Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Metagenomic and Targeted Sequencing of Imported Chikungunya Virus Cases

Lianghui Wei
Yangtze University
Li Zhu
Shenzhen International Travel Health Care Centre
Jianzhong Ye
Shenzhen International Travel Health Care Centre
Ran Zhang
Shenzhen International Travel Health Care Centre
Ying Ye
Shenzhen International Travel Health Care Centre
Chunzhong Zhao
Shenzhen International Travel Health Care Centre
Ye Yang
Yangtze University
Jianan He
hejianan6398@163.com
Shenzhen International Travel Health Care Centre

Article

Keywords: Chikungunya virus, molecular epidemiology, metagenomic NGS, Oxford Nanopore sequencing

Posted Date: March 20th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9103082/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.

Abstract

Objective: To verify the technical feasibility of metagenomic next-generation sequencing (mNGS) and targeted Oxford Nanopore Technology (ONT) sequencing for the genomic surveillance of imported Chikungunya virus (CHIKV), and to evaluate the complementary value of these two strategies in small-sample application scenarios.

Methods: Three serum samples confirmed positive for imported CHIKV by RT-qPCR were selected. Viral genomic sequences were obtained using unbiased mNGS and argeted-amplicon-based ONT sequencing. Technical verification and assessment of the two sequencing strategies were conducted focusing on genomic coverage, sequence identity, assembly continuity, and the consistency of phylogenetic genotyping results.

Results: After consensus sequence correction, near-complete CHIKV genomes were recovered from all samples. Phylogenetic analysis demonstrated that both sequencing strategies yielded consistent results in genotype determination and phylogenetic placement. Under targeted enrichment, ONT achieved continuous genomic coverage, while mNGS maintained stable sequence accuracy within an unbiased detection context.

Conclusion: In small-sample case series, both mNGS and targeted ONT sequencing can support the molecular epidemiological analysis of imported CHIKV. The two strategies exhibit complementary characteristics in practical surveillance, and their result consistency can be ensured through appropriate bioinformatics correction strategies. This study aims at technical verification and does not involve a generalized performance comparison of different sequencing platforms.

1. Introduction

Chikungunya fever (CHIKF), caused by the Chikungunya virus (CHIKV), is an arthropod-borne disease transmitted primarily by *Aedes aegypti* and *Aedes albopictus mosquitoes*. the most common symptoms of CHIKF are fever and joint pain, other symptoms can include rash, headache and debilitating polyarthralgia. Since its re-emergence and global expansion in early 2000s, CHIKF has become a significant public health concern, causing large-scale outbreaks across Asia, Africa, and the Americas[1–2]. Over the past 15 years, more than 5 million cases have been recorded across 119 countries and regions[3]. With the continued global spread of *Aedes albopictus*, the risk of imported epidemics in tropical, and subtropical regions continues to rise, posing increasing challenges for border health control and highlighting the urgent need for genomic surveillance of CHIKV[[4].

Conventional RT-PCR assay is highly sensitive and rapid for CHIKV detection, however,it requirement for known target sequences, means it may fail to detect novel variants or mixed infections and cannot provide comprehensive genomic information[5–6]. Moreover, the genetic heterogeneity among CHIKV strains necessitates full-genome data to identify mutational hotspots, track transmission chains, and elucidate viral evolution[7]. However, traditional assays fall short in resolving genome-wide diversity, structural variation, and in distinguishing vaccine strains from wild-type strains[8].

Metagenomic next-generation sequencing (mNGS) has emerged as a powerful unbiased tool that can detect all nucleic acids for pathogen discovery and genomic characterization[9]. As an untargeted approach, mNGS enables large-scale parallel sequencing of all nucleic acids in a sample, enabling simultaneous detection of CHIKV and co-infecting pathogens while recovering genome-wide variation[10]. This is especially valuable for febrile cases of unknown etiology and for discovering novel variants. However, mNGS also introduces notable limitations, More than 95% of reads typically originate from the host, resulting in insufficient effective sequencing data for pathogen genomes particularly in low-titer samples[11]. Additionally, short-read sequencing (50–300 bp) also complicates de novo assembly, often yielding fragmented contigs and incomplete mutation profiles[12].

Oxford Nanopore sequencing (ONT), a third-generation platform, addresses these limitations by generating ultra-long reads, which enables the reconstruction of near-complete viral genomes without the need for assembly, thereby facilitating direct genotyping and accurate identification of key mutations[13].Moreover, by employing specific probes or primers to enrich CHIKV nucleic acids, this approach significantly reduces host background interference[14–15].

Currently, mNGS and ONT each have strengths in CHIKV research. mNGS is suitable for broad-spectrum pathogen screening and unknown pathogen identification, while ONT excels in high-precision analysis of known viruses[16]. In port surveillance, sample numbers are limited and time is critical, but empirical data on how to strategically combine sequencing platforms for molecular epidemiology are scarce[17]. Therefore, this study performed a case series analysis of imported Chikungunya virus (CHIKV) infections identified during quarantine inspections at ports in Shenzhen, Guangdong Province, China. Integrating mNGS and ONT sequencing to comprehensively examine viral genome features and phylogenetic relationships, with a focus on evaluating cross-platform consistency and complementarity for case tracing and typing[18].

2. Materials and Methods

2.1 Sample Collection

This study included three serum samples from febrile travelers that were previously screened by RT-qPCR for CHIKV detection at Shenzhen Customs between July 2024 and January 2025. All participants presented with fever (> 37°C) and at least one additional symptom (rash or arthralgia).

2.2 Nucleic Acid Extraction and RT-PCR Detection of CHIKV

Viral RNA was extracted from 200 µL of patient serum samples using the Generotex96 automated nucleic acid extractor (Tialong, Xi'an, China) and its dedicated extraction kits, followed by elution in 60 µL of elution buffer. Real-time RT-qPCR was performed on an ABI 7500 Real-Time PCR System using a Chikungunya virus nucleic acid detection kit (Zhuhai Huifei Biotechnology Co., Ltd., China) according to the manufacturer's instructions.

The total reaction volume was 20 µL, consisting of equal parts of 4* Master Mix, reaction buffer, ultrapure water, and 5µL of template RNA. The amplification program was set as follows: reverse transcription at 50°C for 10 min; pre-denaturation at 95°C for 2 min; followed by 5 pre-cycles (95°C for 5 s, 50°C for 30 s) and 36 analysis cycles (95°C for 5 s, 55°C for 30 s). Fluorescence signals were collected via FAM, VIC, and CY5 channels during the 55°C extension phase, with ROX normalization set to "None." This assay targeted highly conserved regions of the viral genome.

2.3 Metagenomic Sequencing of Chikungunya Virus

Untargeted metagenomic sequencing libraries were constructed using the MGIEasy RNA Library Prep Set (MGI Tech, China; Cat. No. 1000006384). Extracted total RNA was incubated at 94°C for 8 min for random fragmentation, followed by first- and second-strand cDNA synthesis and purification.

The purified double-stranded cDNA products underwent end-repair, dA-tailing, and ligation with segmented adapters. PCR amplification was performed according to the recommended protocol for pathogen detection, with the number of cycles set to 15. The purified PCR products were quantified using the Qubit® dsDNA HS Assay Kit. After single-strand circularization and DNA nanoball (DNB) preparation, the libraries were sequenced on the MGISEQ-200 platform in PE150 mode (paired-end 150 bp).

2.4 Third-Generation Targeted Sequencing of Chikungunya Virus

For the Oxford Nanopore Technology (ONT) platform, a sequencing strategy based on tiled targeted amplicons was employed using a rapid detection kit for Chikungunya virus (Guangzhou Qinglan Biotechnology Co., Ltd., China), with library preparation performed using the Rapid Barcoding Kit (SQK-RBK114.24).

The cDNA was synthesized by adding 8 μL of RNA sample to the RT Mix, followed by virus-specific PCR amplification using the bundled Pool 1 and Pool 2 primer sets. The PCR reaction volume was 25 μL, with the following thermal cycling conditions: 98°C for 30 s; followed by 35 cycles of 98°C for 15 s and 65°C for 5 min. The amplified products were purified using magnetic beads, followed by sequencing adapter ligation, and the library was quantified using a Qubit 4.0 fluorometer.

Sequencing was performed on a MinION sequencer using FLO-MIN114 (R10.4.1) flow cells. Basecalling was executed using the Dorado (v0.3.0) super-high accuracy (SUP) model to generate raw sequencing reads in FASTQ format, which served as the input for subsequent bioinformatics analysis.

2.5 Bioinformatics Analysis

Customized bioinformatics pipelines were implemented to accommodate the distinct data characteristics of different sequencing platforms and enhance the accuracy of consensus sequences. For mNGS data, raw reads in FASTQ format were first subjected to quality control using Fastp (v0.23.2) to remove low-quality reads and adapter sequences. Subsequently, de novo assembly was performed using MEGAHIT (v1.2.9) to generate assembled contigs in FASTA format, which served as consensus sequences for downstream analysis.

For ONT data, reads were aligned to the initial consensus sequence using Minimap2 (v2.30). Multiple rounds of iterative polishing were conducted using Racon (v1.5.0) and Medaka (v2.1.1) to correct insertion or deletion (indel) errors frequently occurring in homopolymer regions. For Sample 1, a hybrid assembly approach was further employed using Unicycler (v0.5.1), where high-accuracy MGI short reads were integrated to polish the Nanopore consensus sequence, thereby restoring the biological integrity of the Open Reading Frames (ORFs).

2.6 Molecular Epidemiology and Phylogenetic Analysis

Multiple sequence alignment was performed on the corrected full-genome CHIKV sequences using MAFFT (v7.526) under the --auto mode. Phylogenetic analysis was conducted using IQ-TREE (v3.0.1) to construct a Maximum Likelihood (ML) tree. The optimal nucleotide substitution model (GTR + F+I+G4) was automatically selected via ModelFinder, and branch support was evaluated through 1,000 replicates of the Ultrafast Bootstrap test.

3. Results

3.1 RT-PCR Results

Nucleic acids were extracted from the blood samples and tested for three vector-borne viruses using fluorescent RT-PCR. The results for Dengue virus and Zika virus were negative, while Chikungunya virus was positive. The Ct values for the positive samples were as follows: Sample 1: Ct 20.65; Sample 2: Ct 31.75; Sample 3: Ct 25.06, which can be categorized as high, low, and medium viral load, respectively. The positive control showed expected amplification, and the negative control showed no amplification signal, indicating no contamination during the experiment. The experimental data is reliable.

3.2 Metagenomic Sequencing Results

Paired-end (PE100) mNGS generated high-quality data for all samples (Table 1). Average Q20 and Q30 values were 97.89 ± 0.20% and 94.67 ± 0.32%, respectively, These values significantly exceeded the recommended quality control standards for MGI metagenomic sequencing (Q20 ≥ 90%, Q30 ≥ 85%). In terms of sequence composition, the mean proportion of host-derived sequences in the original sequencing sequences was (95.73 ± 1.63)%, while the mean proportion of valid sequences was only (1.89 ± 1.57)%[19]. This result further confirmed that the high proportion of host sequences in macro-genome sequencing is prone to cause strong interference in the analysis of target sequences.Additionally, it is worth noting that although sample 1 had the lowest Ct value (typically indicating a relatively higher viral load), it only identified 884 sequences of Chikungunya virus. The potential mechanism behind this phenomenon needs further in-depth analysis[20].

Table 1
mNGS results

Sample	Raw Sequences	Q20	Q30	Host Sequences	Host Percentage	Effective Sequences	Effective Percentage	CHIKV Sequences
Sample 1	51482297	97.78%	94.47%	50233032	97.5%	99871	0.19%	884
Sample 2	54024210	97.77%	94.5%	51581597	95.4%	1189375	2.2%	171155
Sample 3	29357713	98.12%	95.04%	27705241	94.3%	965519.5	3.29%	38784

3.3 Third-Generation Targeted Sequencing Results

Third-generation targeted sequencing of Chikungunya virus yielded high-quality data for all samples. In this study, "Quality Data (%)" was defined as the proportion of valid reads passing the platform-recommended quality control threshold (Q ≥ 10) relative to the total raw reads. Sample 1 generated 78.41 Mb of raw sequencing data (1 Mb = 10^6bases), with 94.5% passing quality control. A total of 148.32 K reads (1 K = 10^3 reads) were obtained, of which 92.3% met quality standards, indicating both high sequencing quality and sufficient effective coverage. Sample 2 produced 40.62 Mb of raw data, with 91.8% passing quality control, and 70.2 K reads, 93.9% of which were of acceptable quality. Sample 3 yielded 53.92 Mb of raw data, with 93.1% passing quality control, and 104.45 K reads, 92.7% of which were high-quality (Table 2). The Q score distributions exhibited single, sharp peaks without noticeable fluctuations or multiple modes, reflecting a stable sequencing process, effective removal of low-quality reads, and the absence of sample contamination or chip obstruction. Peaks were predominantly centered around Q20(Figure 1), confirming that the majority of reads had high-quality scores and that the results are reliable[21–22].

Further analysis of the host sequence filtering results revealed that the proportion of host sequences in the three samples was 30.47%, 23.78%, and 30.05%, respectively. After removing the host sequences, the number of effective reads obtained was 136,899, 65,917, and 96,870, accounting for 49.19%, 64.67%, and 49.49% of the total reads, respectively. On this basis, a significant number of Chikungunya virus (CHIKV)-related sequences were successfully identified in all three samples. Specifically, 67,342, 42,627, and 47,940 CHIKV reads were detected in Samples 1, 2, and 3, respectively. These results indicate that the targeted sequencing strategy demonstrates high enrichment efficiency for the target virus in clinical samples, providing reliable data support for subsequent viral genome assembly, mutation analysis, and phylogenetic studies.

Table 2
Results of ONT

Total Raw Data (Mb)	Quality Data (%)	Total Reads (K)	Reads Meeting Quality Standards (%)	Total Raw Data (Mb)	Host Percentage(%)	Effective Sequences	Effective Percentage(%)	CHIKV Sequences
Sample 1	78.41	94.5	148.32	92.3	30.47%	136899	49.19%	67342
Sample 2	40.62	91.8	70.2	93.9	23.78%	65917	64.67%	42627
Sample 3	53.92	93.1	104.45	92.7	30.05%	96870	49.49%	47940

3.4 Genomic Characteristics of Cases

3.4.1 Time and Cost Comparison

The integrated workflows for both second-generation metagenomic sequencing (mNGS) and third-generation Oxford Nanopore Technology (ONT) targeted sequencing were standardized into four stages(Table 3): nucleic acid extraction, pretreatment, library preparation, and sequencing/data analysis. Total RNA was extracted using the Tialong Generotex 96 system (approx. 13 min), followed by platform-specific processing. In the pretreatment stage, mNGS required host rRNA depletion, high-temperature fragmentation (94°C, 8min), and random-primed reverse transcription, totaling approximately 3h (excluding extraction). In contrast, the ONT platform utilized an amplicon-based targeted strategy, requiring only 15 min for reverse transcription without rRNA depletion or physical fragmentation, significantly enhancing efficiency. Notably, while both platforms shared the same sample source, reverse transcription was performed independently to optimize for each technology—using random primers for mNGS unbiasedness and a dedicated RT Mix for ONT long-fragment PCR.

For library preparation, the mNGS workflow (including end-repair, ligation, and DNB preparation) took approximately 6.5 h, while the ONT workflow (centered on 35 cycles of tiled PCR and barcoding) required about 4 h. During the sequencing and analysis phase, mNGS (PE150 mode) required 30 h for data output and 2 h for bioinformatics, whereas ONT (real-time mode) completed data acquisition and consensus error correction (including Medaka) within 8 h. Economically, the cost per sample was approximately $110 for mNGS and $140 for ONT targeted sequencing, the latter reflecting the additional expense of specific primer pools and R10.4.1 flow cells.

Table 3
mNGS and ONT Time and Cost Comparison

Metric	mNGS	ONT
Nucleic Acid Processing Time	3h	0.5h
Library Construction Time	6.5h	4h
Sequencing Analysis Time	32h	8h
Cost per Sample	110$	140$

3.4.2 Genome Features by Sequencing Method

Full-genome sequences were obtained from three imported Chikungunya virus cases using both metagenomic next-generation sequencing (mNGS) and Oxford Nanopore Technology (ONT) sequencing, followed by a side-by-side comparative analysis (Table 4). Both sequencing strategies achieved high genomic coverage across all three samples, ranging from 96.0% to 100.0%. Specifically, for Sample 1 and Sample 3, the genomic coverage reached 96.0% and 99.9% via ONT, compared to 96.5% and 97.6% via mNGS, respectively. For Sample 2, 100.0% coverage was achieved using both platforms. The total genome lengths obtained from both platforms ranged from 11,216 to 11,369 bp, with minimal length variation observed for each sample between the two technologies.

Nucleotide identity analysis revealed that Sample 1 and Sample 2 yielded identical values between the mNGS and ONT results. However, for Sample 3, the nucleotide identity obtained via ONT was 92.9%, higher than that obtained via mNGS (91.3%). A similar trend was observed for amino acid identity (91.2% vs. 89.9%) and overall identity (92.1% vs. 90.6%). Across metrics including genomic coverage, assembly length, and nucleotide/amino acid identity, mNGS and ONT sequencing displayed similar numerical distributions across the three samples.

Table 4
Genome Comparison Table of mNGS and ONT

Sample	Sequencing_Type	Coverage(%)	Total_length	NT_Identity(%)	AA_Identity(%)	Overall_Identity(%)
1	mNGS	96.5%	11216	90.5	89.6	90.0
	ONT	96%	11369	90.5	89.1	79.0
2	mNGS	100.0%	11237	96.5	95.5	96.0
	ONT	100.0%	11237	96.5	95.5	96.0
3	mNGS	97.6%	11237	91.3	89.9	90.6
	ONT	99.9%	11225	92.9	91.2	92.1

3.4.3 Regional Coverage

mNGS achieved high genomic coverage across all functional regions of CHIKV, with overall coverage ranging from 91.2% to 100.0%. Specifically, coverage for the non-structural protein-coding regions (nsP1, nsP2, nsP3, nsP4) was no less than 93.5%, while the structural protein-coding regions (E3, E2, 6K, E1) ranged from 94.1% to 100.0%. The coverage for the 3' untranslated region (3'UTR) was between 91.2% and 100.0% (Fig. 2). In contrast, ONT achieved near or total 100.0% coverage across all functional regions, with minimal variation between different regions. Although some fluctuations in coverage levels were observed in certain functional regions across different cases, the minimum coverage for each region remained consistently above 90.0%.

3.4.4 Key Variant Sites

Analysis of the three clinical samples revealed certain discrepancies in the number of variant sites identified by mNGS and ONT. As illustrated in the Circos plots(Fig. 3), ONT detected 155 and 78 additional variant sites in Sample 1 and Sample 3, respectively, which were not covered by mNGS. In contrast, the variant counts for Sample 2 remained highly consistent between the two platforms. Region-specific statistical analysis indicated that these discrepant sites were predominantly localized within the non-structural protein-coding regions. However, the major variant sites identified in the structural protein regions and other key functional domains showed high concordance between both sequencing methods, suggesting that while ONT targeted sequencing offers higher sensitivity in specific genomic regions, both platforms are reliable for identifying critical functional mutations.

tailored to the complexity of genomic variation in target samples to achieve accurate and comprehensive variant detection.

Figure

3 Comparison of single-nucleotide variant (SNV) detection between mNGS and ONT sequencing across three clinical samples.(Each donut plot represents the total number of variants detected in one sample, with the pink core indicating variants identified by mNGS and the red outer ring indicating variants unique to ONT sequencing. Sample 1 had 662 variants detected by mNGS and 155 unique to ONT; Sample 2 had 398 variants detected by mNGS and 1 unique to ONT; Sample 3 had 728 variants detected by mNGS and 78 unique to ONT.)

3.4.5 Mutation Site Genome Distribution Analysis

By analyzing the density distribution of shared variants and ONT-specific mutation sites, the differences in mutation detection between mNGS and ONT across the CHIKV genome were clearly illustrated(Fig. 4). In Sample 1, shared mutations were evenly distributed in regions such as P1 and nsP2, whereas ONT-specific mutations were notably enriched in the E1 region, suggesting that ONT’s long-read capability enables the detection of complex structural variations that short-read mNGS may miss. In Sample 2, mutation profiles from both platforms were largely comparable, with ONT’s long-read

advantage less pronounced. In Sample 3, ONT-specific mutations exhibited minor peaks in nsP3 and E1. Overall, mNGS provides stable detection of common mutations and consistent genome-wide coverage, while ONT is superior for identifying complex structural regions and specific mutations. The two methods exhibit clear complementarity in characterizing viral genetic diversity.

3.4.6 Source Tracking Analysis

Based on the reconstructed phylogenetic tree(Fig. 5), sequences generated by ONT and mNGS showed a high degree of concordance in clustering patterns for the same samples. Specifically, ONT1 clustered closely with mNGS1, ONT2 with mNGS2, and ONT3 with mNGS3, each forming paired branches with extremely short branch lengths, indicating minimal sequence divergence between the two sequencing approaches. Notably, all analyzed samples were consistently assigned to the selected reference clade without evidence of cross-lineage placement or misclassification. Phylogenetic analysis further demonstrated that all three samples clustered within the Asian lineage, confirming their genetic relatedness to Asian CHIKV strains. These results indicate that both sequencing methods reliably capture the phylogenetic signal of the samples. Overall, evolutionary tree reconstruction showed strong consistency between ONT and mNGS in whole-genome evolutionary inference, supporting their use as complementary approaches. While ONT offers advantages in sequence continuity, mNGS remains robust for high-coverage short-read detection.

4. Discussion

4.1 Core Performance Differences and Application Scenario Adaptability of Sequencing Technologies

As a globally significant arbovirus, rapid genotyping and source tracing of Chikungunya virus (CHIKV) depend on the availability of sequencing strategies capable of generating reliable whole-genome data under practical surveillance conditions[23]. In this study, three imported CHIKV-positive cases collected by Shenzhen Customs were analyzed using two integrated sequencing approaches: an unbiased metagenomic workflow and a targeted amplicon-based workflow utilizing ONT technology. Based on whole-genome sequencing outputs, the molecular epidemiological characteristics of these viruses were investigated to explore how different analytical strategies perform in applied surveillance scenarios and to provide technical evidence for optimizing infectious disease monitoring systems[24].

Both sequencing workflows generated near-complete CHIKV genomes sufficient for genotype determination and phylogenetic analysis in this case series, although differences were observed in genome coverage patterns and sequence continuity[25]. In Case 1, the targeted amplicon-based workflow produced more uniform coverage across predefined genomic regions, including structurally complex non-structural protein–coding regions (such as nsP3 and nsP4), where local coverage gaps were reduced. In contrast, the unbiased metagenomic workflow exhibited region-specific fluctuations in coverage, consistent with the stochastic distribution of viral reads in a host-background context (Fig. 4)[26]. These differences are therefore interpreted as consequences of distinct library preparation and enrichment designs rather than intrinsic advantages of a particular sequencing platform. Collectively, the findings suggest that different sequencing strategies may provide complementary genome-level information.

Differences were also observed between the two sequencing approaches in the detection of potential variant sites across cases. In Case 1, the targeted workflow identified a greater number of candidate variant sites than the metagenomic workflow, with several variants located in the E1 coding region. However, the number of detected variant sites was not used as a metric of comparative sensitivity or accuracy; instead, it was treated as an observational reflection of case-level genomic variation under different analytical conditions. Importantly, these variations did not affect phylogenetic branch assignment or source attribution, indicating that consensus sequences generated through both workflows demonstrate concordance for molecular epidemiological interpretation[27].

From an operational perspective, the two sequencing strategies exhibit different characteristics that may align with distinct surveillance objectives. The targeted amplicon-based workflow increases the relative abundance of viral sequences through PCR enrichment and allows rapid genome recovery, which may be advantageous in time-sensitive settings such as port-of-entry inspections or localized outbreak investigations, including monitoring of key mutation sites (e.g., E1-A226V)[28]. In contrast, the unbiased metagenomic workflow retains pathogen-agnostic detection capability and high-throughput analytical capacity, making it suitable for broader pathogen spectrum surveillance and longitudinal evolutionary analysis[29]. Flexible selection of sequencing strategies—or their combined application—based on specific public health requirements may enhance the adaptability and resilience of infectious disease surveillance systems[30].

4.2 Improving Detection Performance via Technological Complementarity

Case-level analyses indicated that the unbiased metagenomic workflow and the targeted amplicon-based workflow exhibit complementary characteristics in variant site detection. Differences in the number and distribution of candidate variants were more apparent in samples with relatively complex genomic regions. For example, in Case 1, the targeted workflow identified a greater number of potential variant sites than the metagenomic workflow, and a similar pattern was observed in Case 3. These observations are interpreted in the context of differences in enrichment design and coverage architecture rather than as indicators of relative analytical sensitivity.In genomic regions where coverage continuity is influenced by library preparation strategy, reliance on a single sequencing workflow may not fully capture the spectrum of case-associated variation. Integrating data generated from distinct analytical frameworks may therefore facilitate a more comprehensive characterization of variant distribution and reduce the risk of information loss related to workflow-specific coverage patterns.From a methodological perspective, the unbiased metagenomic workflow provides pathogen-agnostic detection and stable base-level concordance under routine conditions, whereas the targeted amplicon workflow increases the relative abundance of predefined genomic regions through PCR enrichment, enabling more continuous coverage in those areas. Consequently, these two sequencing strategies offer complementary perspectives in case-based genomic analysis[31].

4.3Limitations of the Study and Future Directions

This study has certain limitations in the analysis of the association between third-generation ONT targeted sequencing data characteristics and sample Ct values, which should be objectively acknowledged and may inform directions for future optimization. First, the small sample size limits the generalizability of the conclusions. Only three samples were included in this analysis, and although a potential association between Ct values and raw data yield or total read counts was preliminarily observed, results derived from such a limited dataset are susceptible to the influence of inter-sample variability and stochastic errors introduced during library preparation[32]. Consequently, these findings are insufficient to comprehensively characterize sequencing data patterns across different viral load gradients. Future studies should expand the sample size and include samples from diverse sources and across a broader range of viral loads to validate these associations and improve the robustness and applicability of the conclusions[33].

In this study, inconsistencies were observed between Ct values and metagenomic sequencing (mNGS) results in individual samples. For instance, Case 1 exhibited the lowest Ct value in RT-qPCR, indicating a high viral load, yet yielded a relatively low number of CHIKV-specific reads in the mNGS data. This phenomenon may primarily be attributed to fluctuations in the proportion of host background nucleic acids. In an unbiased mNGS workflow, the competitive presence of host sequences can significantly dilute the relative abundance of target viral sequences. Furthermore, the multi-step mNGS library preparation process is susceptible to the combined effects of RNA integrity, reverse transcription efficiency, and stochastic amplification bias. In contrast, the targeted amplification strategy employed by Oxford Nanopore Technologies (ONT) increases the proportion of the target virus in the library through PCR enrichment, demonstrating superior interference resistance against complex host backgrounds[34].

Regarding technical process control, although sample quality was assessed via Qubit quantification and terminal quantification of DNA Nanoballs (DNBs), these findings suggest that the relationship between Ct values and pathogen read counts in mNGS is not simply linear for molecular epidemiological analysis. Future studies should expand the sample size and systematically introduce samples from diverse sources and across various Ct ranges to conduct stratified and multivariate analyses. This will allow for a more comprehensive evaluation of how viral load influences the data characteristics of ONT targeted sequencing. Additionally, it is essential to integrate raw sequencing metrics with downstream analytical outcomes—such as genome assembly completeness, depth of coverage distribution, and functional annotation sufficiency—to assess the practical utility of Ct values in sequencing applications. Moreover, to address host interference and amplification bias, future efforts should focus on optimizing library construction and enrichment strategies to improve the recovery of pathogen sequences from low-viral-load samples, thereby providing a robust technical foundation for ONT targeted sequencing in pathogen detection and genomic characterization[35].

5. Conclusion

This study compared the performance of second-generation metagenomic sequencing (mNGS) and third-generation nanopore sequencing (ONT) in molecular epidemiological analyses based on three imported Chikungunya virus (CHIKV) cases. The results demonstrated that both sequencing strategies were able to generate near-complete CHIKV genome sequences sufficient for case genotyping and phylogenetic source tracing, showing a high level of concordance in phylogenetic placement and source attribution. The observed differences between the two approaches were mainly related to sequencing strategy, sequence continuity, and modes of information representation[36].

In port-of-entry surveillance settings where rapid turnaround is critical, amplicon-based ONT sequencing shows practical potential by shortening sequencing time and improving genome continuity, thereby supporting rapid case-level genotyping and preliminary source tracing. In contrast, mNGS, owing to its untargeted nature, remains valuable for broad-spectrum pathogen screening, detection of mixed infections, and surveillance of unknown or unexpected pathogens. Overall, the case-based analyses indicate that mNGS and ONT provide complementary strengths in CHIKV molecular epidemiological investigations[37]. In routine surveillance practice, flexible selection or combined application of these sequencing strategies according to sample characteristics and analytical requirements may enhance genome reconstruction completeness and result robustness, providing more reliable technical support for the monitoring and source tracing of imported CHIKV cases[38].

Declarations

Ethics approval and consent to participate This study was reviewed and approved by the Shenzhen International Travel Health Care Centre Ethics Committee (Approval No: [BJZX20250008]). Written informed consent was obtained from all participants. The study was conducted in accordance with the principles of the Declaration of Helsinki.

Consent for publication Not applicable. No identifiable personal or clinical details, images, or videos of any individual person are included in this manuscript.

Availability of data and materials The datasets generated and/or analyzed during the current study have been deposited in the Genome Sequence Archive (GSA) under the accession number CRA030563 (submission ID: subCRA049893, title: Chikungunya virus). The data are publicly accessible at https://ngdc.cncb.ac.cn/gsa/browse/CRA030563

Authors Contributions: Lianghui Wei was responsible for article writing and experiments, Li Zhu for data analysis,Jianzhong Ye for reagent purchasing, Ran Zhang and Chunchong Zhao for instrument management, Ying Ye for organizing experimental literature, and Ye Yang and Jian'an He for technical guidance and experimental scheme design.

Funding: This work was supported by (2024YFC2310205): Research and Application of Rapid Multi-pathogen Identification Technology Based on Real-time Sequencing; General Administration of Customs Research Project (2024HK058): Research and Development of Intelligent Rapid Recognition Technology for Port Vectors.

Informed Consent Statement: The informed consent of all the research subjects has been obtained.

Competing Interests: All authors declare that there are no conflicts of interest.

Acknowledgements: Thanks are extended to all participants for their valuable advice and support during this study. This article is © 2026 by Lianghui Wei. All rights reserved.

References

1. Javaid, A. et al. An overview of chikungunya virus molecular biology, epidemiology, pathogenesis, treatment and prevention strategies. *Future Virol.* **17**, 593–606 (2022).
2. Ni, J., Li, Z. F., Hu, X. W., Zhou, H. & Gong, Z. Y. Chikungunya's global rebound and Asia's growing vulnerability: Implications for integrated vector control and pandemic preparedness. *Biosci. Trends*. 10.5582/bst.2025.01239 (2025).
3. Webb, E. M. et al. Expression of anti-chikungunya single-domain antibodies in transgenic reduces vector competence for chikungunya virus and Mayaro virus. *Front. Microbiol.* **14**. (2023).
4. Resck, M. E. B. et al. Spatial-temporal distribution of chikungunya virus in Brazil: a review on the circulating viral genotypes and as a potential vector. *Front. Public. Health* **12**. (2024).
5. Dieng, I. et al. An amplicon-based Illumina and nanopore sequencing workflow for Chikungunya virus West Africa genotype. *medRxiv* 10.1101/2023.12.07.23299611 (2023).
6. Cardona-Trujillo, M. C., Ocampo-Cárdenas, T., Tabares-Villa, F. A., Zuluaga-Vélez, A. & Sepúlveda-Arias, J. C. Recent molecular techniques for the diagnosis of Zika and Chikungunya infections: A systematic review. Heliyon – 8:e10225. (2022).
7. Schmitz, K. S. et al. A measles virus-based vaccine induces robust chikungunya virus-specific CD4 + T-cell responses in a phase II clinical trial. *Vaccine* **41**, 6495–6504 (2023).
8. Belem, L. R. W. et al. Development of Multiplex Molecular Assays for Simultaneous Detection of Dengue Serotypes and Chikungunya Virus for Arbovirus Surveillance. *Curr. Issues. Mol. Biol.* **46**, 2093–2104 (2024).
9. Olearo, F. et al. Diagnostic accuracy of 16S rDNA PCR, multiplex PCR and metagenomic next-generation sequencing in periprosthetic joint infections: a systematic review and meta-analysis. *Clin. Microbiol. Infect.* **31**, 1115–1125 (2025).
10. Bolloré, K. et al. Investigation of the viral causes of febrile jaundice in Burkina Faso through metagenomic sequencing. *J. Infect.* 91. (2025).
11. Zheng, N. et al. Metagenomic next-generation sequencing-based characterization of the viral spectrum in clinical pulmonary and peripheral blood samples of patients. *Front. Cell. Infect. Microbiol.* **15**. (2025).
12. Bai, Y., Li, Y., Liu, W., Liu, L. & Tong, Y. Application of next-generation sequencing in the detection of emerging outbreak of virus infection. *Chin. J. Experimental Clin. Virol.* **36**, 739–747 (2022).
13. Wei-ying Chen, Y., Qin, T., Zhang, J., Zou, J. & Yang Zhen-yong Chen.2025.A chromosome-level genome assembly of Termitomyces fuliginosus using Oxford Nanopore and Hi-C sequencing.Genomics.ISSN 0888–7543.
14. Javaran, V. J. et al. *Grapevine Virology in the Third-Generation Sequencing Era: From Virus Detection to Viral Epitranscriptomics* Vol. 10 (Plants-Basel, 2021).
15. Brinkmann, A. et al. Genome sequencing of the mpox virus 2022 outbreak with amplicon-based Oxford Nanopore MinION sequencing. *J. Virol. Methods* **325**. (2024).
16. Dabernig-Heinz, J. et al. A multicenter study on accuracy and reproducibility of nanopore sequencing-based genotyping of bacterial pathogens. *J. Clin. Microbiol.* **62**. (2024).

17. Stefan, C. P., Hall, A. T., Graham, A. S. & Minogue, T. D. Comparison of Illumina and Oxford Nanopore Sequencing Technologies for Pathogen Detection from Clinical Matrices Using Molecular Inversion Probes. *J. Mol. Diagn.* **24**, 395–405 (2022).
18. Wang, J. et al. Detection of the Chikungunya Virus and Sindbis Viruses in Simulated Clinical Samples Using Metagenomic Nanopore Sequencing. *Chin. J. Virol.* **36**, 377–384 (2020).
19. Takemae, N., Kuba, Y., Oba, K. & Kageyama, T. Direct genome sequencing of respiratory viruses from low viral load clinical specimens using the target capture sequencing technology. *Microbiol. Spectr.* **12**. (2024).
20. Ning, X. H., Xia, B. H., Wang, J. Q., Gao, R. & Ren, H. Host-adaptive mutations in Chikungunya virus genome. *Virulence* **15**. (2024).
21. Da, H., Meng, T. & Xu, Y. H. Application of targeted next-generation sequencing for detecting respiratory pathogens in the sputum of patients with pulmonary infections. *Infect. Genet. Evol.* **128**. (2025).
22. Zhang, S. S. et al. Optimisation and clinical validation of a metagenomic third-generation sequencing approach for aetiological diagnosis in bronchoalveolar lavage fluid of patients with. (2025). Ebiomedicine 116.
23. Govender, K. N. et al. Rapid clinical diagnosis and treatment of common, undetected, and uncultivable bloodstream infections using metagenomic sequencing from routine blood cultures with Oxford Nanopore. *medRxiv* 10.1101/2025.01.08.25320182 (2025).
24. de Souza, L. M. et al. Technical comparison of MinIon and Illumina technologies for genotyping Chikungunya virus in clinical samples. *J. Genetic Eng. Biotechnol.* **21**. (2023).
25. Peng, M. F. et al. Short-Read and Long-Read Whole Genome Sequencing for SARS-CoV-2 Variants Identification. (2025). Viruses-Basel 17.
26. Peng, S. Y. et al. 557. Cost-effectiveness of Introducing Metagenomic Next-generation Sequencing (mNGS) for the Diagnosis of Suspected Neurological Infections Patients in China. *Open. Forum Infect. Dis.* **10**. (2023).
27. Rodino, K. G. & Simner, P. J. Status check: next-generation sequencing for infectious-disease diagnostics. *J. Clin. Invest.* **134**. (2024).
28. Gong, W. et al. Outbreak of Chikungunya Virus with Aedes albopictus-Adaptive Mutations - Guangdong Province, China, 2025. *China CDC Wkly.* **7** (49), 1528–1532 (2025).
29. Chen, Y. Y., Guo, Y., Xue, X. H. & Pang, F. Application of metagenomic next-generation sequencing in the diagnosis of infectious diseases of the central nervous system after empirical treatment. *World J. Clin. Cases*. **10** (22), 7760–7771 (2022).
30. Carter, L. L. et al. Global genomic surveillance strategy for pathogens with pandemic and epidemic potential 2022–2032. *Bull. World Health Organ.* **100** (4), 239–239A (2022).
31. Fuhrmann, L., Langer, B., Topolsky, I. & Beerenwinkel, N. VILOCA: sequencing quality-aware viral haplotype reconstruction and mutation calling for short-read and long-read data. *Nar Genomics Bioinf.* **6**. (2024).
32. Carbo, E. C. et al. A comparison of five Illumina, Ion Torrent, and nanopore sequencing technology-based approaches for whole genome sequencing of SARS-CoV-2. *Eur. J. Clin. Microbiol. Infect. Dis.* **42** (6), 701–713 (2023).
33. Bull, R. A. et al. Analytical validity of nanopore sequencing for rapid SARS-CoV-2 genome analysis. *Nat. Commun.* **11** (1), 6272 (2020).
34. Tyson, J. R. et al. Improvements to the ARTIC multiplex PCR method for SARS-CoV-2 genome sequencing using nanopore. bioRxiv [Preprint]. 2020 Sep 4:2020.09.04.283077.
35. Charre, C. et al. Evaluation of NGS-based approaches for SARS-CoV-2 whole genome characterisation. *Virus Evol.* **6** (2), veaa075 (2020).
36. Bogaerts, B. et al. Closing the gap: Oxford Nanopore Technologies R10 sequencing allows comparable results to Illumina sequencing for SNP-based outbreak investigation of bacterial pathogens. *J. Clin. Microbiol.* **62**. (2024).

37. Salazar, C. et al. Fast and cost- effective SARS- CoV-2 variant detection using Oxford Nanopore full- length spike gene sequencing. *Microb. Genomics* **9**. (2023).
38. Cook, R. et al. Nanopore and Illumina sequencing reveal different viral populations from human gut samples. *Microb. Genomics* **10**. (2024).

Figures

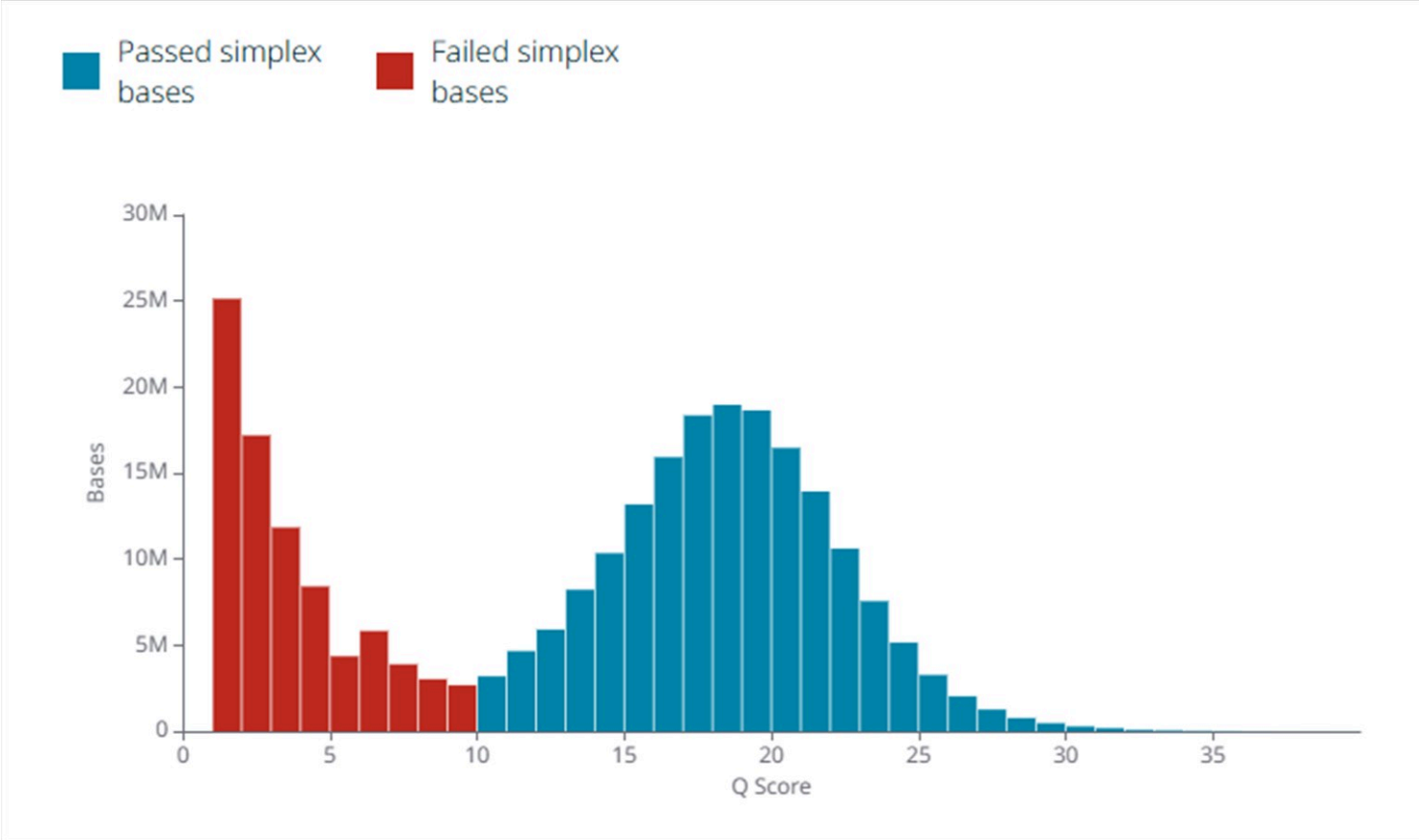


Figure 1

Distribution of passed and failed simplex bases by Q score.(The histogram illustrates the number of simplex bases (y-axis) stratified by their Q score (x-axis), where blue bars represent bases that passed quality control and red bars represent bases that failed quality control)

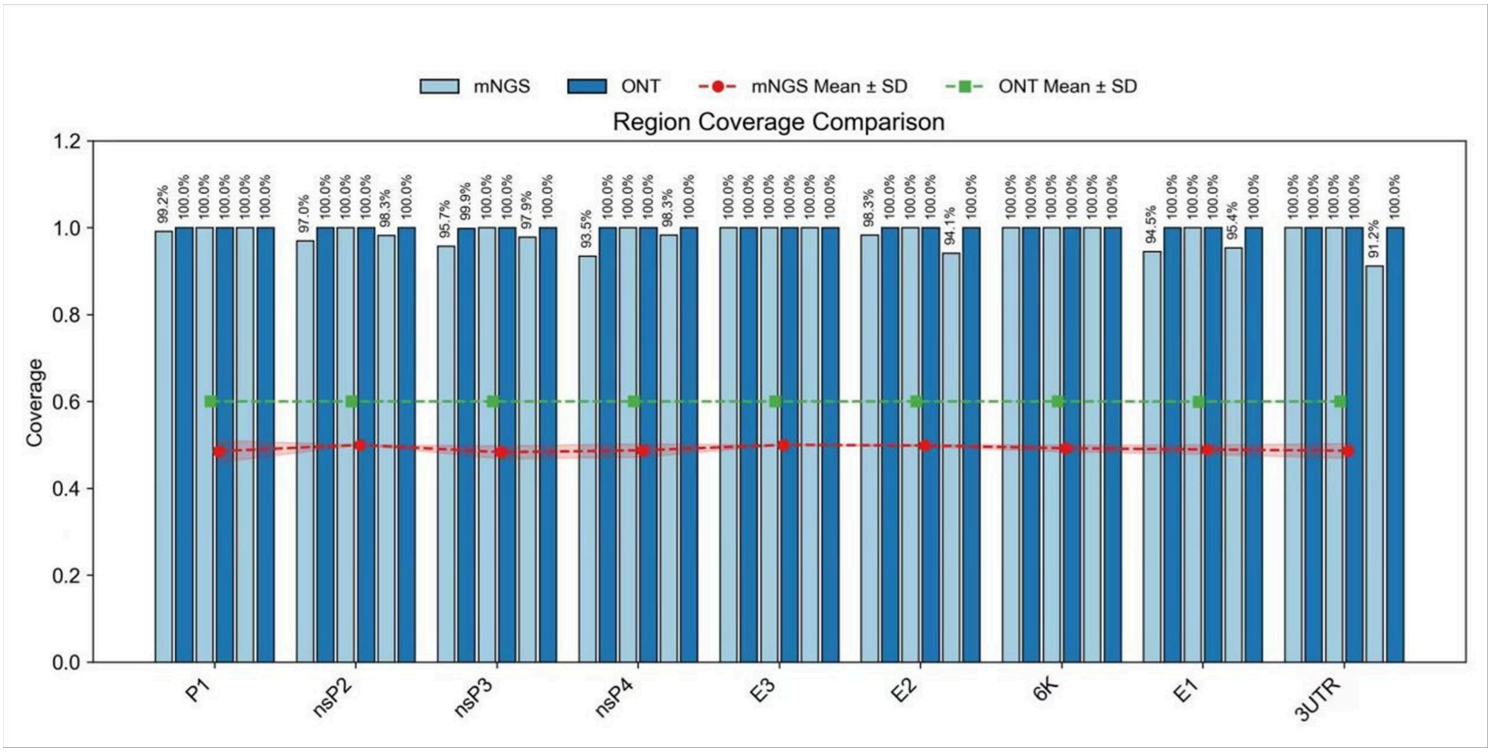


Figure 2

Comparison of regional genome coverage between mNGS and ONT sequencing platforms.(The bar plot shows the coverage of different viral genome regions (x-axis) for both mNGS (light blue bars) and ONT (dark blue bars) sequencing. The red dotted line represents the mean ± standard deviation (SD) of coverage for mNGS, and the green dotted line represents the mean ± SD for ONT.)

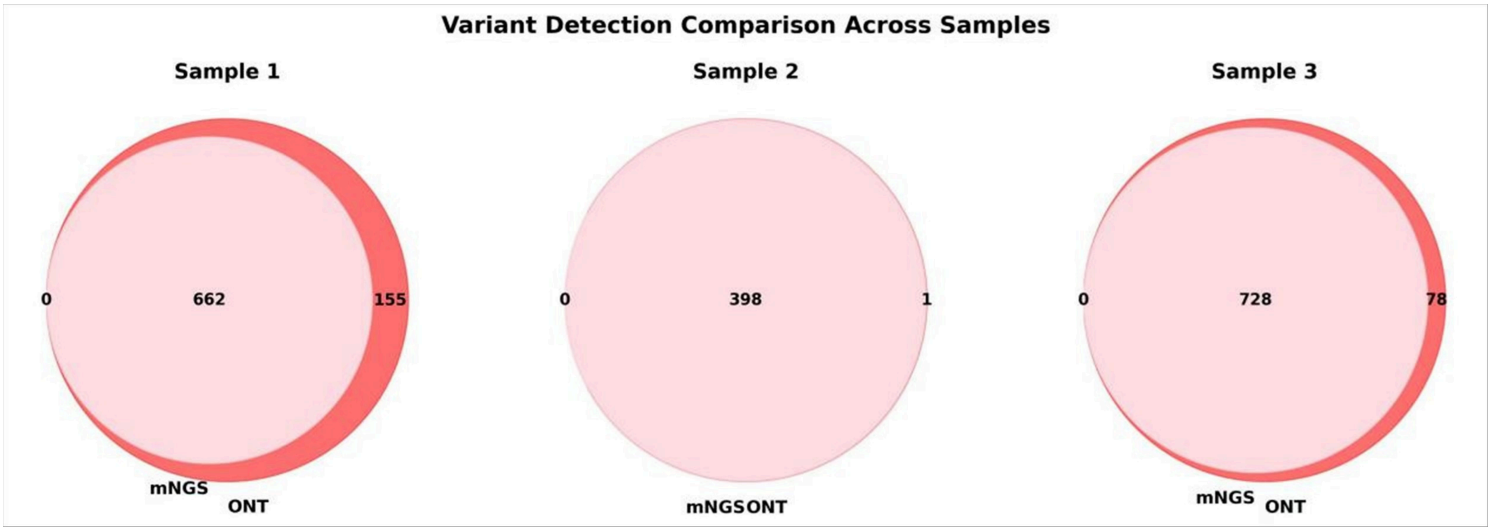


Figure 3

Comparison of single-nucleotide variant (SNV) detection between mNGS and ONT sequencing across three clinical samples.(Each donut plot represents the total number of variants detected in one sample, with the pink core indicating variants identified by mNGS and the red outer ring indicating variants unique to ONT sequencing. Sample 1 had 662 variants detected by mNGS and 155 unique to ONT; Sample 2 had 398 variants detected by mNGS and 1 unique to ONT; Sample 3 had 728 variants detected by mNGS and 78 unique to ONT.)

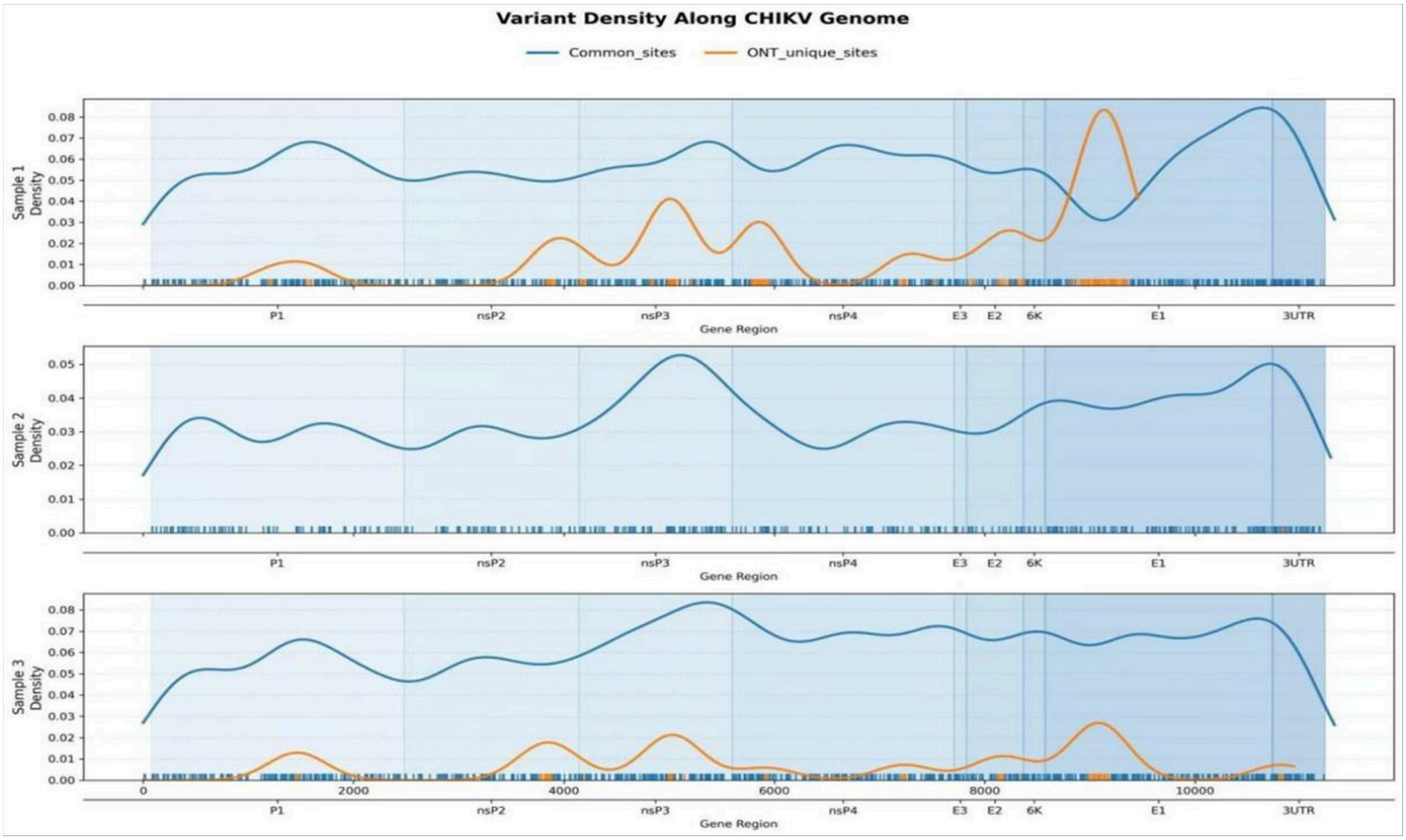


Figure 4

Variant density profiles along the Chikungunya virus (CHIKV) genome across three clinical samples.(Each panel corresponds to one sample, showing the density of shared variants (blue line, Common_sites) and ONT-specific unique variants (orange bars, ONT_unique_sites) across the linear CHIKV genome regions (x-axis, including structural and non-structural gene regions such as P1, nsP2, nsP4, E2, E1). The blue line reflects the overall distribution of variants detected by both sequencing platforms, while the orange bars highlight positions where ONT identified variants not detected by mNGS.)

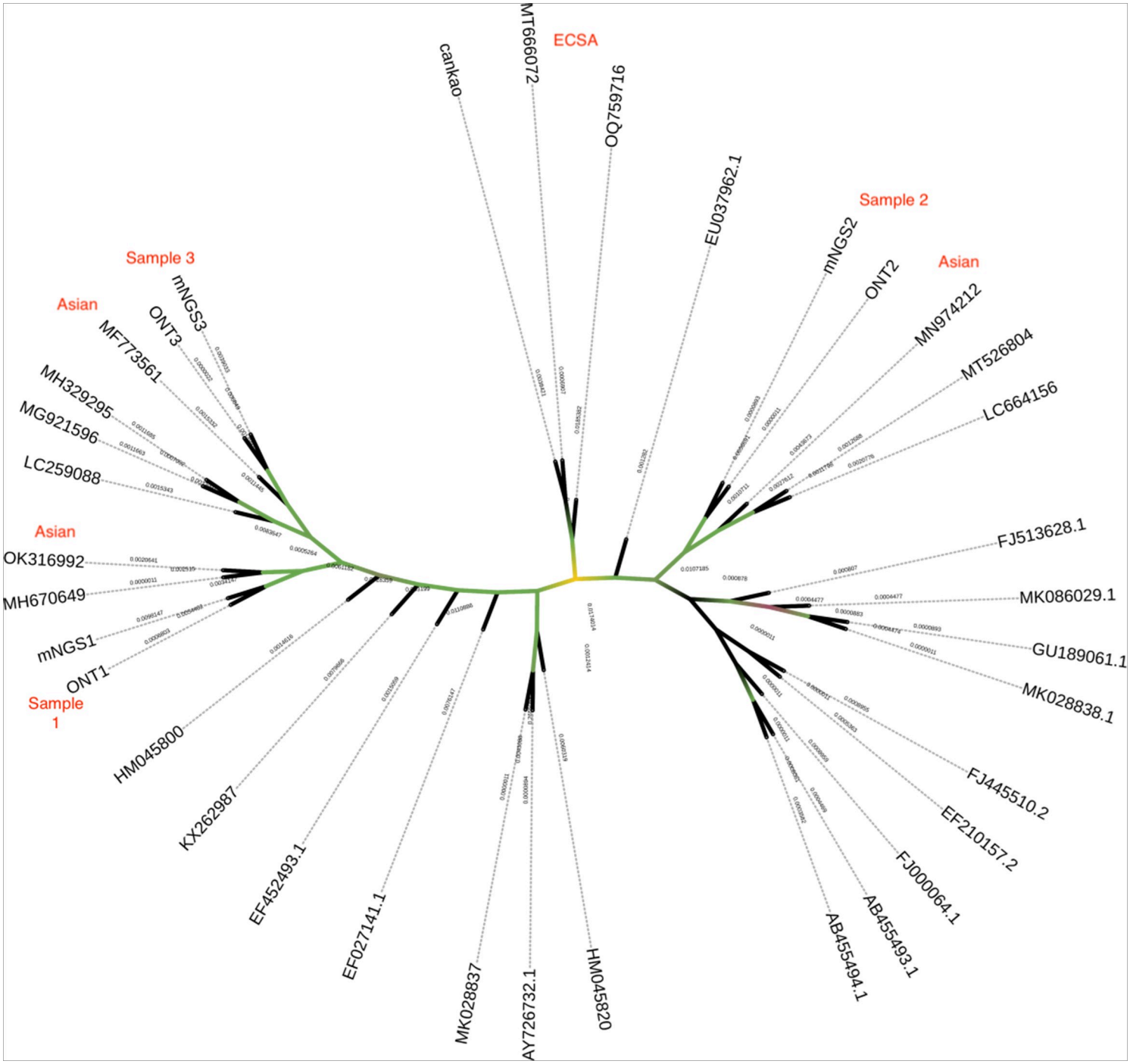


Figure 5

Phylogenetic Tree Source Tracking Analysis.(This maximum-likelihood phylogenetic tree displays the evolutionary relationships of CHIKV strains, with major clades labeled: the Asian clade (red labels) and the ECSA (East/Central/South African) clade. The three study samples (Sample 1, Sample 2, Sample 3) are all clustered within the Asian clade, alongside sequences generated by both mNGS and ONT platforms (marked “mNGS1/2/3” and “ONT1/2”). Black triangles at branch tips represent study-derived sequences, while other tips indicate publicly available reference strains with GenBank accession numbers.)