Journal of General Virology (2008), **89**, 1649–1660 DOI 10.1099/vir.0.2008/001248-0

Accumulation of point mutations and reassortment of genomic RNA segments are involved in the microevolution of Puumala hantavirus in a bank vole (*Myodes glareolus*) population

Maria Razzauti,[1,2] Angelina Plyusnina,[1] Heikki Henttonen[2] and Alexander Plyusnin[1]

Correspondence
Alexander Plyusnin
alexander.plyusnin@helsinki.fi

[1]Department of Virology, Haartman Institute, PO Box 21, FI-00014 University of Helsinki, Finland
[2]Finnish Forest Research Institute, Vantaa Research Unit, PO Box 18, FI-01301 Vantaa, Finland

The genetic diversity of Puumala hantavirus (PUUV) was studied in a local population of its natural host, the bank vole (*Myodes glareolus*). The trapping area (2.5×2.5 km) at Konnevesi, Central Finland, included 14 trapping sites, at least 500 m apart; altogether, 147 voles were captured during May and October 2005. Partial sequences of the S, M and L viral genome segments were recovered from 40 animals. Seven, 12 and 17 variants were detected for the S, M and L sequences, respectively; these represent new wild-type PUUV strains that belong to the Finnish genetic lineage. The genetic diversity of PUUV strains from Konnevesi was 0.2–4.9 % for the S segment, 0.2–4.8 % for the M segment and 0.2–9.7 % for the L segment. Most nucleotide substitutions were synonymous and most deduced amino acid substitutions were conservative, probably due to strong stabilizing selection operating at the protein level. Based on both sequence markers and phylogenetic clustering, the S, M and L sequences could be assigned to two groups, 'A' and 'B'. Notably, not all bank voles carried S, M and L sequences belonging to the same group, i.e. $S_AM_AL_A$ or $S_BM_BL_B$. A substantial proportion (8/40, 20 %) of the newly characterized PUUV strains possessed reassortant genomes such as $S_BM_AL_A$, $S_AM_BL_B$ or $S_BM_AL_B$. These results suggest that at least some of the PUUV reassortants are viable and can survive in the presence of their parental strains.

Received 9 February 2008
Accepted 23 March 2008

INTRODUCTION

Hantaviruses constitute a distinct genus, *Hantavirus*, within the family *Bunyaviridae* (Schmaljohn & Dalrymple, 1983; Nichol *et al.*, 2005). They are the aetiological agents of haemorrhagic fever with renal syndrome (HFRS) in Eurasia and hantavirus pulmonary syndrome in the Americas (Schmaljohn & Hjelle, 1997). Hantaviruses are enveloped, negative-stranded RNA viruses. Their genome consists of three fragments: the small (S) fragment, encoding the nucleocapsid (N) protein; the medium (M) fragment, encoding two surface glycoproteins (Gn and Gc); and the large (L) fragment, encoding the viral RNA-dependent RNA polymerase (Plyusnin *et al.*, 1996b; Plyusnin, 2002). Each hantavirus species is carried by one or a few closely related rodent or insectivore species (Schmaljohn & Hjelle, 1997; Plyusnin & Morzunov, 2001). Hantavirus infection in natural hosts is persistent (Bernshtein *et al.*, 1999; Meyer & Schmaljohn, 2000) and may affect host survival (Kallio *et al.*, 2007). In rodent populations, hantaviruses are transmitted horizontally, either through direct contact or via infectious aerosols generated by contaminated host urine, faeces and saliva (Gavrilovskaya *et al.*, 1990; Bernshtein *et al.*, 1999). The indirect mode of transmission and virus survival outside the host increase the chances for the virus to persist in the host population (Sauvage *et al.*, 2003; Kallio *et al.*, 2006a). The transfer of antibodies from an infected mother to its progeny provides them with a temporary immunity against the infection that lasts up to 3.5 months (Gavrilovskaya *et al.*, 1990; Bernshtein *et al.*, 1999; Kallio *et al.*, 2006b).

Puumala virus (PUUV) (Brummer-Korvenkontio *et al.*, 1980), a major European rodent-borne pathogen, causes a relatively mild form of HFRS, also known as nephropathia epidemica (Vapalahti *et al.*, 2003). The natural host of PUUV is the bank vole, *Myodes glareolus* (previously called *Clethrionomys glareolus*), which belongs to the subfamily Arvicolinae of the family Cricetidae (Wilson & Reeders, 2005). The bank vole is widely distributed in Europe, from the British Isles to the Urals, excluding some northernmost regions and the Mediterranean coast. Its range continues

The GenBank/EMBL/DDBJ accession numbers for the PUUV S, M and L sequences determined in this work are AM980517–AM980552.

eastwards into central Siberia (Mitchell-Jones *et al.*, 1999). The occurrence of nephropathia epidemica in humans depends strongly on the local pattern of the population dynamics of the bank vole (Brummer-Korvenkontio *et al.*, 1982; Niklasson *et al.*, 1995). Similar to other hantaviruses, PUUV infection in bank voles is chronic and asymptomatic; the virus accumulates and is released mainly during the first month of infection (Bernshtein *et al.*, 1999).

It is thought that genetic drift, i.e. a gradual accumulation of point mutations throughout the genome coupled with small deletions and insertions within the non-coding regions of the RNA segments, is the main mechanism generating genetic diversity in hantaviruses (Plyusnin *et al.*, 1996b; Plyusnin, 2002). For PUUV, the estimated evolution rate appears to be low, ranging from 0.7×10^{-7} to 2.2×10^{-6} nt per site per year for the S segment sequences, and from 3.7×10^{-7} to 8.7×10^{-7} nt per site per year for the M segment sequences (Sironen *et al.*, 2001). Synonymous nucleotide substitutions dominate over non-synonymous, reflecting a strong negative selection operating at the protein level and suggesting the neutral mode for hantavirus evolution (Kimura, 1983). In addition to point mutations, the reassortment of genomic RNA segments and homologous recombination seem to be involved in hantavirus evolution (Henderson *et al.*, 1995; Li *et al.*, 1995; Sibold *et al.*, 1999; Chare *et al.*, 2003). There is evidence for both reassortment (Plyusnin *et al.*, 1997; Plyusnina *et al.* 2006) and recombination (Sironen *et al.*, 2001) in PUUV, but their impact on virus evolution remains to be evaluated properly.

The aim of this study was to gain insight into the microevolution of PUUV. In particular, we wanted to look for possible reassortment events and to estimate the frequency of reassortment in a population of wild-type viral genomes. Towards this aim, we recovered PUUV S, M and L segment sequences from bank voles captured within a relatively small study area of 2.5×2.5 km at Konnevesi (Central Finland) in the spring and autumn of 2005 and subjected these sequences to genetic analysis.

METHODS

Sampling of rodents. Rodents were trapped at Konnevesi, Central Finland (62° 34′ N 26° 24′ E), during May and October 2005. The trapping was carried out at 14 sites within an area of 2.5×2.5 km of typical spruce-dominated taiga forest (Fig. 1), each site consisting of 3×3 Ugglan Special live traps (Grahnab AB) at 15 m intervals. The traps were baited with oat seeds. The trapping sites were situated 500–1000 m apart from each other to ensure independence of sampling sites. During the trapping sessions of 3 days, the traps were checked daily. Captured bank voles were taken to the laboratory, bled, euthanized with CO_2, weighed, measured and sexed, and tissue samples were taken and frozen. Blood samples were taken from the retro-orbital sinus with 18 µl capillary tubes (Hematocrit tube; Hirschmann Laborgeräte) and placed on filter paper strips. Heart, lung, liver, spleen and kidney samples were collected and stored at −70 °C. The hearts were first placed in 100 µl 0.1 M PBS (pH 7.2). Every rodent sample was given a code that included the month of capture ('M' for May and 'O' for October), a serial number and the trapping site (in parentheses), for example M88(β) corresponds to bank vole #88 trapped in May at site β.

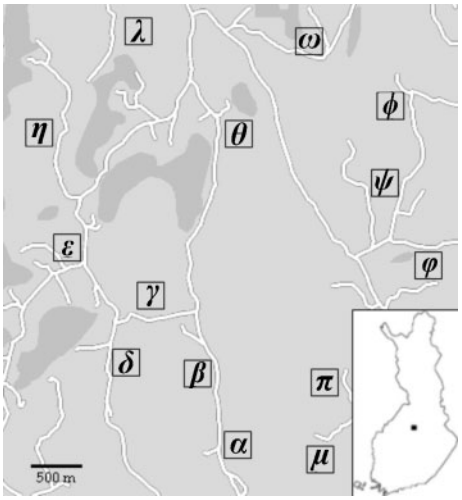


Fig. 1. Map of the Konnevesi trapping area. Dark-grey areas represent lakes and light-grey areas represent land; white lines are roads. The trapping sites were 500–1000 m apart. The location of the region within Finland is shown in the insert.

Screening of rodent samples. All trapped rodents were first screened for the presence of PUUV antibodies using an immunofluorescent assay. Antibody-positive bank voles were analysed further for the presence of PUUV N antigen by Western blotting and for the presence of PUUV S RNA by RT-PCR. Western blotting was performed essentially as described previously (Plyusnin *et al.*, 1995b). Briefly, lung tissue samples (~100 mg) were placed in 500 µl Laemmli buffer and homogenized by sonication. Aliquots of 10 µl were separated by 10 % SDS-PAGE and blotted with rabbit polyclonal antibody raised against PUUV N protein. Swine anti-rabbit antibodies conjugated with horseradish peroxidase (Dako) were used as secondary antibodies.

RT-PCR and sequencing. RNA was extracted from lung tissue samples of N antigen-positive rodents using TriPure reagent (Boehringer Mannheim) according to the manufacturer's instructions. Reverse transcription was performed using SuperScript II reverse transcriptase (Invitrogen/Gibco-BRL) as specified by the manufacturer. The PCR was carried out using AmpliTaq DNA polymerase (Perkin-Elmer). RT-nested-PCR was performed essentially as described previously (Plyusnin *et al.*, 1997). Both reverse transcription and the first PCR were carried out with primers Sa31 and PUU5 for the S segment and with primers A1 and C2 for the M segment. For the L segment, the newly designed primers PUULF1 [5′-CA(A/G)AA(A/G)GGTAATTGTCAATCTGG-3′] and PUULR2 [5′-GTATTTATAGGCCATATC(T/C)CTAG-3′] were used. For the nested PCR, primers PUU2 and Sa5 were used for the S segment, primers B1 and B2 for the M segment (Plyusnin *et al.*, 1997) and the newly designed primers PUULF2 [5′AT(A/C)TCAACACAR-TGGCCTAGTAG-3′] and PUULR2 for the L segment. To generate longer S amplicons, primers PUUSF7 [5′-GAAGGCAGAAGAA-CTCACACC(A/G)GG-3′] and Sa5 were used. The resulting amplicons were 308 bp for the S segment (502 bp for the longer S fragment), 486 bp for the M segment and 594 bp for the L segment. The amplicons were separated by electrophoresis in a

low-melting-point agarose gel (FMC BioProducts) and purified using a QIAquick Gel Extraction kit (Qiagen). Sequencing was performed automatically, using an ABI Prism Dye Terminator sequencing kit (Perkin Elmer/ABI).

Phylogenetic analyses. Multiple sequence alignment was carried out using BioEdit software (Hall, 1999) and CLUSTAL_W version 1.4 with default parameters. For comparison, PUUV genome sequences and sequences of other hantaviruses were retrieved from GenBank. The PHYLIP program (Felsenstein, 1993) was used to create 1000 bootstrap replicates of the sequence data (SEQBOOT). Distance matrices were calculated using the F84 model for nucleotide substitution (DNADIST) and analysed using neighbour-joining (NJ) or Fitch–Margoliash (FM) tree-fitting algorithms. The bootstrap support values for particular branching points were calculated from these trees using CONSENSE. In addition, the SEQBOOT outfiles were analysed using the maximum-likelihood (ML) algorithm (DNAML). ML trees were also reconstructed applying the Bayesian interference with the MrBayes 3 program (Huelsenbeck & Ronquist, 2001). The transition/transversion ratio and nucleotide frequencies were estimated from the dataset. Rate heterogeneity was applied using discrete gamma distribution with eight rate categories, and the shape parameter alpha was estimated from the dataset. The resulting trees were viewed with TreeView version 3.4.0.

RESULTS

Screening of rodent samples

Overall, 47 bank voles were trapped in May and 100 bank voles in October. Lung tissue samples were first screened for the presence of PUUV N antigen by Western blotting. N antigen-positive samples were then analysed by RT-PCR, followed by sequencing. Of 147 bank voles, 44 were found to be PUUV N antigen-positive: 22 voles captured in May and 22 captured in October (Table 1). Thus, the prevalence of N antigen was more than twofold higher among the overwintered rodents than in young ones. Infected bank voles were found at all 14 trapping sites, but their distribution was not equal: sites π, β, μ and ψ gave the largest number of catches and highest antigen prevalence, whilst sites θ, ω, γ and α gave the lowest numbers.

All 44 N antigen-positive bank voles were also found to be viral RNA-positive using RT-PCR with S segment-specific primers. RT-PCR with M segment-specific primers was successful for all 22 antigen-positive animals trapped in October (10 males and 12 females) and for 18 antigen-positive animals captured in May (nine males and nine females). RT-PCR with L segment-specific primers was positive for 41 samples (22 from October and 19 from May). Thus, the M and L segment-specific assays appeared to be less sensitive (or perhaps more dependent on the integrity of the RNA in tissue samples) than the S segment-specific assay. Forty M amplicons, together with 40 S and 40 L amplicons obtained from the same bank voles, were purified and sequenced.

General comparison of Konnevesi strains with other PUUV strains

As expected, all newly recovered S, M and L segment sequences belonged to the PUUV genotype. These sequences were compared with other known PUUV strains and with Hantaan virus (HTNV), Andes virus (ANDV) and Sin Nombre virus (SNV), representing hantaviruses

Table 1. Summary of RT-PCR screening

Trapping site	No. of trapped bank voles/no. of PCR-positive bank voles		S, M and L sequences used for analysis*	
	May	October	May	October
α	4/1	4/0	107†	–
β	7/5	10/2	68†, 78, 88, 89,109	38, 74
γ	4/2	8/2	70, 101	62, 63
δ	0/0	7/2	–	45, 50
ε	2/2	5/0	75, 110	–
η	6/2	2/0	76, 90	–
θ	2/0	11/2	–	27, 78
λ	0/0	4/1	–	15
μ	6/0	10/6	–	6, 8, 9, 12, 13, 14
π	8/4	9/4	82, 94, 99†, 105	56, 57, 59, 93
φ	1/1	2/0	83	–
ϕ	0/0	7/1	–	94
ψ	4/3	11/2	80, 81, 98	19, 22
ω	3/2	10/0	91†, 114	–
Total	47/22	100/22	18	22

*Numbers identify individual bank voles captured in May or October at different trapping sites.
†Partial L and/or M segment sequences were not recovered from these samples. Corresponding S segment sequences were excluded from further analyses.

carried by rodents of the subfamilies Murinae, Sigmodontinae and Neotominae, respectively.

Not surprisingly, the wild-type (wt) PUUV strains from Konnevesi appeared to be most closely related to other strains from the Finnish genetic lineage, which included strains from Finland and Russian Karelia (Asikainen *et al.*, 2000). The S segment sequence of strain Puumala1324 showed the highest identity to Konnevesi strains (95–96 %), whilst the sequence of strain Karhumäki (Russian Karelia, east of the Finnish–Russian border) showed the lowest identity (89–90 %). The S sequences from other PUUV genetic lineages appeared to be more distant, showing identity ranging from 84 (Russian and two Scandinavian lineages) and 83 (Alpe–Adrian lineage) to 80 % (Central European lineage). The other hantavirus species presented lower sequence identities of 60 (SNV and ANDV) and 58 % (HTNV).

Similarly, the M segment sequences of Konnevesi strains were most closely related to strains from the Finnish genetic lineage. Strain Kolodozero (Russian Karelia) appeared to be the most closely related, with a sequence identity of 93–94 %, whilst strain Langemäki showed the lowest sequence identity (88–90 %). PUUV strains from other lineages were more distant, with a sequence identity ranging from 87 (Russian lineage) to 83 % (the two Scandinavian lineages and the Alpe–Adrian lineage) and 81 % (Central European lineage). The other hantavirus species showed sequence identities of 75 (SNV and ANDV) and 67 % (HNTV).

Comparison of the L segment sequences revealed the same pattern. The L sequences of Konnevesi strains were most closely related to the Finnish strain Sotkamo (the only representative of the Finnish lineage), with a sequence identity of 90–93 %, whilst the L sequences from other PUUV genetic lineages were more distant, showing identity ranging from 85 (Russian lineage) to 83 % (Scandinavian lineage). Other hantaviruses showed substantially lower sequence identities of 67 (SNV), 65 (ANDV) and 58 % (HTNV).

Detailed genetic analysis of wt PUUV strains from Konnevesi

The wt PUUV strain recovered from bank vole M114(ω) and designated Konnevesi/MgM114/2005, or M114 for short, was selected as our 'prototype' strain: all S, M and L sequences recovered from other bank voles were compared with those of the M114 strain and corresponding nucleotide substitutions were noted. In fact, the partial S sequence of the M114 strain was one of the first sequences recovered and used for comparison and grouping. It became apparent later that this was an appropriate choice, as the M114 strain was not a reassortant (see below).

Overall, 13 point mutations were found in the partial S segment sequences (nt 844–1082) recovered from 40 bank voles. The mutations were distributed evenly throughout the sequence, with an overall frequency of 5.5 %. The overall diversity of the S sequences was between 0.4 and 4.3 %. Eleven of the 13 observed mutations were transitions (Fig. 2a). All mutations were silent (located at the third position of the codon), suggesting a strong negative selection operating at the N protein level.

In the partial M segment sequences (nt 2180–2610), 29 point mutations were observed. The overall mutation frequency was 6.7 %, close to the value for the S segment sequences, but the mutations were distributed less evenly, forming several clusters, for instance in the region nt 2566–2587 (Fig. 3c). The overall diversity of the M sequences was

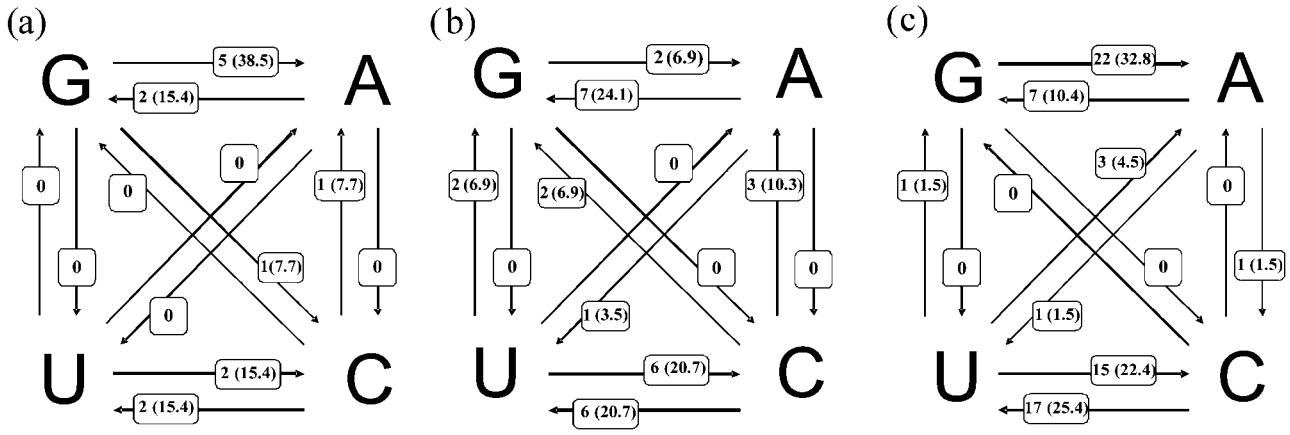


Fig. 2. Summary of the nucleotide substitutions observed in the S (a), M (b) and L (c) RNA segments of PUUV strains from Konnevesi. The substitutions are shown for the negative-sense RNA; the sequences of strain M114 were taken as the reference. Altogether, 13 substitutions were observed in the S segment sequence, 29 substitutions in the M segment sequence and 67 substitutions in the L segment sequence. Numbers in parentheses indicate percentages.

between 0.2 and 4.8 %. As with the S segment mutations, transitions dominated (a total of 21) (Fig. 2b). Twenty-six mutations were located at the third position of the codon and three mutations were found in the first position; all were silent, suggesting strong stabilizing selection working against changes in the viral Gc protein.

A total of 67 point mutations was observed within the partial L segment sequences (nt 502–1036) recovered from 40 bank voles, giving a surprisingly high mutation frequency of 12.5 %, which was approximately twofold higher than the corresponding values for the S and M sequences. The overall genetic diversity of the L sequences was between 0.2 and 9.7 %. As with the S and M segment sequences, transitions dominated (a total of 61) (Fig. 2c). Eight mutations were in the first position of the codon and two in the second position; others were located in the third codon position. Five mutations led to four amino acid substitutions in the deduced sequence of the L protein, three of which were conservative [V251I (in this particular codon, two mutations were observed), R284K and V333I] and one of which was non-conservative (I289T).

Altogether, seven genetic variants of the partial S segment sequence were recognized; these formed two distinct groups designated 'SA' and 'SB' (Fig. 3a). The prototype sequence of strain M114 was designated variant SB5 in the group SB; identical S sequences were recovered from two other bank voles. Other S sequences were assigned to six genetic variants in the SB and SA groups, with increasing diversity from the prototype sequence. The SB group included four more genetic variants; the SA group consisted of two variants. Notably, the grouping of the S sequences appeared to be independent of the site or time of trapping. It should be noted that the S PCR amplicons used to recover the sequences of nt 844–1082 were generated as part of another study, in which the sensitivity of different screening tests for hantavirus markers in bank voles was evaluated (H. Henttonen and others, unpublished data). As we had to keep the RT-PCR detection rate as high as possible, these sequences were rather short. To compensate for a possible bias, a longer sequence of the S segment (nt 640–1082) was recovered for several representatives of each genotype. Fifteen additional point mutations (14 transitions and one transversion) were found in this longer fragment (Fig. 3b), giving a total mutation frequency of 6.6 % and an overall diversity of 0.2–4.9 %, which were very close to the earlier estimates. All additional transitions occupied the third codon position and were silent. The transversion at position 752 (the second codon position) led to homologous R237K substitution. Most importantly, the additional sequence data confirmed our initial S grouping. Furthermore, the analysis of complete S segment sequences recovered from five genetic variants (SA1, SA2, SB1, SB2 and SB5) was in perfect agreement with our initial grouping.

Overall, 12 genetic variants of the partial M segment sequences could be recognized; these formed two distinct groups designated 'MA' and 'MB' (Fig. 3c). Similarly to the S sequences, neither the place nor time of trapping influenced the grouping. A partial M sequence of the prototype strain M114 was assigned to the genetic variant MB6; identical M sequences were recovered from four other bank voles. The MB group included five more variants, whilst the MA group consisted of six variants.

Similarly to the S and M variants, the genetic variants of the partial L segment sequence formed two easily distinguished groups: the 'LA' group was composed of eight variants and the 'LB' group was composed of nine (Fig. 3d). The prototype M114 strain was assigned to the LB9 variant; identical sequences were recovered from three other bank voles. The number of L variants was higher than that of the S or M variants. Consequently, most of the L variants were not represented by a large number of sequences. Seven sequences represented the dominant LB variant (LB5), whilst only three represented the dominant LA variant (LA4).

Phylogenetic analysis of the newly characterized S, M and L genetic variants was performed using a variety of methods: distance matrix methods (NJ and FM algorithms), maximum parsimony and ML (both classical and Bayesian). The corresponding sequences of HTNV, ANDV and SNV were used as outgroups. In agreement with the direct sequence comparison, on the phylogenetic trees (Fig. 4), Konnevesi strains formed a well-supported cluster that was placed within the Finnish genetic lineage. Notably, the grouping made on the basis of genetic markers (mutations) was confirmed using phylogenetic analysis. All of the phylogenetic methods applied confirmed the A/B grouping for the S, M and L segment sequences (the NJ trees are shown in Fig. 4). Interestingly, our analysis suggested that the most recent common ancestor for the cluster of Konnevesi strains and strain Puumala1324 originated from the place where PUUV was first discovered (Plyusnin *et al.*, 1995a). The close phylogenetic ties between Puumala and Konnevesi strains could be explained by a relatively short distance (~150 km) and the absence of geographical obstacles, not even major lakes, for the bank vole, and PUUV gene flow between these two localities.

Evidence for reassortment of PUUV genome segments in wt Konnevesi strains

When the results of (phylo)genetic grouping were compared with the origin of nucleotide sequences, the majority of the Konnevesi strains appeared to possess S, M and L genome segments belonging to the corresponding S, M and L genetic groups. For instance, the prototype strain M114 had the S segment from the SB group (Fig. 3a, b), the M segment from the MB group (Fig. 3c) and the L segment from the LB group (Fig. 3d). Similarly, strains O45 and M80 possessed 'regular' genotypes: $S_AM_AL_A$ and $S_BM_BL_B$, respectively. However, a substantial proportion of the Konnevesi strains (8/40, 20 %) had S, M and L segments belonging to different genetic groups (Fig. 3 and Table 2).

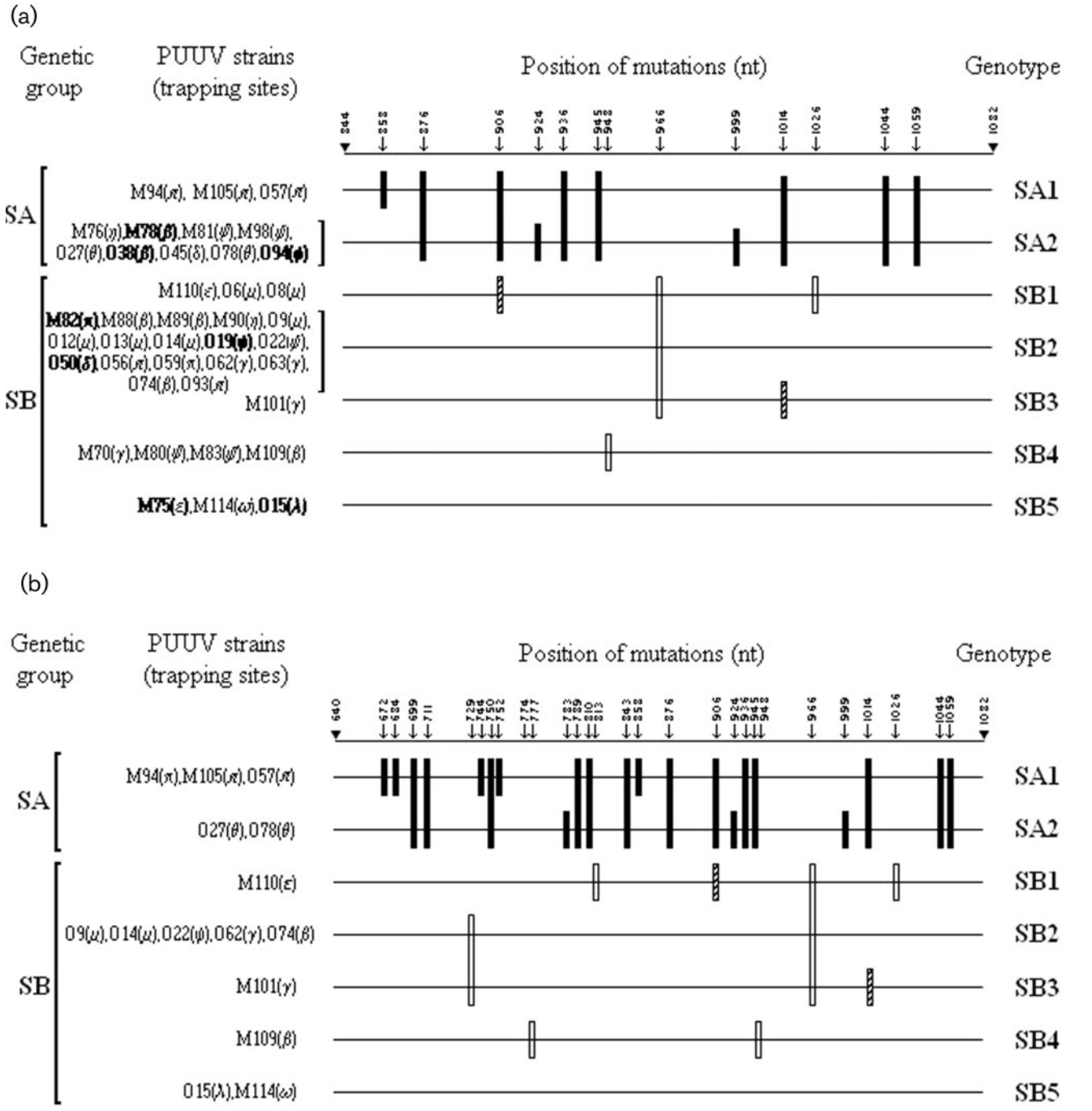
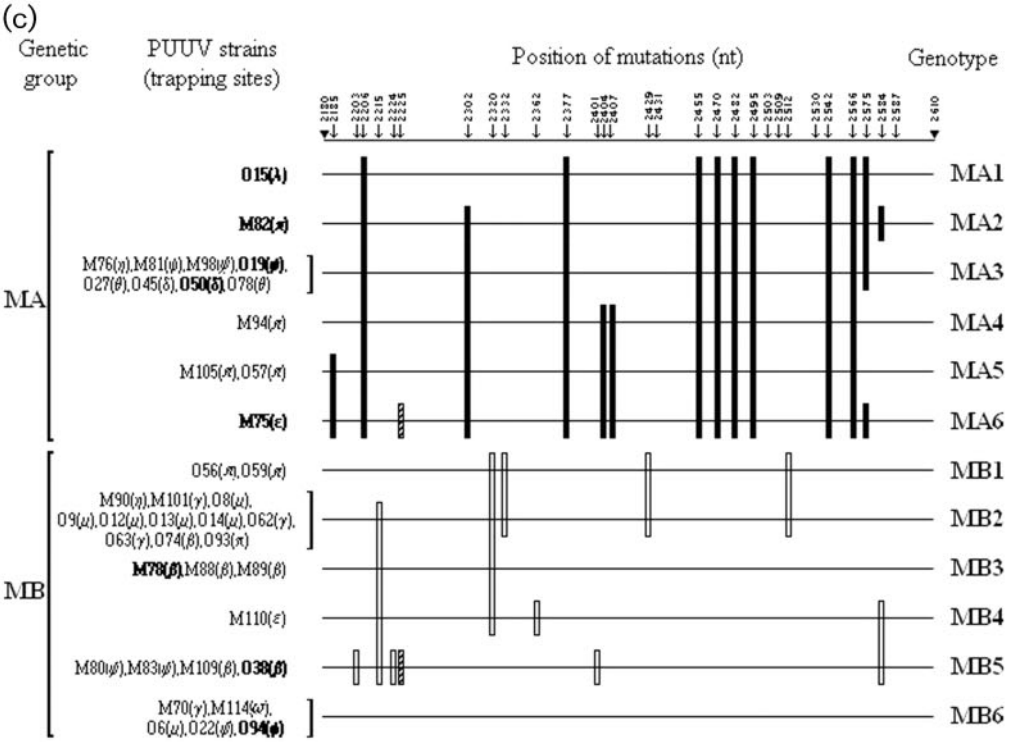
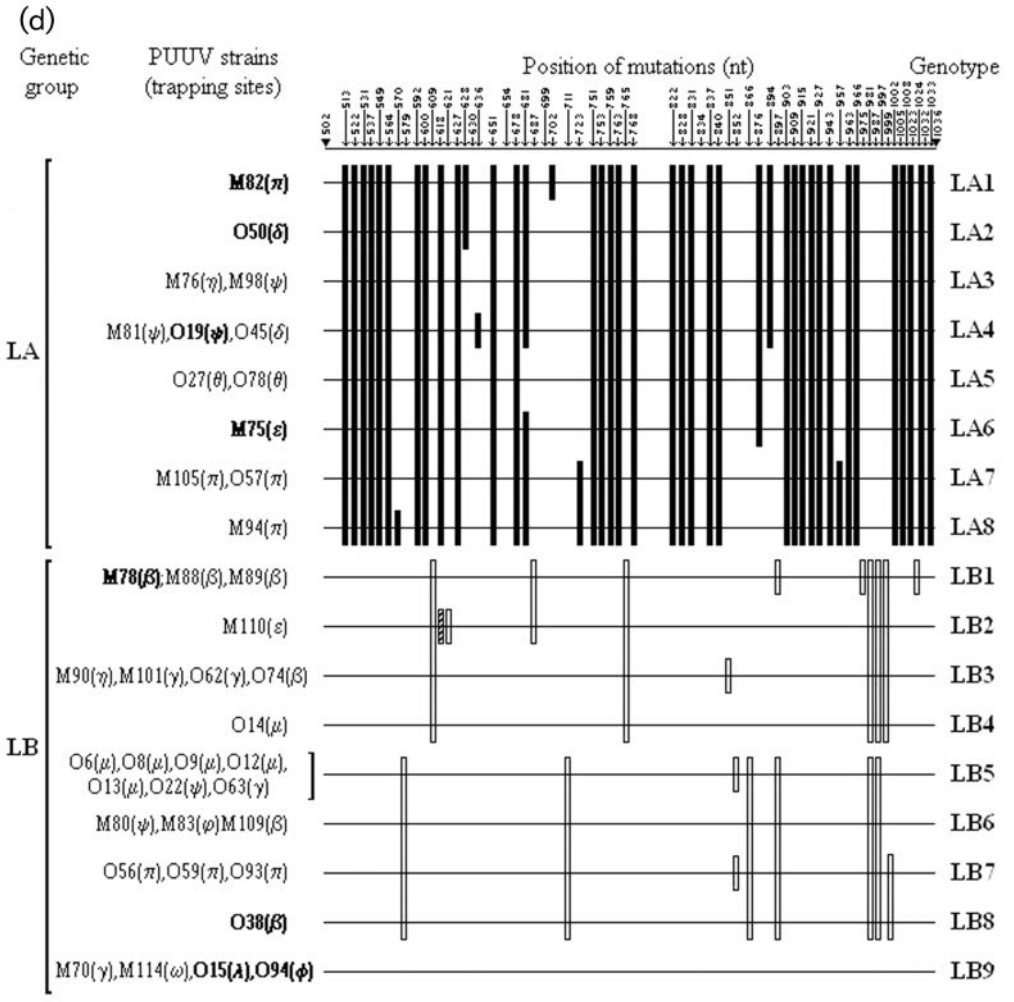


Fig. 3. Mutations observed in the S, M and L segment genetic variants of the Konnevesi PUUV strains. (a) S segment sequences (nt 844–1082). The reference sequence of strain M114 was designated variant SB5 (bottom line). Other S segment sequences were assigned to six genetic variants in the SB and SA groups, with increasing diversity from the prototype sequence. (b) Longer S segment sequences (nt 640–1082), recovered for selected representatives of all seven S genotypes. The reference SB5 sequence differed from the others by mutation G774A (for simplicity, not shown on the figure). (c) M segment sequences (nt 2180–2610). The reference MB6 sequence differed from the others by five substitutions (C2431U, U2503G, C2509U, A2530G and U2687C) and shared one additional mutation (U2185C) with the MB5 sequence. (d) L segment sequences (nt 502–1036). The reference LB9 sequence differed from the others by five substitutions (U630C, U654C, C699U, A834G and U1032C) and shared A579G and U987A with genotypes LB5–LB8, and U981C with genotypes LB1–LB4. In addition, two mutations (C759U and C894U) were shared with the LA genotypes.


(c)
Genetic group
PUUV strains (trapping sites)
Position of mutations (nt)
Genotype
MA
MB
O15(λ)
M82(π)
M76(η),M81(ψ),M98(ψ'),O19(φ),O27(θ),O45(δ),O50(δ),O78(θ)
M94(π)
M105(π),O57(π)
M75(ε)
O56(π),O59(π)
M90(η),M101(γ),O8(μ),O9(μ),O12(μ),O13(μ),O14(μ),O62(γ),O63(γ),O74(β),O93(π)
M78(β),M88(β),M89(β)
M110(ε)
M80(ψ'),M83(ψ'),M109(β),O38(β)
M70(γ),M114(ω),O6(μ),O22(ψ),O94(φ)
MA1
MA2
MA3
MA4
MA5
MA6
MB1
MB2
MB3
MB4
MB5
MB6


(d)
Genetic group
PUUV strains (trapping sites)
Position of mutations (nt)
Genotype
LA
LB
M82(π)
O50(δ)
M76(η),M98(ψ)
M81(ψ),O19(φ),O45(δ)
O27(θ),O78(θ)
M75(ε)
M105(π),O57(π)
M94(π)
M78(β),M88(β),M89(β)
M110(ε)
M90(η),M101(γ),O62(γ),O74(β)
O14(μ)
O6(μ),O8(μ),O9(μ),O12(μ),O13(μ),O22(ψ),O63(γ)
M80(ψ),M83(φ)M109(β)
O56(π),O59(π),O93(π)
O38(β)
M70(γ),M114(ω),O15(λ),O94(φ)
LA1
LA2
LA3
LA4
LA5
LA6
LA7
LA8
LB1
LB2
LB3
LB4
LB5
LB6
LB7
LB8
LB9

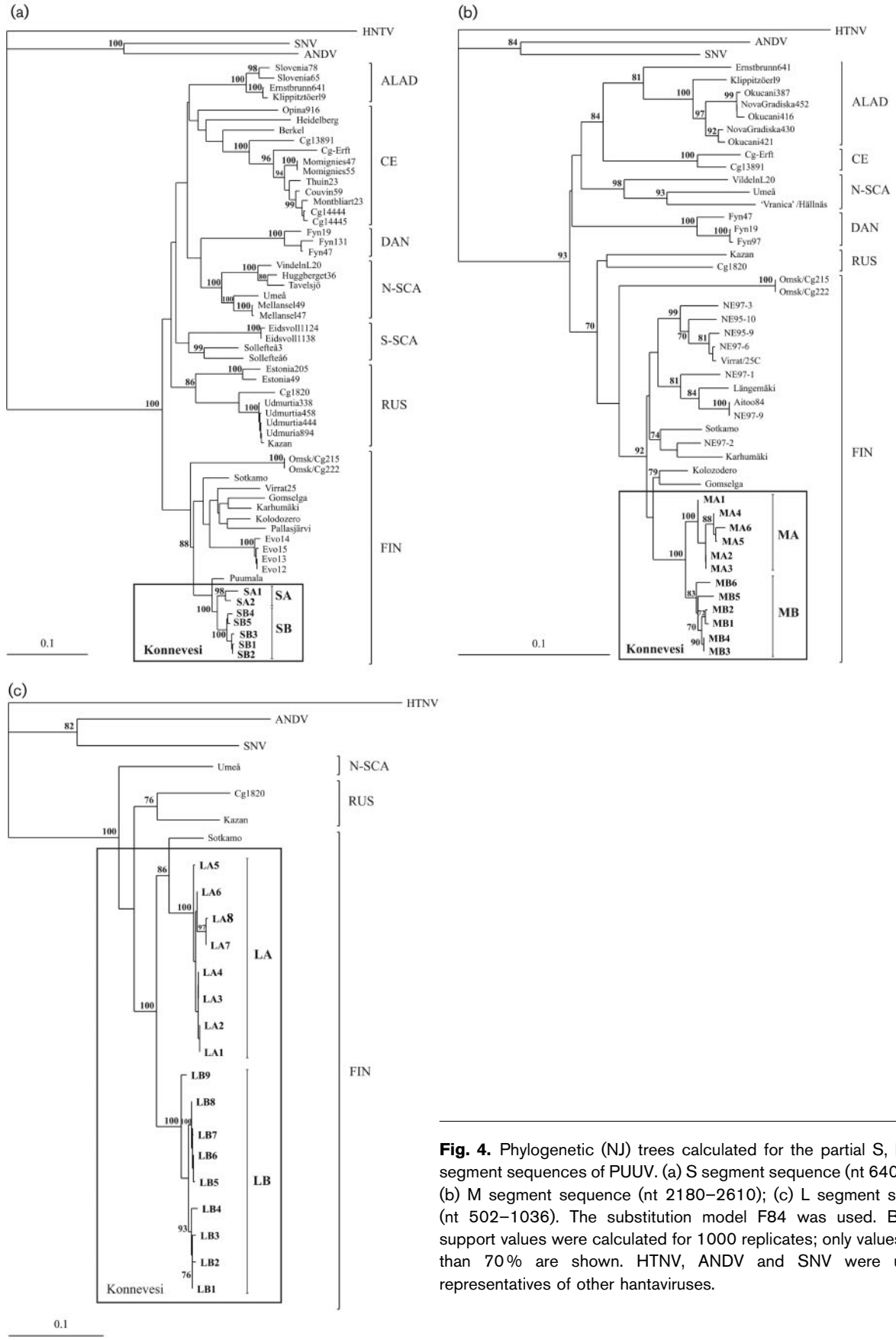


Fig. 4. Phylogenetic (NJ) trees calculated for the partial S, M and L segment sequences of PUUV. (a) S segment sequence (nt 640–1082); (b) M segment sequence (nt 2180–2610); (c) L segment sequence (nt 502–1036). The substitution model F84 was used. Bootstrap support values were calculated for 1000 replicates; only values greater than 70% are shown. HTNV, ANDV and SNV were used as representatives of other hantaviruses.

Table 2. PUUV genotypes found at different trapping sites

Trapping site	Sample	Genotypes*			No. of sequences recovered	No. of genetic variants found			No. of reassortants
		S	M	L		S	M	L	
β	**M78**	**A2**	**B3**	**B1**	6	3	3	4	2
	M88	B2	B3	B1					
	M89	B2	B3	B1					
	M109	B4	B5	B7					
	O38	**A2**	**B5**	**B9**					
	O74	B2	B2	B3					
γ	M70	B4	B6	B5	4	3	2	3	0
	M101	B3	B2	B3					
	O62	B2	B2	B3					
	O63	B2	B2	B6					
δ	O45	A2	A3	A4	2	2	1	2	1
	O50	**B2**	**A3**	**A2**					
ε	**M75**	**B5**	**A6**	**A6**	2	2	2	2	1
	M110	B1	B4	B2					
η	M76	A2	A3	A3	2	2	2	2	0
	M90	B2	B2	B3					
θ	O27	A2	A3	A5	2	1	1	1	0
	O78	A2	A3	A5					
λ	**O15**	**B5**	**A1**	**B9**	1	1	1	1	1
μ	O06	B1	B6	B6	6	2	2	2	0
	O08	B1	B2	B6					
	O09	B2	B2	B6					
	O12	B2	B2	B6					
	O13	B2	B2	B6					
	O14	B2	B2	B4					
π	**M82**	**B2**	**A2**	**A1**	7	2			1
	M94	A1	A4	A8					
	M105	A1	A5	A7					
	O56	B2	B1	B8			5	4	
	O57	A1	A5	A7					
	O59	B2	B1	B8					
	O93	B2	B2	B8					
φ	M83	B4	B5	B7	1	1	1	1	0
ϕ	**O94**	**A2**	**B6**	**B5**	1	1	1	1	1
ψ	M80	B4	B5	B7	5	3			1
	M81	A2	A3	A4					
	M98	A2	A3	A3			3	4	
	O19	**B2**	**A3**	**A4**					
	O22	B2	B6	B6					
ω	M114	B5	B6	B5	1	1	1	1	0
Total					40	7	12	17	8

*Reassortants are shown in bold.

For example, the S segment of strain O50 belonged to group B, whilst the M and L segments belonged to group A. Such conflicting grouping suggested reassortment between the viral genome segments.

Three reassortant genomes were found in the bank voles trapped in May and five in the bank voles trapped in October. As the corresponding numbers of analysed wt strains were 18 and 22, the proportion of reassortants was somewhat higher in October than in May (23 vs 17 %). Reassortants were found at seven of the 14 trapping sites, two at site β and one at each of six other sites (Table 1). Of the six possible S/M/L combinations of the genome segments, only three were found in reassortant PUUV strains from Konnevesi: the $S_BM_AL_A$ genotype was observed in four strains, the $S_AM_BL_B$ genotype in three strains and the $S_BM_AL_B$ genotype in one strain. The proportions of S, M and L variants involved with the

reassortant genomes were similar (57, 58 and 41 %, respectively), and both the most frequently occurring (e.g. SA2 and MA3) and the most rare (e.g. SB5, MA1 and LA1) genetic variants were seen among the reassortants.

DISCUSSION

Ecological aspects of PUUV transmission in bank vole populations

In this study, we performed a detailed genetic analysis of 40 wt PUUV strains circulating within a local bank vole population and have presented here the first solid evidence on reassortment between genetic variants. In earlier studies (Plyusnin *et al.*, 1995a; Hörling *et al.*, 1996), only a few PUUV strains from one locality were analysed. The relatively large numbers of PUUV S, M and L segment sequences recovered from one locality allowed us to address for the first time the question of the frequency of PUUV segment reassortment occurring in nature.

In agreement with previously published data (Olsson *et al.*, 2002), PUUV seroprevalence was higher among the overwintering bank voles (trapped in May) than in the young of the year (trapped in October). There are two probable reasons for this difference. First, the overwintering bank voles were exposed to the virus for a longer time. Second, the young rodents are awkward and hence remain more local, and infected females transfer maternal antibodies to their offspring, providing a temporary protection against the infection for up to 3.5 months (Gavrilovskaya *et al.*, 1990; Bernshtein *et al.*, 1999; Kallio *et al.*, 2006b). As mentioned above, the trapping sites were situated 500–1000 m apart to ensure that the bank voles were not in frequent contact. Nevertheless, we did not see any site-wise clustering of PUUV genetic variants: they appeared to circulate within the whole study area. The majority of bank voles, which do not reach maturity in the first summer of their lives, stay within a small area of about 2000–3000 m^2 until maturation the following spring (Crawley, 1969; our unpublished data). However, breeding and maturing bank voles (males and females), which look for free home ranges, can move long distances, up to 2–3 km (our unpublished data), and thus can spread their genes as well as the virus across a larger area. This spreading is not unlimited, as seen, for example, within the contact zone between two genetically different bank vole lineages in central Sweden (Hörling *et al.*, 1996). Over approximately 10 000 years since the two recolonizing bank vole lineages met, PUUV variants from these two lineages have not spread into another lineage by more than several kilometres (Hörling *et al.*, 1996). To learn more about spatial limits for the spread of local PUUV genetic variants, we plan to analyse the strains circulating in a larger area around Konnevesi in future studies.

Molecular mechanisms of PUUV microevolution

Our data showed that both accumulation of point mutations and reassortment of genomic RNA segments contribute to the generation of PUUV genetic diversity. Point mutations, mostly transitions, occurred at a relatively high frequency, especially in the L segment. However, this should not be interpreted as a contradiction to previously estimated low rates of PUUV evolution (Sironen *et al.*, 2001). High genetic diversity, measured at a given time point, cannot necessarily be converted into a high rate of accumulation of genetic changes over longer periods of time. Of 124 point mutations found in the analysed regions of the three PUUV genome segments, only six were non-synonymous and only one of the deduced amino acid substitutions was non-homologous (the detailed structural/functional organization of the hantaviral L protein is yet to be determined and therefore the functional significance of these amino acid replacements remains unclear). This suggests strong stabilizing selection operating at the protein level and is in agreement with the hypothesis that the evolution of hantaviruses in general, and PUUV in particular, follows the neutral mode (Sironen *et al.*, 2001). Here, one can see a parallel with the high genetic diversity of hantaviral quasispecies co-existing with strong conservation and, consequently, slow evolution of consensus sequences (Plyusnin *et al.*, 1995a, 1996a). Certainly, synonymous mutations could still have an impact on a viral phenotype, by altering RNA folding or codon usage. There is increasing evidence of the importance of synonymous mutations for the evolution of RNA viruses (Novella *et al.*, 2004; Hamano *et al.*, 2007), including PUUV (Sironen *et al.*, 2008). From this point of view, the unexpectedly high level of the L segment diversity observed in the wt PUUV strains from Konnevesi is of special interest: so far, most of the known hantaviral wt sequences are those of the S or M segment. Whether this bias is an essential part of the virus survival strategy remains to be investigated.

The substantial proportion of reassortant genomes (20 %) among PUUV genetic variants in our rather small study area was surprising. It should be emphasized that the reassortants identified cannot merely be mutants that have evolved convergently. To be converted from one genetic group to another, a large number of point mutations in a given genome segment need to be acquired simultaneously: for 443 nt of the S segment, this number is 17. This makes the likelihood of such an event unrealistically low. Previously, natural reassortants of SNV have been observed in infected deer mice (*Peromyscus maniculatus*) in the USA (Henderson *et al.*, 1995). These authors reported that 39 % of all wt SNV strains were reassortants, but their study area was much larger, including the trapping sites from two states, Nevada and California, more than 100 km apart. It is possible that the large number of reassortant PUUV genomes observed in our study is rooted in the high genetic diversity of the S, M and especially the L segment sequences, which has made the detection of reassortants

feasible. The grouping of PUUV genetic variants based on molecular markers (mutations) was in good agreement with the phylogenetic grouping, thus providing solid support for the reassortment scenario. From a practical point of view, it is also important that the initial grouping based on the analysis of short S sequences appeared to be correct. This was confirmed when sequences of approximately twice the length (nt 640–1082) were recovered for each of the S variants, and was further verified by the analysis of complete S segment sequences.

The reassortants, or at least some of them, seem to be able to survive in the presence of their parental strains, otherwise they would not have been detected. However, it remains to be seen whether any of them can successfully compete with the parental strains and survive for several years. Interestingly, the proportion of reassortants was slightly higher in October than in May, but the overall numbers were too small for any definite conclusion on possible seasonal variations.

As mentioned above, only three of the six possible combinations of the genome segments were observed in the PUUV reassortants described here. Notably, in seven of the eight reassortants, the M and L segments belonged to the same group: three reassortants were of the $S_AM_BL_B$ type and four of the $S_BM_AL_A$ type. Even with these low numbers, it seems that PUUV M and L segments reassort together more frequently than the S and L segments. This is in agreement with the data on natural reassortants of the bunyaviruses La Crosse virus, Rift Valley fever virus and Crimean-Congo hemorrhagic fever virus (Urquidi & Bishop, 1992; Sall *et al.*, 1999; Deyde *et al.*, 2006), but is in contrast to what has been observed in SNV natural reassortants, where the majority (14/17) possessed S and L segments from the same phylogenetic group (Henderson *et al.*, 1995).

One observation was of special interest: direct sequencing of the S amplicons recovered from bank voles O06 and O08 revealed double peaks at several positions (data not shown). Interestingly, these positions were the ones that distinguished genetic variants SA2 and SB1. Direct sequencing of the corresponding M and L amplicons revealed only one genetic variant of each: MB6/LB6 variants in bank vole O06 and MB2/LB6 variants in bank vole O08. These data suggest that each of these bank voles was infected with a mixture of two PUUV variants, which possessed two distinct types of S segment and shared the same type of M and L segment.

It is known that, in addition to point mutation and reassortment, recombination plays a role in hantavirus evolution (Sibold *et al.*, 1999; Chare *et al.*, 2003). Some evidence for a recombinant origin of several PUUV strains has been presented (Sironen *et al.*, 2001). Basically, the prerequisites for hantaviral segment reassortment and recombination are the same: co-circulation of several genetic variants of the virus in a local rodent population and co-infection of an individual rodent with two distinct variants. We evaluated the recombination phenomenon on partial S, M and L segment sequences and on complete S segment sequences using special statistical tools (SimPlot and BootScan); however, the analysis did not reveal any traces of recombination, in perfect agreement with the lack of mosaics in the newly recovered sequences (Fig. 3), clearly demonstrating that the frequency of recombination in PUUV microevolution is substantially lower than the frequency of reassortment events.

To conclude, our analysis of the largest collection so far of wt PUUV strains circulating in a local bank vole population showed that the accumulation of point mutations and the reassortment of genome RNA segments were the main mechanisms generating virus genetic diversity. Most of the observed point mutations were synonymous, probably due to strong negative selection.

ACKNOWLEDGEMENTS

We thank Jukka Niemimaa and Liina Voutilainen for their invaluable help in the fieldwork. This work was supported by grants from The Academy of Finland, Sigrid Jusélius Foundation, Center for International Mobility (CIMO, Finland; scholarship to M. R.), and was partially funded by EU grant GOCE-2003-010284 EDEN (http://www.eden-fp6project.net/). This paper has been catalogued by the EDEN steering committee as EDEN0049.

REFERENCES

Asikainen, K., Hänninen, T., Henttonen, H., Niemimaa, J., Laakkonen, J., Andersen, H. K., Bille, N., Leirs, H., Vaheri, A. & Plyusnin, A. (2000). Molecular evolution of Puumala hantavirus in Fennoscandia: phylogenetic analysis of strains from two recolonization routes, Karelia and Denmark. *J Gen Virol* **81**, 2833–2841.

Bernshtein, A. D., Apekina, N. S., Mikhailova, T. V., Myasnikov, Yu. A., Khlyap, L. A., Korotkov, Yu. S. & Gavrilovskaya, I. N. (1999). Dynamics of Puumala hantavirus infection in naturally infected bank voles (*Clethrionomys glareolus*). *Arch Virol* **144**, 2415–2428.

Brummer-Korvenkontio, M., Vaheri, A., Hovi, T., von Bonsdorff, C.-H., Vuorimies, J., Manni, T., Penttinen, K., Oker-Blom, N. & Lähdevirta, J. (1980). Nephropathia epidemica: detection of antigen in bank vole and serological diagnosis of human infection. *J Infect Dis* **141**, 131–134.

Brummer-Korvenkontio, M., Henttonen, H. & Vaheri, A. (1982). Hemorrhagic fever with renal syndrome in Finland: ecology and virology of nephropathia epidemica. *Scand J Infect Dis Suppl* **36**, 131–134.

Chare, E. R., Gould, E. A. & Holmes, E. C. (2003). Phylogenetic analysis reveals a low rate of homologous recombination in negative-sense RNA viruses. *J Gen Virol* **84**, 2691–2730.

Crawley, M. C. (1969). Movements and home-ranges of *Clethrionomys glareolus* Schreber and *Apodemus sylvaticus* L. in north-east England. *Oikos* **20**, 310–319.

Deyde, V. M., Khristova, M. L., Rollin, P. E., Ksiazek, T. G. & Nichol, S. (2006). Crimean-Congo hemorraghic fever virus genomics and global diversity. *J Virol* **80**, 8834–8842.

Felsenstein, J. (1993). PHYLIP (phylogeny inference package), 3.66 version. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle, WA, USA.

Gavrilovskaya, I. N., Apekina, N. S., Bernshtein, A. D., Demia, V. T., Okulova, N. M., Myasnikov, Y. A. & Chumanov, M. P. (1990). Pathogenesis of hemorrhagic fever with renal syndrome virus infection and mode of horizontal transmission of hantavirus in bank voles. *Arch Virol Suppl* **1**, 57–62.

Hall, T. A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser* **41**, 95–98.

Hamano, T., Matsuo, K., Hibi, Y., Victoriano, A. F., Takahashi, N., Mabuchi, Y., Soji, T., Irie, S., Sawanpanyalert, P. & other authors (2007). A single-nucleotide synonymous mutation in the *gag* gene controlling human immunodeficiency virus type 1 virion production. *J Virol* **81**, 1528–1533.

Henderson, W. W., Monroe, M. C., St. Jeor, S. C., Thayer, W. P., Rowe, J. E., Peters, C. J. & Nichol, S. T. (1995). Naturally occurring Sin Nombre virus genetic reassortants. *Virology* **214**, 602–610.

Hörling, J., Lundkvist, Å., Jaarola, M., Plyusnin, A., Tegelström, H., Persson, K., Lehväslaiho, H., Hörnfeldt, B., Vaheri, A. & Niklasson, B. (1996). Distribution and genetic heterogeneity of Puumala virus in Sweden. *J Gen Virol* **77**, 2555–2562.

Huelsenbeck, J. P. & Ronquist, F. (2001). MRBAYES: Bayesian inference of phylogeny. *Bioinformatics* **17**, 754–755.

Kallio, E. R., Klingström, J., Gustafsson, E., Manni, T., Vaheri, A., Henttonen, H., Vapalahti, O. & Lundkvist, Å. (2006a). Prolonged survival of Puumala hantavirus outside the host: evidence for indirect transmission via the environment. *J Gen Virol* **87**, 2127–2134.

Kallio, E. R., Poikonen, A., Vaheri, A., Vapalahti, O., Henttonen, H., Koskela, E. & Mappes, T. (2006b). Maternal antibodies postpone hantavirus infection and enhance individual breeding success. *Proc Biol Sci* **273**, 2771–2776.

Kallio, E. R., Voutilainen, L., Vapalahti, O., Vaheri, A., Henttonen, H., Koskela, E. & Mappes, T. (2007). Endemic hantavirus infection impairs the winter survival of its rodent host. *Ecology* **88**, 1911–1916.

Kimura, M. (1983). *The Neutral Theory of Molecular Evolution*. Cambridge: Cambridge University Press.

Li, D., Schmaljohn, A. L., Anderson, K. & Schmaljohn, C. S. (1995). Complete nucleotide sequences of M and S segments of two hantavirus isolates from California: evidence for reassortment in nature among viruses related to hantavirus pulmonary syndrome. *Virology* **206**, 973–983.

Meyer, B. J. & Schmaljohn, C. S. (2000). Persistent hantavirus infections: characteristics and mechanisms. *Trends Microbiol* **8**, 61–67.

Mitchell-Jones, A. J., Amori, G., Bogdanowicz, W., Krystufek, B., Reijnders, P. J. H., Spitzenberger, F., Stubbe, M., Thissen, J. B. M., Vohralik, V. V. & Zima, J. (1999). *Clethrionomys glareolus*. In *The Atlas of European Mammals*, pp. 212–213. Edited by A. J. Mitchell-Jones. London: T & AD Poyser Natural History.

Nichol, S., Beaty, B. J., Elliott, R. M., Goldbach, R., Plyusnin, A., Schmaljohn, C. S. & Tesh, R. B. (2005). *Bunyaviridae*. In *Virus Taxonomy. Eighth report of the International Committee on Taxonomy of Viruses*, pp. 695–716. Edited by C. M. Fauquet, M. A. Mayo, J. Maniloff, U. Desselberger & L. A. Ball. Amsterdam: Elsevier Academic Press.

Niklasson, B., Hörnfeldt, B., Lundkvist, Å., Björsten, S. & Leduc, J. (1995). Temporal dynamics of Puumala virus antibody prevalence in voles and of nephropathia epidemica incidence in humans. *Am J Trop Med Hyg* **53**, 134–140.

Novella, I. S., Zarate, S., Metzgar, D. & Ebendick-Corpus, B. E. (2004). Positive selection of synonymous mutations in vesicular stomatitis virus. *J Mol Biol* **342**, 1415–1421.

Olsson, G. E., White, N., Ahlm, C., Elgh, F., Verlemyr, A. C., Juto, P. & Palo, P. T. (2002). Demographic factors associated with hantavirus infection in bank voles (*Clethrionomys glareolus*). *Emerg Infect Dis* **8**, 924–929.

Plyusnin, A. (2002). Genetics of hantaviruses: implications to taxonomy. *Arch Virol* **147**, 665–682.

Plyusnin, A. & Morzunov, S. P. (2001). Virus evolution and genetic diversity of hantaviruses and their rodent hosts. *Curr Top Microbiol Immunol* **256**, 47–75.

Plyusnin, A., Vapalahti, O., Lehväslaiho, H., Apekina, N., Mikhailova, T., Gavrilovskaya, I., Laakkonen, J., Niemimaa, J., Henttonen, H. & other authors (1995a). Genetic variation of wild Puumala viruses within the serotype, local rodent populations and individual animal. *Virus Res* **38**, 25–41.

Plyusnin, A., Cheng, Y., Vapalahti, O., Pejcoch, M., Unar, J., Jelinkova, Z., Lehväslaiho, H., Lundkvist, Å. & Vaheri, A. (1995b). Genetic variation in Tula hantaviruses: sequence analysis of the S and M segments of strains from Central Europe. *Virus Res* **39**, 237–250.

Plyusnin, A., Cheng, Y., Vapalahti, O., Lehväslaiho, H. & Vaheri, A. (1996a). Quasispecies in wild Tula hantavirus populations. *J Virol* **70**, 9060–9063.

Plyusnin, A., Vapalahti, O. & Vaheri, A. (1996b). Hantaviruses: genome structure, expression and evolution. *J Gen Virol* **77**, 2677–2687.

Plyusnin, A., Hörling, J., Kanerva, M., Mustonen, J., Cheng, Y., Partanen, J., Vapalahti, O., Kukkonen, S.K., Niemimaa, J. & other authors (1997). Puumala hantavirus genome in patients with nephropathia epidemica: correlation of PCR positivity with HLA haplotype and link to viral sequences in local rodents. *J Clin Microbiol* **35**, 1090–1096.

Plyusnina, A., Aberle, S. W., Aberle, J. H. & Plyusnin, A. (2006). Genetic analysis of Puumala hantavirus strains from Austria. *Scand J Infect Dis* **38**, 512–519.

Sall, A. A., Zanotto, P. M. de A., Sene, O. K., Zeller, H. G., Digoutte, J. P., Thiongane, Y. & Bouloy, M. (1999). Genetic reassortment of Rift Valley fever virus in nature. *J Virol* **73**, 8196–8200.

Sauvage, F., Langlais, M., Yoccoz, N. G. & Pontier, D. (2003). Modelling hantavirus in fluctuating populations of bank voles: the role of indirect transmission on virus persistence. *J Anim Ecol* **72**, 1–13.

Schmaljohn, C. S. & Dalrymple, J. M. (1983). Analysis of Hantaan virus RNA: evidence for a new genus of *Bunyaviridae*. *Virology* **131**, 482–491.

Schmaljohn, C. & Hjelle, B. (1997). Hantaviruses: a global disease problem. *Emerg Infect Dis* **3**, 95–104.

Sibold, C., Meisel, H., Kruger, D. H., Labuda, M., Lysy, J., Kozuch, O., Pejcoch, M., Vaheri, A. & Plyusnin, A. (1999). Recombination in Tula hantavirus evolution: analysis of genetic lineages from Slovakia. *J Virol* **73**, 667–675.

Sironen, T., Vaheri, A. & Plyusnin, A. (2001). Molecular evolution of Puumala hantavirus. *J Virol* **75**, 11803–11810.

Sironen, T., Kallio, E., Vaheri, A., Lundkvist, Å. & Plyusnin, A. (2008). Quasispecies and fixation of synonymous mutation in hantavirus transmission. *J Gen Virol* **89**, 1309–1313.

Urquidi, V. & Bishop, D. H. L. (1992). Non-random reassortment between the tripartite RNA genomes of La Crosse and snowshoe hare viruses. *J Gen Virol* **73**, 2255–2265.

Vapalahti, O., Mustonen, J., Lundkvist, Å., Henttonen, H., Plyusnin, A. & Vaheri, A. (2003). Hantavirus infections in Europe. *Lancet Infect Dis* **3**, 653–661.

Wilson, D. E. & Reeders, D. M. (2005). *Mammal Species of the World. A Taxonomic and Geographic Reference*, 3rd edn, vol. 2. Baltimore: The Johns Hopkins University Press.