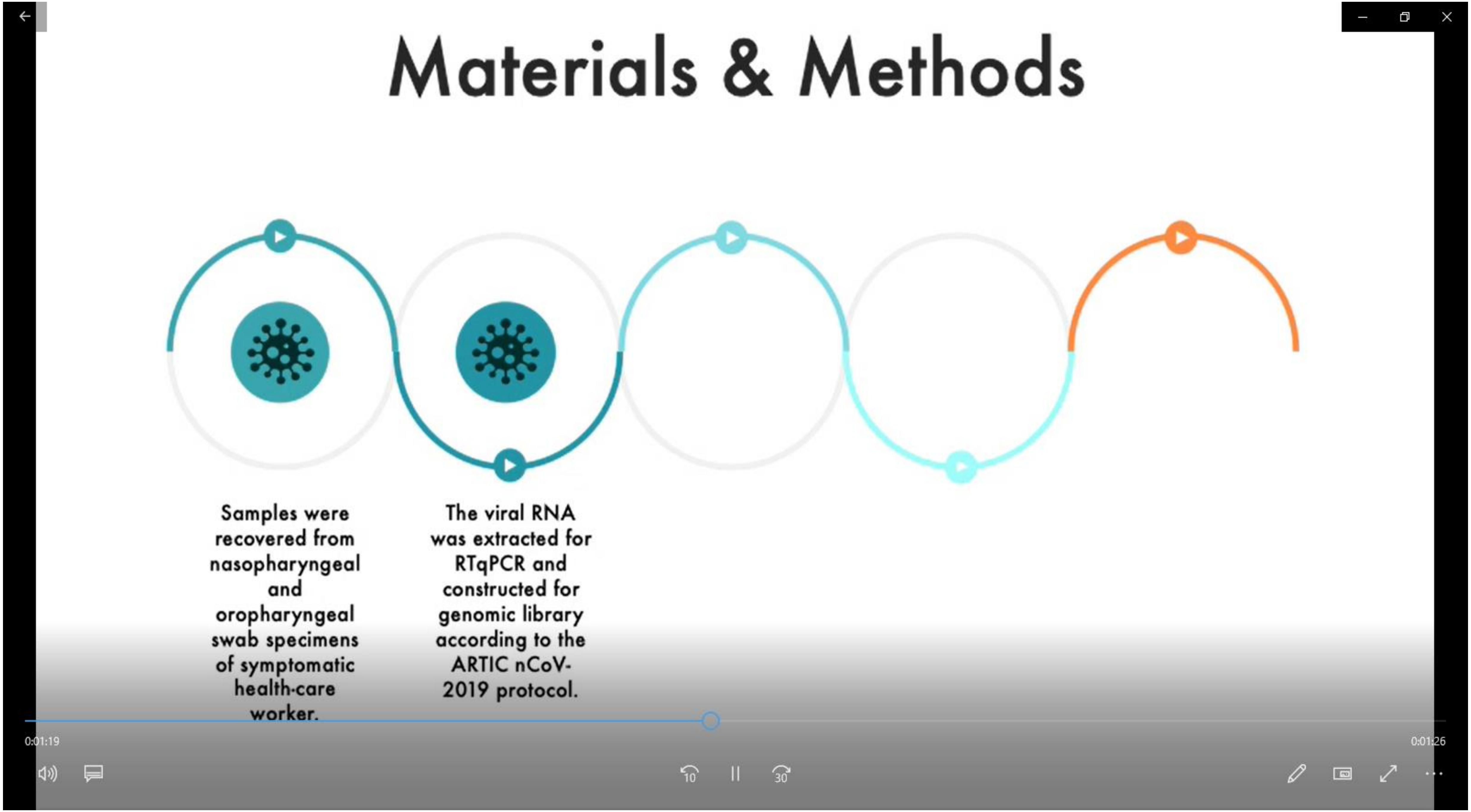


Materials & Methods



Materials & Methods

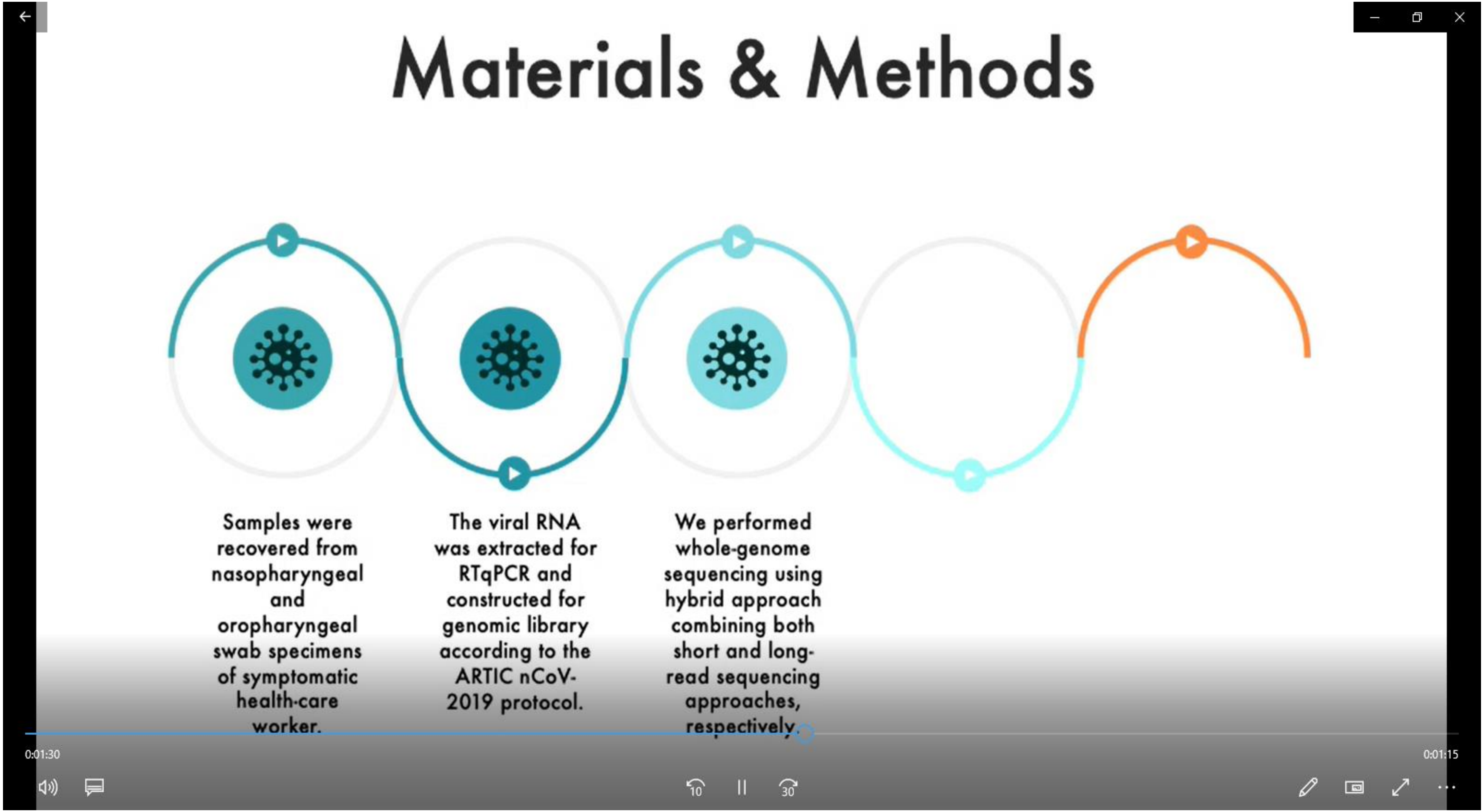
Samples were recovered from nasopharyngeal and oropharyngeal swab specimens of symptomatic health-care worker.



Materials & Methods

Samples were recovered from nasopharyngeal and oropharyngeal swab specimens of symptomatic health-care worker.

The viral RNA was extracted for RTqPCR and constructed for genomic library according to the ARTIC nCoV-2019 protocol.

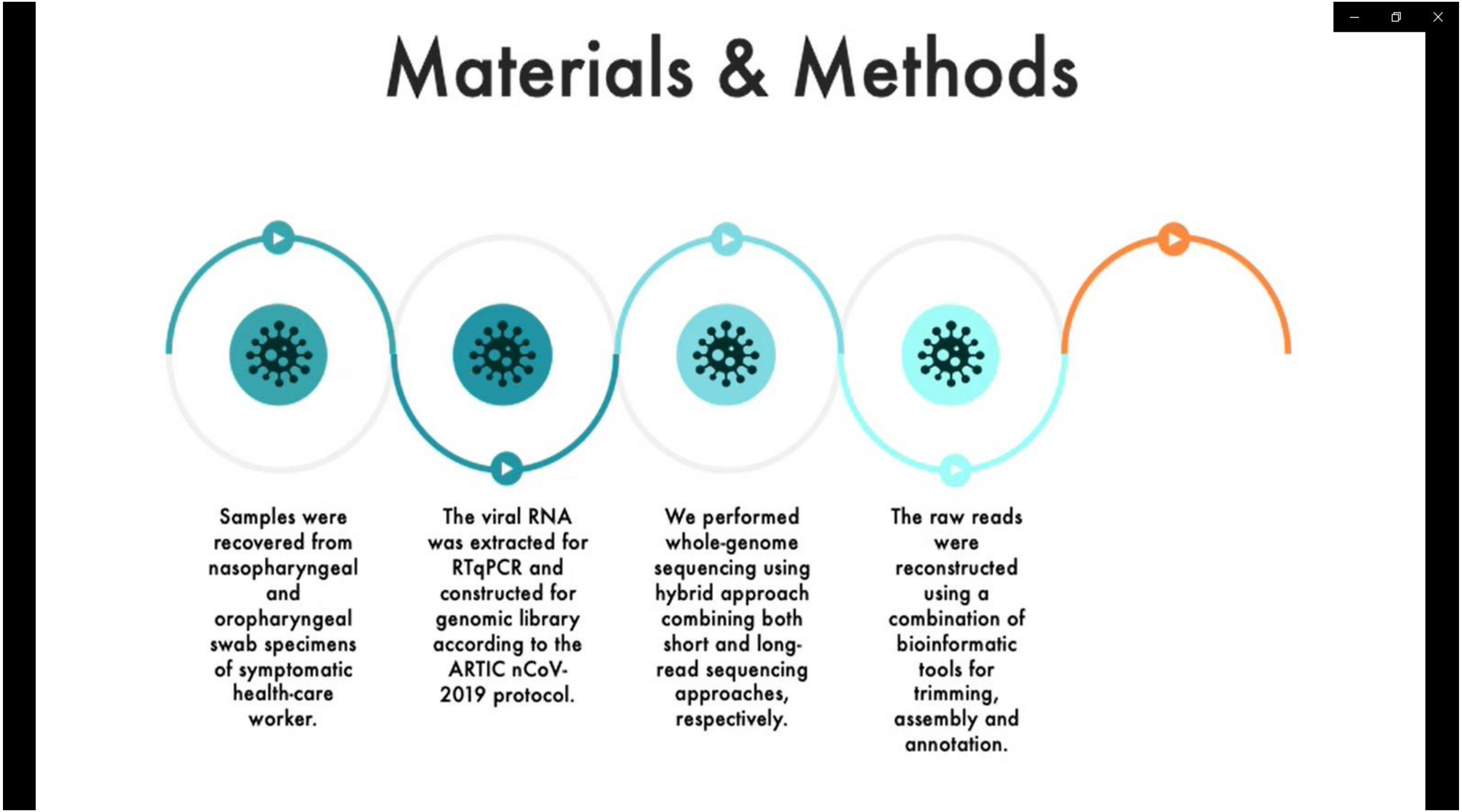


Materials & Methods

Samples were recovered from nasopharyngeal and oropharyngeal swab specimens of symptomatic health-care worker.

The viral RNA was extracted for RTqPCR and constructed for genomic library according to the ARTIC nCoV-2019 protocol.

We performed whole-genome sequencing using hybrid approach combining both short and long-read sequencing approaches, respectively.



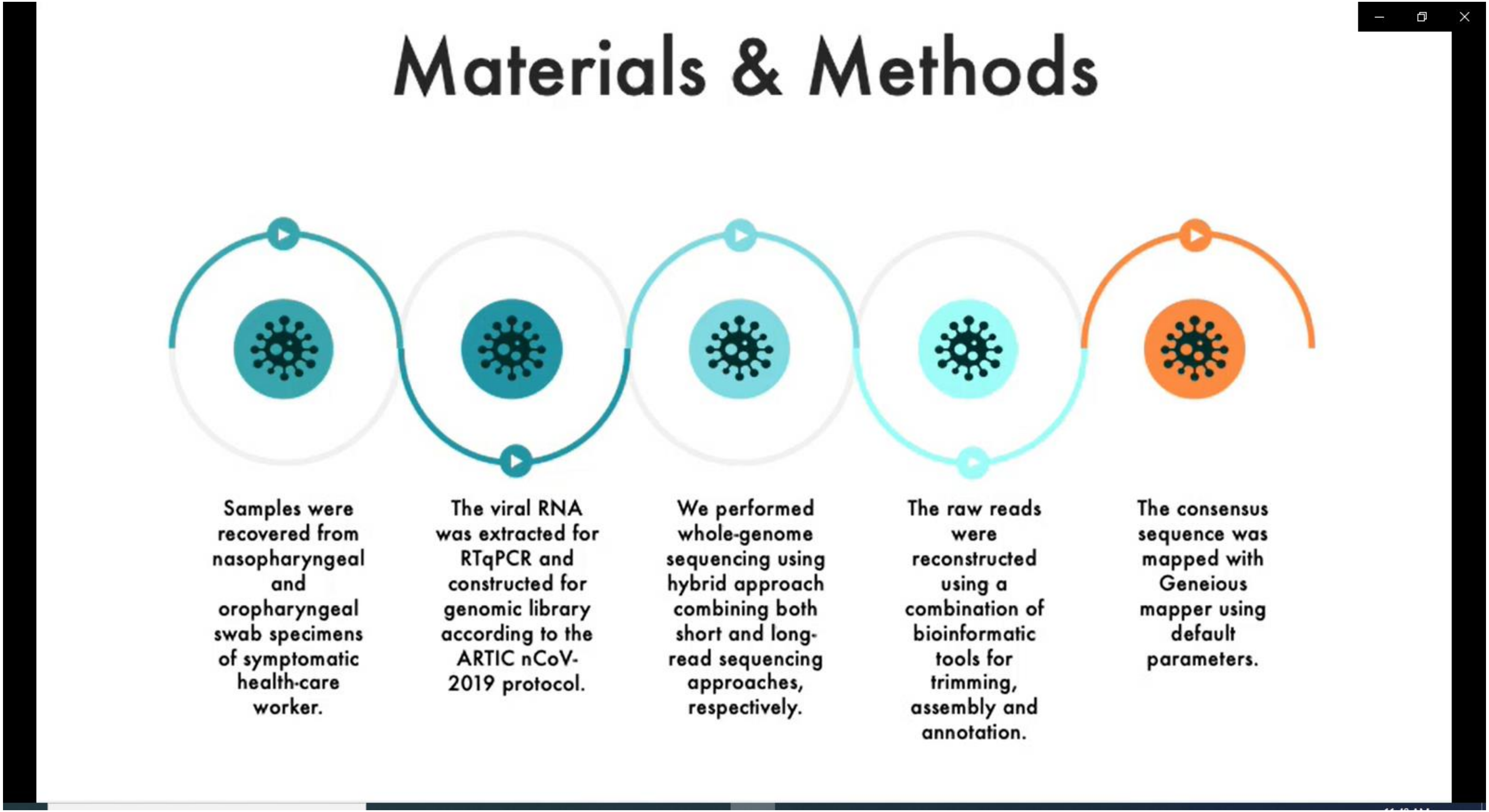
Materials & Methods

Samples were recovered from nasopharyngeal and oropharyngeal swab specimens of symptomatic health-care worker.

The viral RNA was extracted for RTqPCR and constructed for genomic library according to the ARTIC nCoV-2019 protocol.

We performed whole-genome sequencing using hybrid approach combining both short and long-read sequencing approaches, respectively.

The raw reads were reconstructed using a combination of bioinformatic tools for trimming, assembly and annotation.



Materials & Methods

Samples were recovered from nasopharyngeal and oropharyngeal swab specimens of symptomatic health-care worker.

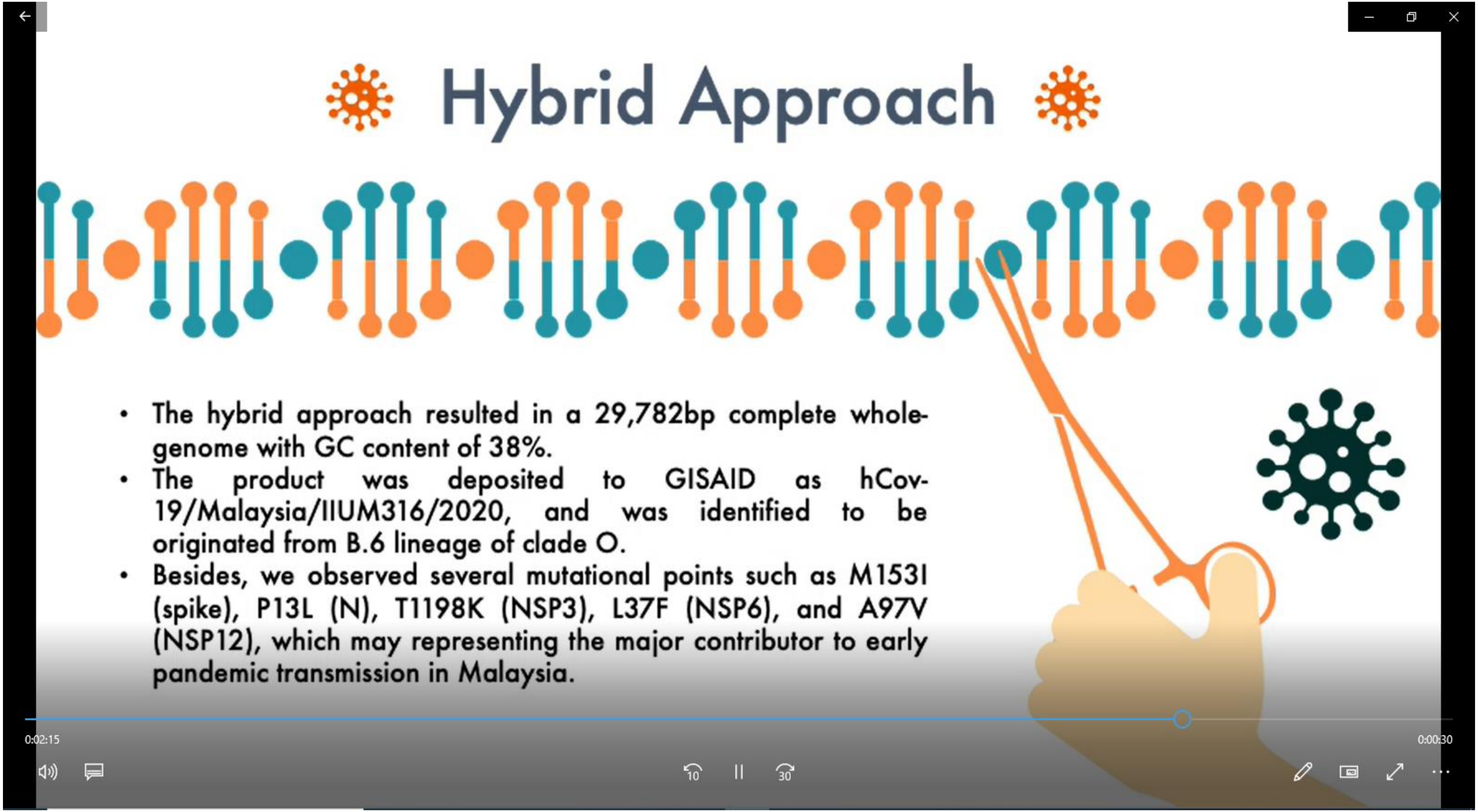
The viral RNA was extracted for RTqPCR and constructed for genomic library according to the ARTIC nCoV-2019 protocol.

We performed whole-genome sequencing using hybrid approach combining both short and long-read sequencing approaches, respectively.

The raw reads were reconstructed using a combination of bioinformatic tools for trimming, assembly and annotation.

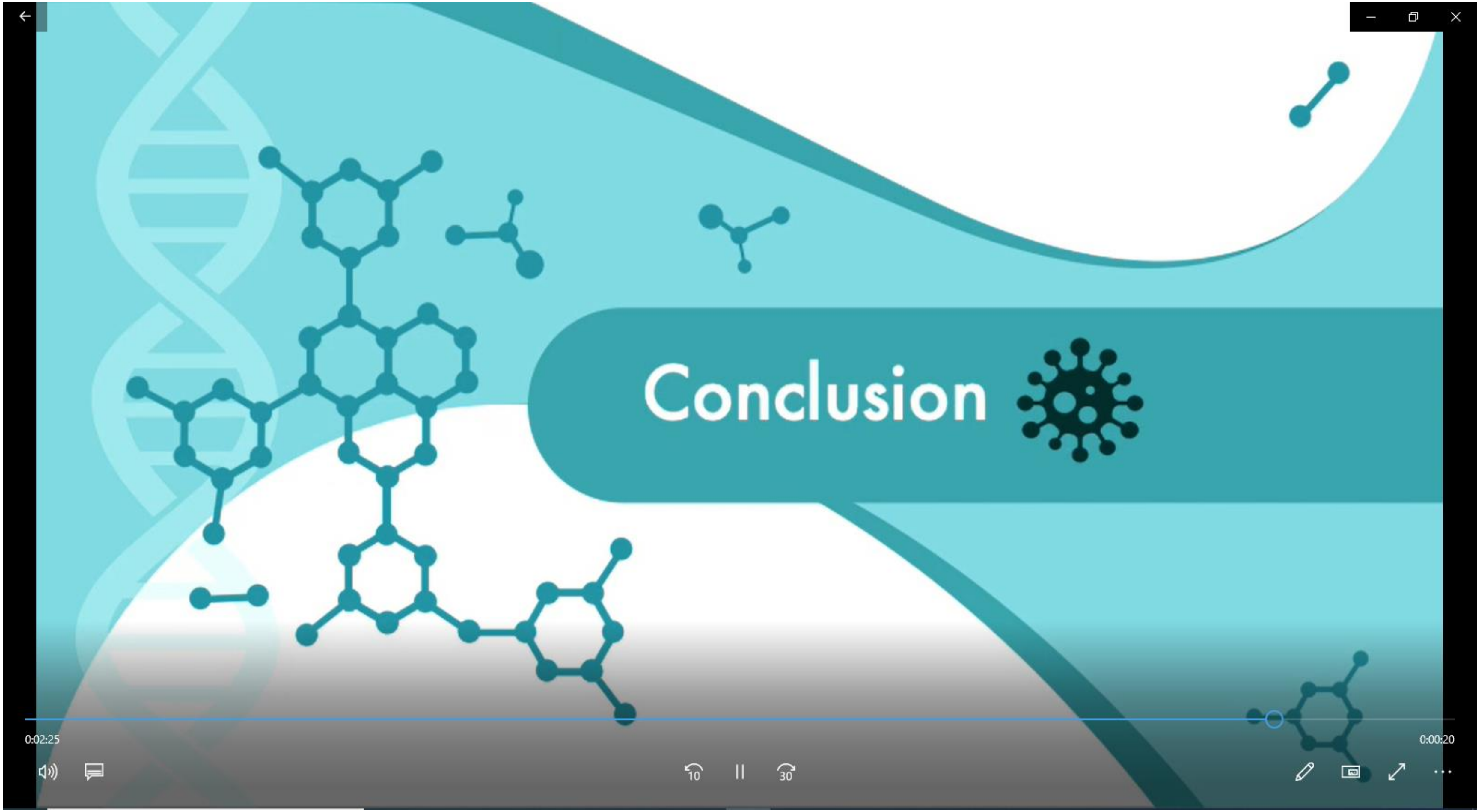
The consensus sequence was mapped with Geneious mapper using default parameters.

Results

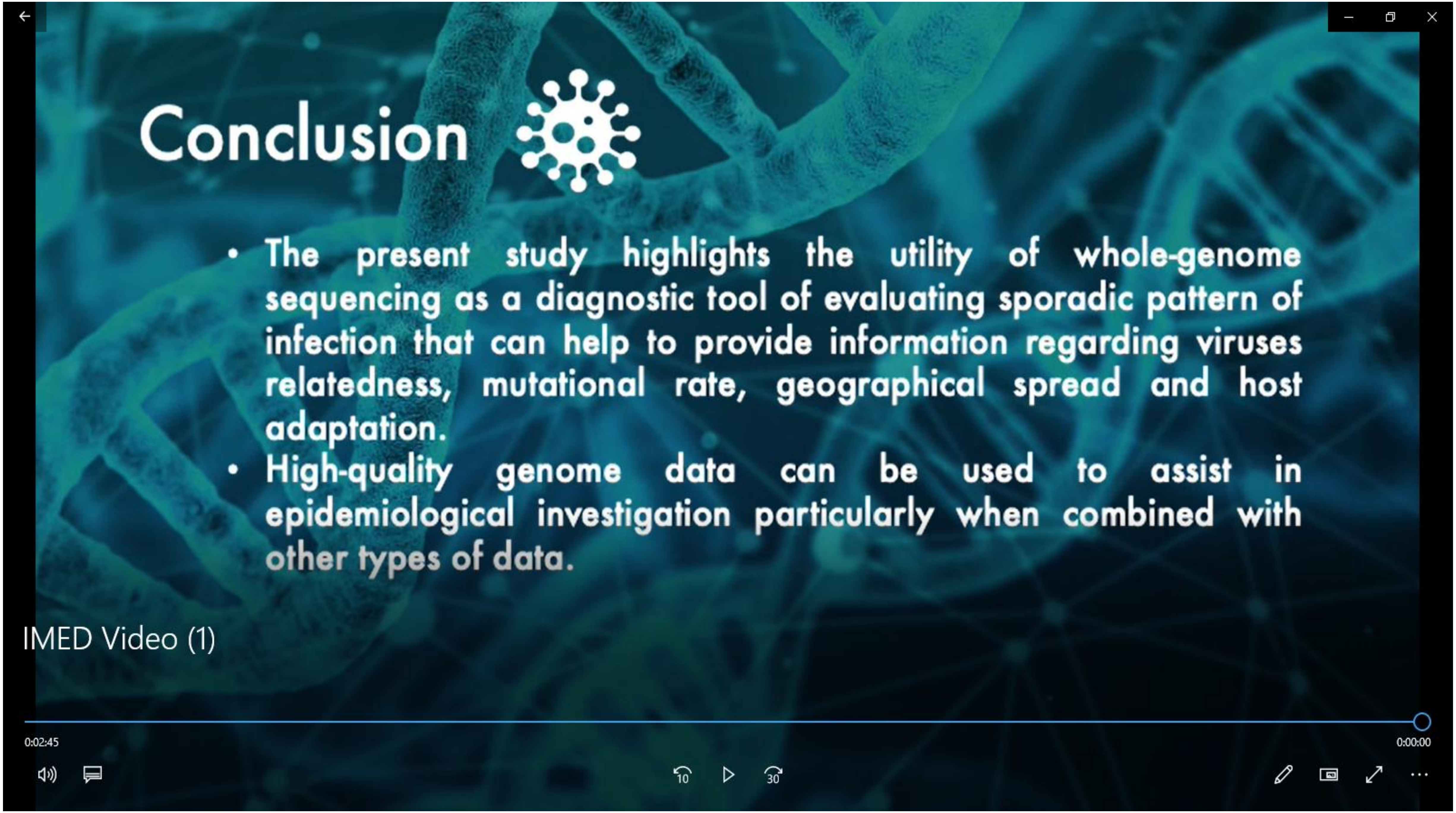


Hybrid Approach

- The hybrid approach resulted in a 29,782bp complete whole-genome with GC content of 38%.
- The product was deposited to GISAID as hCov-19/Malaysia/IIUM316/2020, and was identified to be originated from B.6 lineage of clade O.
- Besides, we observed several mutational points such as M153I (spike), P13L (N), T1198K (NSP3), L37F (NSP6), and A97V (NSP12), which may representing the major contributor to early pandemic transmission in Malaysia.



Conclusion



Conclusion

- The present study highlights the utility of whole-genome sequencing as a diagnostic tool of evaluating sporadic pattern of infection that can help to provide information regarding viruses relatedness, mutational rate, geographical spread and host adaptation.
- High-quality genome data can be used to assist in epidemiological investigation particularly when combined with other types of data.