Recovery of Complete Genome Sequences of Crimean-Congo Haemorrhagic Fever Virus (CCHFV) Directly from Clinical Samples: A Comparative Study Between Targeted Enrichment and Metagenomic Approaches

Jake D'Addiego, Nadina Wand, Babak Afrough, Tom Fletcher, Yohei Kurosaki, Hakan Leblebicioglu, Roger Hewson

PII: S0166-0934(23)00158-1

DOI: https://doi.org/10.1016/j.jviromet.2023.114833

Reference: VIRMET114833

To appear in: *Journal of Virological Methods*

Received date: 26 July 2023
Revised date: 6 October 2023
Accepted date: 20 October 2023

Please cite this article as: Jake D'Addiego, Nadina Wand, Babak Afrough, Tom Fletcher, Yohei Kurosaki, Hakan Leblebicioglu and Roger Hewson, Recovery of Complete Genome Sequences of Crimean-Congo Haemorrhagic Fever Virus (CCHFV) Directly from Clinical Samples: A Comparative Study Between Targeted Enrichment and Metagenomic Approaches, *Journal of Virological Methods,* (2023) doi:https://doi.org/10.1016/j.jviromet.2023.114833

Recovery of Complete Genome Sequences of Crimean-Congo Haemorrhagic Fever Virus (CCHFV) Directly from Clinical Samples: A Comparative Study Between Targeted Enrichment and Metagenomic Approaches

Jake D'Addiego[a,b,*], Nadina Wand[a], Babak Afrough[a], Tom Fletcher[c], Yohei Kurosaki[d], Hakan Leblebicioglu[e], Roger Hewson[a,b,c,d]

[a]UK Health Security Agency, Science Group, Porton Down, Salisbury, United Kingdom

[b]Faculty of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London, UK.

[c]Department of Clinical Sciences, Liverpool School of Tropical Medicine, Liverpool, UK

[d]National Research Centre for the Control and Prevention of Infectious Diseases, Nagasaki University, Japan.

[e]VM Medicalpark Samsun Hospital, Samsun, Turkey

*Corresponding author.

Abstract

Crimean-Congo haemorrhagic fever (CCHF) is the most prevalent human tick-borne viral disease, endemic to the Balkans, Africa, Middle East and Asia. There are currently no licensed vaccines or effective antivirals against CCHF. CCHF virus (CCHFV) has a negative sense segmented tripartite RNA genome consisting of the small (S), medium (M) and large (L) segments. Depending on the segment utilised for genetic affiliation, there are up to 7 circulating lineages of CCHFV. The current lack of geographical representation of CCHFV sequences in various repositories highlights a requirement for increased CCHFV sequencing capabilities in endemic regions.

We have optimised and established a multiplex PCR tiling methodology for the targeted enrichment of complete genomes of Europe 1 CCHFV lineage directly from clinical samples and compared its performance to a non-targeted enrichment approach on both short-read and long-read sequencing platforms.

We have found a statistically significant increase in mapped viral sequencing reads produced with our targeted enrichment approach. This has allowed us to recover near complete S segment sequences and above 90% of the M and L segment sequences for samples with Ct values as high as 31.3. This study demonstrates the superiority of a targeted enrichment approach for recovery of CCHFV genomic sequences from samples with low virus titre.

CCHFV is an important vector-borne human pathogen with wide geographical distribution. The validated methodology reported here adds value to front-line public health laboratories employing genomic sequencing for CCHFV Europe 1 lineage surveillance, particularly in the Balkan and Middle Eastern territories currently monitoring the spread of the pathogen. Tracking the genomic evolution of the virus across regions improves risk assessment and directly informs the development of diagnostics, therapeutics, and vaccines.

Keywords: CCHFV, Targeted Enrichment, Tiling Amplicon Sequencing, Next-Generation Sequencing

1. Introduction

Crimean-Congo haemorrhagic fever (CCHF) is the most prevalent human tick-borne viral disease, endemic to the Balkans, Africa, Middle East and Asia (World Health Organization - WHO, 2022). Outbreaks have been reported in 26 countries spanning three continents from the Far East to Southern Europe and West Africa, with Turkey reporting the highest number of CCHF cases(1) worldwide. Whilst the average case fatality rate for CCHF in symptomatic patients is between 10% and 40%(2), it can be as high as 80% in some nosocomial settings (3).

The disease is caused by Crimean-Congo haemorrhagic fever virus (CCHV), an orthonairovirus within the *Nairoviridae* family (order *Bunyavirales*). The virus is commonly transmitted to humans by a bite of an infected tick, predominantly in the warmer spring and summer months when adult stages of *Hyalomma* ticks and other vector species are seeking their blood meals for maturation(4). Transmission through direct contact with viraemic bodily fluids from infected hosts and patients has also been observed, principally among abattoir and healthcare workers. Whilst wild animals, including small mammals, birds and larger ungulates are susceptible to CCHFV infection, they do not display clinical symptoms, and act as transient reservoirs for the virus(4), for which *Hyalomma* ticks remain the main reservoirs. Despite ongoing global efforts to develop a vaccine against CCHFV(5), there are currently no licensed vaccines or effective treatments against the virus. Although ribavirin has been used to treat CCHF in some endemic regions, data on its clinical efficacy is inconclusive(6), and recent studies have shown only a limited mutagenic effect on CCHFV genomes *in vivo*(7).

CCHFV is one of the most genetically diverse arboviruses(8), with a segmented tripartite negative sense RNA genome consisting of the small (S), medium (M) and large (L) segments encoding respectively the viral nucleoprotein (NP), the precursor to the viral glycoproteins (GPC) and the viral RNA-dependent RNA polymerase(9). The M segment exhibits the highest sequence divergence of 31%, followed by the L segment (22%) and the S segment (20%)(4). Depending on the segment utilised for genetic classification, there are 6 (M and L segments) or 7 (S segment) CCHFV lineages (10) which are traditionally named after their main geographic location: Africa 1, Africa 2 and Africa 3; Asia 1 and Asia 2, and Europe 1 and Europe 2(11).

With recent advances in sequencing technologies, including the release of field-friendly and portable sequencers such as the Oxford Nanopore Technologies (ONT) MinION device, whole genome sequencing has become more affordable, and offers scientists the opportunity to characterise the genetic diversity of viruses beyond what could be accomplished by more traditional sequencing approaches.

Of 3,777 CCHFV sequences deposited in the Bacterial and Viral Bioinformatics Resource Center (BV-BRC)(12), almost 60% belong to viruses from Russia, with Turkish viruses representing only 8.76% of all sequences despite Turkey being a key hotspot of CCHF disease and reporting the highest number of cases annually. Although CCHFV sequences from other geographical areas are gradually being deposited in public repositories, there is still poor representation of complete CCHFV genome sequences, limiting accurate genetic characterisation of CCHFV in disease endemic regions.

Generally, the recovery of partial viral genome sequencing data has found applications in clinics for some time, such as in the detection of drug resistance markers in viruses like HIV(13) and human cytomegalovirus (HCMV)(14) for example. The recovery of complete genome sequences from clinical samples, however, is often hampered by low levels of target virus genomes, which although these can be in the region of millions of copies per microlitre, are often overwhelmed by the host's nucleic acids.

Several approaches have been developed for enriching target viral nucleic acids and/or depleting the host's genetic material to increase the amount of generated microbial sequencing reads.

Among the methods focusing on enriching microbial nucleic acids, targeted and non-targeted approaches have been published. One of the major advantages of non-targeted approaches is that no

sequence information is required *a priori*, making such methods particularly suitable to study whole microbial communities in environmental samples(15, 16). Of such non-targeted approaches, Sequence-Independent, Single-Primer Amplification (SISPA) is a primer-initiated method that requires specific sequence modification to achieve non-specific amplification(17). SISPA has been successfully utilised to sequence complete or near complete genomes of several viruses including Influenza virus(18), SARS-CoV-2(19) Lassa virus(20), CCHFV and Jingmen tick virus(21) directly from clinical samples and tick homogenates.

Though some of the difficulties in recovering complete genome sequences from clinical samples have been addressed by non-targeted approaches(18-21), samples with low virus titre and high cycle-threshold (Ct) values still present a challenge to such methods, with genome coverages reported to decline dramatically above Ct values of 30(19).

Polymerase chain reaction (PCR) amplification sequencing is a desirable approach for targeted enrichment due to its high-specificity, relatively low cost, and straight forward assay design(14). Multiplex tiling PCR approaches have been published for Tick-Borne Encephalitis Virus (TBEV)(22), SARS-CoV-2(23), Zika Virus (ZIKV)(24) and recently Respiratory Syncytial Virus (RSV) A and B(25). These schemes work by amplifying fragments across the entire genome of the target virus in alternate primer pools, with a consensus sequence being called in regions with sequencing depths between 1x and 20x depending on the study(22, 24). Compared to non-targeted enrichment strategies, the study by Quick *et al.* (2017) reported coverages greater than 90% with depths of sequencing in the region of hundreds (nanopore sequencing) or thousands (Illumina sequencing) of reads from samples with Ct values as high as 35.9(24) with a tiling multiplex PCR approach, highlighting the superiority of such approaches for samples with low virus titre. A targeted multiplex PCR enrichment approach has also been successfully utilised to detect emerging SARS-CoV-2 variants of concern (VOC) directly from environmental samples(23), providing further insights into additional applications of such methodologies beyond mere genetic virus characterisation.

In this study, we have developed a targeted, multiplex PCR approach for the amplification of Europe 1 CCHFV genomes directly from Turkish clinical samples and have compared its performance to the non-targeted SISPA approach utilising both long read (ONT) and short read (Illumina) sequencing technologies.

2. Methods

2.1. Primer design for targeted enrichment

Primer pairs targeting 500bp fragments across complete CCHFV genome segments were designed using the online tool Primal Scheme(24). As a reference, unpublished in-house generated S, M and L sequences from clinical isolates of Europe 1 lineage of CCHFV were utilised, most closely related to sequences from a Turkish CCHFV isolate from 2004 (GenBank accession numbers: MF511213, S segment; MF511230, M segment; MF511196, L segment). All designed primer sequences for each pool can be found in Supplementary Table S1 (A and B).

2.2. Sample Collection

Clinical serum samples were collected during a previous study from patients admitted with confirmed CCHF in Turkey, between May 2015 and August 2016 at Ondokus Mayis University (OMU) Hospital, Tokat State Hospital and Tokat University Hospital, and made available for this work. Original diagnosis was confirmed by positive CCHF polymerase-chain reaction (PCR) or IgM enzyme-linked immunosorbent assay. All participants provided written consent with research ethics committee approval from the UK Ministry of Defense Research and Ethics Committee (MoDREC 571) and Ondokus Mayis University Research and Ethics Committee (OMU KAEK 2014/739), Turkey.

2.3. Sample Processing and RNA Extraction

Serum samples were inactivated in a containment level (CL) 4 laboratory; 140 µl of each sample were added to 560 µl of buffer AVL and 560 µl of ethanol. RNA was extracted utilising the QIAamp Viral RNA Mini Kit (Cat. No. 52904, Qiagen) following the manufacturer's instructions. RNA was eluted in 60 µl of nuclease-free water and stored at -80 °C.

2.4. Real-Time PCR

All extracted RNAs were quantified in duplicate utilising an in-house real-time PCR assay as previously described(7). Briefly, 5 µl of extracted RNA were amplified in final 20 µl reactions containing 0.5 µM of forward primer CCHF-F1 (5'-CCC TTT TTA AAC TCC TCA AAC C-3'), 0.5 µM of forward primer CCHF-F2 (5'-CCT TTT TTA AAC TCC TCA AAC C-3'), 1 µM of reverse primer CCHF-R (5'- TCT CAA AGA AAC ACG TGC C-3'), 0.5 µM of probe CCHF-P (5'FAM-ACT CAA GAG AAC ACT GTG GGC GTA AG-3'MGBEQ) with TaqMan™ Fast Virus 1-Step Master Mix (cat. no. 4444436, Thermo Fisher Scientific). Primers/Probe binding sites are provided in Supplementary Table S3. Copies of CCHFV genomic RNA in the extracts were determined based on quantified synthetic S Segment RNA standard curves. Quantitative real-time RT-PCR was performed on a QuantStudio 7 Flex Real-Time PCR System (Applied Biosystems) platform with the following cycling parameters: 50 °C for 10 min, 95 °C for 2 min and 45 cycles of 95 °C for 10 s and 60 °C for 30 s.

2.5. Target Enrichment

cDNA from extracted RNA was synthesised using SuperScript™ IV First-Strand Synthesis System (Thermo Fisher Scientific, cat. 18091050) in final 20 µl reactions. Briefly, 11 µl of extracted RNA were mixed with 50 ng of random hexamers (50 ng/µl) and 0.5 mM dNTP mix and incubated at 65 °C for 5 minutes. RNA and annealed primer mixes were then placed on ice for 1 minute, then mixed with 1x SSIV Buffer, 5 mM DTT, 1 µl of RNase Out™ (Invitrogen, cat. no. 100777019) and 200 units of SSIV RT enzyme before being incubated at 42 °C for 50 minutes and 70 °C for 10 minutes. Two aliquots of cDNA were prepared for each sample. cDNA was stored at -20 °C.

Target enrichment was based on a previously published protocol(26). Two separate primer pools for each CCHFV genome segment are prepared, each containing primers targeting alternate fragments. Details of each primer pool with final primers concentrations are provided in Supplementary Table S2. 2.5 µl of cDNA is amplified in triplicate 25 µl reactions containing either one of two sets of primer pools utilising Q5 High-Fidelity 2X Master Mix (NEB, cat. no. M0492L). A maximum of three replicate reactions were prepared for each sequencing platform to achieve the recommended concentration for sequencing library preparation. Master-mix composition is provided in Supplementary Table S2. cDNA was amplified in a thermocycler with the following cycling parameters: 98 °C for 30 s, 30 cycles of 98 °C for 15 s, 57 °C for 30 s and 70 °C for 30 s and held at 4 °C. PCR amplicons were cleaned up with AMPure XP Reagent (Beckman Coulter, cat. no. A63881) and stored at -20 °C.

2.6. Sample Preparation for Metagenomic Sequencing

22 µl of extracted RNA were treated with TURBO™ DNase (Thermo Fisher Scientific, cat. no. AM2239) in final 50 µl reactions. RNA was mixed with 0.8X TURBO™ DNase buffer and 2 units of TURBO™ DNase enzyme and incubated at 37 °C for 30 minutes.

DNase-treated RNA was purified using the RNA Clean and Concentrator-5 kit (Zymo Research, cat. no. R1016). Randomly amplified cDNA was prepared using a SISPA approach according to a previously described protocol(18). For round A, 4 µl of cleaned-up RNA are annealed with 1 µl of 40 µM Sol-A primer (5'-GTTTCCCACTGGAGGATANNNNNNNNN-3') at 65 °C for 5 minutes, then incubated at 21 °C for 5 minutes. First strand cDNA is synthesised in 10 µl reactions containing 1X SSIV Buffer, 1mM of dNTPs, 10U of SSIV reverse-transcriptase, 5mM of DTT and 5µl of annealed RNA and incubating at

42 °C for 10 minutes. Second strand cDNA is synthesised utilising Sequenase™ Version 2.0 DNA polymerase (cat. no. 70775Y200UN, Thermo Fisher Scientific) by adding 5 µl of 1X Sequenase buffer and 0.39 units of Sequenase Enzyme followed by an incubation at 37 °C for 8 minutes, then by adding 0.6 µl of 3.25 units of Sequenase enzyme diluted in Sequenase dilution buffer and incubating further at 37 °C for 8 minutes. For round B, samples are amplified with AccuTaq™ LA DNA Polymerase (Merck, cat. no. D8045) in 50 µl reactions containing 1X reaction buffer, 2µM Sol-B primer (5'-GTTTCCCACTGGAGGATA-3'), 1 µl of DMSO, 0.5mM dNTPs and 2.5 units of AccuTaq enzyme. Reactions are incubated at 98°C for 30 seconds followed by 30 cycles of 94°C for 15 seconds, 50°C for 20 seconds and 68 °C for 2 minutes and by a final extension cycle at 68 °C for 10 minutes. PCR products were cleaned up with AMPure XP Reagent and eluted with molecular grade water in final 30 µl volumes.

2.7. Sequencing library preparation

PCR amplicons and SISPA B products were cleaned up with AMPure XP Reagent with a 2:1 beads to sample ratio following the manufacturer's protocol, and quantified with Qubit™ fluorometer's dsDNA High-Sensitivity assay (Thermo Fisher Scientific, cat. no. Q32854) following the manufacturer's protocol.

Illumina sequencing libraries were prepared from 1.65 ng cDNA using Nextera XT V2 kit (Illumina, cat. no. FC-131-1096) and sequenced on a 2 x 150 bp paired-end Illumina MiSeq instrument operated by the UK Health Security Agency (UKHSA) Genomics Services Development Unit (Colindale).

Barcoded MinION sequencing libraries were prepared from 65 ng of cDNA with the Ligation Sequencing kit SQK-LSK-110 and Native Barcoding Kit EXP-NBD196 (ONT). End repair and dA tailing was performed with NEBNext® Ultra™ II End Repair/dA-Tailing Module (New England Biolabs - NEB, cat. no. E7546L) in 30 µl reactions containing 65 ng of cDNA, 3.5 µl of Reaction Buffer and 1.5 µl of Enzyme Mix and incubating at room temperature for 10 minutes followed by an incubation at 65 °C for 10 minutes. Ligation of ONT specific barcodes/adapters was done by adding 1X Blunt/TA Ligase Master Mix (cat. no. M0367L, NEB) and 2.5 µl of barcode directly to the end-repaired reactions followed by an incubation at room temperature for 20 minutes and at 65 °C for 10 minutes. Barcoded samples were then cleaned-up in three pools containing a maximum of 22 samples each and normalised to 65 ng in 45 µl. Adapter ligation was performed with NEBNext® Ultra™ II Ligation Module (NEB, cat. E7595L) by adding 5 µl of adapter, 40 µl of Ligation Master Mix and 1 µl of Enhancer (NEB) directly to the barcoded library and incubating at room temperature for 30 minutes. Adapter-ligated libraries were cleaned-up and washed with Short Fragment Buffer (SFB) (SQK-LSK-110 kit) and eluted in 15 µl of Elution Buffer before final quantification.

Libraries were sequenced for 24 h on FLO-MIN106D flow cells using a Mk1C MinION device (ONT).

2.8. Bioinformatic Analysis

Raw Illumina sequencing reads were processed with Trimmomatic with the following parameters: ILLUMINACLIP: NexteraPE-PE.fa:2:30:10:8 LEADING:30 TRAILING:30 SLIDINGWINDOW:10:20 MINLEN:50. Further adapter trimming and quality filtering were performed with BBDuk with the following settings: -Xmx1g in1=R1.fastq.gz in2=R2.fastq.gz out1=R1_clean.fastq.gz out2=R2_clean.fastq.gz ref=nextera.fa.gz ktrim=r k=23 mink=11 hdist=1 tpe tbo qtrim=rl trim=20. Illumina reads were mapped against a reference genome (accession numbers OQ866631-OQ866633 for the S, M and L segments respectively) utilizing BWA MEM with kmer length set to 35. Nanopore sequencing data was filtered based on length (0.4-0.6kb) and quality (minimum 8) directly on the MinION data output options. Nanopore sequencing reads were mapped against the reference utilising minimap2(27). Reference coverage and sequencing depth statistics were derived utilising SAMtools(28) depth, coverage and flagstat functions.

Consensus sequences were generated for a subset of samples with Ct values ranging between 17.5 and 29.9. Sequences with identified frameshifts were excluded from downstream phylogenetic analysis. For targeted enrichment Illumina data, consensus sequences were derived from sorted BAM files with iVar(29). Consensus sequences from Illumina metagenomic data were derived from sorted BAM files using an in-house C++ program QuasiBAM(30), with minimum sequencing depth for consensus calling set to 10x.

For ONT data, sequencing reads were first quality filtered and trimmed using Filtlong(31) prior to mapping. Consensus sequences from metagenomic data were derived from sorted BAM files utilising BCFtools(28), masking regions with coverage below 10x. For targeted enrichment data, primer sequences were first removed from sorted BAM files with BAMClipper(32) prior to consensus generation with BCFtools(28) as described previously. All generated consensus sequences have been deposited in GenBank with the following accession numbers: OQ454691-OQ454750 (Illumina sequencing data) and OQ454777-OQ454836 (ONT sequencing data).

Pairwise identities were calculated with Geneious Prime® (version 2023.0.2).

2.9. Phylogenetic Analysis

Maximum likelihood phylogenies were generated with MEGA(33) (version 7.0.26) applying the general time reversible model. Sequences were aligned using MUSCLE(34). Alignments were trimmed to exclude low coverage regions to a final length of 1572 bp (S segment), 3057 bp (M segment) and 5102 bp (L segment). Bootstrap support values were generated with 1000 replicates.

2.10. Statistical Analysis

Statistical significance was determined with a T Test using Microsoft Office's Excel software. Reported *p* values are significant at the 5% level.

3. Results

3.1. Extracted RNA samples presented with a wide range of cycle threshold (Ct) values.

Extracted RNA from a total of 23 CCHFV positive serum samples was quantified utilising an in-house real-time PCR assay. Average Ct values and RNA quantities expressed as copies per millilitre of serum are presented in Table 1.

The Ct values ranged between 17.5 and 34.8, with one sample having an undetermined result. The recorded RNA quantities were between 5.06×10^3 and 5.12×10^8 copies per millilitre of serum (Table 1). Input number of genome copies for each enrichment methodology ranged between 2.6×10^2 and 2.6×10^7 genome copies.

3.2. Targeted enrichment produced a statistically significant increase in mapped sequencing reads.

Illumina sequencing generated a total of 39.2M sequencing reads which passed quality control. Targeted enrichment generated on average 25.5% more sequencing reads compared to the non-targeted SISPA approach ($P < 0.05$) with a total of 22.5M and 16.8M sequencing reads respectively. A statistically significant difference in mapped sequencing reads was also found between the two methods ($P < 0.05$). SISPA generated on average 69,974 mapped sequencing reads per sample, ranging

between 4 and 748,230 reads (Figure 1A). Targeted enrichment produced on average 608,511 mapped sequencing reads per sample, ranging between 124,535 and 944,278 reads (Figure 1B).

Oxford Nanopore Technologies (ONT) sequencing generated 2.6M reads which passed quality and length filters. The total reads generated with targeted enrichment and SISPA was 53,188 and 51,412 respectively ($P = 0.81$). A statistically significant difference in mapped sequencing reads was found between the two methodologies ($P < 0.05$). SISPA produced on average 7,356 mapped sequencing reads per sample, ranging between 4,004 and 105,325 reads (Figure 1C). Targeted enrichment produced on average 48,145 mapped sequencing reads per sample, ranging between 11,019 and 100,638 reads (Figure 1D).

3.3. Near complete genome reference coverages with increased sequencing depths were obtained with targeted enrichment.

Illumina sequencing data coupled with SISPA produced, on average, reference genome coverages of 68.79%, 53.03%, 34.99% and 10.94% with ranges of 1%-100%, 0%-100%, 0%-99.96% and 0%-98.97% at 1X, 10X, 100X and 1000X sequencing depths respectively across the entire CCHFV genome (Figure 2A). With targeted enrichment, the average reference genome coverages were 93.56%, 86.04%, 78.91% and 63.02% with ranges of 54.45%-99.74%, 19.77%-98.20%, 16.23%-95.33% and 13.15%-87% at 1X, 10X, 100X and 1000X sequencing depths respectively (Figure 2B).

ONT sequencing coupled with SISPA produced on average reference genome coverages of 84.71%, 57.26%, 31.58% and 3.12% with ranges of 5%-100%, 0%-100%, 0%-99.77% and 0%-35.34% at 1X, 10X, 100X and 1000X sequencing depths respectively (Figure 2C). With targeted enrichment, the average reference genome coverages were 98.29%, 92.06%, 74.95% and 42.79% with ranges of 94.59%-99.97%, 63.68%-98.12%, 6.87%-91.63% and 0%-75.11% at 1X, 10X, 100X and 1000X sequencing depths respectively (Figure 2D).

3.4. Targeted enrichment produced on average a 6.55-fold increase in sequencing depth across all CCHFV genomic segments.

The average sequencing depth across all samples recorded with Illumina sequencing following SISPA was 778.61, 384.7 and 401.68 for the S, M and L segment respectively (Figure 3A). With targeted enrichment, the average sequencing depths were 4305.73, 2283.64 and 3194.36 (Figure 3B) for the S, M and L segments respectively, with an average 6.25-fold increase in sequencing depth across all segments.

ONT sequencing data coupled with SISPA produced average sequencing depths of 301.45, 126.6 and 130.94 for the S, M and L segments respectively (Figure 3C). Targeted enrichment produced sequencing depths of 2165.79, 717.20 and 1243.32 (Figure 3D) for the S, M and L segments respectively, with an average 7.38-fold increase in sequencing depth across all segments.

3.5. Phylogenetic analysis of the generated consensus sequences for each genomic segment confirms genetic affiliation to the Europe 1 CCHFV lineage.

Generated consensus sequences cluster with previously published Europe 1 CCHFV sequences (Figure 4). For samples 6, 22, 23 and 29, insufficient sequencing depth and reference coverage were achieved with the SISPA approach to generate consensus sequences.

The average alignment pairwise distances of the generated sequences were 0.81% (S segment), 1.3% (M segment) and 0.6% (L segment), with overall pairwise distances of 11.5%, 19.3% and 7.3% for the S, M and L segments respectively. Pairwise identities between generated consensus sequences for each sample were on average 99.99%, 98.72% and 99.39% for the S, M and L segments respectively, with an overall average pairwise distance of 99.36%.

4. Discussion

Bioinformatic analysis of viral sequencing data relies on either mapping sequencing reads against a reference genome (reference-based assembly) or looking for overlaps between the reads to generate larger fragments (*de novo* assembly)(35), followed by extraction of a consensus sequence which is often based on majority voting(36). Different sequencing platforms have varying levels of error rates, which when taken on their own can resemble genuine variants; nonetheless having several reads mapping to the same region increases confidence in base calling and support for minor variants(37). Very high sequencing depths are generally not required for consensus calling, with most bioinformatic pipelines using cut-offs of 1x(38) or 10x(39) sequencing depths; however for other downstream applications such single nucleotide variant (SNV) calling, depths of 400x have been recommended with a multiplex PCR approach(40), further supporting the requirement of targeted enrichment approaches for samples with low virus titre. Whilst metagenomic approaches have been vastly utilised for virus discovery and molecular surveillance, the large proportion of host sequencing reads generated from clinical samples makes recovery of complete genome sequences challenging with such non-targeted approaches.

Here, we adapted an amplicon tiling strategy previously reported by Quick *et al.*(24) for sequencing Europe 1 (or genotype V) CCHFV which has produced a statistically significant increase in mapped sequencing reads with both Illumina and nanopore technologies ($P < 0.05$, Figure 1). Whilst Illumina technology also resulted in a statistically significant difference in total sequencing reads generated by the two enrichment methodologies possibly due to the different number of PCR cycles between the different methods ($P < 0.05$), nanopore sequencing did not ($P = 0.81$). The lack of a significant difference in total generated sequencing reads is likely due to the limited amount of nanopores available for sequencing which become depleted over time and are therefore only able to produce a finite amount of sequencing reads in a single run, suggesting that higher sequencing depths could be achieved by running fewer samples in a single flow-cell.

Kafetzopoulou *et al.* (2018)(41) have found that whilst a higher proportion of target viral reads are generally generated for samples with lower Ct values, genome reference coverages above 90% of at least 20x sequencing depths were also observed for samples with Ct values as high as 30. The study by Lewandowski *et al.* (2019) however, reported that only 63% of samples with a Ct value below 30 produced >90% complete genomes, with reference coverage dropping dramatically between Ct values of 25 and 30(18), suggesting that significant variation is observed between different sample types.

In our study, with SISPA enrichment coupled with both Illumina and nanopore sequencing, genome coverages >90% with at least 10x sequencing depths were achieved only for samples with Ct values below 27, with reference genome coverage declining beyond this value (Figure 2A and 2C). In contrast our targeted enrichment approach showed greater than 90% reference genome coverage with average sequencing depths of at least 10x were achieved with both sequencing platforms up to Ct value 31.3 (Figure 2B and 2D), supporting the superiority of a targeted enrichment strategy for samples with titres as low as few thousand copies per millilitre of serum.

Figure 2 also shows an increase in reference genome coverage also at 100x and 1000x sequencing depths, which presented with average reference sequence coverage increases of 43.92% and 52.08% for 100x and 1000x depths respectively with Illumina and increases of +44.68% and +38.81% for 100x and 1000x depths respectively with nanopore sequencing. Whilst not strictly required for consensus generation, the study by Liu *et al.* (42) has shown that the false positive rate for variant calling utilising ONT data drops to zero when a minimum sequencing depth of 200x is used, suggesting that higher sequencing depths are preferable to minimise the introduction of errors in generated consensus sequences.

The S and L segments exhibited on average the highest sequencing depths across both sequencing platforms and methodologies (Figure 3). Although the reduced sequencing depth in the M segment compared to the S and the L segments for the targeted enrichment can be attributed to the greater genetic diversity of the M segment, the same cannot be said with the non-targeted approach, suggesting

perhaps a lower starting number of M segment copies in the extracted viral RNA, possibly due to the presence of defective M-segment free particles, speculated on previously(43).

Additional reference coverage could have been achieved by lowering the minimum depth required for consensus calling to 1x, as done by Zakotnik *et al.* in their TBEV study, however, we have maintained a minimum depth of 10x to increase confidence.

Although we have observed a loss of sequencing depth below 10x in certain regions of the M and L segments (Figure 3), our phylogenetic analysis suggests that enough data is retained for accurate lineage discrimination even with partial M and L segment sequences (Figures 4M and 4L), which confirmed affiliation of our generated consensus sequences to the Europe 1 lineage, supported by high bootstrap values. Similar to other reports(44), we have found no evidence of genomic inter-lineage reassortment in our Europe 1 sequences based on phylogenetic analysis. Genetic lineages Asia 1 and Asia 2 (genotype IV) and Africa 3 (genotype III), however, do not form distinct clades by M segment phylogenetic analysis (Figure 4B), with evidence of genetic reassortment between these lineages having been described previously(44, 45), highlighting the importance of a whole genome sequencing approach for accurate genetic characterisation of CCHFV.

Ticks are important vectors of a wide range of viral pathogens affecting humans, livestock and wild animals. The last decades have seen a rise in tickborne disease (TBD) cases, and established populations of medically important tick species have been reported over expanding geographic areas(46). Although part of the increased incidence of tick-borne pathogens can be attributed to advances in diagnostics, increased global travel, trading of livestock and environmental changes have all contributed to the spread of tick-borne pathogens in Europe(47). Evidence of a northward expansion of the range of several tick species has been observed in Europe and the Americas(48-50), with concomitant emergence of new endemic foci(51, 52). The geographical expansion of tick vectors coupled with the fast mutation rate of RNA viruses(53) has made the investigation of tick-borne pathogens a public health priority. Research into tick-borne pathogens will be crucial to not only inform disease prevention and control measures(54), but also to understand their diversity, host and vector ranges, evolutionary histories, as well predict the emergence of new pathogenic species(53).

The SARS-CoV-2 pandemic has, now more than ever, highlighted the threat that rapidly evolving RNA viruses pose to public health. With emerging VOCs having accounted for immune response evasion(55), increased transmissibility(56) and increased virulence(57), it is clear that complete viral genome sequencing will continue to play an important role in the prevention, control and management of infectious diseases.

Since there is still no licensed vaccine against CCHFV, and the efficacy of ribavirin is still questionable(7, 58), there has been increased interest by public health authorities to support active CCHFV surveillance. The WHO has included CCHFV in their 2018 R&D Blueprint as a priority for research and product development for early diagnostics(59).

This study represents the first developed tiling multiplex PCR methodology for CCHFV, for which we have observed a comparable performance to other published schemes. Whilst recovery of a complete genomic sequence from samples with high virus titre is possible by metagenomic sequencing such as SISPA, for samples with low viraemia, a targeted enrichment approach is preferable prior to sequencing to recover meaningful portions of the genome. In this study, we were able to recover 100% of the CCHFV genome in samples with a minimum of 1.41×10^6 copies/ml of serum, with reference genome coverage dropping to at least 53.46% for samples with Ct values above 30.

Future work will look at further optimising the presented targeted enrichment approach. In particular, we will aim to increase the Ct value threshold for successful sequencing and consensus calling. Efforts will also be made to identify additional primers that can be added to the scheme to increase genome reference coverage. The scheme will also be tested on a broader set of samples in other endemic regions, as well as on other complex sample types such as tick homogenates. Whilst the development of a single multiplex PCR strategy for enrichment of all CCHFV genomes is hindered by the high genetic diversity between CCHFV clades, future efforts will also look at developing enrichment schemes for other lineages, to further support sequencing capabilities of CCHFV in endemic regions.

Acknowledgements

The study was funded by the Newton Fund: Institutional Links programme (grant number 527597374) through the British Council and TUBITAK (Turkey) - Katip Çelebi between UK and Turkey (grant number 119N515), the Royal Society (grant number IES\R2\212130) and Grant-in-Aid UK Health Security Agency (grant number 111752). The views expressed are those of the authors and not necessarily those of the funding bodies.

Figure captions

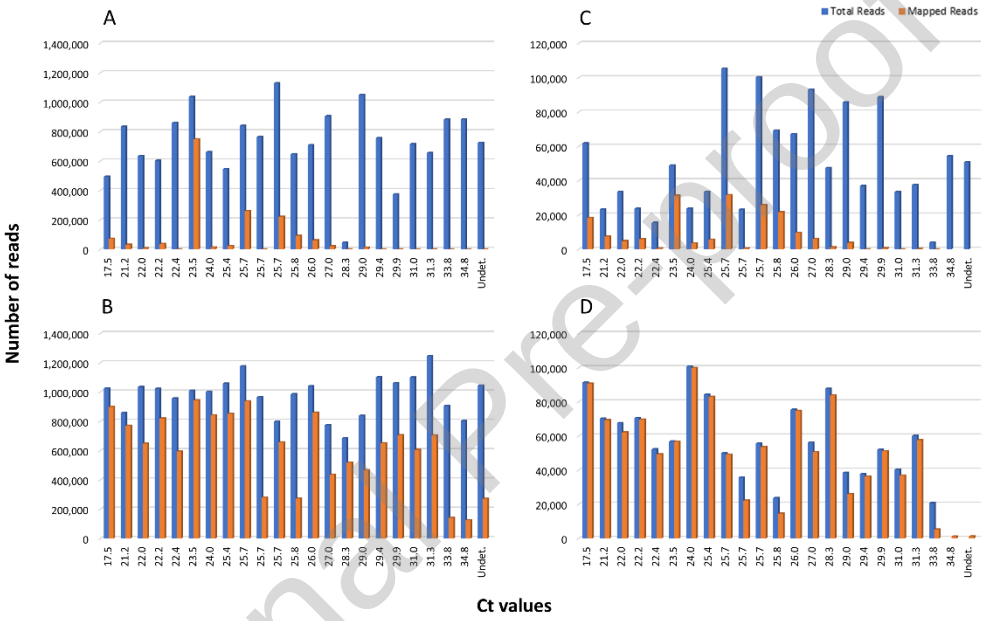

Figure 1 - Total (blue bars) and mapped (orange bars) sequencing reads for (A) SISPA and (B) targeted enrichment with Illumina sequencing and (C) SISPA and (D) targeted enrichment with nanopore sequencing.

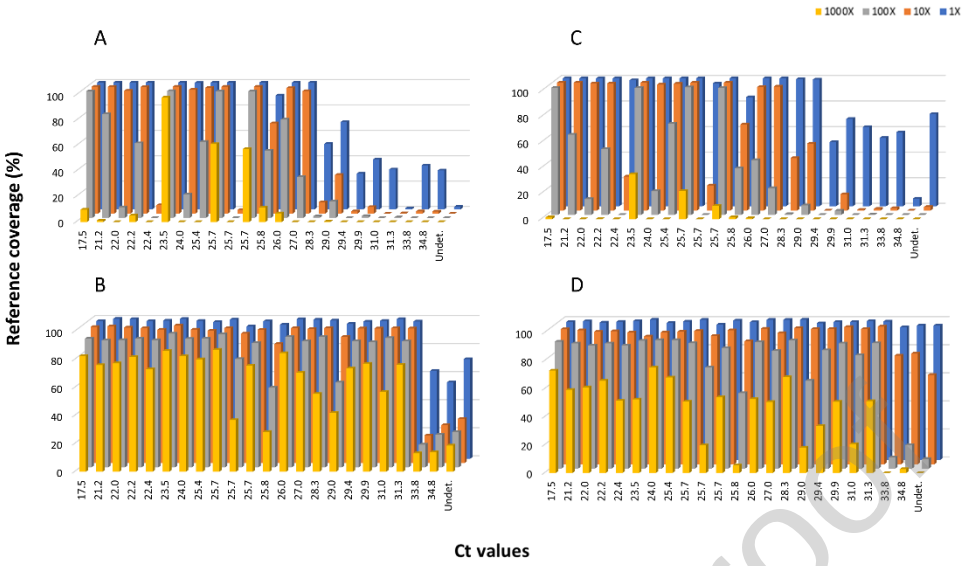

Figure 2 - Reference coverage at sequencing depth thresholds of 1x (blue), 10x (orange) and 1000x (yellow) for (A) SISPA and (B) targeted enrichment with Illumina sequencing, and for (C) SISPA and (D) targeted enrichment with nanopore sequencing.

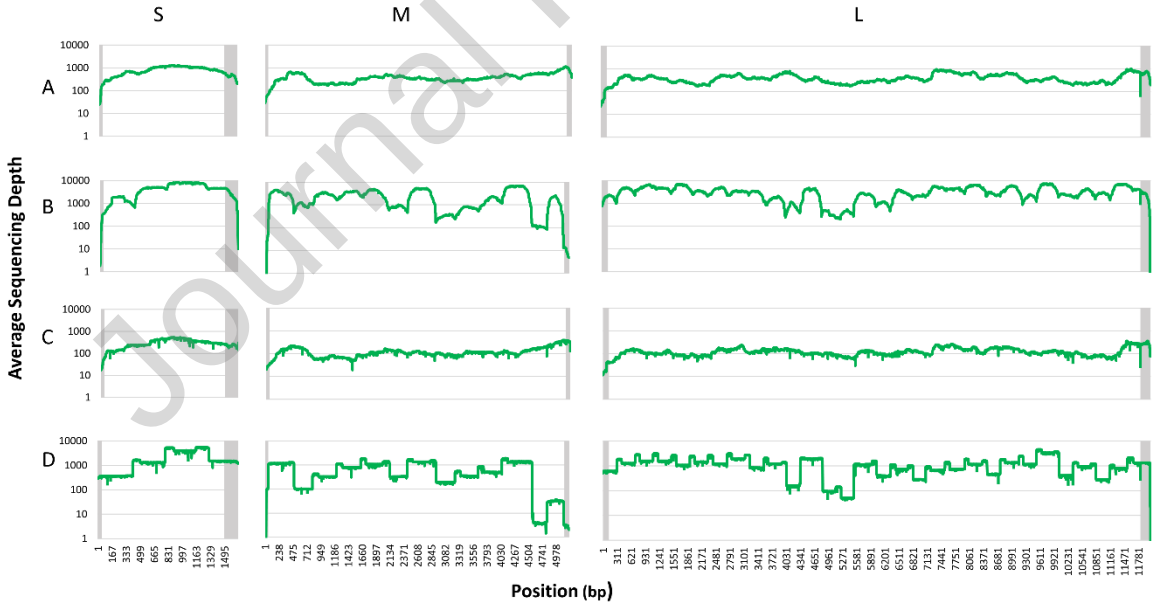

Figure 3 - Average sequencing depth (shown in green) for the Small (S), Medium (M) and Large (L) CCHFV genome segments for (A) SISPA and (B) targeted enrichment with Illumina sequencing and for (C) SISPA and (D) targeted enrichment with nanopore sequencing. 3' and 5' UTR regions are indicated in grey.

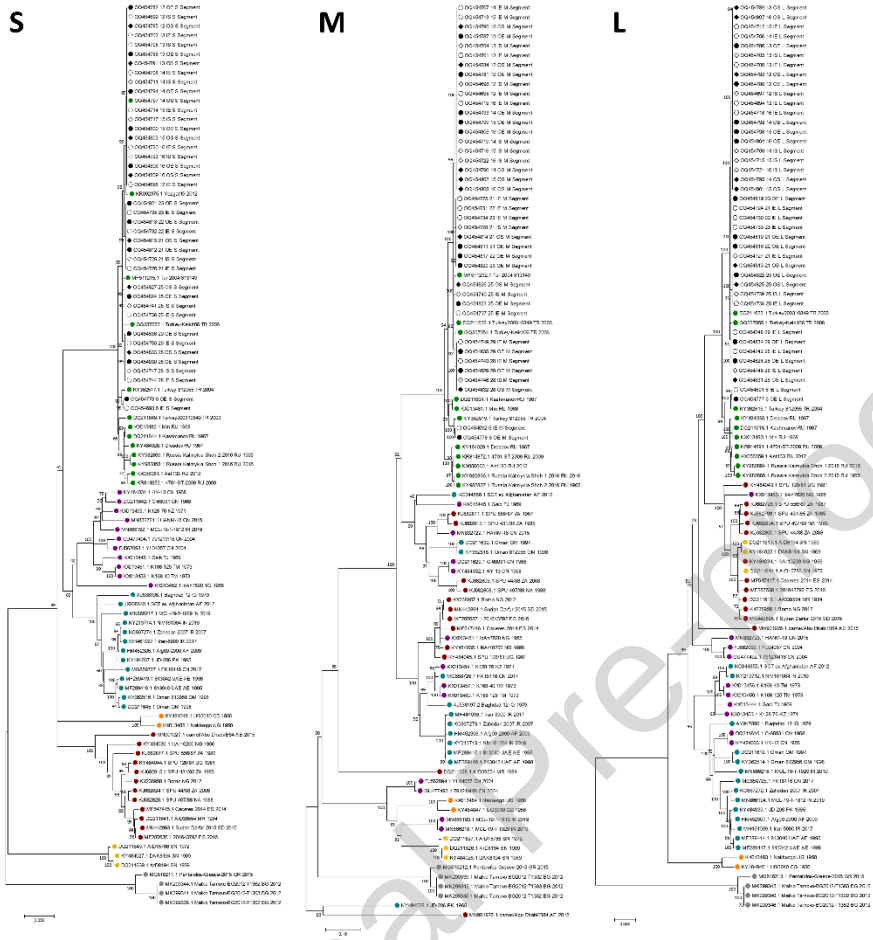

Figure 4 - Maximum likelihood phylogeny of the S (1572 bp), M (3057 bp) and L (5102 bp) segments. Sequencing technology and enrichment strategy are indicated for each sequence as follows: IE (Illumina, targeted enrichment) with a white circle, IS (Illumina, SISPA) with a white rhombus, OE (ONT, targeted enrichment) with a black circle and OS (ONT, SISPA) with a black rhombus. Different genetic lineages are indicated as follows: Asia 1/IV (teil), Asia 2/IV (purple), Europe 1/V (green), Africa 3/III (red), Africa 2/II (orange), Africa 1/I (yellow), AIGV/VI (grey).

References

1. Belhadi D, El Baied M, Mulier G, Malvy D, Mentre F, Laouenan C. The number of cases, mortality and treatments of viral hemorrhagic fevers: A systematic review. PLoS Negl Trop Dis. 2022;16(10):e0010889.

2. Aslam S, Latif MS, Daud M, Rahman ZU, Tabassum B, Riaz MS, et al. Crimean-Congo hemorrhagic fever: Risk factors and control measures for the infection abatement. Biomed Rep. 2016;4(1):15-20.
3. Shahrear S, Islam A. Immunoinformatics guided modeling of CCHF_GN728, an mRNA-based universal vaccine against Crimean-Congo hemorrhagic fever virus. Comput Biol Med. 2021;140:105098.
4. Bente DA, Forrester NL, Watts DM, McAuley AJ, Whitehouse CA, Bray M. Crimean-Congo hemorrhagic fever: history, epidemiology, pathogenesis, clinical syndrome and genetic diversity. Antiviral Res. 2013;100(1):159-89.
5. Tipih T, Burt FJ. Crimean-Congo Hemorrhagic Fever Virus: Advances in Vaccine Development. Biores Open Access. 2020;9(1):137-50.
6. Hawman DW, Haddock E, Meade-White K, Nardone G, Feldmann F, Hanley PW, et al. Efficacy of favipiravir (T-705) against Crimean-Congo hemorrhagic fever virus infection in cynomolgus macaques. Antiviral Res. 2020;181:104858.
7. D'Addiego J, Elaldi N, Wand N, Osman K, Bagci BK, Kennedy E, et al. Investigating the effect of ribavirin treatment on genetic mutations in Crimean-Congo haemorrhagic fever virus (CCHFV) through next-generation sequencing. J Med Virol. 2023.
8. Mears MC, Bente DA. In silico Design of a Crimean-Congo Hemorrhagic Fever Virus Glycoprotein Multi-Epitope Antigen for Vaccine Development. . Zoonoses 2022;2(1).
9. Portillo A, Palomar AM, Santibanez P, Oteo JA. Epidemiological Aspects of Crimean-Congo Hemorrhagic Fever in Western Europe: What about the Future? Microorganisms. 2021;9(3).
10. Papa A, Marklewitz M, Paraskevopoulou S, Garrison AR, Alkhovsky SV, Avsic-Zupanc T, et al. History and classification of Aigai virus (formerly Crimean-Congo haemorrhagic fever virus genotype VI). J Gen Virol. 2022;103(4).
11. Pickin MJ, Devignot S, Weber F, Groschup MH. Comparison of Crimean-Congo Hemorrhagic Fever Virus and Aigai Virus in Life Cycle Modeling Systems Reveals a Difference in L Protein Activity. J Virol. 2022;96(13):e0059922.
12. Bacterial and Viral Bioinformatics Resource Center [Available from: https://www.bv-brc.org/.
13. Hue S, Clewley JP, Cane PA, Pillay D. HIV-1 pol gene variation is sufficient for reconstruction of transmissions in the era of antiretroviral therapy. AIDS. 2004;18(5):719-28.
14. Houldcroft CJ, Beale MA, Breuer J. Clinical and biological insights from viral genome sequencing. Nat Rev Microbiol. 2017;15(3):183-92.
15. Kolundzija S, Cheng DQ, Lauro FM. RNA Viruses in Aquatic Ecosystems through the Lens of Ecological Genomics and Transcriptomics. Viruses. 2022;14(4).
16. Vlok M, Lang AS, Suttle CA. Marine RNA Virus Quasispecies Are Distributed throughout the Oceans. mSphere. 2019;4(2).
17. Reyes GR, Kim JP. Sequence-independent, single-primer amplification (SISPA) of complex DNA populations. Mol Cell Probes. 1991;5(6):473-81.
18. Lewandowski K, Xu Y, Pullan ST, Lumley SF, Foster D, Sanderson N, et al. Metagenomic Nanopore Sequencing of Influenza Virus Direct from Clinical Respiratory Samples. J Clin Microbiol. 2019;58(1).
19. Gauthier NPG, Nelson C, Bonsall MB, Locher K, Charles M, MacDonald C, et al. Nanopore metagenomic sequencing for detection and characterization of SARS-CoV-2 in clinical samples. PLoS One. 2021;16(11):e0259712.
20. Kafetzopoulou LE, Pullan ST, Lemey P, Suchard MA, Ehichioya DU, Pahlmann M, et al. Metagenomic sequencing at the epicenter of the Nigeria 2018 Lassa fever outbreak. Science. 2019;363(6422):74-7.
21. Brinkmann A, Uddin S, Krause E, Surtees R, Dincer E, Kar S, et al. Utility of a Sequence-Independent, Single-Primer-Amplification (SISPA) and Nanopore Sequencing Approach for Detection and Characterization of Tick-Borne Viral Pathogens. Viruses. 2021;13(2).

22. Zakotnik S, Knap N, Bogovic P, Zorec TM, Poljak M, Strle F, et al. Complete Genome Sequencing of Tick-Borne Encephalitis Virus Directly from Clinical Samples: Comparison of Shotgun Metagenomic and Targeted Amplicon-Based Sequencing. Viruses. 2022;14(6).
23. Lin X, Glier M, Kuchinski K, Ross-Van Mierlo T, McVea D, Tyson JR, et al. Assessing Multiplex Tiling PCR Sequencing Approaches for Detecting Genomic Variants of SARS-CoV-2 in Municipal Wastewater. mSystems. 2021;6(5):e0106821.
24. Quick J, Grubaugh ND, Pullan ST, Claro IM, Smith AD, Gangavarapu K, et al. Multiplex PCR method for MinION and Illumina sequencing of Zika and other virus genomes directly from clinical samples. Nat Protoc. 2017;12(6):1261-76.
25. Maloney DM, Fernandes G, Dewar R, O'Toole A, Mphanga C, Williams TC, et al. Preliminary Results from two novel "ARTIC-style" amplicon based sequencing approaches for RSV A and RSV B virological.org2022 [Available from: https://virological.org/t/preliminary-results-from-two-novel-artic-style-amplicon-based-sequencing-approaches-for-rsv-a-and-rsv-b/918#post_1.
26. Quick J. nCov-2019 sequencing protocol (single sample) 2020 [Available from: https://www.protocols.io/.
27. Li H. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics. 2018;34(18):3094-100.
28. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, et al. Twelve years of SAMtools and BCFtools. Gigascience. 2021;10(2).
29. Grubaugh ND, Gangavarapu K, Quick J, Matteson NL, De Jesus JG, Main BJ, et al. An amplicon-based sequencing framework for accurately measuring intrahost virus diversity using PrimalSeq and iVar. Genome Biol. 2019;20(1):8.
30. Penedos AR, Myers R, Hadef B, Aladin F, Brown KE. Assessment of the Utility of Whole Genome Sequencing of Measles Virus in the Characterisation of Outbreaks. PLoS One. 2015;10(11):e0143081.
31. Wick R. Filtlong 2017 [updated 30/07/2021. Available from: https://github.com/rrwick/Filtlong.
32. Au CH, Ho DN, Kwong A, Chan TL, Ma ESK. BAMClipper: removing primers from alignments to minimize false-negative mutations in amplicon next-generation sequencing. Sci Rep. 2017;7(1):1567.
33. Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol Biol Evol. 2016;33(7):1870-4.
34. Edgar RC. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics. 2004;5:113.
35. Khan AR, Pervez MT, Babar ME, Naveed N, Shoaib M. A Comprehensive Study of De Novo Genome Assemblers: Current Challenges and Future Prospective. Evol Bioinform Online. 2018;14:1176934318758650.
36. Maurier F, Beury D, Flechon L, Varre JS, Touzet H, Goffard A, et al. A complete protocol for whole-genome sequencing of virus from clinical samples: Application to coronavirus OC43. Virology. 2019;531:141-8.
37. Sims D, Sudbery I, Ilott NE, Heger A, Ponting CP. Sequencing depth and coverage: key considerations in genomic analyses. Nat Rev Genet. 2014;15(2):121-32.
38. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009;25(16):2078-9.
39. Castellano S, Cestari F, Faglioni G, Tenedini E, Marino M, Artuso L, et al. iVar, an Interpretation-Oriented Tool to Manage the Update and Revision of Variant Annotation and Classification. Genes (Basel). 2021;12(3).
40. Grubaugh ND, Gangavarapu K, Quick J, Matteson NL, De Jesus JG, Main BJ, et al. An amplicon-based sequencing framework for accurately measuring intrahost virus diversity using PrimalSeq and iVar. Genome Biol. 2019;20.
41. Kafetzopoulou LE, Efthymiadis K, Lewandowski K, Crook A, Carter D, Osborne J, et al. Assessment of metagenomic Nanopore and Illumina sequencing for recovering whole genome

sequences of chikungunya and dengue viruses directly from clinical samples. Euro Surveill. 2018;23(50).
42. Liu YX, Kearney J, Mahmoud M, Kille B, Sedlazeck FJ, Treangen TJ. Rescuing low frequency variants within intra-host viral populations directly from Oxford Nanopore sequencing data. Nat Commun. 2022;13(1).
43. Weidmann M, Sall AA, Manuguerra JC, Koivogui L, Adjami A, Traore FF, et al. Quantitative analysis of particles, genomes and infectious particles in supernatants of haemorrhagic fever virus cell cultures. Virol J. 2011;8:81.
44. Lukashev AN, Klimentov AS, Smirnova SE, Dzagurova TK, Drexler JF, Gmyl AP. Phylogeography of Crimean Congo Hemorrhagic Fever Virus. PLoS One. 2016;11(11):e0166744.
45. Negredo A, Sanchez-Arroyo R, Diez-Fuertes F, de Ory F, Budino MA, Vazquez A, et al. Fatal Case of Crimean-Congo Hemorrhagic Fever Caused by Reassortant Virus, Spain, 2018. Emerg Infect Dis. 2021;27(4):1211-5.
46. Madison-Antenucci S, Kramer LD, Gebhardt LL, Kauffman E. Emerging Tick-Borne Diseases. Clin Microbiol Rev. 2020;33(2).
47. Michelet L, Delannoy S, Devillers E, Umhang G, Aspan A, Juremalm M, et al. High-throughput screening of tick-borne pathogens in Europe. Front Cell Infect Microbiol. 2014;4:103.
48. Jore S, Vanwambeke SO, Viljugrein H, Isaksen K, Kristoffersen AB, Woldehiwet Z, et al. Climate and environmental change drives Ixodes ricinus geographical expansion at the northern range margin. Parasite Vector. 2014;7.
49. Porretta D, Mastrantonio V, Amendolia S, Gaiarsa S, Epis S, Genchi C, et al. Effects of global changes on the climatic niche of the tick Ixodes ricinus inferred by species distribution modelling. Parasit Vectors. 2013;6:271.
50. Estrada-Pena A. The climate niche of the invasive tick species Hyalomma marginatum and Hyalomma rufipes (Ixodidae) with recommendations for modeling exercises. Exp Appl Acarol. 2023.
51. Holding M, Dowall SD, Medlock JM, Carter DP, Pullan ST, Lewis J, et al. Tick-Borne Encephalitis Virus, United Kingdom. Emerg Infect Dis. 2020;26(1):90-6.
52. Monsalve Arteaga L, Munoz Bellido JL, Negredo AI, Garcia Criado J, Vieira Lista MC, Sanchez Serrano JA, et al. New circulation of genotype V of Crimean-Congo haemorrhagic fever virus in humans from Spain. PLoS Negl Trop Dis. 2021;15(2):e0009197.
53. Damian D, Maghembe R, Damas M, Wensman JJ, Berg M. Application of Viral Metagenomics for Study of Emerging and Reemerging Tick-Borne Viruses. Vector Borne Zoonotic Dis. 2020;20(8):557-65.
54. Johnson N, Phipps LP, Hansford KM, Folly AJ, Fooks AR, Medlock JM, et al. One Health Approach to Tick and Tick-Borne Disease Surveillance in the United Kingdom. Int J Environ Res Public Health. 2022;19(10).
55. Harvey WT, Carabelli AM, Jackson B, Gupta RK, Thomson EC, Harrison EM, et al. SARS-CoV-2 variants, spike mutations and immune escape. Nat Rev Microbiol. 2021;19(7):409-24.
56. Balint G, Voros-Horvath B, Szechenyi A. Omicron: increased transmissibility and decreased pathogenicity. Signal Transduct Target Ther. 2022;7(1):151.
57. Wu H, Xing N, Meng K, Fu B, Xue W, Dong P, et al. Nucleocapsid mutations R203K/G204R increase the infectivity, fitness, and virulence of SARS-CoV-2. Cell Host Microbe. 2021;29(12):1788-801 e6.
58. Johnson S, Henschke N, Maayan N, Mills I, Buckley BS, Kakourou A, et al. Ribavirin for treating Crimean Congo haemorrhagic fever. Cochrane Database Syst Rev. 2018;6:CD012713.
59. Organization WH. Crimean-Congo haemorrhagic fever 2013 [updated 31 January 2013. Available from: https://www.who.int/news-room/fact-sheets/detail/crimean-congo-haemorrhagic-fever.

Tables

Table 1 - Ct values and viral genome copies per millilitre of patient serum for each extracted RNA sample.

Sample ID	Ct value	Copies/ml	Sample ID	Ct value	Copies/ml
1	25.7	1.76×10^6	16	25.4	2.28×10^6
2	26	1.41×10^6	18	33.8	1.04×10^4
4	31	4.48×10^4	19	34.8	5.06×10^3
5	29.4	1.43×10^5	21	21.2	5.96×10^7
6	29.9	9.57×10^4	22	22.4	2.63×10^7
7	31.3	3.63×10^4	23	25.7	2.76×10^6
8	Undetermined		25	25.7	2.72×10^6
10	25.8	1.69×10^6	26	27	9.85×10^5
12	17.5	5.12×10^8	28	23.5	1.27×10^7
13	22	2.3×10^7	29	28.3	5.01×10^5
14	22.2	2.06×10^7	30	29	2.67×10^5
15	24	5.73×10^6			

Authors' contributions

Conceptualisation: RH, NW, JD, BA, YK. Data Curation: JD, NW, TF, HL. Formal analysis: JD, NW, RH. Funding acquisition: RH. Methodology: JD, NW, BA, YK. Project administration: RH, JD, NW. Writing – original draft – JD. Writing – review and editing: JD, NW, RH, BA. All authors have read and approved the final draft of this manuscript.

Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Highlights

- A tiling, multiplex PCR enrichment strategy has been adapted to Crimean-Congo haemorrhagic fever virus and validated against a panel of clinical samples.
- With a targeted, multiplex PCR tiling enrichment strategy, recovery of near-complete CCHFV genome sequences was possible from samples with low viraemia.

- A statistically significant increase in number of mapped viral sequencing reads was produced with our targeted enrichment methodology compared to a non-targeted, random enrichment approach.
- The presented methodology produced comparable results between short-read (Illumina) and long-read (Oxford Nanopore Technologies) sequencing data, and is amenable to both platforms.