Puumala hantavirus and *Myodes glareolus* in northern Europe: no evidence of co-divergence between genetic lineages of virus and host

Kirill Nemirov,1 Herwig Leirs,2,3 Åke Lundkvist1,4 and Gert E. Olsson1,5

**Correspondence**
Kirill Nemirov
kirillnemirov@hotmail.com

1Department of Virology, Swedish Institute for Infectious Disease Control, SE-171 82 Solna, Sweden
2Danish Pest Infestation Laboratory, University of Aarhus, DK-2800 Kgs Lyngby, Denmark
3Evolutionary Ecology Group, University of Antwerp, BE-2020 Antwerpen, Belgium
4Department of Microbiology, Tumour and Cell Biology, Karolinska Institute, SE-171 77 Stockholm, Sweden
5Department of Wildlife, Fish, and Environmental Studies, Swedish University of Agricultural Sciences, SE-901 83 Umeå, Sweden

The genus *Hantavirus* (family *Bunyaviridae*) includes negative-strand RNA viruses that are carried by persistently infected rodent and insectivore species. Puumala virus (PUUV), carried by bank voles (*Myodes glareolus*), is a pathogenic hantavirus that causes outbreaks of mild haemorrhagic fever with renal syndrome across Europe. In northern Europe, PUUV is represented by several genetic lineages that are maintained by distinct phylogroups of bank voles. The present study describes sequences of new PUUV strains recovered from northern and southern regions of Scandinavia and compares phylogenetic relationships between north-European PUUV strains and *M. glareolus*. This analysis revealed contradictions in phylogenetic clustering and remarkable differences in estimated divergence times between the lineages of PUUV and its host, suggesting that the established PUUV lineages did not co-diverge with the distinct phylogroups of *M. glareolus* that carry them at present.

**INTRODUCTION**

Hantaviruses (genus *Hantavirus*, family *Bunyaviridae*) are zoonotic viruses with a segmented negative-strand RNA genome that cause two clinically distinct human diseases: hantavirus cardiopulmonary syndrome in the Americas and haemorrhagic fever with renal syndrome (HFRS) in Europe and Asia. The hantavirus genome consists of three genomic segments: small (S), encoding the nucleocapsid (N) protein, medium (M) encoding the surface glycoproteins (Gn and Gc) and large (L) segment encoding the RNA-dependent RNA polymerase (Plyusnin et al., 1996).

Distinct hantaviruses are maintained in nature by different rodent or insectivorous species, and can be transmitted to humans through aerosolized host excreta. All rodent species in which hantaviruses have been discovered to date belong to three subfamilies of muroid rodents: Arvicolinae, Murinae and Sigmodontinae (the latter also includes Neotominae, which is sometimes considered a separate subfamily; Wilson & Reeder, 2005). The notion of a substantial similarity in the phylogenetic relationships of rodents and hantaviruses (largely irrespective of their present geographical distribution) has resulted in a theory of hantavirus–host co-evolution (Plyusnin et al., 1996; Plyusnin & Morzunov, 2001). The theory suggests that hantaviruses have co-evolved and co-diverged with their hosts throughout most of their evolutionary history, with infrequent virus transmissions between distantly related host species (i.e. host switching) resulting in a significant co-phylogeny of hantavirus–host relationships (Vapalahti et al., 1999; Plyusnin & Morzunov, 2001; Nemirov et al., 2002). The co-evolution theory is also supported by data suggesting that distinct genetic lineages of certain hantaviruses are associated with phylogenetically distinct populations of host species (Morzunov et al., 1998; Plyusnin & Morzunov, 2001). However, the applicability of the co-evolution theory to these viruses has recently been questioned, and preferential host switching followed by virus speciation has been suggested as an alternative explanation for the convergent phylogenies of hantaviruses and rodents (Ramsden et al., 2009).

The GenBank/EMBL/DDBJ accession numbers for the sequences determined in this study are GQ339473–GQ339503.

Results of saturation tests and details of the hantavirus and rodent sequences used for phylogenetic analysis are available with the online version of this paper.
Puumala virus (PUUV) is a major hantavirus pathogen that causes a milder form of HFRS (also known as nephropathia epidemica) in central and northern Europe and central Russia. Most of the serologically verified HFRS cases caused by this virus occur in northern Europe, with a mean number of 1000 cases in Finland, 350 cases in Sweden and 100 cases in Norway annually (Vapalahti et al., 2003; Bi et al., 2008). The natural host of PUUV, the bank vole (Myodes glareolus), is widely distributed across Europe (Mitchell-Jones et al., 1999). The Scandinavian peninsula in northern Europe represents one of the most interesting geographical regions for studying associations between PUUV and its host, firstly because the recolonization history of this region by animal species after the retreat of Weichselian glaciation [~13–8 thousand years ago (TYA)] is relatively well known (Tegelström, 1987; Jaarola et al., 1999), and secondly, because intensive studies of PUUV in northern Europe have generated a substantial amount of genetic data (Sironen et al., 2001).

Previous phylogeographical studies of M. glareolus have indicated that, in Fennoscandia (Scandinavia and Finland), this species is represented by several genetic lineages distinguishable by mitochondrial DNA (mtDNA) sequences (Tegelström, 1987). Two distinct mtDNA types were described, one resident in the northern parts of Sweden and Finland (northern mtDNA type) and the other in central and southern Fennoscandia (southern mtDNA type). Rodents that belong to different types are divided by a contact zone (~50 km wide) in central Sweden, and by another population border in central Finland. Analysis of mtDNA has demonstrated that, although bank voles from Finland and southern Scandinavia belong to the same mtDNA type, they differ from each other. Such a population structure of M. glareolus in this region was explained by post-glacial recolonization of Fennoscandia from several directions around 13–8 TYA (Tegelström, 1987). One stream of migrating bank voles is thought to have come from the south across the land bridge that existed at that time between the territories of modern Sweden and Denmark, and to have settled in southern Scandinavia. Other streams of migrants arrived from the south-east and north-east directions to settle in central Finland and northern Fennoscandia, respectively. mtDNA of the northern M. glareolus was found to be very similar to that of the red-backed vole (Myodes rutilus), and hybridization between M. glareolus and M. rutilus species during the recolonization process was suggested as a possible origin of that mtDNA type (Tegelström, 1987). The most recent phylogeographical studies based on the mitochondrial cytochrome b gene have demonstrated that bank voles in Europe belong to several distinct phylogenetic groups: western, eastern, Balkan and Ural (Deffontaine et al., 2005; Kotlik et al., 2006).

Previous studies of hantaviruses in northern Europe identified four distinct genetic lineages of PUUV: Finnish (FIN), North Scandinavian (N-SCA), South Scandinavian (S-SCA) and Danish (DAN) (Plyusnin et al., 1994, 1995; Hörling et al., 1996; Lundkvist et al., 1998; Asikainen et al., 2000; Sironen et al., 2002; Johansson et al., 2008). The FIN lineage was recovered from bank voles trapped in southern parts of Finland and Russian Karelia, the DAN lineage originated from the Danish island of Fyn, and the lineages N-SCA and S-SCA were identified in bank vole populations that located to the north and south of the Scandinavian bank vole population border, respectively. As PUUV strains recovered from different sides of the population border were clearly genetically distinct, it was suggested that each mtDNA type of bank voles brought its own genetic variant of PUUV during the recolonization of Fennoscandia (Asikainen et al., 2000). Although such a scenario appears to be the most logical one, no comparative analyses of the virus–host phylogenetic relationships have been performed, and the detailed relationships between PUUV lineages identified in the region remain obscure (Plyusnin & Morzunov, 2001; Sironen et al., 2001). In the present study, we have described new sequences of PUUV recovered from northern and southern regions of Scandinavia, and have performed comparative phylogenetic analysis of the virus–host relationships in order to understand better the role of phylogeographic factors in PUUV evolution and its co-differentiation with the host.

RESULTS

PUUV prevalence in different geographical regions of Sweden

Screening of 957 bank vole specimens collected in the course of this study for PUUV IgG by ELISA identified 48 positive samples (5.0%). The distribution of seropositive animals varied significantly between the northern and southern regions. The highest seroprevalence for PUUV was detected in the northern sampling region, where 34/220 bank voles (15.5%) collected from 14 sampling sites had antibodies against the virus. In central Sweden, only 7/601 bank voles (1.2%) collected at 19 sampling sites along the south-central transect, and 7/68 bank voles (10.3%) sampled near the bank vole contact zone, respectively, had detectable PUUV antibodies. None of 68 bank voles sampled at three different sites in southern Sweden was seropositive for PUUV.

Sample screening and RT-PCR

Rodent serum samples positive for IgG in the ELISA test were screened by a focus reduction neutralization test (FRNT). Twenty-six out of 48 animals were positive in the FRNT and all were subjected to RT-PCR specific for the PUUV S segment. The PCR resulted in 14 complete S segment sequences of the virus that originated from 11 different locations in northern and central Sweden: seven localities in Norrbotten county, two localities in Västernorrland county and two localities in Uppland county (Fig. 1). Additionally, partial sequences of the
PUUV M segment (nt 2164–2651) and partial sequences of the bank vole mitochondrial cytochrome \( b \) gene (nt 102–1080) were amplified from the eight tissue samples that produced the most divergent S segment sequences of PUUV (four from Norrbotten, two from Västernorrland and two from Uppland).

**Phylogenetic analysis of M. glareolus in northern Europe**

Phylogenetic analysis indicated that bank voles currently inhabiting northern Europe belonged to three distinct phylogenetic clades (Fig. 2): (i) rodents from southern Scandinavia belonging to the ‘western’ clade, which were most closely related to bank voles from central–eastern Europe; (ii) rodents from Denmark (originating from the Jutland peninsula and the islands of Zealand and Fyn), which were placed within the ‘eastern’ phylogenetic group together with voles from Finland, Poland, Ukraine and western Russia; and (iii) bank voles from northern Sweden, which were placed in the ‘Ural’ clade, together with \( M. \) glareolus from central and Siberian parts of Russia. The latter phylogroup was more distantly related to the first two, and was more closely related to \( M. \) rutilus than to other \( M. \) glareolus species.

**Comparative analysis of new PUUV strains**

Phylogenetic analysis based on the complete coding sequence of the S segment placed newly recovered PUUV strains within two previously defined phylogenetic lineages: N-SCA and S-SCA.

Nine strains recovered from Norrbotten county in Sweden were placed within the N-SCA lineage forming a distinct sublineage: northern N-SCA (NN-SCA) (Fig. 3). All other strains from northern Sweden sequenced previously formed a separate phylogenetic sublineage: central N-SCA (NC-SCA). Genetic comparison indicated that strains from the northernmost localities of Kiviniemi and Äijäjärvä were most closely related to each other (up to 1% nucleotide diversity between strains) and more distantly related to strains from the eastern localities Kälvudden and Jockfall.

---

**Fig. 1.** Geographical distribution of \( M. \) glareolus and PUUV in northern Europe. Coloured regions indicate the Swedish counties from which PUUV sequences were recovered: purple, Norrbotten (NB); blue, Västerbotten (VB); green, Västernorrland (VN); pink, Uppland (UP). Place names that are underlined indicate PUUV strains described in this study, whilst those that are not underlined indicate PUUV strains described previously.
(up to 5% nucleotide diversity), whilst strains from more southern localities (Gyttjeå, Ljustraåsk and Moskosel) were the most distantly related to other strains of NN-SCA (6–8% nucleotide diversity).

Two strains from central Sweden (Västernorrland) and three strains from southern Sweden (Uppland) were placed within the S-SCA lineage, formed by previously described PUUV sequences originating from Sweden (Sollefteå) and Norway (Eidsvoll) (Lundkvist et al., 1998). Phylogenetic analysis indicated that, with the addition of newly recovered PUUV sequences, S-SCA comprised three sublineages formed by: (i) strains from Uppland (localities Mångelbo and Munga); (ii) strains from Västernorrland (Sollefteå, Bergsjöbo and Fåboviken); and (iii) strains from Norway (Eidsvoll). Two sublineages from Sweden were more closely related to each other (8–9% nucleotide diversity) than to the Norwegian sublineage (12–13% nucleotide diversity).

Phylogenetic analysis of partial M segment sequences demonstrated clustering of newly recovered PUUV strains similar to that of the S segment: sequences from Norrbotten (localities Kiviniemi, Jockfall, Moskosel and Kälvudden) were placed together with strains sequenced earlier from Västerbotten (N-SCA lineage), whilst sequences from Västernorrland and Uppland were placed within the S-SCA lineage together with previously described sequences (see Fig. 5a).

**Phylogenetic relationships between PUUV lineages from northern Europe**

**S segment.** Previous studies demonstrated that phylogenetic trees of PUUV often have a ‘star-like’ topology where evolutionary relationships between different PUUV lineages appear unresolved. Such a topology has always been interpreted as evidence of an early evolutionary split between distinct genetic lineages of the virus (Asikainen et al., 2000; Sironen et al., 2001). However, two alternative explanations may be considered: (i) homologous recombination between strains of PUUV, and (ii) the loss of an essential part of phylogenetic information due to substitution saturation. Recombination analysis performed initially with the Simplot program identified several regions with distinctive similarity patterns in the coding part of the S segment. However, other recombination-detecting methods (implemented in the RDP3 program) indicated that these potential recombination signals were not statistically significant. Saturation of synonymous sites has been reported previously in comparisons of distantly related hantaviruses (Hughes & Friedman, 2000), and we used a substitution saturation test implemented in DAMBE to evaluate the hypothesis of significant saturation in PUUV datasets. The test indicated that nucleotide substitutions might be saturated in large datasets that include sequences of distantly related hantavirus species, but they are not saturated in smaller datasets (e.g. including only PUUV or sequences of closely related hantaviruses) (see Supplementary Table S1, available in JGV Online).

Phylogenetic trees calculated by the maximum-likelihood (ML), Bayesian, maximum-parsimony (MP) and distant matrix (DM) methods [based on alignments of PUUV strains and Hokkaido virus (HOKV)] demonstrated consistent placement of PUUV lineages into well-supported clusters; however, the clustering patterns inferred from the nucleotide and amino acid sequences were different (Fig. 4a, c). Remarkably, the placement of an outgroup sequence also differed between the nucleotide and amino acid trees, suggesting that lineages of PUUV show different patterns of similarity to an outgroup at the nucleotide and amino acid levels. In order to minimize the influence of the outgroup sequence on phylogenetic reconstruction, we repeated the analysis on datasets that included only PUUV strains. Phylogenetic trees based on the nucleotide and amino acid alignments supported bipartition of the dataset into two distinct groups: one including lineages S-SCA and FIN, and the other N-SCA and DAN (Fig. 4b, d).

**M segment.** Phylogenetic trees calculated with the ML method based on the nucleotide and amino acid sequences had similar topology and indicated that the Scandinavian
PUUV lineages N-SCA and S-SCA were most closely related to each other and more distantly related to the DAN lineage, whilst the FIN lineage was the most distinct (Fig. 5a). Initially, only the amino acid tree demonstrated reliable grouping of PUUV lineages; however, removal of the third codon position from the nucleotide alignment significantly improved resolution of the nucleotide-based tree. Whilst other phylogenetic methods also indicated the FIN lineage to be the most distinct one, the relationships between the three remaining lineages were not resolved with certainty. Congruence of the ML, Bayesian and MP methods improved when the outgroup sequence was removed from the nucleotide dataset: all three methods supported bipartition of four PUUV lineages into two distinct groups: one including lineages N-SCA and S-SCA, and the other lineages DAN and FIN (Fig. 5b).

**Co-phylogeny testing**

The testing of the bank vole and PUUV S segment-based phylogenetic trees for congruence was performed with the TreeMap v1.0 and CoRe-PA programs (Fig. 6a). Both programs determined the most parsimonious scenario for reconciliation of host (rodent) and parasite (virus) tree topologies, and used statistical tests that estimate the likelihood of obtaining such reconciliation by chance. The most likely scenario of co-evolution suggested by both programs implied two switching events: (i) host switching of HOKV from *M. rufocanus* to *M. glareolus* that established PUUV; and (ii) transfer of N-SCA lineage between genetic lineages of *M. glareolus*, which established the ancestor of the FIN and S-SCA lineages. Statistical tests indicated that the proposed number of co-speciation events was significantly higher than expected by chance alone, i.e. the dataset demonstrated statistically significant co-phylogeny between host and virus species (Fig. 7) that could be indicative of co-evolution.

**Estimation of the divergence times**

Two scenarios of hantavirus and host evolution that could result in substantial co-phylogeny were considered: (i) viruses have co-diverged and co-speciated with the distinct rodent species/genetic lineages (Sironen et al., 2001); or (ii)
viruses were acquired by *Myodes* rodents in the course of evolution and spread between different hosts via preferential host switching (Ramsden *et al.*, 2008). Mean estimates of the divergence times between North European lineages of *M. glareolus* and lineages of PUUV were obtained from the phylogenetic trees with the BEAST program (Fig. 8). Firstly, phylogenetic trees of *Myodes* rodents and associated hantaviruses were constructed by the ML method and tested for the clock-like pattern of evolution. Both datasets passed the likelihood ratio test for the molecular clock, suggesting that they were fit for estimations of constant evolutionary rates. To calibrate the rodent phylogenetic tree, we used time estimates of one ancient and one recent historic event: (i) the divergence of *M. glareolus* and *M. rutilus* (estimated at 2.5 MYA by Lebedev *et al.*, 2007); and (ii) the isolation of *M. rufocanus* on Hokkaido island (Japan) by the rising sea levels, approximately 11 TYA (Ohdachi *et al.*, 1997; Bogatov *et al.*, 2006). Two distinct priors resulted in two very different sets of estimates. The ‘ancient’ calibration suggested that the western and eastern mitochondrial lineages of *M. glareolus* diverged around 800 TYA, and that the Ural mitochondrial lineage was formed approximately 420 TYA (Fig. 8). Estimates of the same divergence events based on the ‘recent’ calibration date were significantly smaller: 19 and 10 TYA, respectively. Such poorly reconcilable estimates may indicate that one of the priors is incorrect; however, the timing of both calibration events was supported by currently available data. The evolutionary rate estimated from the first tree ($1.6 \times 10^{-8}$) was closer to the evolutionary rates commonly assumed for mammalian cytochrome *b* sequences than that of the second tree ($7.95 \times 10^{-7}$) (Brown *et al.*, 1979; Lanave *et al.*, 1984). On the other hand, the estimates of divergence times produced by the second tree were more consistent with the data available on post-glacial migration and geographical isolation of rodent populations. More specifically: (i) the mean divergence of *M. rufocanus* and *M. rutilus* populations on Sakhalin Island (estimated at 7.4 and 5.9 TYA, respectively) correlated with the isolation of Sakhalin from the mainland (around 7 TYA) by the rising sea levels (Ohdachi *et al.*, 1997; Bogatov *et al.*, 2006); (ii) the mean divergence of *M. rutilus* on Hokkaido was estimated at 9.9 TYA; and (iii) the mean divergence of *M. rutilus* from the Russian Far East and Alaska (separated in the course of rodent migration over the Beringian land bridge) was estimated at around 13.4 TYA (Cook *et al.*, 2005). Although currently available paleogeographic data do not allow a decisive preference of one set of estimates over the other, the estimates based on the ‘recent’ calibration event may be used as an approximation of the fastest theoretically possible divergence of rodent lineages. As phylogenetic trees based on the ‘ancient’ and ‘recent’ calibrations had similar proportions of branch lengths and time estimates, only the estimates from the ‘recent’ tree were used to compare divergence times of host and virus lineages.

The host switching of HOKV to *M. glareolus* suggested by co-phylogeny analysis is unlikely to have occurred later than 11 TYA, when Hokkaido island became isolated from the

---

**Fig. 4.** Clustering of PUUV lineages on the phylogenetic trees with and without the outgroup. Trees were calculated using TREE-PUZZLE (ML) based on the complete coding region of the S segment (a, b) or the complete amino acid sequence of the N protein (c, d). Dividing lines indicate separation of PUUV lineages in phylogenetically distinct groups. Values supporting each clustering pattern are underlined and given in the order: ML, Bayesian, DM, MP. Bars, 0.1 nucleotide (a, b) or 0.01 amino acid (c, d) substitutions per site.
Asian mainland by the rising sea (Bogatov et al., 2006). Assuming the most recent possible time period for the host switching (during the last glaciation peak, 22–11 TYA; Bogatov et al., 2006), divergence time between HOKV and PUUV can be estimated as a normal distribution with a range of 22–11 TYA (mean estimate 16.5 TYA). Effectively, this implies that the hantavirus phylogenetic tree can be calibrated with the same date as the tree of rodent species. Such a calibration provides the following mean estimates for the divergence of virus lineages: 14.8 TYA (separation of HOKV and PUUV), 12.7 TYA (DAN lineage), 11.6 TYA (N-SCA lineage) and 9.5 TYA (S-SCA and FIN lineages) (Fig. 8).

Comparison of the divergence times estimated from the virus and host phylogenetic trees (Fig. 6b) further indicates that the lineages of PUUV could not have diverged at the same time as the lineages of the host species.

**DISCUSSION**

Reconstruction of virus evolutionary history has a general problem associated with the lack of geological records. In previous studies, evolution of hantaviruses was either assumed to follow the evolution of their hosts (Sironen et al., 2001) or was estimated based on mutation rates of other RNA viruses and the genetic variability observed in natural foci (Ramsden et al., 2008). Both approaches may present potential shortcomings: the first is almost exclusively based on the observed co-phylogeny of virus and host species, whilst the second approach uses data on the variability of hantavirus quasispecies in individual hosts and mutation rates obtained in vitro to estimate the rate of evolution in nature. Calibration of a hantavirus phylogenetic tree with known dates of sampling (tip-date analysis) may lead to an overestimation of the mutation rate if virus strains collected at different time points from distinct natural foci are compared.

In the present study, we tested the applicability of the co-evolution theory on the intraspecies level by analysing the phylogenetic relationships of *M. glareolus* and PUUV in northern Europe and comparing divergence times of host and virus genetic lineages.

Phylogenetic analysis of bank voles demonstrated that populations of *M. glareolus* from central-southern parts of
Scandinavia and Finland belonged to the western and eastern phylogroups, respectively. Remarkably, bank voles from Denmark also belonged to the eastern phylogroup, i.e. they were more closely related to the bank voles from central Finland than to the bank voles currently present in Scandinavia. Bank voles inhabiting northern Sweden (and northern Finland, as shown recently by Razzauti et al., 2009) belonged to the Ural phylogroup, which carries mtDNA of red-backed voles and is the most distinct from all other European bank vole lineages. Representatives of the Ural phylogroup were also discovered in the central part of Russia and in Ural and Siberian regions, suggesting that its evolutionary origin dates back much further than suggested previously (Deffontaine et al., 2005).

Phylogenetic relationships of PUUV strains from northern Europe reconstructed from the complete coding sequences of the S segment and partial sequences of the M segment did not mirror the interlineage relationships of M. glareolus. Analysis of the S segment sequences indicated that the lineages FIN and DAN, carried by the eastern phylogroup of bank voles, were only distantly related to each other. Instead, FIN was most closely related to the S-SCA lineage, carried by the western phylogroup of bank voles. The pattern of lineage clustering derived from the M segment-based phylogenetic trees was different from both rodent and S segment-based trees, and suggested the closest genetic ties between lineages S-SCA and N-SCA, carried by the western and Ural phylogroups of M. glareolus, respectively. Contradictory results of the analyses based on S and M segment sequences can theoretically be explained by reassortment of viral genomic segments or recombination between the ancestors of modern PUUV lineages. However, lack of congruence between viral and rodent phylogenetic trees (which has also been noticed previously; T. Sironen and A. Plyusnin, personal communication) could indicate that lineages of PUUV did not co-diverge with the lineages of the host.

Testing of host and virus trees for co-phylogeny suggested that they are only partly incongruent and could be reconciled with the assumption that two switching events occurred during the evolution of PUUV: (i) host-switching of HOKV from M. rufocanus to M. glareolus that established PUUV; and (ii) transfer of the N-SCA lineage from the Ural phylogroup of M. glareolus to a common ancestor of the western and eastern phylogroups. Statistical testing indicated that such a scenario of reconciliation is compatible with the idea of co-evolution. Establishment of PUUV via host switching of HOKV from M. rufocanus to M. glareolus...
appears highly likely, as HOKV is more closely related to PUUV than to Fusong virus (FUSV) and Muju virus (MUJV), carried by *M. rufocanus* in China and *Myodes regulus* in Korea, respectively. However, the reconciliation scenario proposed for PUUV does not take into account the origin of the Ural mitochondrial lineage of bank voles. As shown by this and earlier studies, mtDNA of that lineage was derived from another rodent species (*M. rutilus*) at some point during the evolution, and therefore its most ancestral position to other *M. glareolus* lineages on phylogenetic trees is somewhat artificial. Divergence estimates of rodent lineages presented here suggest that the transfer of mtDNA from *M. rutilus* that established the Ural phylogroup of *M. glareolus* occurred when the western and eastern phylogroups of *M. glareolus* were already established. The lack of information concerning the origin of the Ural phylogroup makes reconciliation analysis difficult, as its precise evolutionary relationship to other phylogroups of bank voles (and the evolutionary history of the N-SCA PUUV lineage) cannot be determined with certainty.

In order to compare the timing of divergence events in the evolution of PUUV and its host, we used the time estimate of the same historic event in the Late Pleistocene (separation of Hokkaido Island from the mainland). Our estimates indicated that genetic lineages of *M. glareolus* diverged much earlier than the lineages of PUUV they carry. An alternative calibration of the host phylogenetic tree with the ancient historic event suggested that divergence of the host lineages may have happened much earlier, making the difference between host and virus divergence times even more apparent. These data support the idea that PUUV has not co-evolved with *M. glareolus* throughout its evolutionary history, but was acquired from *M. rufocanus* approximately 15 TYA. Our estimates also indicate that PUUV is unlikely to have evolved with the fast mutation rates suggested previously (Ramsden *et al.*, 2008, 2009). The mean substitution rate estimated from the PUUV clock-like phylogenetic tree (2.38 × 10⁻⁵) is approximately 20–30 times lower than the estimates proposed for PUUV and other hantavirus species in those studies.

Interestingly, although the divergence times of PUUV lineages cannot be reconciled with those of the host, they are placed within the time frame of post-glacial rodent migration into northern Europe (Tegelström, 1987; Jaarola *et al.*., 1999). Moreover, the timing of rodent migration into southern Scandinavia (12–8 TYA) is consistent with the estimated mean divergence of PUUV lineages S-SCA and FIN at 9.5 TYA. At present, *M. rufocanus* is sympatric with *M. glareolus* in northern Europe and central Russia (Stenseth, 1985), and the transmission (or possibly spill-over) of PUUV between these species has been documented (Dekonenko *et al.*, 2003); therefore, one cannot exclude the possibility that host switching of HOKV from *M. rufocanus* to *M. glareolus* occurred later than has been presumed in this study. However, such a scenario would require the presence of HOKV in north European and/or central Asian populations of *M. rufocanus* after 11 TYA, and currently there are no data to suggest this. One interesting question that could not be answered by this study is how PUUV (assuming the relatively recent divergence of its genetic lineages) was transmitted between distinct genetic lineages of the host. Phylogroups of *M. glareolus* presumably recolonized Europe from several glacial refugia, some of which were located far from the possible contact zones of *M. rufocanus* and *M. glareolus* (e.g. the western phylogroup, which is thought to have survived the last glaciation in central Europe; Deffontaine *et al.*, 2005). One likely possibility is that transmission of the viruses occurred when rodents of different phylogroups (e.g. western and eastern) came into contact during the recolonization process.

The discussion concerning the evolutionary history and relationships of PUUV lineages is going to continue and...
may in fact reveal a more complex picture than appears from our analysis. For instance, comparison of all known lineages of PUUV indicates that lineages from eastern and central Russia, not the S-SCA lineage, are the closest relatives of the FIN lineage (Sironen et al., 2001; Dekonenko et al., 2003). However, the comparative analysis of PUUV and host lineages from northern Europe suggests that they did not co-diverge. Instead, the co-phylogeny analysis and estimation of the divergence times suggest that PUUV originated via viral host switching from *M. rufocanus* to *M. gareolus* around 15 TYA and was transmitted between distinct phylogenetic lineages of the new host during the post-glacial recolonization.

**METHODS**

**Rodent trapping.** Rodent sampling was carried out during 2005 in three regions of Sweden: the north (within 67° 37′–65° 51′ N, 19° 54′–20° 48′ E), the south (56° 15′–55° 34′ N, 13° 33′–13° 37′ E) and in a transect running east to west over the southern-central part of the country, i.e. the Limes Norrlandicus (59° 34′–60° 17′ N, 12° 02′–17° 23′ E). Each region was sampled over a 5-day period in autumn under similar climatic conditions. Rodents were collected by the use of snap traps that were set out for two consecutive nights and checked on a daily basis (Olsson et al., 2005). For comparisons between sampling regions, trapping results were expressed as the indexed number of voles captured per 100 trap nights (the trap index), as a reflection of the relative population size on each sampling occasion (Hanski et al., 1994). PUUV screening was also performed on archived bank vole specimens collected in 1998 in central Sweden (63° 10′–63° 17′ N, 17° 32′–17° 49′ E) close to the previously described bank vole contact zone (Hörlining et al., 1996). Lungs and kidneys were dissected from the captured animals and kept frozen at −70°C until RNA extraction. Blood for serological analysis was collected using Nobuto blood filter strips (Toyo Roshi Kaisha). The strips were dried and subsequently eluted with 1000 µl PBS.

**Serological screening of rodent sera.** Initial screening was performed by a standard ELISA using native inactivated viral antigen (PUUV, strain CG-1820) and recombinant antigens (PUUV, strain Kazan) for detection of IgG antibodies, as described previously.

---

**Fig. 8.** Chronogram of the rodent (cytochrome *b*, nt 110–995) and hantavirus (S segment, complete coding region) sequences. The divergence times correspond to the mean posterior estimate of their age in TYA. Grey bars represent the 95% highest posterior density intervals. The timing of the paleogeographic events used to calibrate phylogenetic trees of viruses and rodents is shown in grey boxes. Divergence estimates of the rodent species based on the ‘ancient’ calibration are shown in parentheses.
(Klingström et al., 2002). Hantavirus antibody-positive sera were confirmed and serotyped by a FRNT as described previously (Lundkvist et al., 1997). An 80% reduction in the number of foci was used as the criterion for virus neutralization titres.

**RT-PCR, cloning and sequencing.** Kidney tissue fragments (50–100 mg) of FRNT-positive bank voles were disrupted in a TissueLyser (Qiagen) and used for extraction of total RNA and DNA with TriPure reagent (Roche) as recommended by the manufacturer. cDNA synthesis was performed with SuperScript III (Invitrogen) and all amplifications were carried out with Taq DNA polymerase (Applied Biosystems) using gene-specific primers. Complete sequences of the 5’ segment were reverse transcribed and amplified with primers based on 21–22 conserved terminal nucleotides of PUUV (strain Kazan). Amplified products were separated in agarose gels, purified with a MinElute Gel Extraction kit (Qiagen) and cloned into pCR2.1-TOPO vector using a TOPO-TA kit (Invitrogen). Cloned fragments were sequenced on an ABI 3100 (Applied Biosystems) with a Big Dye Terminator cycle sequencing kit v3.1. Partial sequences of the M segment were recovered using a nested PCR with primers based on the consensus of complete PUUV M segment sequences. Amplification resulted in a 536 bp product that could be compared with other partial M segment sequences of PUUV deposited in GenBank. For the analysis of bank vole mitochondrial lineages, total DNA was extracted from all hantavirus RT-PCR-positive rodents and used to amplify part of the mitochondrial cytochrome b coding region with primers 5’-ATCATCAACACCKATTATG-3’ and 5’-GGTGAATGTATTACATCGC-3’. Sequences determined in this study have been deposited in GenBank under accession numbers GQ339473–GQ339503. Other sequences used in this study (retrieved from GenBank) are listed in Supplementary Table S2 (available in JGV Online).

**Sequence analysis.** Several methods of phylogenetic reconstruction were used for the analysis. To generate phylogenetic trees based on the ML criterion, we used the quartet-puzzling algorithm for the best tree search implemented in TREE-PUZZLE v5.2 (Schmidt et al., 2002) and the Bayesian approach implemented in MrBayes v3.1.2 (Huelsenbeck & Ronquist, 2001). The TREE-PUZZLE program was set to calculate 10 000–25 000 puzzling steps with a transitions/transversions ratio and nucleotide frequencies estimated from the calculated distribution of the optimal trees in MrBayes were performed with two independent Metropolis-coupled Markov chain Monte Carlo runs. Both runs started from a random tree for each of four simultaneous chains and were set to run for several million generations (most often 3–5 million) until the runs converged. The calculation was continued until the number of trees sampled after reaching the convergence was equal to 75% of all samples (burn-in of 25%). Trees obtained after the convergence point were summarized. Substitution models and model parameters best fitting each dataset were estimated with ModelTest v3.7 (Posada & Crandall, 1998). Phylogenetic inference according to the ML approach was performed with neighbour-joining and Fitch–Margoliash algorithms implemented in PHYLIP v3.4 (Felsenstein, 2005) as described previously (Nemirov et al., 1999). MP trees were calculated with PAUP v3.1 (Swofford, 1991) using a parsimony ratchet approach (Sikes & Lewis, 2001). A consensus tree was based on ten independent ratchet runs of 500 repetitions each (with a permutation of 25% informative characters) that were filtered for the trees with the best parsimony score. Phylogenetic trees calculated with all methods were visualized using TreeView v3.2 (Page, 1996). Substitution saturation tests were performed with DAMBE (Xia, 2000). Simplot v3.2 (Lole et al., 1999) and REL3D (Martin, 2009) were used to search for evidence of recombination in PUUV sequences. Co-phylogeny analysis of hantavirus and rodent sequences was performed with TreeMap v1.0 (Page, 1994) and CoRe-PA (Merkle et al., 2010). Estimation of the divergence times and evolutionary rates of rodent and virus species was performed with BEAST v1.5.2 (Drummond & Rambaut, 2007) using the available data on the node ages. The most suitable nucleotide substitution models were determined with ModelTest, and an uncorrelated lognormal relaxed molecular clock with the Yule (birth–death) demographic model was used as a prior. Each program run was continued (with sampling every 1000 generations) until the effective sampling size of all parameters was greater than 200. The resulting phylogenetic trees were summarized in TreeAnnotator (using a 10% burn-in) and visualized using FigTree (http://beast.bio.ed.ac.uk).

**ACKNOWLEDGEMENTS**

We thank Peter Levin for kindly providing the bank vole samples from the contact zone, and Elizabeth Gustavsson and Nina Lagerqvist for excellent technical support. We also thank Tarja Sironen, Alexander Plyusnin and Heikki Henntonen for sharing sequences of PUUV strains Kuhmo/X5 and Kuhmo/X11 prior to publication and for helpful comments. This work was funded by the Swedish Emergency Management Agency (KBM), EU grant GOC-2003-010284 EDEN and grants from the Swedish Medical Research Council (project 12177). The paper is catalogued by the EDEN Steering Committee as EDEN0162 (http://www.eden-fp6project.net/).

**REFERENCES**


Sikes, D. S. & Lewis, P. O. (2001). PAUP* implementation of the parsimony ratchet (beta software, version 1). Distributed by the authors. Department of Ecology and Evolutionary Biology, University of Connecticut, Storrs, USA.


