

Article

# Prediction of the Spatial Origin of Puumala Virus Infections Using L Segment Sequences Derived from a Generic Screening PCR †

Sabrina Weiss <sup>1,\*</sup>, Boris Klempa <sup>1,2</sup>, Beate Tenner <sup>1</sup>, Detlev H. Kruger <sup>1</sup> and Jörg Hofmann <sup>1</sup>

<sup>1</sup> Institute of Virology, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, 10117 Berlin, Germany

<sup>2</sup> Institute of Virology, Biomedical Research Center, Slovak Academy of Sciences, 84505 Bratislava, Slovakia

\* Correspondence: Sabrina.weiss@charite.de or sabrinanweiss@gmail.com; Tel.: +49 30 450 625 089

† This work is dedicated to our good friend and excellent co-worker Brita Auste (1965–2019) who died before one's time.

Received: 30 May 2019; Accepted: 27 July 2019; Published: 30 July 2019



**Abstract:** To screen diagnostic specimens for the presence of hantavirus genomes or to identify new hantaviruses in nature, the pan-hanta L-PCR assay, a broadly reactive nested reverse transcription polymerase chain reaction (RT-PCR) assay targeting the L segment, is highly preferred over other assays because of its universality and high sensitivity. In contrast, the geographic allocation of Puumala virus strains to defined outbreak regions in Germany was previously done based on S segment sequences. We show that the routinely generated partial L segment sequences resulting from the pan-hanta L-PCR assay provide sufficient phylogenetic signal to inform the molecular epidemiology of the Puumala virus. Consequently, an additional S segment analysis seems no longer necessary for the identification of the spatial origin of a virus strain.

**Keywords:** Hantavirus; Puumala virus; phylogeography; molecular epidemiology

## 1. Introduction

Hantaviruses (family *Hantaviridae*) are tri-segmented negative stranded RNA viruses that are carried by small mammals (rodents, insectivores, and bats). Human infections are thought to occur through the inhalation of rodent excretions. In Germany, two pathogenic hantaviruses have been circulating, the Dobrava-Belgrade virus (DOBV), restricted to northern and eastern Germany, and Puumala virus (PUUV), endemic in the southern and western parts of the country. Most human hantavirus disease cases in Germany are caused by PUUV which usually manifests as a mild form of hemorrhagic fever with renal syndrome (HFRS), sometimes also named as nephropathia epidemica [1–3]. Annual case numbers vary considerably depending on the various environmental factors [4] associated with the changes in the population densities of bank voles (*Myodes glareolus*), the PUUV reservoir hosts. Since hantavirus disease became notifiable in Germany in 2001, several outbreaks have occurred over the years with up to 2825 cases being recorded, the latest one in 2017 with 1731 notified cases [5].

Initial hantavirus disease diagnoses in Germany are based on clinical and serological data, but these do not allow for reliable typing of the virus species or even virus lineage. Reverse transcription polymerase chain reaction (RT-PCR)-based diagnostic assays are not routinely used because they are not required for case notification and virus RNA in endemic PUUV infections is detectable in patient samples only for a short period of time after onset of the first, mostly unspecific symptoms. Therefore, patient-derived PUUV genomic sequences were previously rare, and most of the PUUV genomic data was available from the reservoir hosts. Nevertheless, the recent outbreaks over the years in Germany (2007, 2010, 2012, and 2017) boosted public and medical awareness of hantaviruses, resulting in more

early diagnoses and a significant build-up of PUUV genome sequence data from patients, in particular of the S segment encoding the nucleocapsid protein [6–8]. Combination of these sequences with the sequence data obtained from bank voles, lack of human-to-human transmission, local distribution of bank voles, and a high degree of identity between viruses of human and rodent origin allowed allocation of certain patient-derived virus sequences to defined geographic regions, thus identifying high-risk areas [7].

On the other hand, pan-hanta L-PCR assay, a broadly reactive nested RT-PCR assay, targeting the L segment [9] has recently become more widely used because of its universality and high sensitivity. Consequently, partial L segment sequences as the outcome of this screening assay are accumulating in recent years but cannot be combined with the older, S segment-based datasets [10,11].

Using sequences from the 2017 outbreak in Germany, together with human- and rodent-derived PUUV sequences from Genbank, we aimed at investigating the possible temporal and spatial changes of known geographic clusters. Furthermore, we evaluated whether the routinely generated L segment sequences provide sufficient phylogenetic signal and are suitable for studying the molecular epidemiology of the Puumala virus in Germany.

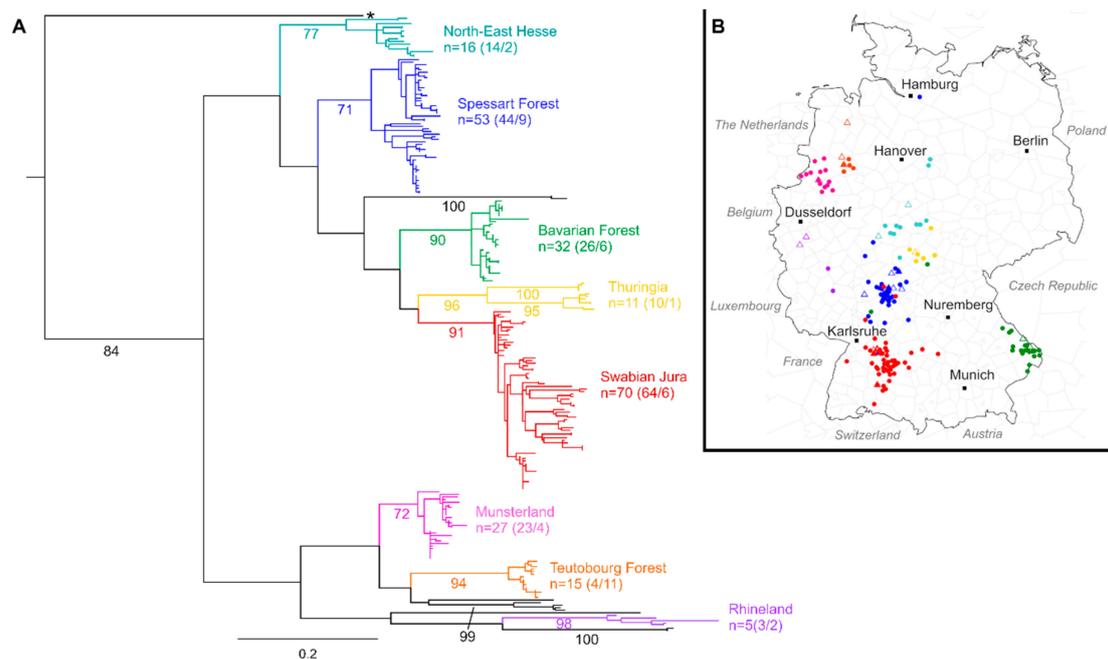
## 2. Material and Methods

As part of a nationwide alert network, serum samples from patients with acute hantavirus disease are referred to the German National Consultation Laboratory (NCL) for hantaviruses. Samples collected during 2017 until August of 2018 were tested with two different nested PCR assays, targeting the L and S segment [9,12]. Amplicons were Sanger-sequenced and processed using the MUSCLE (MULTiple Sequence Comparison by Log-Expectation) and gblocks algorithms as implemented in SeaView [13–15]. All sequences were analysed in the context of available sequences from previous years. A subset of samples for which both, S and L segment sequences could be generated, was analysed separately to allow for direct comparison. All maximum likelihood (ML) phylogenetic trees were inferred using the PhyML web server, including the implemented SMS (Smart Model Selection) algorithm to define the best-fit model [16,17]. The sequence analysis for the preparation of the boxplot and tanglegram, and the maps were done in RStudio using the packages ape, dendextend, and maps [18–21].

## 3. Results

During the hantavirus outbreak in 2017 in Germany and the first eight months in 2018, a total of 492 samples from patients with a clinical and serological diagnosis of hantavirus disease were referred to the NCL for molecular diagnostics and surveillance purposes. In our generic diagnostic PCR assay (addressing the L segment), 73/492 (14.8%) tested positive for hantavirus RNA. Of these, 61 (83.6%) also tested positive in a PCR targeting the S segment. All the resulting sequences were analysed in the context of respectively available sequences from previous years. Novel sequences were submitted to Genbank under accession numbers (MN026167–MN026264).

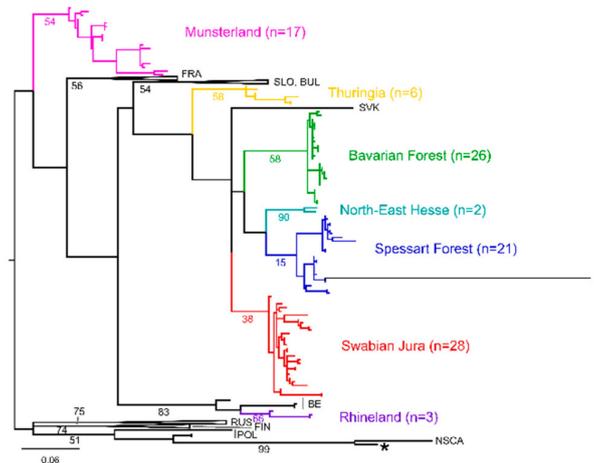
When analysing PUUV S segment sequences obtained from patients during 2017 and 2018 together with sequences available in Genbank, 194/206 (94.2%) human-derived sequences fall into well-supported phylogeographic clusters (Figure 1A). The regional allocations are supported by sequences derived from bank voles captured in the respective areas, and are mostly in concordance with the putative place of infection as determined by residence or known travel history of the patients (Figure 1B and Supplementary Information S1) [5]. In each cluster, sequences from multiple years are interspersed, showing no temporal signal. For clarity, phylogeographic clusters are only colored when comprised of five or more sequences. A detailed phylogenetic tree of S segment sequences can be found in the Supplementary Figure S2.



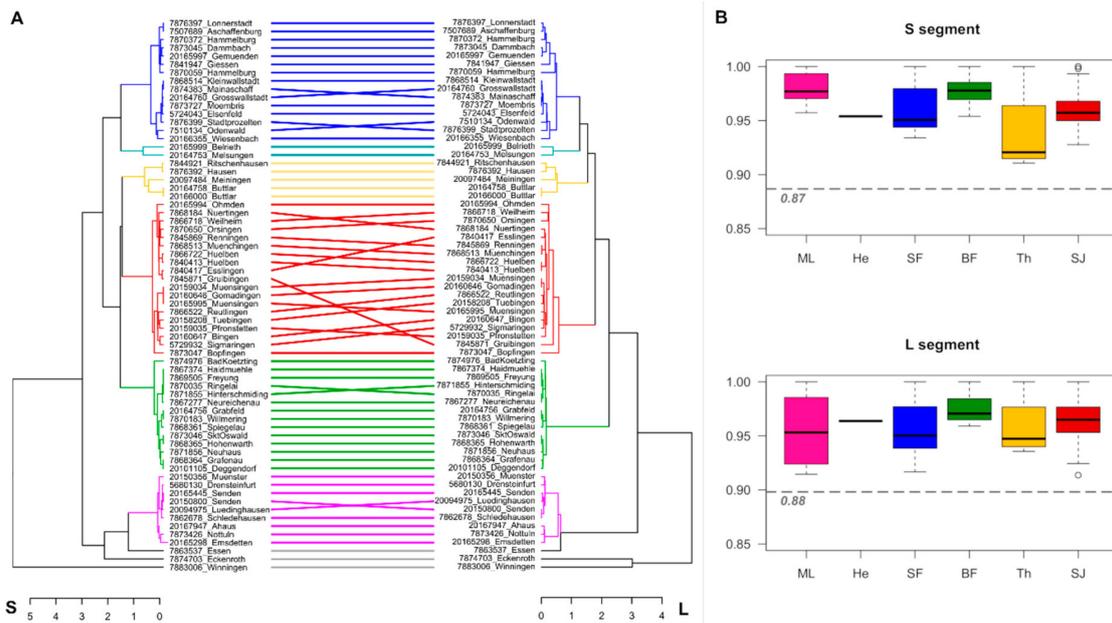
**Figure 1.** (A)—Phylogenetic analysis of Puumala virus S segment sequences. Human sequences stem from samples referred to the German National Hantavirus Consultation Laboratory, 2004–2018. Rodent sequences were downloaded from Genbank. A detailed list of included sequences is given in Supplementary Information S2. Maximum likelihood tree based on a 504 nt alignment, calculated using the HKY85 +G +I model of nucleotide substitution is shown. Scale bar indicates nucleotide substitutions per site. Bootstrap values (500 iterations) are given in percent for relevant clades. The tree was rooted by a human-derived sequence from Finland (accession number EU833888, not shown). Colors indicate phylogeographic clusters; black branches refer to sequences not yet assigned to any clear-defined phylogeographic cluster. n = number of sequences within the respective cluster (human origin/rodent origin). The sequence marked with a star stems from a patient with a travel history and known rodent exposure abroad. (B) Map showing the allocation of Puumala virus sequences in Germany. Circles indicate human-derived sequences; triangles indicate rodent-derived sequences. Multiple sequences originating from the same location are depicted only once for visibility. Colors are according to clusters in the phylogenetic tree (A). Black squares indicate major German cities to facilitate orientation.

We then analyzed sequences derived from the L segment and observed a similar picture. Of all the available PUUV L sequences obtained from patients in Germany, 91/97 (93.8%) fall into well-supported phylogenetic clusters that are concordant to those seen in the S segment based analysis, without any sign of temporal clusters (Figure 2).

To allow for a direct comparison, we used all samples from 2015–2018 for which both sequences (L and S segment derived) were available and compared the resulting phylogenetic trees. While the overall topologies of the ML trees slightly differ, the previously identified phylogenetic and geographic clusters remain stable, as shown in Figure 3A. The pairwise identity (PWI) of sequences within those phylogeographic clusters averages to  $\geq 93\%$  for both, S and L segment derived sequences, whereas the average PWIs between sequences from distinct clusters are  $\leq 88\%$ , as illustrated in Figure 3B.



**Figure 2.** Phylogenetic analysis of Puumala virus L segment sequences. Colors indicate phylogeographic clusters in Germany. The maximum likelihood (ML) tree based on a 288 nt alignment, calculated using the GTR +G model of nucleotide substitution is shown. The tree is midpoint rooted. The scale bar indicates nucleotide substitutions per site. Bootstrap values (500 iterations) are given in percent for relevant clades. Sequences marked with a star and hash stem from patients with reported travel history abroad.



**Figure 3.** (A)—Tanglegram of human-derived PUUV sequences based on phylogenetic trees of the S (left side) and L (right side) segments. (B) Boxplots showing the pairwise identities (PWI) within defined phylogeographic clusters. Gray dashed line, and the number below indicates the mean PWI between all sequences in the dataset (87% and 88% for S and L segment alignment, respectively). Colors indicate phylogeographic regions: Pink = Muensterland, turquoise = North-East Hesse, blue = Spessart Forest, green = Bavarian Forest, yellow = Thuringia, red = Swabian Jura.

**4. Discussion**

Hantavirus disease caused by Puumala virus belongs to the most important notifiable virus diseases endemic in Germany. Analyses of patient’s sequences obtained in Germany between 2004 and 2018 revealed very low longitudinal variations within defined outbreak regions. In agreement

with epidemiological data, this indicates a stable circulation of distinct PUUV variants in defined geographic areas [5].

Since multiple hantavirus species have been circulating in Germany and also imported infections have to be taken into consideration, it is highly advantageous that the diagnostic RT-PCR assay targets all so-far known hantavirus species. The L segment sequences obtained as the outcome of the pan-hanta L-PCR assay are usually used only for molecular screening and identification of the hantavirus species. For further phylogenetic studies at the species level, additional S segment sequences had been assumed to be essential. Our study indicates that this is no longer strictly necessary.

Moreover, in this study, S segment sequences were not obtained for all pan-hanta L-PCR positive samples. Consequently, relying on S segment sequences for further analysis would neglect that 16% for which only the L segment sequences are available. Here we showed that despite a very large degree of amino acid conservation (median identity between all sequences 0.98) on the L segment sequence, the nucleotide differences (Figure 3B) seem to be sufficient for meaningful phylogenetic analyses. While the tree topologies slightly differ, the main phylogeographic clusters remain identical with those identified with S segment analyses, albeit with lower statistical support. This is in accordance to data from Slovenia, where a correlation of phylogenetic and geographical clustering based on PUUV L sequences has also been described [22].

It was previously shown that sequences from viruses circulating outside Germany could be identified using data from the S segment [5]. When analyzing available patient L segment sequences, we observe two outliers within our dataset, represented by long branches compared to the rest (Figure 2; sequences marked with a star and a hash). Both patients reported travel histories abroad with possible rodent exposure. One patient (sequence marked with a star) returned from Swedish Lapland prior to developing symptoms. Accordingly, L and S segment sequences obtained from this patient cluster within the North-Scandinavian PUUV lineage (Figure 2 and Supplementary Figure S2, respectively). For one patient (marked with a hash in Figure 2), the place of infection remains elusive. The sequence does not cluster with any available sequence from Genbank, and no further information could be obtained on the reported travel destination. These results show that the L segment data is sufficient to distinguish between imported cases and the autochthonous PUUV infections in Germany.

Direct comparison of the two datasets indicates that the analysis of the L segment-based data is sufficient to allocate the patient-derived sequences to the defined phylogeographic clusters representing the endemic regions in Germany. Additional efforts to obtain the S segment sequences are not strictly necessary. Traversal lines observed upon direct comparison, and varying PWIs within clusters (Figure 3) could indicate reassortment events as previously described for PUUV [23]. However, full-length sequences would be necessary to confirm whether this is truly the case here. While short genomic segments seem sufficient for the purpose of molecular surveillance, it is always preferable to obtain sequence information of longer genomic stretches and from all genome segments. This not only enables more reliable and precise estimation of evolutionary relationships but also allows for studying the occurrence of segment reassortments and its relevance for the PUUV emergence dynamics [24].

In summary, we have shown that molecular epidemiologic studies of Puumala virus can be performed using exclusive data obtained from the routine (L segment-based) molecular diagnostic assay. This leads to a significant reduction of the time required from sample receipt to first results, which can be crucial for timely public health interventions in acute outbreak situations. Despite the recent progress in understanding PUUV epidemiological situation in Germany, continuous molecular surveillance remains a necessary tool to detect new virus variants and putative new outbreak regions and to face new challenges related with global climate and social changes.

**Supplementary Materials:** The following are available online at <http://www.mdpi.com/1999-4915/11/8/694/s1>.

**Author Contributions:** Conceptualization—S.W. and J.H.; Investigation—S.W., J.H., and B.T.; Visualization—S.W.; Writing—original draft—S.W.; Writing—review & editing—B.K. and D.H.K.

**Funding:** This study was supported by the Robert Koch Institute with funds from the Federal Ministry of Health in Germany (grant no. 1369-382). B.T. was supported by the Federal Ministry of Education and Research (grant

no. 01KI1721C). S.W. was supported by the Federal Ministry of Education and Research (grant no. 031L0096B). We acknowledge support from the German Research Foundation (DFG) and the Open Access Publication Fund of Charité—Universitätsmedizin Berlin.

**Acknowledgments:** We would like to thank all laboratories and clinicians that sent us samples for molecular surveillance and Christine Stephan, Susanne Schwarz and Monika Just for excellent technical assistance.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Latus, J.; Schwab, M.; Tacconelli, E.; Pieper, F.-M.; Wegener, D.; Dippon, J.; Müller, S.; Zakim, D.; Segerer, S.; Kitterer, D.; et al. Clinical Course and Long-Term Outcome of Hantavirus-Associated Nephropathia Epidemica, Germany. *Emerg. Infect. Dis.* **2015**, *21*. [[CrossRef](#)] [[PubMed](#)]
2. Krüger, D.H.; Ulrich, R.G.; Hofmann, J. Hantaviruses as Zoonotic Pathogens in Germany. *Dtsch. Arztebl. Int.* **2013**, *110*, 461–467. [[CrossRef](#)] [[PubMed](#)]
3. Vaheri, A.; Henttonen, H.; Voutilainen, L.; Mustonen, J.; Sironen, T.; Vapalahti, O. Hantavirus Infections in Europe and Their Impact on Public Health. *Rev. Med. Virol.* **2013**, *23*, 35–49. [[CrossRef](#)] [[PubMed](#)]
4. Reil, D.; Binder, F.; Freise, J.; Imholt, C.; Beyrer, K.; Jacob, J.; Krüger, D.H.; Hofmann, J.; Dreesman, J.; Ulrich, R.G. Hantaviren in Deutschland: Aktuelle Erkenntnisse Zu Erreger, Reservoir, Verbreitung Und Prognosemodellen Hantaviruses in Germany: Current Knowledge on Pathogens, Reservoirs, Distribution and Forecast Models. *Berliner und Münchener Tierärztliche Wochenschrift* **2018**, *131*, 453–464. [[CrossRef](#)]
5. Faber, M.; Krüger, D.H.; Auste, B.; Stark, K.; Hofmann, J.; Weiss, S. Molecular and Epidemiological Surveillance of Human Puumala and Dobrava-Belgrade Hantavirus Infections, Germany, 2001–2017. *Eurosurveillance* **2019**, (in press).
6. Schilling, S.; Emmerich, P.; Klempa, B.; Auste, B.; Schnaith, E.; Schmitz, H.; Krüger, D.H.; Günther, S.; Meisel, H. Hantavirus Disease Outbreak in Germany: Limitations of Routine Serological Diagnostics and Clustering of Virus Sequences of Human and Rodent Origin. *J. Clin. Microbiol.* **2007**, *45*, 3008–3014. [[CrossRef](#)] [[PubMed](#)]
7. Ettinger, J.; Hofmann, J.; Enders, M.; Tewald, F.; Oehme, R.M.; Rosenfeld, U.M.; Sheikh Ali, H.; Schlegel, M.; Essbauer, S.; Osterberg, A.; et al. Multiple Synchronous Outbreaks of Puumala Virus, Germany, 2010. *Emerg. Infect. Dis.* **2012**, *18*, 1461–1464. [[CrossRef](#)] [[PubMed](#)]
8. Hofmann, J.; Meisel, H.; Klempa, B.; Vesenbeckh, S.M.; Beck, R.; Michel, D.; Schmidt-Chanasit, J.; Ulrich, R.G.; Grund, S.; Enders, G.; et al. Hantavirus Outbreak, Germany, 2007. *Emerg. Infect. Dis.* **2008**, *14*, 850–852. [[CrossRef](#)]
9. Klempa, B.; Fichet-Calvet, E.; Lecompte, E.; Auste, B.; Aniskin, V.; Meisel, H.; Denys, C.; Koivogui, L.; Ter Meulen, J.; Krüger, D.H. Hantavirus in African Wood Mouse, Guinea. *Emerg. Infect. Dis.* **2006**, *12*, 838–840. [[CrossRef](#)]
10. Weber de Melo, V.; Sheikh Ali, H.; Freise, J.; Kühnert, D.; Essbauer, S.; Mertens, M.; Wanka, K.M.; Drewes, S.; Ulrich, R.G.; Heckel, G. Spatiotemporal Dynamics of Puumala Hantavirus Associated with Its Rodent Host, *Myodes Glareolus*. *Evol. Appl.* **2015**, *8*, 545–559. [[CrossRef](#)]
11. Castel, G.; Couteaudier, M.; Sauvage, F.; Pons, J.-B.; Murri, S.; Plyusnina, A.; Pontier, D.; Cosson, J.-F.; Plyusnin, A.; Marianneau, P.; et al. Complete Genome and Phylogeny of Puumala Hantavirus Isolates Circulating in France. *Viruses* **2015**, *7*, 5476–5488. [[CrossRef](#)] [[PubMed](#)]
12. Bowen, M.D.; Gelbmann, W.; Ksiazek, T.G.; Nichol, S.T.; Nowotny, N. Puumala Virus and Two Genetic Variants of Tula Virus Are Present in Austrian Rodents. *J. Med. Virol.* **1997**, *53*, 174–181. [[CrossRef](#)]
13. Edgar, R.C. MUSCLE: Multiple Sequence Alignment with High Accuracy and High Throughput. *Nucleic Acids Res.* **2004**, *32*, 1792–1797. [[CrossRef](#)] [[PubMed](#)]
14. Talavera, G.; Castresana, J. Improvement of Phylogenies after Removing Divergent and Ambiguously Aligned Blocks from Protein Sequence Alignments. *Syst. Biol.* **2007**, *56*, 564–577. [[CrossRef](#)] [[PubMed](#)]
15. Gouy, M.; Guindon, S.; Gascuel, O. SeaView Version 4: A Multiplatform Graphical User Interface for Sequence Alignment and Phylogenetic Tree Building. *Mol. Biol. Evol.* **2010**, *27*, 221–224. [[CrossRef](#)] [[PubMed](#)]
16. Guindon, S.; Dufayard, J.-F.; Lefort, V.; Anisimova, M.; Hordijk, W.; Gascuel, O. New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. *Syst. Biol.* **2010**, *59*, 307–321. [[CrossRef](#)]

17. Lefort, V.; Longueville, J.-E.; Gascuel, O. SMS: Smart Model Selection in PhyML. *Mol. Biol. Evol.* **2017**, *34*, 2422–2424. [[CrossRef](#)] [[PubMed](#)]
18. RStudio Team. *RStudio: Integrated Development Environment for R*; RStudio, Inc.: Boston, MA, USA, 2016; Available online: <http://www.rstudio.com> (accessed on 2 July 2018).
19. Paradis, E.; Schliep, K. Ape 5.0: An Environment for Modern Phylogenetics and Evolutionary Analyses in {R}. *Bioinformatics* **2018**, *35*, 526–528. [[CrossRef](#)]
20. Galili, T. Dendextend: An R Package for Visualizing, Adjusting, and Comparing Trees of Hierarchical Clustering. *Bioinformatics* **2015**, *31*, 3718–3720. [[CrossRef](#)]
21. Becker, R.A.; Wilks, A.R.; Brownrigg, R. Maps: Draw Geographical Map. 2018. Available online: <https://cran.r-project.org/package=maps> (accessed on 10 May 2019).
22. Korva, M.; Knap, N.; Rus, K.R.; Fajs, L.; Grubelnik, G.; Bremec, M.; Knapič, T.; Trilar, T.; Županc, T.A. Phylogeographic Diversity of Pathogenic and Non- Pathogenic Hantaviruses in Slovenia. *Viruses* **2013**, *5*, 3071–3087. [[CrossRef](#)]
23. Razzauti, M.; Plyusnina, A.; Henttonen, H.; Plyusnin, A. Microevolution of Puumala Hantavirus during a Complete Population Cycle of Its Host, the Bank Vole (*Myodes glareolus*). *PLoS ONE* **2013**, *8*, e64447. [[CrossRef](#)] [[PubMed](#)]
24. Szabó, R.; Radosa, L.; Ličková, M.; Sláviková, M.; Heroldová, M.; Stanko, M.; Pejčoch, M.; Osterberg, A.; Laenen, L.; Schex, S.; et al. Phylogenetic Analysis of Puumala Virus Strains from Central Europe Highlights the Need for a Full-Genome Perspective on Hantavirus Evolution. *Virus Genes* **2017**, *53*, 913–917. [[CrossRef](#)] [[PubMed](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).