Hybrid capture-based next-generation sequencing of new and old world Orthohantavirus strains and wild-type Puumala isolates from humans and bank voles

William Rosenbaum, Erik Bovinder Ylitalo, Guillaume Castel, Andreas Sjödin, Pär Larsson, Julia Wigren-Byström, Mattias N.E. Forsell, Clas Ahlm, Lisa Pettersson, Anne Tuiskunen Bäck

To cite this version:

William Rosenbaum, Erik Bovinder Ylitalo, Guillaume Castel, Andreas Sjödin, Pär Larsson, et al.. Hybrid capture-based next-generation sequencing of new and old world Orthohantavirus strains and wild-type Puumala isolates from humans and bank voles. Journal of Clinical Virology, 2024, 172, pp.105672. 10.1016/j.jcv.2024.105672 . hal-04546344

HAL Id: hal-04546344
https://hal.inrae.fr/hal-04546344

Submitted on 15 Apr 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

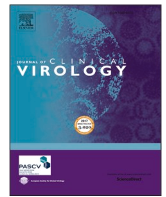
Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

Journal of Clinical Virology 172 (2024) 105672

Contents lists available at ScienceDirect

Journal of Clinical Virology

journal homepage: www.elsevier.com/locate/jcv



Hybrid capture-based next-generation sequencing of new and old world *Orthohantavirus* strains and wild-type Puumala isolates from humans and bank voles

William Rosenbaum [a], Erik Bovinder Ylitalo [b], Guillaume Castel [c], Andreas Sjödin [d], Pär Larsson [e], Julia Wigren Byström [f], Mattias N.E. Forsell [f], Clas Ahlm [f], Lisa Pettersson [f], Anne Tuiskunen Bäck [f,*]

[a] *Department of Medical Biosciences, Umeå University, SE-90185, Umeå, Sweden*
[b] *Department of Diagnostics and Intervention, Umeå University, SE-90185, Umeå, Sweden*
[c] *CBGP, INRAE, CIRAD, Institut Agro, IRD, Univ Montpellier, Montpellier, France*
[d] *CBRN Security and Defence, Swedish Defence Research Agency - FOI, Umeå, Sweden*
[e] *Clinical Genomics Umeå, Umeå University, SE-90185, Umeå, Sweden*
[f] *Department of Clinical Microbiology, Umeå University, SE-90185, Umeå, Sweden*

ARTICLE INFO

Keywords:
Targeted sequencing
Whole-genome sequencing
Puumala virus
Orthohantaviruses
Hemorrhagic fever with renal syndrome
Diagnostics

ABSTRACT

Orthohantaviruses, transmitted primarily by rodents, cause hemorrhagic fever with renal syndrome (HFRS) in Eurasia and hantavirus pulmonary syndrome in the Americas. These viruses, with documented human-to-human transmission, exhibit a wide case-fatality rate, 0.5–40 %, depending on the virus species, and no vaccine or effective treatment for severe Orthohantavirus infections exists. In Europe, the Puumala virus (PUUV), carried by the bank vole *Myodes glareolus*, causes a milder form of HFRS. Despite the reliance on serology and PCR for diagnosis, the three genomic segments of Swedish wild-type PUUV have yet to be completely sequenced.

We have developed a targeted hybrid-capture method aimed at comprehensive genomic sequencing of wild-type PUUV isolates and the identification of other Orthohantaviruses. Our custom-designed panel includes >11,200 probes covering the entire *Orthohantavirus* genus. Using this panel, we sequenced complete viral genomes from bank vole lung tissue, human plasma samples, and cell-cultured reference strains. Analysis revealed that Swedish PUUV isolates belong to the Northern Scandinavian lineage, with nucleotide diversity ranging from 2.8 % to 3.7 % among them. Notably, no significant genotypic differences were observed between the viral sequences from reservoirs and human cases except in the nonstructural protein.

Despite the high endemicity of PUUV in Northern Sweden, these are the first complete Swedish wild-type PUUV genomes and substantially increase our understanding of PUUV evolution and epidemiology. The panel's sensitivity enables genomic sequencing of human samples with viral RNA levels reflecting the natural progression of infection and underscores our panel's diagnostic value, and could help to uncover novel Orthohantavirus transmission routes.

1. Background

Orthohantaviruses (genus *Orthohantavirus*, family *Hantaviridae*) are enveloped, negative single-stranded RNA viruses consisting of three circular genomic segments: small (S, ~1.8 kb), medium (M, ~3.7 kb), and large (L, ~6.5 kb). These segments encode five proteins, including a nucleocapsid (NP) and nonstructural protein (NS), two glycoproteins (Gc and Gn), and an RNA-dependent RNA polymerase (RdRP) [1]. Recognized as airborne zoonoses of global significance, Orthohantaviruses pose a threat due to their high mortality and morbidity rates, coupled with the absence of effective vaccines and antiviral strategies. Concerns have been raised about the potential use of Orthohantaviruses as bioterrorism weapons [2]. Human diseases associated with Orthohantaviruses can manifest as either hantavirus pulmonary syndrome (HPS) or hemorrhagic fever with renal syndrome (HFRS), depending on the virus species. HPS prevails in North and South America (the New

* Corresponding author.
E-mail address: anne.tuiskunen@umu.se (A. Tuiskunen Bäck).

https://doi.org/10.1016/j.jcv.2024.105672
Received 18 December 2023; Received in revised form 20 March 2024;
Available online 30 March 2024
1386-6532/© 2024 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

World), with mortality rates reaching up to 40 %. In contrast, HFRS occurs endemically in Europe (including northern Scandinavia) and Asia (the Old World), exhibiting a milder course with a mortality rate between 0.5 % and 12 % depending on virus species [3,4]. HFRS is characterized by a febrile illness, thrombocytopenia, and transient renal failure, with severe cases necessitating intensive care [5].

Puumala virus (PUUV), transmitted by the bank vole (*Myodes glareolus*), is endemic in the northern parts of Sweden, which ranks second to Finland in PUUV incidence with over 40 cases per 100,000 inhabitants (national statistics, 2022). Human infections occur through accidental inhalation of infectious virus particles present in rodent saliva, urine, and feces. High-risk scenarios include handling wood and working in forested areas, placing individuals in northern Sweden at elevated risk of PUUV infection. The seroprevalence in northern Sweden in 2009 was 13.4 %, signifying a 50 % increase compared with that two decades earlier [6].

Despite PUUV's significant impact on public health, Swedish isolates have not undergone extensive study in terms of virus genetics. Mapping the genetic landscape of Swedish PUUV would provide a foundation to better understand the genetic diversity of PUUV within the country and its correlation with geographical prevalence, transmission patterns, and disease presentation. We have previously shown a correlation between low titers of neutralizing antibodies and high viremia in a fatal case of HFRS [7]. A tool for complete genomic sequencing would be useful in analyzing genetic viral traits and their potential correlation with human disease phenotypes, and ultimately guide the development of effective antiviral strategies. Whole-genome sequencing has been hampered due to the low number of copies of the virus genome in clinical specimens, a challenge acknowledged by other research groups [8,9]. Molecular diagnosis of PUUV relies on PCR and serological assays, leaving the detailed genetic sequence of individual isolates unknown. PUUV incidence peaks every 3–4 years [10] and genetically characterizing PUUV in an area with one of the highest human incidence rates in the world would contribute to improved outbreak control and diagnostic applications. In addition, there is a need for an effective surveillance tool for other Orthohantaviruses, given their global spread and increased prevalence.

2. Objectives

The primary objective of this study was to establish an encompassing protocol for target-based sequencing for complete genomic sequencing of PUUV and the identification of related Orthohantaviruses, utilizing a panel comprising over 11,200 probes. Successful sequencing was achieved for both reference strains of both New and Old World Orthohantaviruses and wild-type PUUV isolates from humans and bank voles. Genomic sequencing of clinical specimens provides a valuable resource, enhancing our comprehension of the genetic composition of these viruses and fortifying national preparedness against potential *Orthohantavirus* pandemics.

3. Study design

3.1. Trapping of bank voles and preparation of samples

Bank vole (*Myodes glareolus*) samples were collected near human HFRS cases in Umeå, Northern Sweden (Fig. 1). Snap-traps baited with dry apple and rapeseed oil-soaked cotton thread were used for capture. Lung tissue, collected and snap-frozen, was stored in LN_2. A FastPrep-24 with steel beads was used to homogenize the tissue (2 × 20 s at 6.5 m/s). RNA was extracted by use of a Nucleospin RNA kit (Macherey-Nagel), following the manufacturer's instructions, with elution in 60 µl RNase-free water.

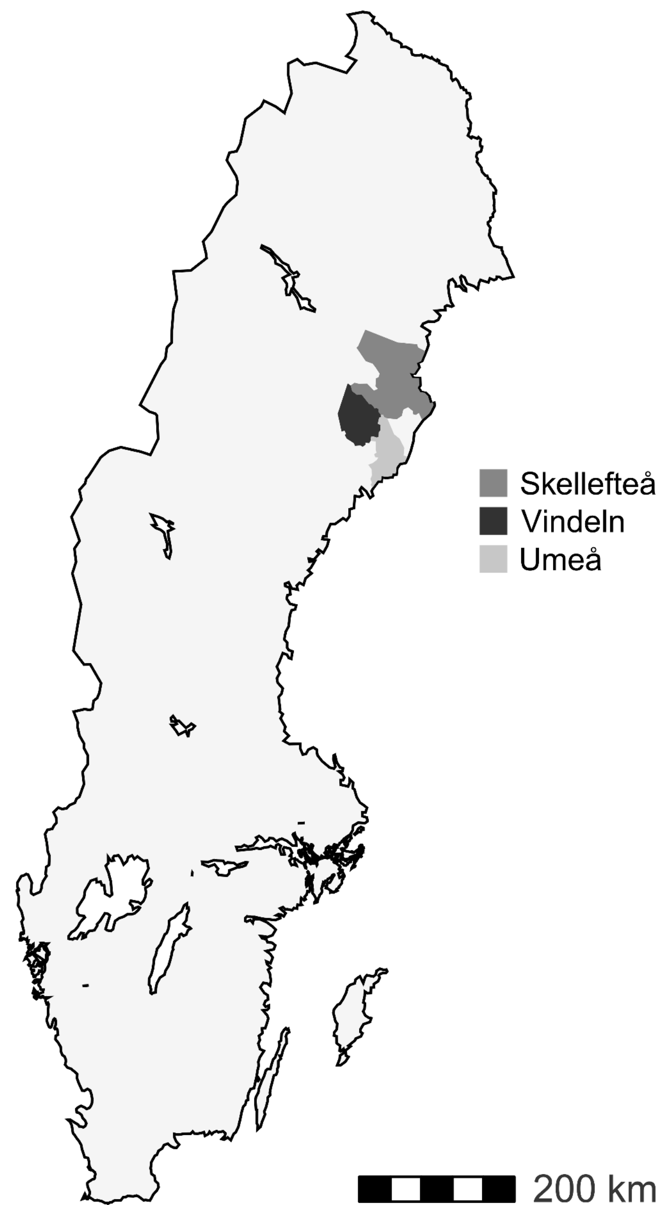


Fig. 1. Map of Sweden with the municipalities Umeå, Skellefteå, and Vindeln highlighted.

3.2. Inclusion of patients and preparation of human samples

Patients with acute HFRS were included and bled after informed consent. The diagnosis was verified by a commercially available PUUV-specific IgG and IgM enzyme immunoassay (EIA) (Reagena, Finland, product codes 114,301 and 1142). Patient characteristics are described in Table 1, and the geographical location in Fig. 1. Total RNA was extracted and purified from plasma using a QiaAmp Viral RNA kit (Qiagen) according to the manufacturer's instructions.

3.3. In vitro grown viruses

We used TRIzol-inactivated viral stocks propagated in Vero E6 cells for the following hantaviruses: Hantaan virus (HTNV) strain 76–118, Seoul virus (SEOV) strain R22 NPCV #749, Prospect Hill virus (PHV) strain PH-1, Dobrava virus (DOBV) strain H119/99 NPCV #1128, Thottopalayam virus (TPMV) strain NPCV #1165, and Tula virus (TULV) strain NPCV #656. Additionally, we included PUUV strain Umeå/hu, which was not TRIzol inactivated.

Table 1
PUUV isolates from humans and bank voles collected at different time points and locations within the same region. *F* = female. N/A = not applicable. P.i = post-infection. CT = Cycle threshold.

Accession No. (segment)	Sex	Age	Sample date	Days p.i	Hospitalized	CT-value	Location
OR915064 (S) OR915065 (M) OR915066 (L)	*F*	47	November 2022	9	Yes	32.9	Vindeln
OR915067 (S) OR915068 (M) OR915069 (L)	*F*	59	June 2022	6	No	33.1	Umeå
OR915058 (S) OR915059 (M) OR915060 (L)	*F*	54	March 2023	8	Yes	27.4	Skellefteå
OR915061 (S) OR915062 (M) OR915063 (L)	N/A (Bank vole)		July 2016	N/A	N/A	20.4	Umeå
OR915070 (S) OR915071 (M) OR915072 (L)	N/A (Bank vole)		July 2016	N/A	N/A	17.2	Umeå

3.4. Quantification of viral RNA

In a 20 μl volume, each qRT-PCR analysis included 5 μl RNA, 200 nM of each primer (PuKzSF2: GAARTGGACCCGGATGACGTTAAC, PuKzSR2: CKGGACACAGCATCTGCCATTC, and GAPDH reference gene from Quantitect #QT00079247), 11 μl qPCR Bio SYBR Green 1-step GO High ROX kit master mix (PCR Biosystems), and water. The reactions were performed with a StepOnePlus Real-Time PCR system (Applied Biosystems) with the following conditions: 10 min at 45 °C for cDNA synthesis, 2 min at 95 °C for enzyme activation, followed by 40 cycles of 5 s at 95 °C and 25 s at 60 °C for denaturation and annealing/extension, respectively. Amplification was followed by a melt-curve analysis [11].

3.5. Design of the Orthohantavirus enrichment panel

To construct the *Orthohantavirus* enrichment panel, FASTA files representing the S, M, and L segments of the *Orthohantavirus* genus were obtained from the BV-BRC database [12]. We excluded partial sequences that were less than 1000 base pairs in length to ensure comprehensive genomic coverage. The dRep tool was employed to select a set of representative sequences, effectively capturing the nucleotide diversity within the genus [13]. This process resulted in a total of 1281 FASTA entries. These entries served as the foundational dataset for designing the *Orthohantavirus* enrichment panel using CATCH within the Twist platform [14]. A probe set was specifically designed to identify *Orthohantavirus* species, with meticulous exclusion of probes that bind to other viral genomes, *Homo sapiens*, or *Myodes glareolus.* Probes containing low-complexity sequences were excluded, resulting in a final collection of 11,217 unique probes. These probes were meticulously designed to optimize the detection and analysis of *Orthohantavirus* genetic material.

3.6. Sequencing

Libraries were created using the Twist EF Library Preparation Kit (cat. no. 100,573, Twist Bioscience), along with the Target Enrichment Kit, Hybridization kit, Twist Universal Adapter System, Twist Universal Blockers, Binding and Purification Beads, Wash Buffers, and Equinox Library Amp mix (cat. no. 104,187, Twist BioScience), employing the custom *Orthohantavirus* enrichment panel. According to the manufacturer's protocol, first strand synthesis was performed using ProtoScript II First Strand cDNA synthesis kit with 15 ul RNA per sample (≤50 ng RNA), followed immediately by second-strand synthesis using the NEBNext Ultra II Non-Directional RNA Second Strand Synthesis kit. According to the Twist BioScience protocol, the subsequent steps included DNA purification via bead-based washing before DNA fragmentation. Each samples contained 25 ng of DNA, followed by end repair and dA-tailing. Twist Universal Adapters were then ligated to the dA-tailed DNA fragments and purified with DNA-binding beads. Subsequently, adapted libraries were amplified and purified. The average fragment length was confirmed to be within the recommended range of 250–450 base pairs. Batches of eight equally pooled libraries, each using 187.5 ng, were combined and multiplexed in the capture reaction, followed by purification with Streptavidin binding beads. A post-capture PCR was conducted using the Streptavidin bead slurry with the following conditions: 45 s at 98 °C initialization, followed by 16 cycles of 15 s at 98 °C for denaturation, 30 s at 60 °C for annealing, and 30 s at 72 °C for extension, and a final extension step for 60 s at 72 °C. A final washing with DNA binding beads was performed. Prior to sequencing the libraries, fragment length was confirmed to be within the recommended range of 375–425 base pairs. cDNA quantification utilized the Qubit dsDNA HS Assay kit (Thermo Fisher Scientific). The cDNA quality was assessed with an Agilent 2200 TapeStation using the High Sensitivity D1000 ScreenTape Assay (Agilent).. Sequencing was performed on the Illumina MiSeq platform (Illumina, San Diego, CA, USA) in paired-end mode (PE, 2 × 75 bp). Complete genome sequences were deposited in Genbank under Accession Number (Acc. No.) OR915058 to OR915072. To determine the lower limit for complete genomic sequencing, a ten-fold dilution series of Umeå/Hu RNA was conducted and sequenced.

3.7. Quality control and de novo assembly of MiSeq paired-end reads

Reads underwent adapter removal and quality filtering at the base-pair level using fastp (version 0.23.2) [15]. To eliminate known host sequences, adapter-trimmed reads were aligned to the human reference genome GRCh38 (https://www.ncbi.nlm.nih.gov/datasets/genome/GCF_000001405.40/) (accessed January 2023) with bwa (version 0.7.17) [16]. Unaligned reads were extracted via samtools (version 1.6) [17] and then *de novo* assembled using MEGAHIT (version 1.2.9) [18] or SPAdes (version 3.15.5) [19]. Blastn analysis of all contigs employed a custom database comprising viruses, mitochondrial DNA, and mammalian rRNA from Refseq and GenBank (accessed September 2023) [20]. Contigs exceeding 90 % of the length of the *Orthohantavirus* S, M, or L segment were retained and polished with Pilon [21]. Coverage information for each segment was generated by aligning filtered reads to the most similar segment as determined by blastn, using bwa, and extracting coverage data with samtools mpileup and idxstats subcommands. Resulting files were used to create figures with in-house scripts. The analysis was performed using the Python workflow manager Snakemake, and can be found in github (https://github.com/willros/puuv_publication).

To determine the percentage of reads originating from viruses, reads were classified using Kraken2 (version 2.1.2) [22] and Kaiju (version

1.9.0) [23]. Using the prebuilt Standard Kraken database (https://genome-idx.s3.amazonaws.com/kraken/k2_standard_20240112.tar.gz) and the prebuilt Refseq Kaiju database (https://kaiju-idx.s3.eu-central-1.amazonaws.com/2023/kaiju_db_refseq_2023-05-23.tgz)

3.8. Phylogenetic analyses

Phylogenetic analyses were performed on three datasets composed of complete (or nearly complete) coding regions of S, M, and L segment sequences of PUUV recovered in this study and from sequences from different geographical areas, representative of the eight known PUUV lineages and available in GenBank. Sequences of Fusong, Muju, and Hokkaido Orthohantaviruses were used as outgroups. Multiple sequence alignments were generated with the Clustal Omega alignment program implemented in SeaView v4.6.1. Phylogenetic analyses were then performed using the Maximum Likelihood method implemented in PhyML 3.0 (available online at http://www.atgc-montpellier.fr/phyml/) with a statistical approximate likelihood ratio test (aLRT) for branch support. The optimal substitution models were identified as the GTR + R model for all datasets, using the “Automatic model selection by the Smart Model Selection (SMS) program” option implemented in PhyML. Phylogenetic trees were then edited with FigTree v1.4.3.

Base and amino acid differences per site were determined for all samples and segments, employing mean-diversity calculations (Eq. 12.73 in reference [24]). Standard error estimates, obtained through a bootstrap procedure with 250 replicates, are presented in the second column in Supplementary Table 1. Rate variation among sites was modeled using a gamma distribution (shape parameter = 1). The analysis included 5 amino acid sequences for each segment, with the exclusion of positions containing gaps and missing data (complete deletion option). For the large segment’s coding sequence (CDS), there were a total of 6471 nucleotide and 2156 amino acid positions; for the medium segment CDS, a total of 3447 nucleotide and 1148 amino acid positions; and for the small segment CDS, a total of 1312 nucleotide and 412 amino acid positions. Codon positions considered were 1st + 2nd + 3rd + noncoding. The evolutionary analyses were conducted using MEGA11 [25].

4. Results

4.1. Whole-genome sequencing from low viral load samples

Five wild-type PUUV isolates from two bank voles and three patients were sequenced, covering the complete open reading frame of all three genomic segments. Cell-propagated reference strains of seven Orthohantaviruses were also sequenced, including Umeå/Hu, PHV, SEOV, DOBV, HTNV, TULV, and TPMV. Complete genomes were obtained for all viruses except for TPMV, where half of the L segment was missing. A ten-fold dilution series of Umeå/Hu RNA revealed the lower limit for complete genomic sequencing to be a cycle threshold (CT) of 35 (Supplementary Figure S3).

For the human-derived wild-type PUUV isolates, approximately 5–12 % of the reads originated from the virus, while the remaining reads were derived from human DNA. Additionally, an overlap of reads aligning to both the human reference and the custom virus database was observed (Supplementary Figure 2). In the case of bank vole–derived isolates, 88–89 % of the reads were of viral origin, with the remainder belonging to the bank vole genome. Concerning the *in vitro* grown reference strains, 6–81 % of the reads aligned to virus genomes, with a higher yield observed for strains more closely related to PUUV (Table 2). This trend was also evident in the read depth (Fig. 2), showing a higher read depth for virus sequences in samples with a larger proportion of virus reads. Notably, no specific genomic segment was favored in terms of the number of reads per segment (Fig. 3).

Table 2
Sample sequencing alignment overview. Table presenting read counts for each sequencing run, accompanied by the corresponding percentages of reads classified as viruses by Kaiju.

Sample (Acc.No or strain)	Host	No. of reads	Virus reads (%)
OR915064, OR915065, OR915066	Human	3 603 706	6.10
OR915067, OR915068, OR915069		3 637 942	8.49
OR915058, OR915059, OR915060		3 437 752	8.94
OR915061, OR915062, OR915063	Bank vole	486 784	88.61
OR915070, OR915071, OR915072		1 584 298	89.17
DOBV	Cell culture	229 038	73.96
HTNV		1 034 983	80.42
PHV		142 372	65.59
Umeå/Hu		30 482 746	81.04
SEOV		460 120	79.75
TPMV		37 710	6.03
TULV		93 476	62.49

4.2. Phylogenetic characterization of new PUUV isolates

The five Swedish isolates (Acc. No. OR915058 to OR915072) sequenced in this study clustered within the Northern Scandinavian (N-SCA) lineage of PUUV, together with the reference Swedish strain Umeå/hu (S, M, L segments) and with previously sequenced strains from Sweden (Fig. 4). The new isolates are closely related, with nucleotide differences of 2.8 % (0.7 % in the internal open reading frame encoding NS), 3.5 %, and 3.7 % in the S, M, and L segments, respectively (Supplementary Table 1). Amino acid diversity of the NP was 0.18 %, while the glycoproteins and RdRP exhibited 0.66 % and 0.33 %, respectively (Supplementary Table 1). Notably, the NS protein displayed significant variation, with amino acid differences at four positions, including two unique amino acids in each of the bank vole isolates (Trp42Lys and Thr29Ala, respectively), and the amino acid diversity was 2.2 % (Supplementary Table 1). Besides these variations, no significant differences were observed between PUUV isolates derived from bank voles and humans.

5. Discussion

Sweden has the second-highest incidence of human HFRS cases globally, trailing only Finland. Two PUUV lineages have been identified in Sweden [26], yet the genetic landscape of PUUV remains inadequately explored, with a lack of complete genomic sequences from wild-type isolates, especially of M and L segments. To date, only one human PUUV isolate from Sweden has been sequenced without cell-culture propagation, albeit missing the L segment [7]. In this study we aimed to develop a sequencing tool for the complete genomic sequencing of wild-type PUUV isolates and identification of remaining Orthohantaviruses. Pilot experiments using the Twist Comprehensive Virus Research Panel yielded inadequate coverage and depth for complete genomic segment enrichment of Puumala virus (unpublished results). Therefore, we opted to pursue a hantavirus-specific approach. Our targeted hybrid-capture panel, featuring over 11,200 carefully selected probes, provided comprehensive genomic coverage for all three segments of five new PUUV isolates from Sweden (human and bank vole). These sequences constitute the first complete PUUV genomes from Sweden with the exception of the historical cell-propagated reference strain Umeå/hu and are therefore a very important contribution to future monitoring of the evolution of PUUV in Sweden. The CT values of the viral RNA from the human isolates reflected the natural course of infection, underscoring the panel’s effectiveness as a diagnostic tool. As

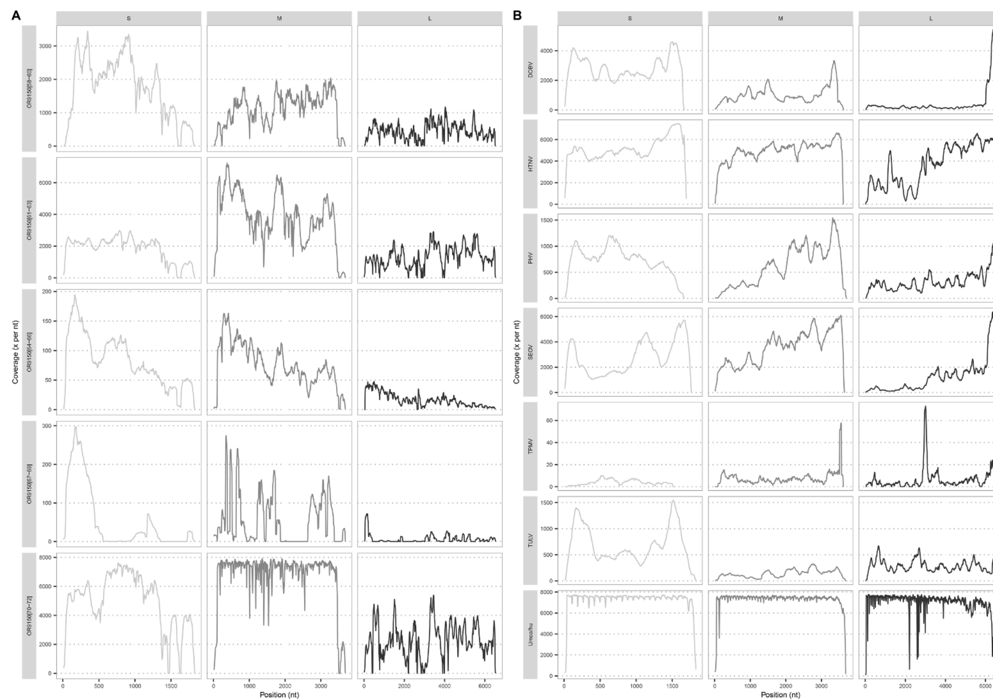


Fig. 2. Coverage details for trimmed reads aligned to the most similar segment, as determined by blastn. The x-axis represents the position in nucleotides, while the y-axis indicates the alignment coverage using a rolling window of 10 nucleotides. (A) Patients and vole samples. (B) Isolates.

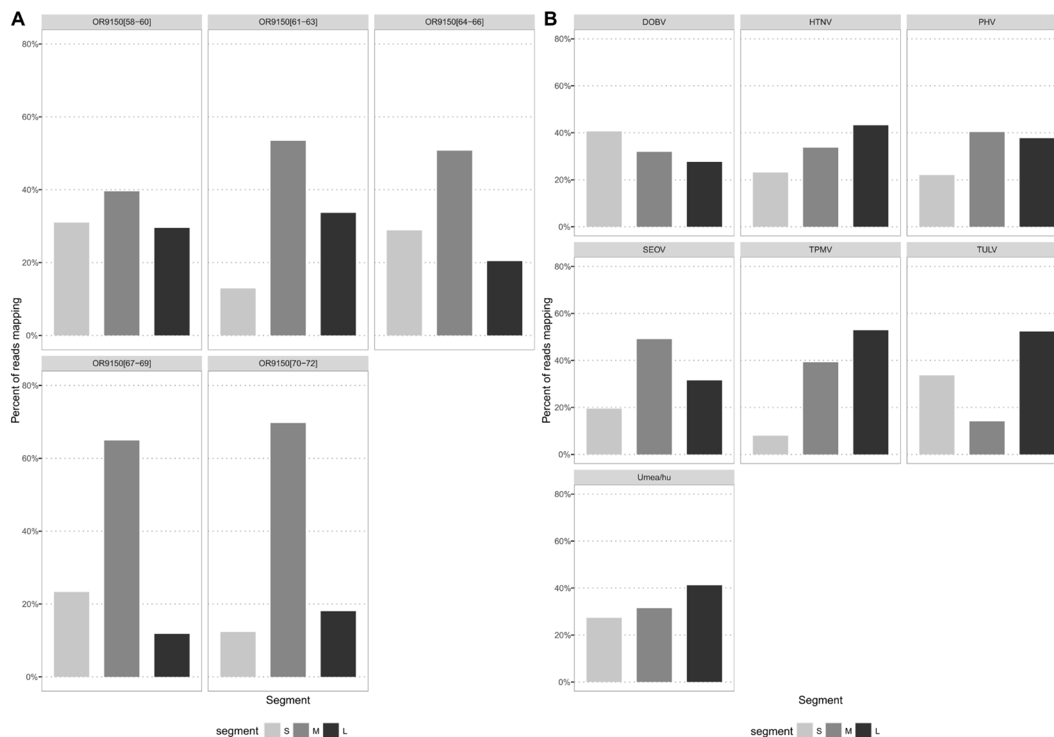


Fig. 3. Trimmed reads were aligned to a file containing the three most similar segments for each sample, as determined by blastn. The Y-axis represents the percentage of mapped reads (relative to the total mapped to all three segments), while the X-axis indicates the segments for all samples. The number above each bar indicates the count of reads aligned to each segment. (A) Patients and bank vole samples. (B) Isolates.

expected, the sequence depth varied, with the highest depth observed in the bank vole isolates and cell-propagated Swedish reference strain Umeå/hu. Although the probes were designed to avoid hybridizing with the host genome, most captured sequences originated from human DNA, suggesting the potential benefit of a host depletion step, albeit at the cost of increased complexity of the protocol. However, we observed an overlap of reads aligning to both the human reference and the custom virus database (Supplementary Figure 2). This could indicate sequences homologous to proviruses in the human DNA, or the existence of homologous regions in reads capable of aligning to both the human

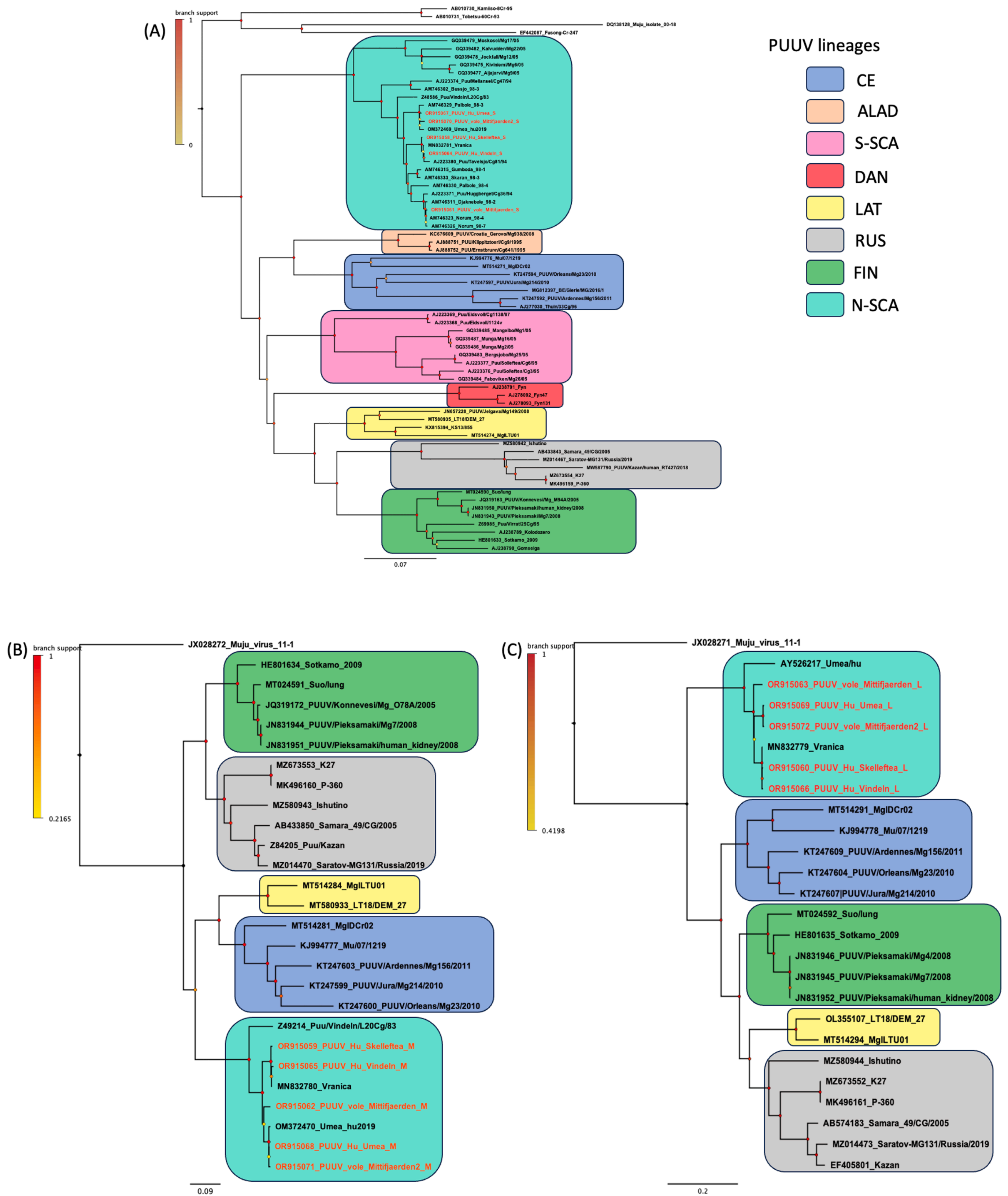


Fig. 4. Phylogenetic trees of PUUV strains generated with PhyML 3.0 by the maximum-likelihood method on the complete coding region of the S (A), M (B), and L (C) segments with a GTR + R nucleotide substitution model. New PUUV isolates sequenced in this study are indicated in red. PUUV lineages are indicated (CE, Central European; FIN, Finnish; ALAD, Alpe-Adrian; RUS, Russian; LAT, Latvian; N-SCA, North Scandinavian; S-SCA, South Scandinavian; DAN, Danish). Branch support values as determined by an aLRT test are represented by a colored dot at each node. Scale bars indicate the number of substitutions per site.

reference and viral reference genomes. In a study conducted in 2022, Szóstak et al. simulated reads from a bacterial background and observed a similar occurrence (10). Using the default settings in bwa, approximately 25 % of the simulated reads aligned to the human reference, with a significant proportion simultaneously classified as of bacterial origin by Kraken2 [27].

The PUUV isolates in this study, collected over 7 years within a highly endemic region (Fig. 1), all belonged to the same N-SCA PUUV lineage and exhibited relatively discrete sequence variation, ranging from 2.8 % to 3.7 % at the nucleotide level, and 0.18–0.66 % for amino acids (Supplementary Table 1). This is in line with previous findings reporting a 0–9 % difference in nucleotide sequences between PUUV strains within the same lineage [8]. The S segment, which encodes the NP—the most abundant viral protein during infection and formerly considered the most heterogeneous—exhibited the highest level of conservation (2.8 % difference at the nucleotide level, Supplementary Table 1). The L segment, in contrast, showed the highest level of diversity (3.7 % difference at the nucleotide level, Supplementary Table 1). Among this limited number of isolates, PUUV seems to exhibit conservation over time within this highly endemic region. To validate these observations, sequences from a larger number of isolates and a broader geographical area would be essential. Phylogenetic analyses show a genetic link between PUUV isolates associated with human HFRS cases and local wild-type PUUV strains from bank voles, as previously shown in Finland [28].

The 90-amino acid long NS protein exhibited 0.7 % diversity at the nucleotide level, while the amino acid diversity reached 2.2 %, suggesting a potential positive selection pressure. The NS protein is recognized for its influence on the interferon response during infection [29]. However, it lacks structural properties and would potentially be prone to a more rapid evolutionary drift without compromising essential functions critical for the virus's fitness. A significantly elevated amino acid divergence in the NS protein has also been documented by others [30]. A noteworthy observation pertains to the two amino acid substitutions within the NS segment that are the primary difference between the bank vole isolates and the human isolates in this study. Functional studies are necessary to assess the phenotypic significance of these genetic findings.

Successful sequencing of seven *Orthohantavirus* genus reference strains demonstrates the panel's utility in identifying and characterizing other Orthohantaviruses. Notably, the probe panel could facilitate the detection and sequencing of SEOV in humans, which has not been investigated in Sweden despite findings of anti-SEOV antibodies in rats in Stockholm [31]. It would be intriguing to sequence additional, more distantly related genera of Orthohantaviruses to evaluate the panel's versatility.

6. Conclusions

We have created a hybrid capture–based next-generation sequencing panel designed for the comprehensive genomic characterization of PUUV isolates derived from both the natural reservoir and humans. This innovative panel facilitates molecular epidemiological studies to unravel the dynamics of PUUV, laying the groundwork for a deeper comprehension of the pathophysiology of HFRS and, ultimately, the formulation of effective antiviral strategies. Moreover, the panel holds potential significance as a valuable tool for identifying other Orthohantaviruses. It can serve as a means of surveillance, contributing to enhanced outbreak control measures.

Funding

This work was supported by Region Västerbotten (RV-970009 and RV-982503) to ATB, the Kempe Foundation (SMK21-0039) to ATB, OE and Edla Johansson's Scientific Foundation to ATB, the Lars Hierta Memorial Foundation (FO2021-0251) to ATB, the Swedish Laboratory Network In Microbiology (00304-2019) to LP, and National Institutes of Health (U19AI142777) and Swedish Research Council (2020-06235) to MF.

Institutional review board statement

The study was conducted in accordance with the Declaration of Helsinki and approved by the Regional Ethics Committee of Umeå University (project identification codes 04-37-32, dated 2015-03-05, and 07-162, dated 09-01-2009) for studies involving humans. Trapping of animals was approved by the Swedish Environmental Protection Agency (permission: NV-01124-15) and the Animal Ethics Committee in Umeå (permission: A 39-14), and all applicable institutional and national guidelines for the use of animals were followed.

Informed consent statement

All subjects gave their informed consent for inclusion before they participated in the study.

Disclaimer/Publisher's note

The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of Elsevier and/or the editor(s). Elsevier and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

CRediT authorship contribution statement

William Rosenbaum: Writing – review & editing, Visualization, Validation, Software, Investigation, Formal analysis, Data curation. **Erik Bovinder Ylitalo:** Writing – review & editing, Validation, Investigation. **Guillaume Castel:** Writing – review & editing, Visualization, Software, Formal analysis. **Andreas Sjödin:** Writing – review & editing, Software, Methodology, Data curation, Conceptualization. **Pär Larsson:** Writing – review & editing, Visualization, Software, Formal analysis, Data curation. **Julia Wigren Byström:** Writing – review & editing, Resources, Investigation. **Mattias N.E. Forsell:** Writing – review & editing, Resources, Conceptualization. **Clas Ahlm:** Writing – review & editing, Resources, Project administration, Conceptualization. **Lisa Pettersson:** Writing – review & editing, Resources, Project administration, Conceptualization, Methodology. **Anne Tuiskunen Bäck:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The authors declare no conflict of interest.

Acknowledgments

The authors would like to thank Dr. Göran Bucht and Dr. Jonas Näslund from the Swedish Defense Research Agency for the kind donation of laboratory reference Orthohantavirus.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jcv.2024.105672.

References

[1] P. Johansson, et al., Complete gene sequence of a human Puumala hantavirus isolate, Puumala Umea/hu: sequence comparison and characterisation of encoded gene products, Virus Res. 105 (2) (2004) 147–155.

[2] M.H. D`Souza, T.R. Patel, Biodefense implications of new-world hantaviruses, Front. Bioeng. Biotechnol. 8 (2020) 925.

[3] C.B. Jonsson, L.T. Figueiredo, O. Vapalahti, A global perspective on hantavirus ecology, epidemiology, and disease, Clin. Microbiol. Rev. 23 (2) (2010) 412–441.

[4] A. Papa, Dobrava-Belgrade virus: phylogeny, epidemiology, disease, Antiviral Res. 95 (2) (2012) 104–117.

[5] M. Hjertqvist, et al., Mortality rate patterns for hemorrhagic fever with renal syndrome caused by Puumala virus, Emerg. Infect. Dis. 16 (10) (2010) 1584–1586.

[6] K. Bergstedt Oscarsson, et al., Human Puumala hantavirus infection in northern Sweden; increased seroprevalence and association to risk and health factors, BMC Infect. Dis. 16 (1) (2016) 566.

[7] A. Tuiskunen Back, et al., Clinical and genomic characterisation of a fatal Puumala orthohantavirus case with low levels of neutralising antibodies, Infect. Dis. (Lond) 54 (10) (2022) 766–772.

[8] J.S. No, et al., Comparison of targeted next-generation sequencing for whole-genome sequencing of Hantaan orthohantavirus in Apodemus agrarius lung tissues, Sci. Rep. 9 (1) (2019) 16631.

[9] M. Kramski, et al., Detection and typing of human pathogenic hantaviruses by real-time reverse transcription-PCR and pyrosequencing, Clin. Chem. 53 (11) (2007) 1899–1905.

[10] G.E. Olsson, et al., Human hantavirus infections, Sweden, Emerg. Infect. Dis. 9 (11) (2003) 1395–1401.

[11] J. Wigren Bystrom, et al., Quantification and kinetics of viral RNA transcripts produced in Orthohantavirus infected cells, Virol. J. 15 (1) (2018) 18.

[12] R.D. Olson, et al., Introducing the bacterial and viral bioinformatics resource center (BV-BRC): a resource combining PATRIC, IRD and ViPR, Nucleic Acids Res. 51 (D1) (2022) D678–D689.

[13] M.R. Olm, et al., dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication, ISME J. 11 (12) (2017) 2864–2868.

[14] H.C. Metsky, et al., Capturing sequence diversity in metagenomes with comprehensive and scalable probe design, Nat. Biotechnol. 37 (2) (2019) 160–168.

[15] S. Chen, et al., fastp: an ultra-fast all-in-one FASTQ preprocessor, Bioinformatics 34 (17) (2018) i884–i890.

[16] H. Li, R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform, Bioinformatics 25 (14) (2009) 1754–1760.

[17] H. Li, et al., The sequence alignment/map format and SAMtools, Bioinformatics 25 (16) (2009) 2078–2079.

[18] D. Li, et al., MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph, Bioinformatics 31 (10) (2015) 1674–1676.

[19] A. Bankevich, et al., SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing, J. Comput. Biol. 19 (5) (2012) 455–477.

[20] S. McGinnis, T.L. Madden, BLAST: at the core of a powerful and diverse set of sequence analysis tools, Nucleic Acids. Res. 32 (2004) W20–W25 (Web Server issue).

[21] B.J. Walker, et al., Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement, PLoS ONE 9 (11) (2014) e112963.

[22] D.E. Wood, J. Lu, B. Langmead, Improved metagenomic analysis with Kraken 2, Genome Biol. 20 (1) (2019) 257.

[23] P. Menzel, K.L. Ng, A. Krogh, Fast and sensitive taxonomic classification for metagenomics with Kaiju, Nat. Commun. 7 (2016) 11257.

[24] M.N.a.S. Kumar, Molecular Evolution and Phylogenetics, Oxford University Press, 2000.

[25] K. Tamura, G. Stecher, S. Kumar, MEGA11: molecular evolutionary genetics analysis version 11, Mol. Biol. Evol. 38 (7) (2021) 3022–3027.

[26] G. Castel, et al., Phylogeography of Puumala orthohantavirus in Europe, Viruses 11 (8) (2019) 679.

[27] N. Szostak, et al., The standardisation of the approach to metagenomic human gut analysis: from sample collection to microbiome profiling, Sci. Rep. 12 (1) (2022) 8470.

[28] A. Plyusnina, et al., Analysis of complete Puumala virus genome, Finland, Emerg. Infect. Dis. 18 (12) (2012) 2070–2072.

[29] K.M. Jaaskelainen, et al., Tula and Puumala hantavirus NSs ORFs are functional and the products inhibit activation of the interferon-beta promoter, J. Med. Virol. 79 (10) (2007) 1527–1536.

[30] F. Binder, et al., Spatial and Temporal Evolutionary Patterns in Puumala Orthohantavirus (PUUV) S Segment, 9, Pathogens, 2020.

[31] J. Verner-Carlsson, et al., First evidence of Seoul hantavirus in the wild rat population in the Netherlands, Infect. Ecol. Epidemiol. 5 (2015) 27215.