scientific reports

Check for updates

OPEN

Application of nanopore adaptive sampling for metagenomic detection of tick-borne RNA viruses

Hye-ryung Byun[1], Seong-Ryeong Ji[1], Lexi E. Frank[2], Evan J. Kipp[2], Peter A. Larsen[2] & Joon-Seok Chae[1]✉

Nanopore sequencing is a powerful tool for real-time pathogen detection and genomic characterization; however, its application to individual ticks is limited by abundant host-derived nucleic acids and low viral RNA levels. In this study, we applied nanopore adaptive sampling (NAS) to sequence viral RNA from individual *Haemaphysalis* (*H.*) ticks collected in the Republic of Korea (ROK). By combining NAS with long-read sequencing, high-resolution genome assembly can be achieved from samples containing low-abundance viral RNA and relatively short complementary DNA (cDNA) fragments generated during library preparation. These results indicate that NAS remains effective under suboptimal fragment-size conditions and improves genome assembly compared to conventional nanopore workflows. Phylogenetic analyses revealed that the detected Dabieshan tick virus (DTV) sequences were clustered with isolates from China and Japan, suggesting regional circulation facilitated by the widespread distribution of *H. longicornis*. Unlike previous studies relying on pooled samples without selective sequencing, NAS allowed high-resolution viral genome assembly from single ticks. These findings confirm the presence and genotypes of DTV for the first time in the ROK and demonstrate NAS as a practical, scalable approach for tick-borne RNA virus surveillance in single ticks, improving genomic assembly and supporting the monitoring of emerging tick-borne viruses in endemic regions.

Ticks are distributed worldwide and inhabit a wide variety of environments. As members of the subclass Acari (acarines), they are distinguished from insects by their unsegmented bodies and generally spherical shape[1]. Ticks carry a diverse range of bacterial, viral, and protozoan pathogens, and differ from dipteran vectors in their intrinsic sensitivity to weather and climate[2]. As tick populations increase and their geographic ranges expand, tick-borne diseases remain underdiagnosed despite advances in molecular and immunological detection methods, highlighting the importance of regional awareness for timely diagnosis[3].

Nanopore sequencing, a third-generation sequencing platform developed by Oxford Nanopore Technologies (ONT), enables the generation of massive amounts of high-quality metagenomic data that exceed the capabilities of short-read sequencing and offer powerful long-read approaches[4]. Nanopore long-read sequencing works by passing DNA or RNA molecules through a membrane-embedded, nanometer-scale protein pore such that nucleic acids moving through the pore alter its electrical resistance. These changes in electrical current are measured and processed into sequential data in real time[5,6]. Nanopore adaptive sampling (NAS) is a software-controlled approach that allows selective sequencing of target molecules. Using a programming interface known as ReadUntil, individual nanopores can be dynamically controlled, enabling the software to reject unwanted molecules during sequencing. In this process, the initial few hundred bases of a strand are sequenced, after which the system determines whether the molecule should continue sequencing or be rejected based on target-molecule matching[7,8].

Recent studies have emphasized the potential of NAS as a novel metagenomic strategy for pathogen surveillance across diverse arthropod vectors, enabling the simultaneous detection of multiple bacterial pathogens directly from ticks[9]. In the present study, we used NAS to detect tick-borne RNA viruses circulating in *Haemaphysalis* ticks, which are predominant in Asia. Tick-borne viruses such as severe fever with thrombocytopenia syndrome virus (SFTSV), Dabieshan tick virus (DTV), Alongshan virus, and Yezo virus have been reported in East Asia, reflecting the expanding diversity of tick-borne viruses in the region and underscoring the need for continuous

[1]Laboratory of Veterinary Internal Medicine, BK21 FOUR Future Veterinary Medicine Leading Education and Research Center, Research Institute for Veterinary Science and College of Veterinary Medicine, Seoul National University, 1 Gwanak-ro, Gwanak-gu, Seoul 08826, Republic of Korea. [2]Department of Veterinary and Biomedical Sciences, College of Veterinary Medicine, University of Minnesota-Twin Cities, St. Paul, MN 55108, USA. ✉email: jschae@snu.ac.kr

surveillance[10–13]. By employing host-depletion strategies, we aimed to enrich viral reads and improve sequencing depth across viral genome segments, thereby enabling comprehensive genomic characterization of tick-borne RNA viruses from vector species. This study provides a framework for advancing tick-borne RNA virus research by demonstrating the application of NAS for metagenomic detection of RNA viruses in ticks.

Results

Tick sampling

Four female, six male, four nymphal *Haemaphysalis* (*H.*) *longicornis* ticks, as well as one male *H. flava*, were collected from vegetation and wild animals, including Siberian roe deer (*Capreolus pygargus*), Korean water deer (*Hydropotes inermis*), Northern Goshawk (*Accipiter gentilis*), and Amur hedgehog (*Erinaceus amurensis*). All ticks used in this study had previously tested positive for DTV in our earlier surveillance using nested RT-PCR. NAS was performed across four runs. Run 1 included a single tick, whereas runs 2, 3, and 4 consisted of individually barcoded nanopore libraries prepared from single ticks. All sequencing experiments were generated from individual ticks rather than pooled specimens (Table 1).

Sequencing output and host depletion efficiency

Fourteen tick samples were sequenced using the ONT PromethION platform, generating 2,978,403–7,237,233 reads per sample. The mean read length was relatively short (346–588 bp), with read N50 length values ranging from 337 to 656 bp. The shorter read lengths observed in NAS datasets are an expected consequence of the adaptive sampling process, as non-target reads are actively rejected, and reversed from the pores.

The average read quality varied from 12.1 to 15.3 (Table 2). Representative NanoPlot profiles from the highest-coverage DTV-positive samples (ticks 8 and 14) are shown, indicating nanopore cDNA read length and quality distributions (Fig. 1).

Following adaptive sampling, host depletion efficiency varied among sequencing runs depending on the reference genome used. Host depletion statistics were obtained from the adaptive sampling summary generated by Dorado (v7.9.8) software, which represent the proportion of reads rejected during NAS due to matches with the provided host reference genomes.

The host depletion rate ranged from 37% to 69%, with the highest depletion observed in *H. longicornis* (run 1) when using the *H. longicornis* and the Siberian roe deer genomes. In contrast to the reduced diversity observed in runs 1 and 2, a distinct set of reference genomes, including *Dermacentor silvarum* and *Rhipicephalus microplus* were used in run 3 for comparison. These species were selected because sequences related to these ticks were detected in the random metagenomic reads from runs 1 and 2. Because multiple barcoded tick libraries were sequenced simultaneously within each run, host depletion performance was evaluated at the sequencing run level. However, lower depletion efficiencies (37%) were obtained when multiple tick reference assemblies were used (Table 3). Single-tick run using species-matched reference genomes exhibited higher depletion rates than runs incorporating multiple reference genomes.

Mapping and coverage distribution

Reads were basecalled using Dorado and subsequently mapped to the tick-borne RNA virus reference genomes using minimap2 with the map-ont preset (-ax map-ont) and default parameters optimized for ONT reads (Table 4). A total of nine tick samples were mapped against the reference viral genomes, among these, only DTV was detected, and no significant read enrichment or genome assembly was observed for any other tick-borne

Sample	Tick species	Tick developmental stages	Host animals (scientific name)	Collection date	Run	Notes
Tick 1	*Haemaphysalis longicornis*	Female	Siberian roe deer (*Capreolus pygargus*)	03-July-2022	Run 1	DTV PCR positive
Tick 2	*Haemaphysalis longicornis*	Female	Siberian roe deer	07-Jun-2022	Run 2	DTV PCR positive
Tick 3	*Haemaphysalis longicornis*	Nymph	Korean water deer (*Hydropotes inermis*)	03-Jun-2022		DTV PCR positive
Tick 4	*Haemaphysalis longicornis*	Male	Korean water deer	14-July-2022		DTV PCR positive
Tick 5	*Haemaphysalis longicornis*	Nymph	Northern Goshawk (*Accipiter gentilis*)	03-Jun-2022		DTV PCR positive
Tick 6	*Haemaphysalis longicornis*	Nymph	Northern Goshawk	03-Jun-2022		DTV PCR positive
Tick 7	*Haemaphysalis longicornis*	Male	Siberian roe deer	01-July-2022		DTV PCR positive
Tick 8	*Haemaphysalis longicornis*	Male	Korean water deer	14-July-2022	Run 3	DTV PCR positive
Tick 9	*Haemaphysalis flava*	Male	Amur hedgehog (*Erinaceus amurensis*)	03-Jun-2022		DTV PCR positive
Tick 10	*Haemaphysalis longicornis*	Male	Siberian roe deer	01-July-2022		DTV PCR positive
Tick 11	*Haemaphysalis longicornis*	Nymph	–	15-August-2025		Flagging
Tick 12	*Haemaphysalis longicornis*	Female	–	15-August-2025		Flagging
Tick 13	*Haemaphysalis longicornis*	Female	Siberian roe deer	01-July-2022	Run 4	DTV PCR positive
Tick 14	*Haemaphysalis longicornis*	Male	Korean water deer	14-July-2022		DTV PCR positive
Total	Two species	Two stages	Four animal species	03-Jun-2022 to 15 August 2025	Four runs	

Table 1. Summary of ticks and host animals used for nanopore adaptive sampling.

Sample	No. of reads	Total length (bp)	Read length N50	Mean read quality	Mean read length (bp)
Tick 1	2,978,403	1,031,608,789	337	12.1	346
Tick 2	4,081,299	1,688,573,945	393	15.3	376
Tick 3	3,065,283	1,253,294,224	400	15.3	409
Tick 4	4,700,065	1,934,631,887	413	14.9	412
Tick 5	4,372,163	1,886,372,544	396	15.3	432
Tick 6	5,792,925	2,285,381,697	388	15.3	395
Tick 7	7,237,233	2,814,660,017	398	14.8	389
Tick 8	4,788,296	1,863,617,156	396	14.6	389
Tick 9	4,749,691	2,050,648,409	440	14.5	432
Tick 10	4,203,012	1,700,641,844	423	14.7	405
Tick 11	4,570,167	1,728,130,102	437	13.9	378
Tick 12	3,335,354	1,202,503,101	415	13.6	361
Tick 13	2,399,162	1,024,496,172	389	14.7	427
Tick 14	3,949,226	2,323,416,437	656	15.1	588

Table 2. Summary of raw nanopore adaptive sampling results generated from 14 tick samples.

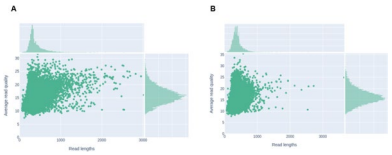

Fig. 1. Read length and quality profiles of the highest-coverage Dabieshan tick virus (DTV)-positive tick sample. (**a**), Tick 8; (**b**), Tick 14. NanoPlot was used to visualize the raw nanopore adaptive sampling (NAS) data generated from the highest-coverage DTV-positive tick sample. Scatter plots show the relationship between read length (x-axis) and average read quality (y-axis), with marginal histograms illustrating the distribution of read lengths (top) and read quality scores (right).

Run	Reference species for depletion	Accession	Data base	Host depletion (%)	Barcoding	Notes
1	*Haemaphysalis longicornis*	GCA_048455015.1	GenBank	69%	No	Tick 1
	Capreolus pygargus	GCA_012922965.1				
2	*Haemaphysalis longicornis*	GCA_048455015.1	GenBank	66%	Yes	Ticks 2–7
	Capreolus pygargus	GCA_012922965.1				
3*	*Dermacentor silvarum*	GCF_013339745.2	Refseq	37%	Yes	Ticks 8–12
	Rhipicephalus microplus	GCF_043290135.1				
4	*Haemaphysalis longicornis*	GCA_048455015.1	GenBank	37%	Yes	Ticks 13 and 14
	Haemaphysalis flava	MT013252				

Table 3. Tick species used for depletion and corresponding host depletion rates. *Note: Run 3 used a distinct set of reference genomes independent from those used in Runs 1 and 2.

viruses included in the reference set. The number of reads ranged from 223 to 2,920 and L segment coverage ranged from 12 to 153x. Genome-wide coverage was calculated using SAMtools and visualized on IGV on a log10 scale. IGV visualization of DTV L segment mapping results from two individual tick samples (Ticks 8 and 14) revealed that both samples displayed consistent read alignment across large portions of the 6,550 bp segment, with higher read depth and more uniform coverage observed in Tick 14 than in Tick 8 (Fig. 2).

Consensus sequence assembly

Consensus sequences were generated and polished using Medaka (v2.1.1) following alignment sorting and indexing with SAMtools. For samples with >11x coverage, near-complete consensus sequences of the DTV L segment were reconstructed. The resulting L segment consensus sequences exhibited 1.1–14.7% N-masking relative to the reference DTV genome, indicating a moderate level of sequence completeness. Among all samples, Ticks 8 and 14 showed the lowest N-masking ratios, 0.4% (24/6,552) and 1.1% (73/6,549), respectively (Table 5). In contrast, the DTV S segment exhibited substantially lower read coverage across most tick samples, resulting in consensus sequences with high levels of N-masking (30.2–85.8%) (Table 5).

Virus	Segment	Accession number	Data base
Severe fever with thrombocytopenia syndrome virus (SFTSV)	L, M, S	GCA_003087695.1	GenBank
Crimean–Congo haemorrhagic fever virus (CCHFV)	L, M, S	GCF_000854165.1	Refseq
Tick-borne encephalitis virus (TBEV)	One segment	GCF_000863125.1	Refseq
Langya virus	One segment	GCA_031319335.1	GenBank
Yezo virus	L, M, S	GCF_029888495.1	Refseq
Jingmen tick virus (JTV)	Four segments	GCF_000919875.1	Refseq
Dabieshan tick virus (DTV)	L, S	GCF_013086875.1	Refseq
Heartland virus	L, M, S	GCF_013086545.1	Refseq
OZ virus	Six segments	GCF_004117175.1	Refseq
Alongshan virus	Four segments	GCA_027256855.1	GenBank
Thogoto virus	Six segments	GCF_000854245.1	Refseq
Kyasanur forest disease virus	One segment	GCF_002820625.1	Refseq
Alkhurma hemorrhagic fever virus	One segment	GCA_031116565.1	GenBank
Bhanja virus	L, M, S	GCF_001019855.1	GenBank
Okutama tick virus	L, S	LC753204.1, LC753204.1	GenBank
Tacheng tick virus	L, M, S	MN379437, NC_031285, MN379436	GenBank
Dugbe virus	L, M, S	GCF_000850985.1	Refseq

Table 4. Reference tick-borne RNA virus complete genomes and their accession numbers used for nanopore adaptive sampling.

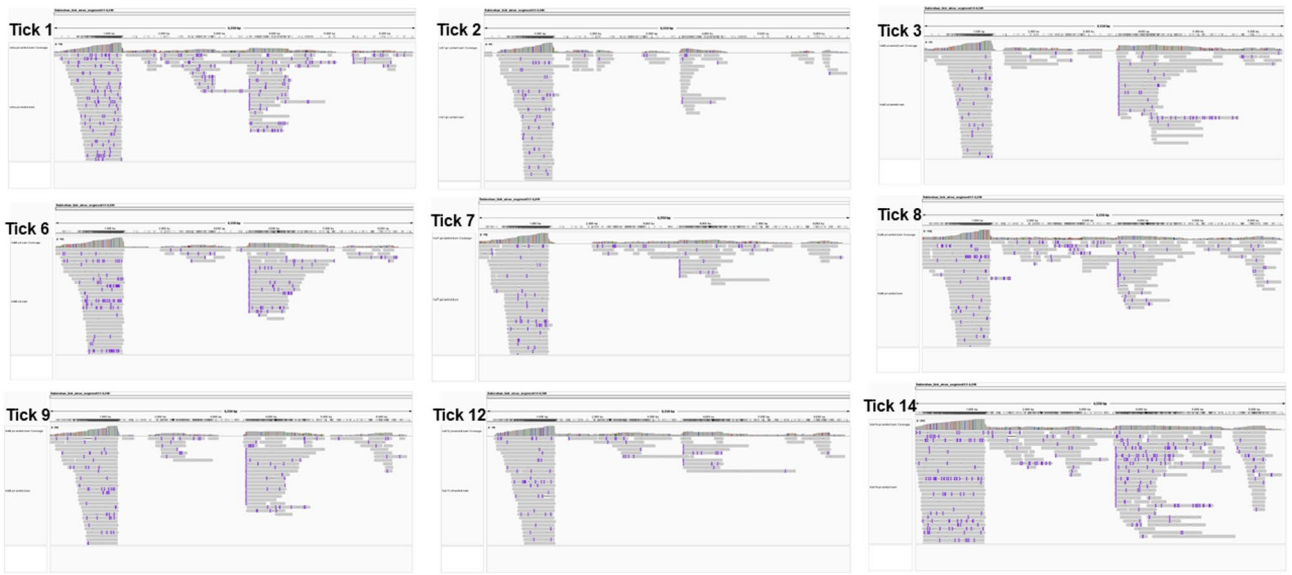

Fig. 2. IGV visualization of Dabieshan tick virus (DTV) L-segment coverage and read alignment in nine DTV-positive tick samples detected by nanopore adaptive sampling (NAS). Mapping profiles are shown across the 6,550-bp reference L segment. The upper tracks display per-base read depth and coverage distribution, while the lower panels depict individual nanopore read alignments. Gray bars represent aligned reads, and purple marks indicate mismatches relative to the reference DTV genome.

Phylogenetic relationship

A total of 45 DTV L segment sequences, including nine sequences generated in this study from the Republic of Korea (ROK), were used for phylogenetic analysis. The final alignment consisted of 6,473 nucleotide sites, of which 5,408 sites (83.55%) were conserved. A total of 345 sites were parsimony-informative, and the dataset contained 648 distinct site patterns, indicating sufficient phylogenetic signal for robust tree reconstruction. ModelFinder identified GTR + F+I+R2 as the best-fitting nucleotide substitution model based on the Bayesian Information Criterion (BIC). The best-scoring tree had a log-likelihood value of -18776.138. The maximum likelihood tree was rooted using the Hubei-derived strain (OP682840), which served as the outgroup and represented the basal lineage of DTV (designated as Group A). The nine ROK sequences generated in this study were distributed across two distinct genotype groups (Groups B and D). Several ROK strains clustered closely with China-derived sequences, suggesting a shared evolutionary origin between the two regions. Japan-

	Isolate name	GenBank accession	Segment	No. of reads	Coverage (x)	Min length (bp)	Max length (bp)	Consensus length (bp)	N count (%)
Tick 1	GW11T	PX636690	DTV L segment	882	35	170	1,817	6,544	332 (5.1%)
	–	–	DTV S segment	20	3	280	599	1,787	1,127 (63.1%)
Tick 2	GW6T	PX660579	DTV L segment	665	21	165	1,818	6,550	332 (5.1%)
	–	–	DTV S segment	19	2	256	852	1,784	1,200 (67.3%)
Tick 3	GW19T	PX677678	DTV L segment	326	15	160	2,442	6,542	524 (8%)
	–	–	DTV S segment	6	0.6	258	509	1,787	1,330 (74.3%)
Tick 6	GW10T	PX660580	DTV L segment	621	27	169	2,801	6,546	743 (11.3%)
	–	–	DTV S segment	92	8	195	879	1,783	886 (49.7%)
Tick 7	GW19T	PX636689	DTV L segment	850	35	179	2,191	6,539	938 (14.3%)
	–	–	DTV S segment	9	0.8	308	701	1,788	1,392 (77.8%)
Tick 8	GW3T	PX590545	DTV L segment	1764	64	166	2,765	6,552	24 (0.4%)
	-	–	DTV S segment	13	1	256	751	1,789	1,204 (67.3%)
Tick 9	GW10T	PX677679	DTV L segment	889	33	171	3,230	6,547	960 (14.7%)
	–	–	DTV S segment	5	0.6	247	771	1,787	1,534 (85.8%)
Tick 12	GW10T	PX660581	DTV L segment	223	12	182	3,393	6,543	227 (3.5%)
	–	–	DTV S segment	13	2	330	1,095	1,783	1,061 (59.5%)
Tick 14	GW39T	PX660578	DTV L segment	2920	153	162	4,203	6,548	73 (1.1%)
	–	–	DTV S segment	38	7	302	1,596	1,779	538 (30.2%)

Table 5. Detection of Dabieshan tick virus (DTV) in nine ticks, including read counts, coverage, genome length, consensus length, and N-masking results. Note: Ticks sharing the same isolate name were collected from the same host animal; consensus lengths represent untrimmed consensus sequences, with low-coverage regions retained as ambiguous bases (N).

associated lineages (Groups B and C) exhibited heterogeneous evolutionary patterns, with some sequences interspersed among the ROK and Chinese strains, while others formed an independent, well-supported clade. Bootstrap support values were consistently high across major nodes, confirming the robustness and reliability of the inferred genotype structure. All sequences generated in this study have been deposited in National Center for Biotechnology Information (NCBI) GenBank (accession numbers: PX677678, PX677679, PX660578-PX660581, PX636689, PX636690, and PX590545) (Fig. 3; Table 5).

Discussion

This study characterized DTV genomes directly from individual *Haemaphysalis* ticks using ONT NAS. By integrating NAS with long-read sequencing, viral sequences can be enriched, enabling viral detection and genome assembly even in samples with low viral abundance. Previous studies have suggested that short DNA molecules may limit NAS enrichment efficiency because of the proportionally longer time required for basecalling, alignment, and rejection relative to read duration[8]. In the present study, we demonstrate that NAS can still be effective in assembly near-complete DTV genomes, even when applied to samples with relatively short mean cDNA fragment sizes (346–588 bp). This observation can be explained by the NAS mechanism, in which non-target host reads are rejected during sequencing, thereby reducing host-derived background reads and allowing viral reads to accumulate to sufficient depth for genome assembly. Nevertheless, relatively short cDNA fragments may reduce the efficiency of host depletion. When fragments are too short, they may pass through the nanopore before a rejection decision can be made during adaptive sampling, reducing the efficiency of host read rejection. Although previous studies have suggested that the utility of NAS may be limited in complex samples, our results demonstrate that NAS can still be an effective approach for enriching viral genomes from host-dominated tick samples. These results suggest that NAS remains effective under suboptimal fragment size conditions and may provide advantages for viral enrichment in complex tick-derived metatranscriptomics.

A recent study demonstrated the versatility of NAS for adaptive sampling in vector biology, showing that NAS can recover full-length 16 S rRNA sequences and improve pathogen detection efficiency in arthropod vectors[9]. Similarly, a previous host-depleted nanopore sequencing approach achieved a 9–50% reduction in host reads and enabled near-complete viral genome assembly from clinical samples[14]. Our findings extend these observations by demonstrating that NAS can be applied to individual ticks and remains effective even when host depletion is moderate. In this study, NAS facilitated the assembly of near-complete DTV genomes from single-tick samples without pooling, highlighting its potential for high-resolution virome surveillance and genomic characterization of emerging tick-borne viruses.

Recent studies on nanopore-based tick analyses have focused on microbial profiling, including bacterial endosymbionts and tick-borne pathogens such as *Anaplasma*, *Borrelia*, *Rickettsia*, and *Francisella*, and have typically relied on pooled tick samples without adaptive sampling[15–17]. Nanopore sequencing has also enabled the genomic detection of several zoonotic tick-borne viruses, including SFTSV, Crimean–Congo hemorrhagic fever virus (CCHFV), and Jingmen tick virus, in both ticks and clinical samples[18–20]. However, these approaches are limited in their ability to efficiently deplete host genomes from single ticks. In contrast, our study demonstrates

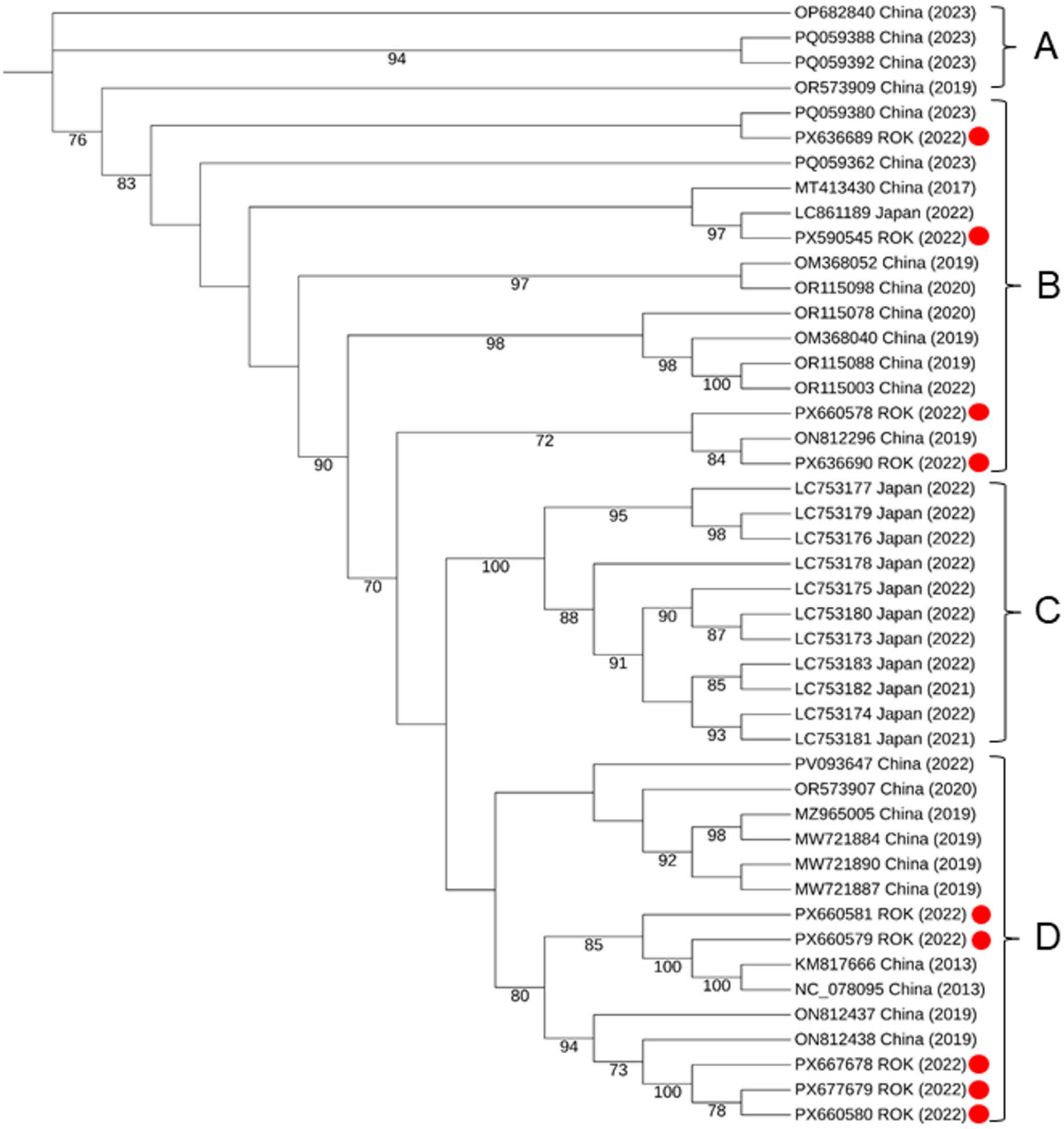

Fig. 3. Maximum likelihood phylogenetic tree was constructed using 45 Dabieshan tick virus (DTV) L-segment (6,473 bp) nucleotide sequences, including nine sequences generated in this study (bold). The tree was rooted using a Hubei-derived strain (OP682840), which served as the outgroup and is designated as Group A. Genotypes were delineated based on monophyletic clustering and are indicated by distinct colors. Bootstrap support values were calculated using 1,000 ultrafast replicates and are shown on major nodes. Major clades were defined based on bootstrap support values ≥ 70%, and values > 90% indicate strongly supported nodes. Korean strains are indicated by red dots.

that NAS selectively enriched viral RNA while reducing host-derived reads, enabling near-complete genome assembly for individual ticks without the need for pooling.

In addition to methodological advances such as NAS, sequencing platform characteristics are also relevant to metaatranscriptomic analyses. Second- and third-generation sequencing platforms are widely used for metatranscriptomic analyses but differ in sequencing characteristics. Second-generation platforms such as Illumina generate highly accurate short reads, whereas third-generation platforms such as ONT allow real-time

long-read sequencing[21,22]. In this study, NAS was applied to reject host-derived sequences in real time during sequencing, thereby enriching viral reads directly from individual tick samples.

Phylogenetic analysis of the complete L segment revealed that the DTV strains identified in this study were closely related to strains previously detected in China. Pairwise nucleotide identities among the DTV L-segment sequences obtained in this study ranged from 97.6% to 98.8%.

Given the wide distribution of *H. longicornis* across East Asia, these findings suggest its potential role in regional viral circulation and broader cross-transmission within tick populations[23]. Since DTV was first detected in *H. longicornis* in China, the virus has been reported in diverse regions, including Zhoushan, Jilin, Hubei, Japan, and the ROK, indicating a broader geographic distribution than initially recognized[11,24–28]. Although the pathogenicity and transmission mechanisms of DTV remain unclear, its expanding detection range suggests that it may pose a potential health concern for animals and humans, underscoring the need for continued genomic surveillance[28].

To better understand the evolutionary relationships among regional DTV strains, the phylogeny was rooted using the earliest reported DTV strain from Hubei, China, which is considered the most ancestral lineage based on a previous study[29]. Using this strain as an outgroup, the results revealed a clear phylogeographical structure across East Asia. The Korean sequences consistently clustered with Chinese strains, suggesting that DTV circulating in the ROK likely shares a common evolutionary origin with mainland Chinese lineages and may have been introduced through historical or ongoing ecological connectivity between regions. Japanese sequences appeared interspersed with Korean and Chinese strains near basal nodes, but ultimately formed a distinct, well-supported clade, indicating that although DTV may have entered Japan via shared ancestral lineages, it has since undergone localized diversification and is now circulating independently. Collectively, these findings support cross-border viral movement followed by country-specific evolutionary trajectories in East Asia.

This study has several limitations. First, only a limited number of samples were analyzed, and all ticks pre-confirmed as DTV-positive by nested RT-PCR were subjected to NAS, which may have introduced selection bias. In addition, viral load differences among tick samples were not quantified using RT-qPCR due to insufficient residual RNA, and variation in DTV titers may therefore have influenced genome recovery efficiency. Second, although NAS was effective, relatively short cDNA fragments may reduce host depletion efficiency. Because adaptive sampling requires an initial sequence length for basecalling and alignment, fragments shorter than this decision may pass through the nanopore before a rejection decision can be made, thereby decreasing the efficiency of host read rejection. Third, only the L segment was consistently assembled across samples, whereas the S segment exhibited very low coverage. This discrepancy may reflect differences in viral abundance, selective degradation, or amplification biases during cDNA synthesis. Future studies incorporating parallel sequencing using conventional ONT or Illumina approaches would help to further evaluate the enrichment efficiency of NAS in environmental metagenomic samples.

Overall, our results demonstrate that NAS is a practical and rapid approach for detecting tick-borne RNA viruses directly from individual ticks. This study provides a methodological framework that can support future large-scale virome surveillance and facilitate the early detection of emerging tick-borne viruses. Furthermore, integrating NAS with ecological or host-associated metadata could offer new opportunities to investigate virus–vector interactions, transmission pathways, and environmental factors shaping viral diversity within tick populations. Continued refinement of depletion references, library preparation strategies, and sequencing depth will likely enhance NAS efficiency and expand its utility for comprehensive tick virome studies.

Materials and methods

Ethical approval

This study was approved by the Institutional Animal Care and Use Committee (IACUC) of Seoul National University (SNU; IACUC No. SNU-220708-4 and SNU-220708-4-1).

Tick collection and identification

A total of 14 ticks were collected from wild animals rescued at the Gangwon Wildlife Rescue Center between July and September 2022. Using a flagging method, in which vegetation was swept with a white cotton flag (1.0 × 1.0 m) attached to a wooden stick, on August 15, 2025. All collected ticks were rinsed with distilled water and 70% ethanol and were stored at -80 °C until further analysis. Two stereomicroscopes were used for tick identification, an OLYMPUS SZ microscope with WF20X/12 mm eyepieces (Japan) and an OLYMPUS SZH10 research stereomicroscope with GWH10X-CD eyepieces (Japan). Tick species and developmental stages were determined based on morphological characteristics, as previously described[30].

RNA extraction

Individual ticks were homogenized in lysis buffer using a TissueLyser II® (Qiagen, Hilden, Germany) equipped with two 5-mm stainless-steel beads at a frequency of 30 Hz for 7 min. After homogenization, samples were centrifuged at 13,000 × g for 1 min at 4 ℃. The supernatant was used for genomic RNA extraction, which was performed using Intron Patho Gene-spin™ DNA/RNA Extraction Kit (iNtRON®, Seongnam, Korea) according to the manufacturer's instructions.

Complementary DNA (cDNA) synthesis and PCR clean-up

Extracted RNA was reverse transcribed into cDNA using the SMARTer® PCR cDNA Synthesis Kit (Takara Bio, Shiga, Japan) according to the manufacturer's instructions. The resulting cDNA was amplified using the Advantage® 2 PCR Kit (Takara Bio, Shiga, Japan) for 24 cycles, following the manufacturer's instructions, to obtain sufficient total cDNA for downstream library preparation. The amplified products were purified using the Nucleospin Gel and PCR Clean-up Kit (Macherey-Nagel, Düren, Germany) to remove residual primers

and enzymes. Prior to library preparation, purified cDNA concentrations were quantified using a Qubit 4 fluorometer (Invitrogen, Carlsbad, United States).

Library preparation and PromethION sequencing

Libraries were prepared using the Native Barcoding Kit 24 v14 (SQK-NBD114-24) according to the manufacturer's instructions. A total of 14 individual tick samples were used for library preparation, with 1 µg of total cDNA used as input for each library prior to native barcoding for multiplex sequencing. Sequencing was performed in four independent runs on the PromethION platform using one R10.4.1 flow cell (FLO-PRO114M) per run. Run 1 consisted of a single tick library, whereas Runs 2–4 included multiplexed libraries generated using native barcoding (Run 2: Ticks 2–7; Run 3: Ticks 8–12; Run 4: Ticks 13–14). For adaptive sampling, indexed reference files in .mmi format, containing the complete genome sequences of 17 emerging or novel tick-borne RNA viruses, were used as target references to enrich viral reads while depleting host-derived sequences (NAS depletion approach). To prioritize removal of the dominant host background, tick genomes were provided as depletion references for adaptive sampling prior to sequencing. Sequencing was initiated and monitored using the MinKNOW software (v25.05.12 and v25.05.14). Raw data were generated in pod5 format and basecalled using Dorado (https://github.com/nanoporetech/dorado) with a 400-bps super-accurate model. Quality filtering (average Phred quality score ≥ 9), adapter trimming, and conversion from pod5 to FASTQ format were performed using Dorado. The filtered data were automatically stored in the pod5pass directory. Each sample yielded approximately 1.5 Gb of sequencing data, providing near-complete genome coverage. The library preparation and PromethION sequencing were performed at the National Instrumentation Center for Environmental Management (NICEM).

Bioinformatic processing

Reads with an average Phred quality score ≥ 9 were used for the analysis. All bioinformatics analyses were conducted in a Windows Ubuntu environment v24.04.02 (Windows subsystem for Linux, WSL). Quality statistics from NAS depletion nanopore sequencing experiments were assessed using NanoPlot v1.42.0 (https://github.com/wdecoster/NanoPlot). Concatenated FASTQ files from each tick sample were processed using seqkit v2.3.0, minimap2 v2.26-r1175, and SAMtools v1.19.2 software packages to evaluate read mapping and coverage statistics[31–33]. Filtered reads were mapped to reference genomes of representative tick-borne RNA viruses used for NAS (Table 4). For DTV-positive samples, consensus sequences for the L segment were generated using the DTV reference genome listed in Table 4. Consensus sequences were generated and polished from mapped reads using Medaka v2.1.1 (https://github.com/nanoporetech/medaka). Sequences with an average sequencing depth ≥ 10× were used for consensus generation, while low-coverage regions were represented as ambiguous bases (N) in the final consensus sequences. Only primary alignments were retained for downstream analyses, whereas secondary and supplementary alignments were excluded using the same tools. Read depth, coverage, and alignment distributions were visualized using Integrative Genomics Viewer[34](IGV, v2.19.5).

Phylogenetic relationships

Phylogenetic trees were constructed using the final consensus sequences generated and polished from mapped reads. Nucleotide sequences of the DTV L segments were aligned using BioEdit v.7.2. ClustalW multiple alignment. Phylogenetic trees were constructed using IQ-TREE v3.0.1 with the maximum likelihood method[35]. The best-fitting substitution model was selected using ModelFinder[36]. Branch support was evaluated using 1,000 ultrafast bootstrap replicates[37]. For tree rooting, a DTV strain (accession number: OP682840) previously identified in Hebei, China was used as the outgroup[24]. The resulting phylogenetic trees were visualized and edited using iTOL v7.

Data availability

All genome sequences generated in this study are available in the NCBI GenBank database (accession numbers: PX677678, PX677679, PX660578–PX660581, PX636689, PX636690, and PX590545). The raw nanopore sequencing data generated in this study have been submitted to the NCBI Sequence Read Archive (SRA) under accession numbers SRR36764186-SRR36764199 (BioSample: SAMN54554392-SAMN54554405; BioProject: PRJNA1400646).

Received: 5 January 2026; Accepted: 22 April 2026
Published online: 26 April 2026

References

1. Boulanger, N., Boyer, P., Talagrand-Reboul, E. & Hansmann, Y. Ticks and tick-borne diseases. *Med. Mal Infect.* **49**, 87–97 (2019).
2. Lee, J. S. & Chung, S. Y. The threat of climate change on tick-borne infections: rising trend of infections and geographic distribution of climate risk factors associated with ticks. *J. Infect. Dis.* **227**, 295–303 (2023).
3. Madison-Antenucci, S., Kramer, L. D., Gebhardt, L. L. & Kauffman, E. Emerging tick-borne diseases. *Clin. Microbiol. Rev.* **33**, e00083–e00018 (2020).
4. Albertsen, M. Long-read metagenomics paves the way toward a complete microbial tree of life. *Nat. Methods*. **20**, 30–31 (2023).
5. Wick, R. R., Judd, L. M. & Holt, K. E. Performance of neural network basecalling tools for Oxford nanopore sequencing. *Genome Biol.* **20**, 129 (2019).
6. Noakes, M. T. et al. Increasing the accuracy of nanopore DNA sequencing using a time-varying cross membrane voltage. *Nat. Biotechnol.* **37**, 651–656 (2019).
7. Loose, M., Malla, S. & Stout, M. Real-time selective sequencing using nanopore technology. *Nat. Methods*. **13**, 751–754 (2016).
8. Martin, S. et al. Nanopore adaptive sampling: A tool for enrichment of low abundance species in metagenomic samples. *Genome Biol.* **23**, 11 (2022).

9. Kipp, E. J. et al. Metagenomic surveillance for bacterial tick-borne pathogens using nanopore adaptive sampling. *Sci. Rep.* **13**, 10991 (2023).
10. Yu, X. J. et al. Fever with thrombocytopenia associated with a novel bunyavirus in China. *N Engl. J. Med.* **364**, 1523–1532 (2011).
11. Li, C. X. et al. Unprecedented genomic diversity of RNA viruses in arthropods reveals the ancestry of negative-sense RNA viruses. *eLife* **4**, e05378 (2015).
12. Wang, Z. D. et al. A new segmented virus associated with human febrile illness in China. *N Engl. J. Med.* **380**, 2116–2125 (2019).
13. Kodama, F. et al. A novel nairovirus associated with acute febrile illness in Hokkaido, Japan. *Nat. Commun.* **12**, 5539 (2021).
14. Han, X. & Xia, Z. Application of host-depleted nanopore metagenomic sequencing in the clinical detection of pathogens in pigs and cats. *Anim. (Basel).* **13**, 3838 (2023).
15. Carrera-Faja, L. et al. Microbiome-based surveillance of zoonotic tick-borne pathogens from urban wild boars in Barcelona, 2022–2023. *One Health*. **20**, 101022 (2025).
16. Du, L. F. et al. Genome-resolved metagenomics reveals microbiome diversity across 48 tick species. *Nat. Microbiol.* **10**, 2631–2645 (2025).
17. White, M. L. et al. A full-length 16S survey using nanopore sequencing to uncover the bacterial microbiome in *Ixodes ricinus* ticks from a single UK woodland, collected across three springs (2019–2021). *Ticks Tick. Borne Dis.* **16**, 102506 (2025).
18. Lee, J. et al. Whole-genome sequencing and genetic diversity of severe fever with thrombocytopenia syndrome virus using multiplex PCR-based nanopore sequencing, Republic of Korea. *PLOS Negl. Trop. Dis.* **16**, e0010763 (2022).
19. Ergunay, K. et al. Impact of nanopore-based metagenome sequencing on tick-borne virus detection. *Front. Microbiol.* **14**, 1177651 (2023).
20. Prayitno, S. P. et al. Amplicon-based MinION sequencing complements severe fever with thrombocytopenia syndrome (SFTS) diagnosis via real-time RT-PCR in patients with suspected SFTS. *J. Korean Med. Sci.* **40**, e69 (2025).
21. Slatko, B. E., Gardner, A. F. & Ausubel, F. M. Overview of next-generation sequencing technologies. *Curr. Protoc. Mol. Biol.* **122**, e59 (2018).
22. Garcia-Pedemonte, D. et al. Comparison of nanopore and synthesis-based next-generation wequencing platforms for SARS-CoV-2 variant monitoring in wastewater. *Int. J. Mol. Sci.* **24**, 17184 (2023).
23. Zhao, L. et al. Distribution of *Haemaphysalis longicornis* and associated pathogens: Analysis of pooled data from a China field survey and global published data. *Lancet Planet. Health*. **4**, e320–e329 (2020).
24. Xu, X. et al. Discovery and vertical transmission analysis of Dabieshan Tick Virus in *Haemaphysalis longicornis* ticks from Chengde, China. *Front. Microbiol.* **15**, 1365356 (2024).
25. Zhu, C. et al. Distribution and phylogenetic analysis of Dabieshan tick virus in ticks collected from Zhoushan, China. *J. Vet. Med. Sci.* **82**, 1226–1230 (2020).
26. Pérez-Sautu, U. et al. Novel viruses in hard ticks collected in the Republic of Korea unveiled by metagenomic high-throughput sequencing analysis. *Ticks Tick. Borne Dis.* **12**, 101820 (2021).
27. Mekata, H., Kobayashi, I. & Okabayashi, T. Detection and phylogenetic analysis of Dabieshan tick virus and Okutama tick virus in ticks collected from Cape Toi, Japan. *Ticks Tick. Borne Dis.* **14**, 102237 (2023).
28. Han, X. et al. Spatial risk of *Haemaphysalis longicornis* borne Dabieshan tick virus (DBTV) in China. *J. Med. Virol.* **96**, e29373 (2024).
29. Liu, N. et al. Molecular evolution and spatial transmission of Dabieshan tick virus in Jilin Province, northeastern China. *IScience*. 28, 112838 (2025).
30. Yamaguti, N., Tipton, V. J., Keegan, H. L. & Toshioka, S. Ticks of Japan, Korea, and the Ryukyu Iislands. Brigham Young University Science Bulletin, Biological Series, **15** 1–227 (1971). Brigham Young Univ.
31. Li, H. et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
32. Li, H. Minimap2: Pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094–3100 (2018).
33. Shen, W., Le, S., Li, Y., Hu, F. & SeqKit A cross-platform and ultrafast toolkit for FASTA/Q file manipulation. *PLOS One*. **11**, e0163962 (2016).
34. Robinson, J. T. et al. Integrative genomics viewer. *Nat. Biotechnol.* **29**, 24–26 (2011).
35. Nguyen, L. T., Schmidt, H. A., von Haeseler, A. & Minh, B. Q. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **32**, 268–274 (2015).
36. Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A. & Jermiin, L. S. ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nat. Methods*. **14**, 587–589 (2017).
37. Hoang, D. T., Chernomor, O., von Haeseler, A., Minh, B. Q. & Vinh, L. S. UFBoot2: Improving the ultrafast bootstrap approximation. *Mol. Biol. Evol.* **35**, 518–522 (2018).

Acknowledgements

We thank the Gangwon Wildlife Rescue Center for providing the samples essential for this study.

Author contributions

H.R.B. conceptualized and designed the sequencing experiments, performed the data analysis, and wrote the original draft of the manuscript.S.R.J., L.E.F., and E.J.K. assisted with the experiments and data analysis.P.A.L. and J.S.C. co-supervised the study.All authors reviewed and approved the final version of the manuscript.

Funding

This work was supported by the Seoul National University Office of Research Affairs research grant (Grant ID 550-20250009).

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to J.-S.C.