Microevolution of Puumala hantavirus in its host, the bank vole (*Myodes glareolus*)

Maria Razzauti Sanfeliu

Research Programs Unit
Infection Biology Research Program
Department of Virology, Haartman Institute
Faculty of Medicine, University of Helsinki
and the Finnish Forest Research Institute

Academic dissertation
Helsinki 2012

To be presented for public examination, with the permission of the Faculty of Medicine, University of Helsinki, in Lecture Hall 2, Haartman Institute on February the 24th at noon
Supervisors
Professor, docent Alexander Plyusnin
Department of Virology, Haartman Institute
P.O.Box 21, FI-00014 University of Helsinki, Finland
e-mail: alexander.plyusnin@helsinki.fi

Professor Heikki Henttonen
Finnish Forest Research Institute (Metla)
P.O.Box 18, FI-01301 Vantaa, Finland
e-mail: heikki.henttonen@metla.fi

Reviewers
Docent Petri Susi
Biosciences and Business
Turku University of Applied Sciences,
Lemminkäisenkatu 30, FI-20520 Turku, Finland
e-mail: petri.susi@turkuamk.fi

Professor Dennis Bamford
Department of Biological and Environmental Sciences
Division of General Microbiology
P.O.Box 56, FI-00014 University of Helsinki, Finland
e-mail: dennis.bamford@helsinki.fi

Opponent
Professor Herwig Leirs
Dept. Biology, Evolutionary Ecology group
Groenenborgercampus, room G.V323a
Groenenborgerlaan 171, B-2020
University of Antwerpen, Belgium
e-mail: herwig.leirs@ua.ac.be

ISBN 978-952-10-7688-6 (PDF)
© Maria Razzauti Sanfeliu, 2012
Mitochondrial DNA amplification ................................................................................... 48
Genetic variation and phylogenetic analyses .............................................................. 48
RESULTS AND DISCUSSION .......................................................................................... 49
Microevolution of PUUV related to host population dynamics ..................................... 49
PUUV lineages diversity and their post-glacial history ................................................... 59
CONCLUDING REMARKS AND FUTURE PROSPECTS .............................................. 63
ACKNOWLEDGEMENTS ................................................................................................. 65
REFERENCES .................................................................................................................. 67
REPRINTS OF ORIGINAL PUBLICATIONS ................................................................ 88
LIST OF ORIGINAL PUBLICATIONS

The present thesis is based on the following publications, which are referred to in the text by roman numerals.


IV. Maria Razzauti, Angelina Plyusnina, Heikki Henttonen & Alexander Plyusnin. Microevolution of Puumala hantavirus during a complete population cycle of its host, the bank vole (*Myodes glareolus*). Manuscript.

The published articles have been reprinted with the permission of the respective copyright holders.
# ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>aa</td>
<td>amino acid</td>
</tr>
<tr>
<td>Ab</td>
<td>antibody</td>
</tr>
<tr>
<td>AD</td>
<td>Anno Domini</td>
</tr>
<tr>
<td>Ag</td>
<td>antigen</td>
</tr>
<tr>
<td>ALAD</td>
<td>Alpe-Adrian genetic lineage of PUUV</td>
</tr>
<tr>
<td>cDNA</td>
<td>complementary DNA</td>
</tr>
<tr>
<td>CE</td>
<td>Central European genetic lineage of PUUV</td>
</tr>
<tr>
<td>ca.</td>
<td>circa</td>
</tr>
<tr>
<td>CR</td>
<td>coding region</td>
</tr>
<tr>
<td>Cyt b</td>
<td>cytochrome b</td>
</tr>
<tr>
<td>DAN</td>
<td>Danish genetic lineage of PUUV</td>
</tr>
<tr>
<td>DM</td>
<td>distance matrix</td>
</tr>
<tr>
<td>EDTA</td>
<td>ethylene diamine tetra-acetic acid</td>
</tr>
<tr>
<td>ELISA</td>
<td>enzyme-linked immunosorbent assay</td>
</tr>
<tr>
<td>ER</td>
<td>endoplasic reticulum</td>
</tr>
<tr>
<td>ERGIC</td>
<td>endoplasic reticulum-Golgi intermediate compartments</td>
</tr>
<tr>
<td>FIN</td>
<td>Finnish genetic lineage of PUUV</td>
</tr>
<tr>
<td>FITC</td>
<td>fluorescein isothiocyanate</td>
</tr>
<tr>
<td>FM</td>
<td>Fitch-Margoliash</td>
</tr>
<tr>
<td>FRNT</td>
<td>focus reduction neutralization test</td>
</tr>
<tr>
<td>GTR</td>
<td>general time reversible</td>
</tr>
<tr>
<td>HCPS</td>
<td>hantavirus cardiopulmonary syndrome</td>
</tr>
<tr>
<td>HFRS</td>
<td>haemorrhagic fever with renal syndrome</td>
</tr>
<tr>
<td>HLA</td>
<td>human leukocyte antigen</td>
</tr>
<tr>
<td>ICTV</td>
<td>International Committee on Taxonomy of Viruses</td>
</tr>
<tr>
<td>IFA</td>
<td>immunofluorescence assay</td>
</tr>
<tr>
<td>IFN-β</td>
<td>interferon beta</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>IgG</td>
<td>immunoglobulin G</td>
</tr>
<tr>
<td>IgM</td>
<td>immunoglobulin M</td>
</tr>
<tr>
<td>IRF-3</td>
<td>interferon regulatory factor-3</td>
</tr>
<tr>
<td>Ka</td>
<td>kiloannum</td>
</tr>
<tr>
<td>kb</td>
<td>kilobase</td>
</tr>
<tr>
<td>kDa</td>
<td>kiloDalton</td>
</tr>
<tr>
<td>L</td>
<td>large genome segment of <em>Bunyaviridae</em></td>
</tr>
<tr>
<td>LAT</td>
<td>Latvian genetic lineage of PUUV</td>
</tr>
<tr>
<td>LGM</td>
<td>Last Glacial Maximum</td>
</tr>
<tr>
<td>M</td>
<td>medium genome segment of <em>Bunyaviridae</em></td>
</tr>
<tr>
<td>Ma</td>
<td>mega-annum</td>
</tr>
<tr>
<td>MCMC</td>
<td>Markov chain Monte Carlo</td>
</tr>
<tr>
<td>MHC</td>
<td>major histocompatibility complex</td>
</tr>
<tr>
<td>MJ</td>
<td>median-joining</td>
</tr>
<tr>
<td>ML</td>
<td>maximum likelihood</td>
</tr>
<tr>
<td>MP</td>
<td>maximum parsimony</td>
</tr>
<tr>
<td>mRNA</td>
<td>messenger RNA</td>
</tr>
<tr>
<td>mtDNA</td>
<td>mitochondrial DNA</td>
</tr>
<tr>
<td>NCR</td>
<td>non-coding region</td>
</tr>
<tr>
<td>NE</td>
<td>nephropathia epidemica</td>
</tr>
<tr>
<td>NF-κB</td>
<td>nuclear factor kappa B</td>
</tr>
<tr>
<td>NJ</td>
<td>neighbour-joining</td>
</tr>
<tr>
<td>nt</td>
<td>nucleotide(s)</td>
</tr>
<tr>
<td>N-SCA</td>
<td>North-Scandinavian genetic lineage of PUUV</td>
</tr>
<tr>
<td>OD</td>
<td>optical density</td>
</tr>
<tr>
<td>ORF</td>
<td>open reading frame</td>
</tr>
<tr>
<td>P bodies</td>
<td>cytoplasmic processing bodies</td>
</tr>
<tr>
<td>PBS</td>
<td>phosphate buffered saline</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>PCR</td>
<td>polymerase chain reaction</td>
</tr>
<tr>
<td>pNPP</td>
<td>p-nitrophenyl phosphate</td>
</tr>
<tr>
<td>PUUV</td>
<td>Puumala virus</td>
</tr>
<tr>
<td>RdRp</td>
<td>RNA-dependent RNA polymerase</td>
</tr>
<tr>
<td>rN</td>
<td>recombinant nucleocapsid protein</td>
</tr>
<tr>
<td>RT</td>
<td>reverse transcription</td>
</tr>
<tr>
<td>S</td>
<td>small genome segment of <em>Bunyaviridae</em></td>
</tr>
<tr>
<td>SD</td>
<td>standard deviation</td>
</tr>
<tr>
<td>SDS-PAGE</td>
<td>sodium dodecyl sulfate polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>S-SCA</td>
<td>South-Scandinavian genetic lineage of PUUV</td>
</tr>
<tr>
<td>TNF</td>
<td>tumor necrosis factor</td>
</tr>
<tr>
<td>RUS</td>
<td>Russian genetic lineage of PUUV</td>
</tr>
<tr>
<td>vRNA</td>
<td>viral RNA</td>
</tr>
<tr>
<td>WB</td>
<td>western blot</td>
</tr>
</tbody>
</table>

A list of hantaviruses and their abbreviations can be found in Table 1 on page 13.
SUMMARY

Puumala hantavirus (PUUV) is a zoonotic virus that causes nephropathia epidemica (NE) in humans, a mild form of haemorrhagic fever with renal syndrome. An average of 10 000 cases are reported annually in Europe, many of which occur in Fennoscandia. The incidence of NE is connected to the distribution and population density of the the bank vole (Myodes glareolus), the main host of the virus. In Fennoscandia, high incidences of NE occur at 3-4 year intervals due to the characteristic population cycles of this woodland rodent.

This study aimed to improve our understanding of PUUV microevolution by examining genetic features of the virus in several bank vole populations of Finland and Latvia.

Genetic variation in PUUV circulating in a bank vole population at Konnevesi in Central Finland was examined and monitored over five-years throughout a complete bank vole cycle, including two peak-phases in 2005 and 2008 and two population declines in 2006 and 2009 (i.e., viral bottlenecks). Altogether, 1369 bank voles were captured and 26.3% were detected PUUV-infected. Partial sequences of the three viral genome segments (Small, Medium and Large) were inspected from 365 PUUV genomes. Genetic diversity was 6.2% for the S segment, 4.8% for the M segment, and a surprisingly high 10.1% for the L segment. Each genome segment had accumulated mutations as a separate gene pool. The majority of nucleotide substitutions were synonymous and most of the deduced amino acid substitutions were conservative, suggesting a strong stabilizing selection operating at the protein level. Genetic markers found along the genome segments allowed for the recognition of two PUUV genogroups co-circulating in the host population. Even though one of the genogroups presented a higher genetic diversity, no signs of competition were observed between them. Nearly 80% the variants exhibited a transient existence, and frequently occurring variants were integrated by the most abundant segment genotypes suggesting a viral mutational robustness. A substantial portion (19.1%) of genomes appeared to be reassortants, with S and M typically being exchanged. Reassortant variants did not outcompete parental variants and were commonly transient. Reassortment was seasonal, occurring more frequently in autumn when infection risk increases. An imperceptible intra-genogroup reassortment could contribute to the steady state of the viral population, counteracting the effects of Muller’s ratchet.

Co-circulation and interaction of two distinct PUUV lineages (Finnish and North-Scandinavian) was monitored in a bank vole population at Pallasjärvi in Northern Finland. To date, seven genetic lineages have been detected, and they exhibit geographic structure within the host distribution. Here, we present new evidence of
two lineages circulating in the same bank vole phylogroup (Ural clade). Genetic diversity within each PUUV lineage was modest (up to 1.7%) and most substitutions were synonymous. However, genetic differences between the two lineages were as high as 18.9%. Phylogenetic analysis revealed that these distinct lineages naturally reassort with a frequency to that genogroups circulating at Konnevesi, i.e., 32%. In contrast to Konnevesi, only M segment was exchanged between PUUV lineages at Pallasjärvi.

Two distinct PUUV lineages were also found to co-circulate in Latvia. One (Russian) has been previously described and the other awaits formal description. The novel Latvian lineage is considerably divergent from other PUUV lineages and several amino acid markers made it easily distinguishable. Phylogenetic analysis suggested an independent evolutionary history for the segments of the Latvian lineage. Similar to Pallasjärvi, both Russian and Latvian lineages were found in a single bank vole phylogroup (Carpathian clade), confirming earlier observations that PUUV lineages are not limited to a single host phylogroup.
REVIEW OF THE LITERATURE

History of hantaviruses

Haemorrhagic fever with renal syndrome (HFRS) is a major public health problem and has been occurring throughout Eurasia for hundreds of years [WHO]. HFRS encompasses a group of clinically similar illnesses previously known as “Korean hemorrhagic fever”, “epidemic nephritis”, “war nephritis”, “trench nephritis”, “field nephritis”, “hemorrhagic nephrosonephritis”, “virus glomerulonephritis”, “epidemic hemorrhagic fever” and “nephropathia epidemica”. In 1982, the WHO recommended fixing “HFRS-like diseases” as the name in widespread use. Hantavirus infections have a long history; ancient Chinese writings dating back to 960 AD describe the symptoms of a haemorrhagic fever syndrome. More recently, outbreaks of “field nephritis” were reported in over 60 000 combatants during wars of the past 150 years: 14 187 participants of the American Civil War (1861-56); ~12 000 soldiers in World War I (1914-18); ~12500 soldiers of Japanese troops in Manchuria, ~8 000 of Soviet troops in the Far East, ~10 000 of German and Finnish troops in Lapland (1942), and ~6 000 German prisoners of war in Yugoslavia during the World War II (1939-45); and 3 256 members of the United Nation Command in the Korean War (1951-54) [Casals et al., 1969; Lee, 1996]. The continuous occurrence of HFRS in the field provided an early indication that the aetiological agent was a zoonotic microbe.

Clinical manifestations of the disease were studied during 1951-54 [Traub & Wisseman, 1978]. In 1976, Dr. Ho-Wang Lee and his colleagues suspected that the etiological agent was a virus borne in rodents. Subsequently, they found an antigen in the lungs and kidneys of the striped field mouse (Apodemus agrarius) collected in endemic foci. The newly discovered virus was named Hantaan virus (HTNV) after the Hantan River in South Korea, where the strain originated [Lee & Lee, 1976]. Lung tissues from Apodemus agrarius coreae presented a specific immunofluorescent reaction with sera from patients recovering from “Korean hemorrhagic fever” [Lee et al., 1978].

Early attempts to infect wild rodents or establish a cell culture system for replication of the presumptive virus failed. In 1981, a prototype virus strain was adapted for growth in cultured human lung carcinoma cells (A-549) [French et al., 1981]. Later, monkey epithelial kidney cells (Vero E6) were shown to support viral growth better than A-549 cells [Schmaljohn et al., 1985], an soon after, the first electron micrographs of the virus were obtained by two groups working independently [McCormick et al., 1982; White et al., 1982]. Although the morphology resembled that of a bunyavirus, Hantavirus was not accepted as a new genus of the Bunyaviridae until Hantaan-like viruses were shown to possess three RNA genomic segments that circularize in vivo as a result of terminal complementary nucleotides that help to fold them into a “hairpin” structure [Schmaljohn & Dalrymple, 1983].
In 1979 Dr. Markus Brummer-Korvenkontio and colleagues demonstrated that lung tissue of bank voles (*Myodes glareolus*) reacts with sera of Finnish HFRS-like patients [Brummer-Korvenkontio et al., 1980, 1982]. Immunofluorescent assays revealed a virus related to but distinct from HTNV. The newly recognized aetiological agent was called *Puumala virus* (PUUV) named after the small town in South-eastern Finland where it was first detected. PUUV causes a mild form of HFRS, called nephropathia epidemica (NE) [Brummer-Korvenkontio et al., 1982]. Antibodies reacting with aetiological agents of HFRS were found in sera of patients in Europe and Asia [Lee, 1982] and led to the discovery of other hantaviruses. *Seoul virus* (SEOV) was found in rats lungs (*Rattus rattus* and *R. norvegicus*) from urban areas of Seoul, and viral antigen was established in laboratory rats [Lee et al., 1982]. In Maryland, another virus named *Prospect Hill* (PHV) was recovered from the meadow vole (*Microtus pennsylvanicus*), although no association with acute human disease was observed [Lee et al., 1985]. Previously considered arbovirus, *Thottapalayam virus* (TPMV) was found in the Asian musk shrew (*Suncus murinus*) in India and subsequently reclassified [Carey et al., 1971; Song et al., 2007]. And *Dobrava-Belgrade virus* (DOBV) was isolated from the yellow-necked mouse (*Apodemus flavicollis*) from Slovenia [Avšić-Županc et al., 1992].

In 1993, hantaviruses became a concern in the Americas after an acute respiratory distress outbreak occurred among the Navajo Nation in "The Four Corners", an area including adjacent parts of Arizona, New Mexico, Colorado and Utah [Nichol et al., 1993]. The newly-described disease was referred as "unexplained adult respiratory distress syndrome" (ARDS). It was ultimately identified as a virus of the *Hantaviridae* family. The Four Corners outbreak occurred during unusual environmental conditions due to the El Niño in 1991-1992 that created a warm winter and a rainy spring in 1993. This contributed to the explosive growth of vegetation, providing food and cover for a burgeoning rodent population. The South-western USA experienced a tenfold increase in the deer mouse (*Peromyscus maniculatus*) population, the species recognized as the virus reservoir [Childs et al., 1994]. The virus was given the name of *Muerto Canyon virus* but, at the request of the Navajo community, its name was eventually changed to *Sin Nombre virus* (SNV) (in Spanish, "nameless virus"), and the linked disease became known as hantavirus cardiopulmonary syndrome (HCPS).

Numerous viruses in the Old and New World have since been classified within *Hantavirus*, but only some have been shown to cause human diseases. A detailed list of known hantaviruses is presented in Table 1.
**Taxonomy of hantaviruses**

*Bunyaviridae* consists of five genera: *Tospovirus, Nairovirus, Phlebovirus, Orthobunyavirus* and *Hantavirus*. Hantaviruses present a tri-segmented single-stranded RNA genome of negative-polarity. Even though the other genera within *Bunyaviridae* are mainly arthropod-borne hantaviruses are transmitted to humans through inhalation of aerosolized rodent excretions. Thus far, no human infections have been associated with hantaviruses borne by insectivores.

Rodent-borne hantaviruses are known as “RoBo-viruses” and insectivore-borne hantaviruses as “InBo-viruses”, and the inclusive “RaInBo-viruses” has recently been adopted. Hantaviruses are also commonly known as Old and New World viruses depending on the geographic distribution of their hosts. Hantaviruses present a firmly established host association, yet host-switching events have been reported in the literature [Morzunov et al., 1998; Vapalahti et al., 1999; Nemirov et al., 2003]. Hantaviruses have been detected in murid and cricetid rodents of subfamilies Murinae (Old World rats and mice), Arvicolinae (voles and lemmings), Neotominae and Sigmodontinae (New World rats and mice). InBo-viruses have been detected in soricine and crocidurine soridics (shrews), as well as moles (Talpidae) (Table 1 and Fig. 8).

**Table 1.** List of hantaviruses. Hantavirus species currently recognized by the International Committee on Taxonomy of Viruses are shown in bold. HFRS stands for haemorrhagic fever with renal syndrome, NE stands for nephropathia epidemica, and HCPS stands for hantavirus cardiopulmonary syndrome.

<table>
<thead>
<tr>
<th>Virus</th>
<th>Abbreviation</th>
<th>Human disease</th>
<th>Host</th>
<th>Geographic range</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Hantaviruses carried by rodents in the family Muridae, subfamily Murinae</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thailand</td>
<td>THAIV</td>
<td>?</td>
<td><em>Bandicota indica</em> (great bandicoot rat)</td>
<td>S Asia</td>
<td>Xiao et al., 1994</td>
</tr>
<tr>
<td>Seoul</td>
<td>SEOV</td>
<td>HFRS</td>
<td><em>Rattus rattus</em> (black rat), <em>Rattus norvegicus</em> (brown rat)</td>
<td>Global</td>
<td>Elwell et al., 1985</td>
</tr>
<tr>
<td>Serang</td>
<td>SERV</td>
<td>?</td>
<td><em>Rattus tanezumi</em> (Asian house rat)</td>
<td>SE Asia</td>
<td>Plyusnina et al., 2009</td>
</tr>
<tr>
<td>Da Bie Shan</td>
<td>DBSV</td>
<td>?</td>
<td><em>Niviventer confucianus</em> (Chinese white-bellied rat)</td>
<td>E Asia</td>
<td>Wang et al., 2000</td>
</tr>
<tr>
<td>Hantaan</td>
<td>HTNV</td>
<td>HFRS</td>
<td><em>Apodemus agrarius</em> (striped field mouse)</td>
<td>E Eurasia</td>
<td>Lee &amp; Lee, 1978</td>
</tr>
<tr>
<td>Saaremaa</td>
<td>SAAV</td>
<td>HFRS</td>
<td><em>Apodemus agrarius</em> (striped field mouse)</td>
<td>W and C Eurasia</td>
<td>Plyusnin et al., 1997, Nemirov et al., 1999</td>
</tr>
</tbody>
</table>
### Review of the literature: Taxonomy of hantaviruses

#### Dobrava-Belgrade
- **DOBV**
- **HFRS**
- **Apodemus flavicollis** (yellow-necked mouse)
- **Europe**
- **Avšić-Županc et al., 1992, 1995**

#### Amur/Soochong
- **ASV**
- **HFRS**
- **Apodemus peninsulae** (Korean field mouse)
- **E Asia**
- **Liang et al., 1994, Yashina et al., 2001, Lokugamage et al., 2002, Baek et al., 2006**

#### Sangassou
- **SANGV**
- ?
- **Hylomyscus simus** (African wood mouse)
- **Côte d'Ivoire, Africa**
- **Klempa et al., 2006**

### Hantaviruses carried by rodents in the family Cricetidae, subfamily Arvicolinae

<table>
<thead>
<tr>
<th>Location</th>
<th>Code</th>
<th>Virus Code</th>
<th>Type</th>
<th>Host Species</th>
<th>Distribution</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Puumala</td>
<td>PUUV</td>
<td>HFRS (NE)</td>
<td><strong>Myodes glareolus</strong> (bank vole)</td>
<td>Europe to W Siberia</td>
<td>Brummer-Korvenkontio et al., 1980</td>
<td></td>
</tr>
<tr>
<td>Hokkaido</td>
<td>HOKV</td>
<td>?</td>
<td><strong>Myodes rufocanus</strong> (grey-sided vole)</td>
<td>N Eurasia</td>
<td>Kariwa et al., 1995</td>
<td></td>
</tr>
<tr>
<td>Muju</td>
<td>MUJV</td>
<td>?</td>
<td><strong>Myodes regulus</strong> (royal vole)</td>
<td>Korea</td>
<td>Song et al., 2007</td>
<td></td>
</tr>
<tr>
<td>Prospect Hill</td>
<td>PHV</td>
<td>-</td>
<td><strong>Microtus pennsylvanicus</strong> (meadow vole)</td>
<td>North America</td>
<td>Lee et al., 1982, 1985</td>
<td></td>
</tr>
<tr>
<td>Tula</td>
<td>TULV</td>
<td>-</td>
<td><strong>Microtus arvalis</strong> (European common vole)</td>
<td>Europe to C Asia</td>
<td>Plyusnin et al., 1994a</td>
<td></td>
</tr>
<tr>
<td>Topografov</td>
<td>TOPV</td>
<td>?</td>
<td><strong>Lemmus sibiricus</strong> (Siberian brown lemming)</td>
<td>Palaearctic tundra</td>
<td>Plyusnin et al., 1996a</td>
<td></td>
</tr>
<tr>
<td>Isla Vista</td>
<td>ISLAV</td>
<td>?</td>
<td><strong>Microtus californicus</strong> (Californian vole)</td>
<td>W North America</td>
<td>Song et al., 1995</td>
<td></td>
</tr>
<tr>
<td>Bloodland Lake</td>
<td>BLLV</td>
<td>?</td>
<td><strong>Microtus ochrogaster</strong> (prairie vole)</td>
<td>C North America</td>
<td>Song et al., 1995</td>
<td></td>
</tr>
<tr>
<td>Khabarovsk</td>
<td>KHAV</td>
<td>?</td>
<td><strong>Microtus maximowiczii</strong> (Maximowicz's vole)</td>
<td>NE Asia</td>
<td>Hörling et al., 1996</td>
<td></td>
</tr>
<tr>
<td>Vladivostok</td>
<td>VLAV</td>
<td>?</td>
<td><strong>Microtus fortis</strong> (reed vole)</td>
<td>NE Asia</td>
<td>Kariwa et al., 1999</td>
<td></td>
</tr>
<tr>
<td>Yuanjiang</td>
<td>YUJV</td>
<td>?</td>
<td><strong>Microtus fortis</strong> (reed vole)</td>
<td>NE Asia</td>
<td>Zou et al., 2008</td>
<td></td>
</tr>
<tr>
<td>Luxi</td>
<td>LUXV</td>
<td>HFRS</td>
<td><strong>Eothenomys miletus</strong> (Yunnan red-backed vole)</td>
<td>SW China</td>
<td>Zhang et al., 2011</td>
<td></td>
</tr>
</tbody>
</table>

### Hantaviruses carried by rodents in the family Cricetidae, subfamily Neotominae

<table>
<thead>
<tr>
<th>Location</th>
<th>Code</th>
<th>Virus Code</th>
<th>Type</th>
<th>Host Species</th>
<th>Distribution</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sin Nombre</td>
<td>SNV</td>
<td>HCPS</td>
<td><strong>Peromyscus maniculatus</strong> (deer mouse)</td>
<td>North America</td>
<td>Nichol et al., 1993</td>
<td></td>
</tr>
<tr>
<td>Monongahela</td>
<td>MGLV</td>
<td>HCPS</td>
<td><strong>Peromyscus maniculatus</strong> (deer mouse)</td>
<td>North America</td>
<td>Song et al., 1996</td>
<td></td>
</tr>
<tr>
<td>New York</td>
<td>NYV</td>
<td>HCPS</td>
<td><strong>Peromyscus leucopus</strong> (white-footed mouse)</td>
<td>E North America</td>
<td>Hjelle et al. 1995a, 1995b</td>
<td></td>
</tr>
<tr>
<td>Blue River</td>
<td>BRV</td>
<td>?</td>
<td><strong>Peromyscus leucopus</strong> (white-footed mouse)</td>
<td>E North America</td>
<td>Morzunov et al., 1998</td>
<td></td>
</tr>
<tr>
<td>Montano</td>
<td>MTNV</td>
<td>?</td>
<td><strong>Peromyscus beatae</strong> (Orizaba deermouse)</td>
<td>Central America</td>
<td>Kariwa et al., 2011</td>
<td></td>
</tr>
</tbody>
</table>
### Review of the literature: Taxonomy of hantaviruses

<table>
<thead>
<tr>
<th>Location</th>
<th>Code</th>
<th>Species</th>
<th>Location</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>El Moro Canyon</td>
<td>ELMC</td>
<td><em>Reithrodontomys megalotis</em> (western harvest mouse)</td>
<td>W North America</td>
<td>Hjelle et al., 1994</td>
</tr>
<tr>
<td>Limestone Canyon</td>
<td>LSCV</td>
<td><em>Reithrodontomys megalotis</em> (western harvest mouse)</td>
<td>W North America</td>
<td>Sanchez et al., 2001</td>
</tr>
<tr>
<td>Huitzilac</td>
<td>HUIV</td>
<td><em>Reithrodontomys megalotis</em> (western harvest mouse)</td>
<td>W North America</td>
<td>Kariwa et al., 2011</td>
</tr>
<tr>
<td>Rio Segundo</td>
<td>RIOSV</td>
<td><em>Reithrodontomys mexicanus</em> (Mexican harvest mouse)</td>
<td>Central America</td>
<td>Hjelle et al., 1995c</td>
</tr>
<tr>
<td>Carrizal</td>
<td>CARV</td>
<td><em>Reithrodontomys sumichrasti</em> (Sumichrast's harvest mouse)</td>
<td>Central America</td>
<td>Kariwa et al., 2011</td>
</tr>
</tbody>
</table>

### Hantaviruses carried by rodents in the family Cricetidae, subfamily Sigmodontinae

<table>
<thead>
<tr>
<th>Location</th>
<th>Code</th>
<th>Species</th>
<th>Location</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Black Creek Canal</td>
<td>BCCV</td>
<td><em>Sigmodon hispidus</em> (hispid cotton rat)</td>
<td>S North America</td>
<td>Ravkov et al., 1995, Rawlings et al., 1996</td>
</tr>
<tr>
<td>Muleshoe</td>
<td>MULV</td>
<td><em>Sigmodon hispidus</em> (hispid cotton rat)</td>
<td>S North America</td>
<td>Rawlings et al., 1996</td>
</tr>
<tr>
<td>Cano Delgadito</td>
<td>CADV</td>
<td><em>Sigmodon alstoni</em> (cane mouse)</td>
<td>N South America</td>
<td>Fulhorst et al., 1997</td>
</tr>
<tr>
<td>Playa de Oro</td>
<td>OROV</td>
<td><em>Sigmodon mascotensis</em> (Jaliscan cotton rat)</td>
<td>Central America</td>
<td>Chu et al., 2008</td>
</tr>
<tr>
<td>Catacamas</td>
<td>CATV</td>
<td><em>Oryzomys couesi</em> (Coues' rice rat)</td>
<td>Central America</td>
<td>Milazzo et al., 2006</td>
</tr>
<tr>
<td>Bayou</td>
<td>BAYV</td>
<td><em>Oryzomys palustris</em> (rice rat)</td>
<td>SE North America</td>
<td>Morzunoz et al., 1995</td>
</tr>
<tr>
<td>Rio Mamore</td>
<td>RIOMV</td>
<td><em>Oligoryzomys microtis</em> (small-eared pygmy rice rat)</td>
<td>N South America</td>
<td>Hjelle et al., 1995a</td>
</tr>
<tr>
<td>Andes</td>
<td>ANDV</td>
<td><em>Oligoryzomys longicaudatus</em> (long-tailed pigmy rice rat)</td>
<td>S South America</td>
<td>Lopez et al., 1996</td>
</tr>
<tr>
<td>Bermejo</td>
<td>BMJV</td>
<td><em>Oligoryzomys chacoensis</em> (Chacoan pygmy rice rat)</td>
<td>WC South America</td>
<td>Levis et al., 1997</td>
</tr>
<tr>
<td>Lechiguana</td>
<td>LECV</td>
<td><em>Oligoryzomys flavescens</em> (yellow pygmy rice rat)</td>
<td>NC South America</td>
<td>Levis et al., 1997</td>
</tr>
<tr>
<td>Choclo</td>
<td>CHOV</td>
<td><em>Oligoryzomys costaricensis</em> (Costa Rican pygmy rice rat)</td>
<td>NC South America</td>
<td>Vicent et al., 2000, Hanson et al., 2011</td>
</tr>
<tr>
<td>Maporal</td>
<td>MAPV</td>
<td><em>Oligoryzomys delicatus</em> (delicate pygmy rice rat)</td>
<td>NC South America</td>
<td>Fulhosrt et al., 2004, Hanson et al., 2011</td>
</tr>
<tr>
<td>Oran</td>
<td>ORNV</td>
<td><em>Oligoryzomys longicaudatus</em> (long-tailed pigmy rice rat)</td>
<td>S South America</td>
<td>Levis et al., 1997, Bohlman et al., 2002</td>
</tr>
<tr>
<td>Castelo dos sonhos</td>
<td>CASV</td>
<td><em>Oligoryzomys eliurus</em> (Brazilian pygmy rice rat)</td>
<td>EC South America</td>
<td>Johnson et al. 1999</td>
</tr>
<tr>
<td>Itapua</td>
<td>ITPV</td>
<td><em>Oligoryzomys nigripes</em> (black-footed pygmy rice rat)</td>
<td>N South America</td>
<td>Chu et al., 2003</td>
</tr>
<tr>
<td>Neembucu</td>
<td>HCPS</td>
<td><em>Oligoryzomys chacoensis</em> (Chacoan pygmy rice rat)</td>
<td>N South America</td>
<td>Chu et al., 2003</td>
</tr>
</tbody>
</table>
### Review of the literature: Taxonomy of hantaviruses

<table>
<thead>
<tr>
<th>Location</th>
<th>Code</th>
<th>Species</th>
<th>Host Family</th>
<th>Species Description</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Juquitiba</td>
<td>JUQV</td>
<td><em>Oligoryzomys fornesi</em> (Fornes' colillargo)</td>
<td>Rodent</td>
<td>N South America</td>
<td>Vasconcelos et al., 1997</td>
</tr>
<tr>
<td>Anajatuba</td>
<td>ANAJV</td>
<td><em>Oligoryzomys fornesi</em> (Fornes' colillargo)</td>
<td>Rodent</td>
<td>N South America</td>
<td>Rosa et al., 2005</td>
</tr>
<tr>
<td>Araucaria</td>
<td>ARAUV</td>
<td><em>Oligoryzomys nigripes</em> (black-footed pygmy rice rat)</td>
<td>Rodent</td>
<td>NC South America</td>
<td>Raboni et al., 2005</td>
</tr>
<tr>
<td>Jaborá</td>
<td>JABV</td>
<td><em>Akodon montensis</em> (montane grass mouse)</td>
<td>Rodent</td>
<td>C South America</td>
<td>Goodin et al., 2009</td>
</tr>
<tr>
<td>Ape Aime-Itapua</td>
<td>AAIV</td>
<td><em>Akodon montensis</em> (montane grass mouse)</td>
<td>Rodent</td>
<td>C South America</td>
<td>Chu et al., 2003</td>
</tr>
<tr>
<td>Pergamino</td>
<td>PRGV</td>
<td><em>Akodon azarae</em> (Azara’s grass mouse)</td>
<td>Rodent</td>
<td>N South America</td>
<td>Levis et al., 1998, Bohlman et al., 2002</td>
</tr>
<tr>
<td>Laguna Negra</td>
<td>LANV</td>
<td><em>Calomys laucha</em> (vesper mouse)</td>
<td>Rodent</td>
<td>N South America</td>
<td>Johnson et al., 1997</td>
</tr>
<tr>
<td>Maciel</td>
<td>MCLV</td>
<td><em>Bolomys obscurus</em> (dark bolo mouse)</td>
<td>Rodent</td>
<td>S South America</td>
<td>Levis et al., 1998, Bohlman et al., 2002</td>
</tr>
<tr>
<td>Araraquara</td>
<td>ARAV</td>
<td><em>Bolomys lascius</em> (hairy-tailed bolo mouse)</td>
<td>Rodent</td>
<td>EC South America</td>
<td>Johnson et al., 1999</td>
</tr>
<tr>
<td>Alto Paraguay</td>
<td>ALPV</td>
<td><em>Holochilus chacarius</em> (Chacoan marsh rat)</td>
<td>Rodent</td>
<td>S South America</td>
<td>Chu et al., 2003</td>
</tr>
<tr>
<td>Rio Mearim</td>
<td>RIMEV</td>
<td><em>Holochilus sciuereus</em> (Amazonian marsh rat)</td>
<td>Rodent</td>
<td>N South America</td>
<td>Rosa et al., 2005</td>
</tr>
<tr>
<td>Calabazo</td>
<td>CALV</td>
<td><em>Zygodontomys brevicauda</em> (short-tailed cane mouse)</td>
<td>Rodent</td>
<td>N South America</td>
<td>Vicent et al., 2000</td>
</tr>
<tr>
<td>Paranoa</td>
<td>HPCS</td>
<td>?</td>
<td>Rodent</td>
<td>N South America</td>
<td>Melo-Silva et al., 2007</td>
</tr>
</tbody>
</table>

### Hantaviruses carried by insectivores in the family Soricidae, subfamily Soricinae

<table>
<thead>
<tr>
<th>Location</th>
<th>Code</th>
<th>Species</th>
<th>Host Family</th>
<th>Species Description</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seewis</td>
<td>SWSV</td>
<td><em>Sorex araneus</em> (Eurasian common shrew)</td>
<td>Soricidae</td>
<td>Eurasia</td>
<td>Song et al., 2007</td>
</tr>
<tr>
<td>Altai</td>
<td>ALTV</td>
<td><em>Sorex araneus</em> (Eurasian common shrew)</td>
<td>Soricidae</td>
<td>Eurasia</td>
<td>Yashina et al., 2008a</td>
</tr>
<tr>
<td>Artybash</td>
<td>ARTV</td>
<td><em>Sorex caecutiens</em> (masked shrew)</td>
<td>Soricidae</td>
<td>Taiga and tundra zones from N Europe to E Siberia</td>
<td>Yashina et al., 2008b</td>
</tr>
<tr>
<td>Lena River</td>
<td>LNAV</td>
<td><em>Sorex caecutiens</em> (masked shrew)</td>
<td>Soricidae</td>
<td>Taiga and tundra zones from N Europe to E Siberia</td>
<td>Arai et al., 2008</td>
</tr>
<tr>
<td>Jemez Springs</td>
<td>JMSV</td>
<td><em>Sorex monticolus</em> (dusky shrew)</td>
<td>Soricidae</td>
<td>W North America</td>
<td>Arai et al., 2008</td>
</tr>
<tr>
<td>Ash River</td>
<td>ARRV</td>
<td><em>Sorex cinereus</em> (cinereus shrew)</td>
<td>Soricidae</td>
<td>North America</td>
<td>Arai et al., 2008</td>
</tr>
</tbody>
</table>
The *Bunyaviridae* study group of the International Committee on Taxonomy of Viruses (ICTV) [Plyusnin et al., 2011] suggests the following criteria define hantavirus species: i) they engage a unique ecological niche, i.e., primary reservoir species; ii) the amino acid (aa) difference of the complete glycoprotein precursor and the nucleocapsid proteins must exceed 7%; iii) they manifest at least a four-fold difference in two-way cross-
neutralization test; and iv) they cannot naturally reassort with other hantavirus species. The ICTV policy on defining species considers viruses as polythetic entities. Several hantaviruses do not meet all four criteria, yet, they are recognized as distinct species [Plyusnin et al., 2002]. The classification of known hantaviruses, their hosts and geographic distribution is detailed in Table 1.

Genome organization and virion structure of hantaviruses

Hantaviruses possess a tri-segmented single-stranded genome of negative polarity. Each of the three segments has a consensus 3'-terminal nucleotide sequence (3'-AUCAUCAUCUG), which is complementary to the 5'-terminal sequence and distinct from those of the other four genera in the Bunyaviridae. These sequences contribute to forming a panhandle structure through imperfect hydrogen bonding that likely plays a role in replication [Schmaljohn & Dalrymple, 1983] (Figure 1).

Figure 1. Panhandle-forming nucleotides of each genomic RNA segment of PUUV: small (S), medium (M) and large (L). Vertical lines reflex base pairs and colons the non-canonical pairing. The corkscrew model of Flick et al. [2002] is shown for the S viral RNA on the right.

The large (L) segment is approximately 6500 nucleotides (nt) in length and encodes the L protein, which has viral RNA-dependent RNA polymerase (RdRp) function [Schmaljohn, 1990]. The 3'-terminal non-coding region (NCR) is shorter (100 nt in average) than the other two segments. The RdRp presents five conserved aa regions [Poch et al., 1989]; these motifs have proven helpful in the detection of novel lineages. The L segment is typically the most conserved part of the genome; the diversity for more distantly related viruses is up to 40%. The medium (M) segment, approx. 3600 nt in length, encodes the Gn and Gc glycoproteins in a single open reading frame (ORF) [Schmaljohn et al., 1986, 1987]. Its 3'-NCR is 200-250 nt in length. The cleavage site for
two glycoproteins Gn and Gc is highly conserved (WAASA motif), yet glycoproteins are considerably variable (up to 60%). The small (S) segment from 1600 to 2004 nt in length encodes the nucleocapsid (N) protein [Schmaljohn et al., 1986]. The length of the 3′-NCR for this segment varies extensively between different hantaviruses (~220-800 nt). There is a highly variable domain within 700-900 nt region with more conserved flanking regions. N protein sequences diversity is up to 57% (Fig. 2). Cricetid-borne hantaviruses are exceptional in that they contain an evolutionary conserved non-structural (NSs) protein in an overlapping second ORF similar to those of other orthobunyaviruses [Spiropoulou et al., 1994; Plyusnin, 2002]. The NSs protein in arvicoline-borne viruses is 88-95 aa residues long, but shorter (~60 aa) in sigmodontine- and neotomine-borne hantaviruses [Plyusnin & Elliott, 2011]. The expression of NSs inhibits interferon beta (IFN-β), nuclear factor kappa B (NF-κB) and interferon regulatory factor-3 (IRF-3) [Jääskeläinen et al., 2007, 2008]. Each of the viral ribonucleoproteins (RNPs) is encapsidated with the N protein [Dahlberg et al., 1977].

**Figure 2.** Structure and genome organization of hantaviruses. The hantavirus virion (ø 80-120 nm) is enveloped and the surface encompasses a glycoproteins (Gn and Gc) layer. Single-stranded RNA molecules of negative-polarity of Puumala hantavirus are represented on the right side. The Small genome segment encodes the nucleocapsid (N) protein that is structurally associated with viral RNAs. The Medium genome segment encodes the Gn and Gc glycoproteins associated in the viral envelope. The Large genome segment encodes the L protein. [Modified from the Swiss Institute of Bioinformatics].

In nature, hantavirus virions are commonly spherical and vary in size from 80 to 120 nm [McCormick et al., 1982]. Pleiomorphism and size variation have been suggested to result from encapsidation of additional genome segments into the nucleocapsid [Rodriguez et al., 1998]. The two glycoproteins Gn and Gc are embedded in the outer lipid membrane. Ultrastructural studies suggest that the surface has a grid-like pattern and
glycoproteins appear as fuzzy surface projections, approx. 7 nm in length [Martin et al., 1985]. Recently, biochemical studies resolved the hantaviral glycoprotein complex, and showed that it consists of Gn tetramers interconnected with Gc dimers [Hepojoki et al., 2010] (Fig. 3). In 1996, Hutchinson and colleagues estimated the amounts of L, M and S genome segments in infected Vero E6 cells and the plateau values were $3.2 \times 10^8$, $6.5 \times 10^8$ and $1.24 \times 10^9$ RNA copies/0.1 mg RNA, respectively. This equals a L:M:S ratio of 1 : 1.9 : 3.9; the inverse numbers are 1 : 0.52 : 0.26, which is very close to the size ratio of the genome segments (1 : 0.56 : 0.31). This suggests that the viral messenger RNAs (mRNA) are produced in amounts that are inversely proportional to their length and that elongation is the rate-limiting step in mRNA transcription [Hutchinson et al., 1996].

Figure 3. Electron cryotomography of Tula hantavirus. Particles vary in shape from tubular to spherical. Some spiked (white triangles) and naked (arrowheads) areas of the surface are indicated in panel A. Small spherical particles devoid of any RNP density are indicated with asterisks. Straight, rod-shaped densities distinct from the RNP densities are indicated with arrows in panel B. Scale bar, 50 nm. [Huiiskonen and co-workers, 2010; reproduced with permission].

Life cycle of hantaviruses

Hantaviruses infect endothelial and epithelial cells, follicular dendritic cells as well as macrophages and lymphocytes without causing any direct cytopathic effect [Pensiero et al., 1992; Temonen et al., 1993; Zaki et al., 1995]. Surface glycoproteins mediate the attachment to host-cell receptors [Mackow et al., 2001]. Receptors for cellular entry are associated with viral pathogenicity. $\beta_1$ integrins (or the ligand fibronectin) are
implicated in the infection of non-pathogenic hantaviruses while pathogenic species bind to β3 integrins (or the ligand vitronectine) [Gavrilovskaya et al., 1998].

Virus entry occurs via clathrine-mediated endocytosis, in which viruses move from early to late endosomes and/or lysosomes. The conformational changes in the Gc glycoprotein induced by a low pH lead to fusion of the viral membrane with the endosomal membrane. The virus then decapsulates in endolysosomal compartments to liberate the three RNPs into the cytoplasm [Jin et al., 2002].

Following penetration of the host cells, RNPs are transported to the perinuclear region [Ravkov & Compans, 2001] where the viral L protein transcribes negative-sense RNA segments into functional mRNAs of the S, M and L segments. Initiation of viral RNA (vRNA) transcription relies on a “cap-snatching” mechanism, in which an endonuclease encoded in the L segment generates 5’-capped oligonucleotides (10-14 nt long) from host cell mRNA [Bouloy et al., 1978]. Initiation of genomic RNA synthesis of the L protein uses a “prime-and-realign” mechanism. This consists of short-capped primers annealing via single G-C base pairing a few nucleotides upstream of the 3′ terminus of the vRNA template. After a brief elongation, the extended primer is realigned with the viral template so that the 3′-terminal nucleotide of the capped primer is at position -1 [Garcin et al., 1995]. N binds the vRNA panhandle (with higher affinity in trimeric than in monomeric conformation) unwinding it, and remains attached to the 5′ terminus, which leaves the 3′ terminus accessible to the L protein [Mir et al., 2006]. Degraded cellular mRNAs accumulate in cytoplasmic processing bodies (P bodies) that serve as a pool of primers during the initiation of viral mRNA synthesis. The N protein enhances transcription by sequestering 5′ mRNA caps stored in P bodies [Mir et al., 2008].

The N is the first protein to be synthesized and accumulates rapidly following infection [Severson et al., 2001]. Its main role is to protect the vRNA from degradation by nucleases, but it also has several key functions in the life cycle, namely formation of RNPs, encapsidation of genomic (negative-sense) and antigenomic (positive-sense) RNA [Patterson & Kolakofsky, 1984], binding of panhandles during transcription initiation [Mir & Panganiban, 2006], interaction with the ribosomal protein RPS19 to facilitate the loading of the 40S ribosomal subunit onto the virus mRNA [Cheng et al., 2011], and in virus assembly [Ravkov & Compans, 2001]. Furthermore, recent studies suggest that the N protein can modulate the host immune response [Taylor et al., 2009ab]. The L protein acts as an RNA transcriptase, replicase and endonuclease [Reguera et al., 2010].

During vRNA replication, antigenomic RNA segments are used as a template for the production of new RNAs. These are full-length complements of the genomic vRNA and can serve as a template for mRNA or as the genome precursor at later stages. The mechanism by which the virus switches from transcription to replication remains unclear but is thought to involve increasing concentrations of N protein leading to more efficient RNA encapsidation [Johnsson et al., 2001].
Translation of the S and L mRNAs occurs on free ribosomes. The M mRNA is translated on membrane-bound ribosomes to generate the glycoprotein precursor (GPC), which is proteolytically cleaved into Gn and Gc proteins at the conserved WAASA-motif during import into the endoplasic reticulum (ER) [Lober et al., 2001]. Interaction of Gn and Gc proteins in the ER is essential for their vesicular transport to the Golgi complex [Ruusala et al., 1992]. Accumulation of glycoproteins at the Golgi complex is thought to be responsible for viral maturation and subsequent assembly [Spiropoulu, 2001].

Figure 4. Hantavirus life cycle. Glycoproteins integrated in the viral envelope attach to host cell surface receptors. Penetration occurs via receptor-mediated endocytosis and uncoating prior to release of the viral genome. Transcription of complementary RNA (cRNA) from the viral RNA (vRNA) genome uses host-derived primers. Translation of the S and L mRNAs into viral proteins using host machinery occurs on free ribosomes, whereas translation of M mRNA takes place on membrane-bound ribosomes. Replicated and amplified vRNA is transported to the Golgi apparatus to be assembled with the N protein. Assembly of all components occurs at the Golgi apparatus or at the plasma membrane in New World hantaviruses. Egress follows the fusion of Golgi vesicles harbouring mature virion particles with the plasma membrane. [Modified from Johnsson et al., 2010].
The RNP complexes travel to the ER-Golgi intermediate compartments (ERGIC) via dynein microtubules [Ramanathan et al., 2007], and then to the Golgi complex containing the embedded viral Gn and Gc glycoproteins. Unlike many other negative-strand RNA viruses, hantaviruses do not encode a matrix protein. Thus, the cytoplasmic tail of the Gn protein mediates the interaction of the glycoprotein multimers with the N protein during viral assembly [Jonhsson et al., 2001]. The Gn cytoplasmic tail also contains a late domain motif (YXXL) that in other viruses has been shown to interact with cellular factors that facilitate virus budding from cells [Spiropoulu et al., 2003]. Recently, a direct interaction of the Gn cytoplasmic tail with the N protein and RNA has been reported for Old World hantaviruses [Wang et al., 2010; Strandin et al., 2011]. Hantavirus particles are transported in vacuoles to the plasma membrane where they egress by exocytosis [Pettersson & Melin, 1996]. In contrast, electron microscopy has suggested that the cell surface may act as an alternative maturation and budding site for New World hantaviruses [Ravkov et al., 1997; Spiropoulou, 2001]. A schematic of the hantavirus infection cycle is shown in Figure 4.

Transmission of hantaviruses

Hantavirus outbreaks are promoted by environmental factors (e.g., weather and food availability), host population fluctuations, woodland disturbance by humans and/or domestic animals, anthropogenic factors (e.g., deforestation, agricultural development, urbanization, etc.) and ecological changes [Brummer-Korvenkontio et al., 1982; Engelthaler et al., 1999; Buceta et al., 2004; Yan, et al., 2007; Clement et al., 2009; Kallio et al., 2009; Klemppa, 2009; Tersago et al., 2009]. The main factor associated with human epidemics appears to be the host population density, in that the number of cases increases when rodents are abundant. A high host population density increases the number of infected animals. However, although rodent density plays an important role, viral prevalence often follows a seasonal cycle (Fig. 5) [Kallio et al., 2009]. Bank vole population dynamics, and consequently NE epidemiology, differ between biomes in Europe. In temperate forests, masting (heavy seed crops of deciduous trees) increases bank vole densities [Tersago et al., 2009] while in boreal forests population cycles of bank voles are more influenced by predation [Henttonen, 1985; Henttonen et al., 1985].
Hantaviruses infect their hosts chronically. The main route for transmission between rodent/insectivore reservoirs is via aggressive behaviour and/or exposure to aerosolized contaminated droppings (Fig. 6) [Glass et al., 1988; Hinson et al., 2004; Kallio et al., 2006a]. Under laboratory conditions, an infected rodent can transmit the virus horizontally to another rodent within the same cage or through infected bedding [Lee et al., 1981a; Kallio et al., 2006b; Hutchinson et al., 2000]. Chronically infected animals have high levels of neutralizing antibodies (Ab). However, infected rodents excrete virus for a limited time shortly after infection, the duration of which appears to be different for each hantavirus/host species [Lee et al., 1981ab; Hutchinson et al., 2000; Kallio et al., 2006a; Harderstam et al., 2009]. Hantaviruses are not transmitted vertically and maternal Abs acquired through breast-feeding protect the offspring for up to 80 days [Kallio et al., 2006b]. Such acquired immunity plays a role in the infection dynamics of natural populations. Breast-feeding bank voles are abundant in October but when lactation ends, naive individuals (≥15 g of weight) become susceptible to hantavirus infection. This suggests that maternal Abs protection may affect the infection dynamics in bank vole populations [Kallio et al., 2010].
Figure 6. Hantavirus transmission cycle. Predation and weather conditions modulate food and water resources that regulate rodent population densities. These parameters among others influence the prevalence of hantaviruses. Hantaviruses are horizontally transmitted among rodents or insectivores through aggressive behaviour and/or exposure to aerosolized contaminated droppings. Humans are dead-end hosts and infected by breathing aerosolized rodent excreta containing the virus.

Humans become incidental hosts when they are exposed to hantavirus-containing excreta of rodent reservoirs. Typically, humans are infected by inhaling aerosolized urine or faeces released by an infected rodent. Infection may also occur if contaminated material or dust gets into broken skin or onto a mucous membrane. Hantaviruses can also be transmitted via a bite from an infected rodent. Person-to-person transmission has only been demonstrated for ANDV in Argentina [Wells et al., 1997].

Hantaviruses are susceptible to drying conditions, but remain viable for several days when protected by moist organic material. Moreover, hantaviruses can survive up to 14 days at room temperature and probably much longer in cold and moist conditions, such as those experience during the Finnish winter [Kallio et al., 2006a]. Hantaviruses can be chemically inactivated with methanol, paraformaldehyde, 1% sodium hypochlorite, 2% glutaraldehyde, 70% ethanol, acetone, methanol, detergents containing lysis buffer, acids lower than pH 5.0, as well as UV irradiation [Kraus et al., 2005].
Hantaviral infections

Clinical features of hantaviral diseases

The incubation period for HFRS is typically 2-3 weeks long, although a 6-week incubation period has recently been described for PUUV [Kramski et al., 2009]. Classically, HFRS occurs in five indiscrète phases: (i) febrile, (ii) hypotensive, (iii) oliguric, (iv) polyuric and (v) convalescent phase. The initial febrile (i) stage typically lasts 3-6 days and is characterized by fever, chills, facial flushing, headache, photophobia, blurry vision, hypotension and malaise. Haemoconcentration, thrombocytopenia and leukocytosis can be observed in the laboratory. Clinical symptoms of the febrile stage are eventually augmented in the (ii) hypotensive phase (2-3 days long) that includes thirst, restlessness, dizziness, nausea, vomiting and petechial rash. One-third of patients during this stage develop shock and mental confusion. Symptoms of oliguric (iii) phase (<400 ml of urine/day) can last up to 16 days and have been responsible for half of all fatal cases. Manifestations of oliguria are vascular leakage, abdominal pain, backache, hypertension tachycardia, acidosis, proteinuria, elevated creatinine, hematuria, and metabolic disturbances related to renal failure. Persistent oliguria is associated with epistaxis and severe gastrointestinal, genitourinary, retroperitoneal and central nervous system bleeding in about one-third of patients. At this stage, patients are at risk for pulmonary oedema and complications of renal insufficiency. Dialysis is required in the most severe cases. These phases can be hard to distinguish, especially in milder cases. The diuretic phase (iv, polyuric) commences 3-7 days after the oliguric phase and heralds the beginning of clinical recovery. This phase can last from 1-2 weeks and is mainly characterised by polyuria that can lead to severe fluid and electrolyte imbalance. Complete recovery during convalescent phase (v) may take 2-3 three months. Although sequelae are rare for HFRS, anaemia and hyposthenuria may persist for several months in some cases [Lee, 1989]. Fatality rates observed for HFRS patients infected with HTNV varied substantially from 14.2% in 1969 to 5.6% in 1981. Since then, fatalities have substantially fallen to ~1% [Zhang et al., 2009]. The fatality rate for SEOV is less than 1% and rises to 12% for DOBV infections.

In general, the prognosis of NE caused by PUUV is positive. In the clinical manifestations of NE, renal failure predominates and haemorrhagic manifestations are generally lacking. Typically, NE has a sudden onset with high fever, headache, abdominal and back pain and gastrointestinal symptoms. Transient thrombocytopenia is a typical finding in the early phase of the disease. Renal involvement manifests as initial oliguria and later as remarkable polyuria. Many NE cases may be subclinical and it has been estimated that perhaps only 20% of infections lead to acute NE [Vapalahti et al., 2003]. Severe complications occur sporadically with a case-fatality rate of <0.1% [Settergren, 2000].
HCPS is a more severe disease than HFRS with an average case-fatality rate of 30-40% [Hjelle & Torres-Pérez, 2010]. Haemorrhagic manifestations are unusual and most of the pathogenic events occur in the thoracic cavity [Hjelle et al., 1995; Schmaljohn & Hjelle, 1997]. The incubation period varies from 9 to 33 days. The onset of HCPS is abrupt with prodromal symptoms of fever, myalgia, headache, nausea, vomiting, abdominal pain and sometimes dizziness. After 3-6 days, the cardiopulmonary phase leads to coughing, dyspnoea, tachycardia, fever and hypotension. Haemoconcentration, thrombocytopenia, left-shift leukocytosis and proteinuria are frequently observed. A rapid development of pulmonary oedema requires intubation and mechanical ventilation. Severe cardiopulmonary dysfunctions correlate with a poor prognosis.

**Risk factors of acquiring hantaviral diseases**

Several human risk factors have been suggested to promote hantavirus infection: Sex of the patient; the male-female ratio of NE cases varies from 2 to 5:1 [Hjertqvist et al., 2010; Klein et al., 2011]. Age of the patient; case fatality rates were higher among women between the ages of 20-39 and more than 50 years old [Klein et al., 2011]. Damage to the respiratory tract due to smoking [Vapalahti et al., 2010], and genetic predisposition. Besides biological factors, several behavioural and societal components are suspected to be involved.

A genetic predisposition to suffer from a severe course of hantaviral diseases has been described for the adaptive immune response involving the human leukocyte antigen (HLA) system. Patients with haplotype HLA-B*8, -DRB1*03:01, -DRB1*13, -C4A*Q0 or -DQ2 have been observed to have a significantly higher risk for a severe course of NE [Mustonen et al., 1996] and HLA-B*35:01 and -DRB1*14:02 alleles have been associated with a severe courses of HCPS [Kilpatrick et al., 2004]. In contrast, HLA-B*27 and HLA-B*07 haplotypes have been associated with a more benign clinical course of NE [Mustonen et al., 1998; Korva et al., 2011]. HLA-DRB1*15 alleles were significantly more common in a group of patients with a mild course of ANDV [Ferrer et al., 2007].

**Diagnosis of hantaviral infections**

The diagnosis of HFRS and HCPS is based on clinical and epidemiological data. Serological tests for viral antigens (Ag) are used to confirm a hantaviral infection. Immunoglobulin M (IgM) antibodies are detected in the early phase of the disease and may persist in sera for several months. Immunoglobulin G (IgG) antibodies appear slightly later than IgM, remain detectable for life [Settergren et al., 1991], and may provide immunity against secondary infections. Classically, serological tests such as immunofluorescence assay (IFA) [Brummer-Korvenkontio et al., 1980; Settergren et al., 1987; Hedman et al., 1991], enzyme-linked immunosorbent assay (ELISA) [Vapalahti et al., 1996; Kallio-Kokko et al., 1998], or immunoblot assay [Hjelle et al., 1997; Schubert et
Review of the literature: Hantaviral infections

Incidence and seroprevalence of hantaviral infections

Hantaviruses are the etiological agents of two diseases in humans: the New World hantavirus cardiopulmonary syndrome (HCPS) and the Old World haemorrhagic fever with renal syndrome (HFRS) [Schmaljohn & Hjelle, 1997]. The geographic distribution of hantaviruses and epidemiology of their associated infections are linked to the distribution of the rodent hosts. The impact of hantavirus infection on human health worldwide is significant; an average of 200 000 HFRS patients are hospitalized each year throughout the world [Lee, 1996] and approx. 200 cases of HCPS are reported annually in the Americas. Although cases of HCPS are less numerous than HFRS, the average case fatality is higher (30-40%) [Lednicky, 2003].

In Asia, clinical cases of HFRS have been reported mainly in China, the Republic of Korea, and the Far East Federal District of Russia. In the recent years, China has accounted for 70 to 90% of HFRS cases worldwide with 40 000 to 60 000 cases being reported annually [Zhang et al., 2004]. In Eastern Russia, up to 200 HFRS cases are reported each year [Yashina et al., 2000], 3% of all cases in the Federation [Tkachenko et al., 2007]. Serological studies have indicated that there is strong evidence for human hantaviral infections in Thailand, Laos, Viet Nam, Hong Kong, Taiwan, Malaysia, Israel, Philippines, Singapore, Indonesia, Kuwait, Mongolia, India and Sri Lanka [Rollin et al., 1986; Shortridge et al., 1987; Wong et al., 1989; Kao et al., 1996; George et al., 1998; Quelapio et al., 2000; Lam et al., 2001; Pacsa et al., 2002; Groen et al., 2002; Zhang et al., 2009; Chandy et al., 2009; Gamage et al., 2011]. Hantaviral prevalence in rodents has also been measured in Japan, Thailand and Cambodia [Arikawa et al., 1985; Reynes et al., 2003; Chandy et al., 2009].

In Europe, NE was described in Sweden as early as 1934 [Myhrman, 1934; Zeetterholm, 1934] and is the most prevalent hantaviral disease in Western and Central Europe. In 1979, PUUV, the causative agent of NE, was first isolated in Finland [Brummer-Korvenkontio et al., 2001] were widely used and effective diagnostics tools for hantavirus. The most reliable serological methods for the detection of PUUV-specific IgM Abs are based on either native or recombinant full-length N-Ag [Brus Sjölander et al., 1997]. Commercial ELISA methods, in which different hantaviral Ags are used, are available from Progen Diagnostics (Heidelberg, Germany) for PUUV and HTNV, and from Focus Technologies (Cypress, CA, USA) for SEOV and SNV. A commercial immunoblot assay is available from Mikrogen (Martinsried, Germany) for PUUV, HTNV, DOBV and SEOV. A rapid immunochromatographic test is available from Reagena (Erilab Ltd., Kuopio, Finland) for DOBV, HTNV and PUUV. Successful Ag detection by immunohistochemical methods has been reported in biopsies and post-mortem samples from patients with severe HFRS or HCPS [Poljak et al., 1994; Zaki et al., 1995]. Nowadays, the use of RT-PCR for the detection of vRNA in body fluids and tissues of patients is preferred for a rapid differentiation of viral strains.
al., 1980]. By the end of 2006, 35 424 European cases of NE had been reported, 95% of which were after 1990 [Heyman et al., 2008, 2009]. However, mild cases can be misdiagnosed as a flu-like disease, meaning the actual number is likely higher. Nowadays, around 10 000 cases are reported each year in Europe. In Fennoscandia, there is a peak of NE every 3–4 years due to the population cycle of bank voles. During the population peaks of 1999, 2002, 2005 and 2008, the number of cases diagnosed in Finland were 2300, 2603, 2526 and 3500, respectively [National Infectious Disease Registry (http://www3.ktl.fi/)]. Since the introduction of PUUV diagnostics in Finland laboratories, 35 000 cases have been confirmed [Heyman et al., 2011]. It has been estimated that PUUV seroprevalence among Finns is 5% [Brummer-Korvenkontio et al., 1999]. In Sweden, a large NE outbreak occurred in 2006-2007 with an incidence of 313 cases per 100 000 inhabitants in the most affected areas [Evander & Ahlm, 2009]. Prior to 2010, 1234 cases of NE-like disease were reported in Norway, at an average rate of 50 cases per year [Lundkvist et al., 1998]. In European Russia, 10 000 – 12 000 clinical cases of PUUV and DOBV infection are reported annually [Klempa et al., 2008]. In southern Baltic countries, PUUV seroprevalence is 0.5% in Lithuania [Sandmann et al., 2005], 1.5% in Latvia [Lundkvist et al., 2002] and 5.1% in Estonia. Also, a neutralizing Ab specific reaction for SAAV has been observed at a rate of 3.4% in Estonia [Golovljova et al., 2002]. In the Netherlands, Belgium, Luxembourg, France, Germany and Austria, NE outbreaks in 2005 and 2007 were the most significant since 1990, with a total of 1114 and 2106 confirmed cases, respectively [Heyman et al., 2007, 2011]. In the Czech Republic since 1992, antibody seroprevalence against hantaviruses peaked up 1.4% [Pejcoch et al., 2003]. In the Balkans, particularly the former Socialist Federal Republic of Yugoslavia, HFRS outbreaks have been recorded since the early 1950’s with a case fatality rate of up to 10%, where DOBV and PUUV were the predominant pathogens [Lukac et al., 1990]. In Croatia, 317 cases of HFRS were diagnosed during an epidemic in 2002 [Mulić & Ropac, 2002]. Around the same time, a large epidemic occurred in Serbia and Montenegro in 2002 with 128 laboratory-confirmed cases [Papa et al., 2006]. In Greece, 210 HFRS cases have been diagnosed since the first detection in 1984, with a case fatality rate of 9% [Papa et al., 2001]. In Hungary, 342 clinical cases of hantavirus were confirmed between 1992 and 2010 [Heyman et al., 2011]. In Bulgaria during 1954-1986, 399 cases of HFRS were registered with a fatality rate of 15.7% [Chumakov et al., 1998]. Clinical infections of hantaviruses have also been recorded in Romania [Manasia et al., 1977], Albania [Eltari et al., 1978], Slovenia [Avšić-Županc et al., 1992], the United Kingdom [Pether et al., 1993], Denmark [Asikainen et al., 2000; Nemirov et al., 2004], Slovakia [Sibold et al., 2001], Poland [Knap et al., 2006], Switzerland [Fontana-Binard et al., 2008], and Turkey [Ertek & Burgan, 2009]. Sero-epidemiological surveys have demonstrated the presence of hantavirus Ab in humans in the Republic of Moldova [Mikhailenko et al., 1994], Spain [Lledó et al., 2003], Portugal [Vapalahti et al., 2003] and Italy [Kallio-Kokko et al., 2006] although no clinical cases have been reported there.
In the New World, hantavirus infections became a major concern after the first outbreak of acute respiratory distress in the Four Corners area of the United States in 1993 [Nichol et al., 1993]. The newly recognized disease (i.e., HCPS) was linked to newly discovered SNV. After the description of HCPS, numerous clinical cases were also confirmed in Canada, Mexico, Honduras, Costa Rica, Panamá, Colombia, Venezuela, Peru, Brazil, Bolivia, Paraguay, Uruguay, Chile and Argentina [Johnson et al., 1999; Padula et al., 2000]. Even though cases of HCPS in Canada were sporadic, the case fatality rate peaked at 26% [Verity et al., 2000]. CDC-USA confirmed 560 cases of HCPS prior to 2010 in the US, of which 36% were fatal [http://www.cdc.gov/hantavirus/surveillance/index.html]. During 2001-2007, antibody prevalence suggested an incidence of Hantavirus in Panamá of 16.5-60.4% [Armien et al., 2011]. In Venezuela, antibody prevalence was only 1.7% [Ribas et al., 2003]. In Bolivia, 10 HCPS cases were identified through 2002, six of which were fatal, and 36 cases had been reported in the country by the end of 2004. In Uruguay, the first evidence of hantavirus circulation originated from a study of serum specimens collected from blood donors between 1985 and 1987 [Weissembacher et al., 1996]. In Paraguay, 99 cases were recorded prior to 2004 [Carroll et al., 2005]. Human infections have also been reported in Mexico [Suzan et al., 2001], Costa Rica [Hjelle et al., 1995c], Colombia [Máttar et al., 2004] and Peru [Powers et al., 1999]. In Brazil, 855 cases of HCPS were reported between 1993 and 2006 with a 39.3% case fatality [Da Silva, 2007]. In Chile, since the first identification of HCPS in 1995, 352 cases were reported prior to August 2006 [http://epi.minsal.cl/epi/html/boletes/reportes/ Hantavirus/Hantavirus.pdf], with a case fatality rate of 33%. In Argentina, the first case of HCPS caused by ANDV was confirmed in 1995 [Lopez et al., 1996]. In contrast to other hantaviruses, ANDV has been associated with human-to-human transmission [Wells et al., 1997].

In Africa, serological evidence of hantaviral infections exists for humans and rodents in Benin, Burkina Faso, Central African Republic, Gabon [Gonzalez et al., 1984], Madagascar [Rollin et al., 1986]. Later, human hantavirus infections were also demonstrated in Senegal, Nigeria, Djibouti, and Egypt through serological survey [Saluzzo et al., 1985; Tomori et al., 1986; Rodier et al., 1993; Botros et al., 2004]. The first African hantavirus was recently identified in Guinea as SANGV [Klempa et al., 2006]. Since then, two other hantaviruses called TGNV [Arai et al., 2007] and AZGV [Kang et al., 2011] have been described from insectivores. Other hantaviruses are expected to be isolated and identified in the near future.

Prophylaxes for hantaviral diseases

To date, no specific treatment for hantavirus infection is available. General supportive measures and careful monitoring of electrolyte, fluid and acid-base balances during hospitalization are recommended treatments. Diuretics may also be used in patients with marked fluid retention and dialysis may be employed in cases of acute renal failure.
Analgesics, tranquilizers and antibiotics for secondary bacterial infections may also be administered. In the People's Republic of China, intravenous injection of ribavirin in the early stages of the disease was found to reduce the risk of mortality 7-fold [Huggins et al., 1991]. However, ribavirin has failed to show efficacy in the treatment of HCPS, the treatment of which requires hospitalization and intensive care, fluid therapy, administration of vasopressors, antibiotics and close monitoring of oxygenation [Hjelle et al., 1995].

Recently, extensive research has sought to develop a safe vaccine against hantaviruses using different techniques such as recombinant live vaccines, virus-like particles, naked DNA and recombinant proteins. Encouraging results have been obtained from in vivo studies; formalin-inactivated HTNV vaccine is commercially available only in China and Korea. Hantavax® (Korea Green Cross, Seoul, Korea) [Cho & Howard, 1999], a suckling mouse brain-derived vaccine, was shown to induce high levels of IgG-specific Ab in almost 100% of human volunteers after three vaccinations periods, and approx. 80% of test individuals produced neutralizing Ab. However, the Ab titters declined rapidly within months, and Hantavax® elicited protection rates of only between 30 and 50% over longer periods [Cho & Howard, 1999, 2002; Hjelle, 2002]. Although Hantavax® is well tolerated, field trials have not been conducted and a case-control study did not have sufficient statistical power to demonstrate efficacy [Park et al., 2004]. As such, it is currently used in only a few countries. To date, seven distinct hantavirus vaccines have been tested in humans, and five of those are based on inactivated-viruses and the other two use recombinant vaccine virus expressing M and S segments of HTNV or plasmid DNA delivered by a gene gun [Schmaljon, 2009]. Mass vaccination is currently under investigation in China by Professor Dexin Li [Li, 2010].

Prevention and control of hantaviruses

Despite progress in vaccine development, the most effective strategy to control infectious diseases including hantaviruses involves public health education combined with host monitoring and control. Improving general awareness of the pathogen sources, virus transmission, rodent control measures, and general hygiene are the most effective and economic ways to prevent hantaviral diseases. The likelihood of contracting hantaviral diseases can be decreased by attenuating human exposure to viral reservoirs and their shedding by building rodent-proof housing and minimizing their access to shelter, food and water around residential areas. Snap-trapping of rodents is strongly recommended, because fresh aerosols are easily spread from live traps. Rodenticides can also be used as a preventative measure to eliminate trap-shy rodents, but local legislation may prevent the use of such agents. Rodent droppings should not be broom or vacuumed, because such actions increase the chance of aerosol inhalation. Faecal pellets should first be saturated in 10% bleach and then collected with paper towels and double bagged prior to incineration or disposal as household waste [Boren & Valdez, 2007]. Monitoring of
hantaviral loads in rodents provides an early warning system to limit human infections. This also provides epidemiological data of hantaviruses that can be used to study outbreaks and their associated factors.

**Evolution of hantaviruses**

The main mechanisms of hantavirus evolution are genetic drift and shift. Genetic drift in hantaviruses applies to the continuous accumulation of nucleotide substitutions, in addition to small insertions and/or deletions most commonly occurring in the NCR [Plyusnin et al., 1995; Rowe et al., 1995; Lundkvist et al., 1998; Razzauti et al., 2008]. Genetic shift in hantaviruses occurs essentially via reassortment of the genome segments [Li et al., 1995; Henderson et al., 1995; Rodriguez et al., 1998; Kang et al., 2002; Klempa et al., 2003; Rivanov et al., 2004; Chu et al., 2006; Zou et al., 2008; Lin et al., 2008; Razzauti et al., 2008, 2009; Black et al., 2009; Handke et al., 2010; Kirsanovs et al., 2010], although evidence for recombination has also been described [Sibold et al., 1999; Sironen et al., 2001; Plyusnin et al., 2002; Chare et al., 2003; Klempa et al., 2003].

Hantaviruses, being RNA-based, face considerably higher replication error rates than DNA viruses, which leads to the formation of viral quasispecies [Plyusnin et al., 1995, 1996c; Lundkvist et al., 1997; Feuer et al., 1999; Nemirov et al., 2003; Chung et al., 2007; Sironen et al., 2008]. The intrinsic error rate, or fidelity, of the replicase protein determines the mutation rate and the range of genetic variation over which selection can operate. Based on the assumption of ancient adaptation and codivergence with their respective hosts, the PUUV molecular evolution rate was estimated to be approximately $10^{-7}$ nucleotide substitutions per site per year (substitutions/site/year) [Hughes & Friedman, 2000; Sironen et al., 2001]. Recently, a Bayesian coalescent method was used to estimate the mutation rates of ARAV, DOBV, PUUV and TULV. These viruses infect three rodent subfamilies: Sigmodontinae, Murinae and Arvicolinae. Controversially, results indicated that substitution rates exhibit $10^{-3}$ to $10^{-4}$ substitutions/site/year, which is comparable to other RNA viruses [Ramsdem et al., 2008]. This range implies that most genomes within the viral population will differ by one or more nucleotides from the relative consensus sequence representative of the entire population [Eigen & Schuster, 1979]. Genomic heterogeneity is thought to allow the population to adapt rapidly to environmental changes. Since RNA viruses live at the edge of maximal variability, an increase in the mutation rate is likely to force the virus beyond the tolerable mutation frequency into “error catastrophe” [Domingo & Holland, 1994; Domingo & Holland, 1997; Domingo, 2000]. In a constant environment where a steady state is reached, the viral pool is presented by a dominant mutant with related but often extremely heterogeneous
genomes [Domingo & Holland, 1988]. This has been demonstrated to occur for TULV at the host level [Plyusnin et al., 1996c; Sironen et al., 2008].

The high ratio of synonymous vs. non-synonymous substitutions (dS/dN) suggests a neutral mode of evolution in hantaviruses [Gojobori et al., 1990]. Furthermore, mechanisms of purifying selection are plausible since high heterogeneity at the nt level contrasts with low variation at the aa level [Hjelle et al., 1995b; Sironen et al., 2001]. Most probably, the majority of non-synonymous substitutions are deleterious and only the aa changes that do not adversely affect the hantaviral fitness may be retained [Kimura, 1983]. Directional selection was shown to play a role to fixation of nt substitutions in the NCR of the S segment in PUUV adapted to cell culture [Lundkvist et al., 1997].

Genetic heterogeneity is comparable among hantaviral species found in different rodent subfamilies. Nucleotide variability ranges from 30-40% over all genomic segments, while variation among deduced aa sequence differs for each segment: 15-40% for the N nucleocapsid, 20-50% for the Gn/Gc glycoproteins and 10-30% for the L protein [Plyusnin et al., 1996b]. Nt diversity among hantavirus species is variable and probably influenced by the number of strains recognized. Diversity of PUUV between lineages may be as high as 38% for the S segment [our unpublished data], but sequence heterogeneity along the genome is unevenly distributed. The hypervariable region of the S segment (233-275 aa) expresses epitopes recognized by both monoclonal antibodies and human patient sera [Lundkvist et al., 1995; Vapalahti et al., 1995]. Furthermore, NCRs are greatly variable due to an abundance of indels and aligning such regions of several hantaviral strains can be difficult, especially the S segment. For this reason, NCRs are routinely excluded from phylogenetic analysis of distantly related hantaviruses. At finer scales, they have been used to help clarify relationships among lineages within a species [Plyusnin, 2002].

Prolonged shedding periods and long survival of hantaviruses outside their host enable co-infection of naïve hosts. This may lead to segment reassortment and/or intra-segment recombination between viral variants co-infecting and individual cell. RNA viruses escape Muller’s ratchet (i.e., the accumulation of deleterious mutations in the genome of an asexual population) using such mechanisms [Chao et al., 1997]. Reassortment provides the means for large changes across the fitness landscape and explain rapid adaptation to new ecological niches. Figure 7 shows a schematic reassortment for a bunyavirid-like virus and the possible combination of segments in the offspring.

Evidence for reassortment can be observed from significant incongruence among phylogenies inferred for each of the genomic segments separately. In hantaviruses, reassortment was first observed between genetic lineages of SNV circulating in California [Li et al., 1995]. Since then, many other cases of natural reassortment have been reported for distinct virus species within the same family of rodent hosts [Klempa et al., 2003; Chu et al., 2006], for hantaviruses of distinct genetic lineages [Hendersson et al., 1995; Zou et
of negative-strand RNA viruses was first described for hantaviruses. Analysis of TULV genetic variants from Slovakia revealed at least two recombination points in the S segment (nt 400-415 and around 1200). Slovakian strains of TULV were shown to have a mosaic of sequences closely related to those of either Russian or Czech-Slovak lineages [Sibold et al., 1999]. Although in vitro generation of a recombinant TULV corroborated these data [Plyusnin et al., 2002], and evidence for recombination in PUUV [Sironen et al., 2001], HTNV [Chare et al., 2003] and ANDV [Medina et al., 2009] has been
described. Nevertheless, recombination seems to be a rare event in hantavirus evolution. Similar to reassortment, recombination can be detected using phylogenetic analysis; particularly by conflicting phylogenies based on different regions of a genome segment. A mosaic structure of sequences supports the hypothesis of homologous recombination between genotypes.

**Hantaviruses and their hosts**

To great extend hantaviruses infect a single host species. Based on strong host specificity and similar phylogenies largely irrespective of their present geographical distribution, a coevolution of hosts and their hantaviruses has been suggested [Plyusnin et al., 1996b; Hughes & Friedman, 2000; Plyusnin & Morzunov, 2001; Kang et al., 2009]. However, an alternative hypothesis recently suggested that host switching and local species-specific adaptation account for similar phylogenies of hosts and their hantaviruses [Ramsden et al., 2008; Plyusnina et al., 2008; Ramsden et al., 2009; Nemirov et al., 2010; Kang et al., 2011]. Some sympatric rodent species occasionally serve as reservoirs for the same hantavirus and host switching has clearly occurred during the evolution of hantaviruses [Plyusnin et al., 1996a; Vapalahti et al., 1999; Klingström et al., 2002; Nemirov et al., 2003; Schlegel et al., 2009; Schmidt-Chanasit et al., 2010]. While spillover infections are generally assumed to be irrelevant in such host populations, infections in spillover or secondary species can vary. For example, in certain cases antibodies are produced but the virus is never found, or both antibodies and vRNA are found but shedding is questionable, or both antibodies and vRNA are found and shedding occurs at low quantity compared to that in the main host [Verhagen et al., 1987; Klingström et al., 2002].

As introduced in the Taxonomy section (Table 1) hantaviruses are traditionally classified according to the taxonomy of their hosts. Accordingly, hantaviruses have been described in two mammalian orders, and only in two families of each order: Soricomorpha (Talpidae and Soricidae) and Rodentia (Muridae and Cricetidae). Generally, each host family carries its “endemic” hantavirus clade with the exception of talpids, which host extrinsic hantaviruses [Kang et al., 2011] (Fig. 8 and 9).

Cricetid rodents carrying hantaviruses belong to subfamilies Arvicolinae, Neotominae and Sigmodontinae, and each is infected by a different clade of hantaviruses. Arvicoline-borne hantaviruses have been found in *Myodes, Microtus, Eothenomys* and *Lemmus*, habitual in Eurasia and in North-America (Table 1). Three clear associations are observed for this subfamily: *Myodes*-(PUUV, HOKV, MUJV), *Microtus/Eothenomys*- (TULV, PHV, ISLAV, BLLV, LUXV) and *Microtus/Lemmus*-borne viruses (KHAV, YUJV, VLAV, TOPV) (Fig. 8).
Figure 8. Host phylogeny according to mitochondrial gene cytochrome b, annotated with the hantaviruses infecting each species. The tree topology, its support and branch lengths were estimated using the GTR model in MaxML [Stamatakis et al., 2008]. A rooted maximum clade credibility tree is shown with mean branch lengths (substitutions per site) and non-parametric bootstrap percentage for each node. See Table 1 for the abbreviation of hantaviruses.

Neotomine-borne hantaviruses are found in *Peromyscus* and *Reithrodontomys*, both of which are endemic to North and Central America. *Peromyscus*-carried (SNV, NYV, MGLV, BRV, MTNV) and *Reithrodontomys*-carried (RIOSV, ELMCV, LSCV, CARV, HUIV) viruses cluster according the rodent host genera with one exception; MTNV is closely related to *Reithrodontomys*-carried viruses. Sigmodontine-borne hantaviruses are found in *Sigmodon* (CADV, BCCV, MULV, OROV), *Oryzomys* (BAYV, CATV, OROV), *Oligoryzomys* (RIOMV, ANDV, BMJV, LECV, CHOV, MAPV, ORNV, CASV, ITPV, JUQV, ANAJV, ARAUV, Neembucu), *Oxymycterus* (ARAUV), *Akodon* (JABV, AAIV, ARAUV, PRGV), *Bolomys* (MCLV, ARAV), *Calomys* (LANV), *Zygodontomys* (CALV), and *Holochilus* (RIEMV, ALPV) genera, mostly endemic to South America with a few exceptions from southern parts of North America. Hantaviruses in the Neotominae subfamily are not congruent with rodent host genera (Fig. 9).
Figure 9. Hantavirus phylogeny based on the S-segment coding region. The tree topology, its supports and branch lengths were estimated using the GTR model in MaxML [Stamatakis et al., 2008]. A maximum clade credibility tree with an arbitrary root is shown with mean branch lengths (substitutions per site), and non-parametric bootstrap percentages are shown at each node. See Table 1 for the abbreviation of hantaviruses.

Hantaviruses in the Muridae have only been found in the Murinae and recovered from Apodemus, Rattus, Bandicota, Hylomyscus and Niviventer. These are endemic to Eurasia and Africa, although Rattus is now prevalent as an introduced species around the world. These hantaviruses are divided into two clades, one is strictly composed of viruses carried by rats (THAIV, SEOV, SERV, GOUV) and the other by mice (HTNV, ASV, DBSV, DOBV, SAAV) (Fig. 9). Cricetid- and murid-borne hantaviruses are well distinguished in that cricetid-borne viruses contain an insertion of 5 aa between nt 784-798 of the S segment ORF.
Hantaviruses are also present in Talpidae and Soricidae of the Soricomorpha. Representatives of this order have a worldwide distribution. Soricid-borne viruses have been recovered from Soricinae and Crocidurinae. Soricine-borne hantaviruses have been found in *Sorex* (SWSV, ALTV, ARRV, JMSV, ARTV, LNAV, ARRV, KKNV, FXCV, PWBV, TLNV, Qiandoe Lake virus), *Blarina* (RPLV, IAMV) and *Anourosorex* (CBNV), and crocidurine-borne hantaviruses in *Suncus* (TPMV) and *Crocidura* (MJNV, TGNV, AZGV). Talpid-borne viruses are in *Talpa* (NVAV), *Urotrichus* (ASAV), *Neurotrichus* (OXBV) and *Scalopus* (RKPV). Neither talpid- nor soricid-borne viruses show cluster according to their hosts. Soricomorph-borne hantaviruses assort in several phylogroups of random distribution (Fig. 9).

**Puumala virus and its host**

*Puumala virus* (PUUV), a major European rodent-borne pathogen, is the etiological agent of NE, a relatively mild form of HFRS [Vapalahti et al., 2003]. It was first discovered in Finland in 1979 [Brummer-Korvenkontio et al., 1980] and has since been detected or implicated in 26 European countries [Olsson et al., 2010].

The host of PUUV is the bank vole, *Myodes glareolus* (formerly *Clethrionomys glareolus*) (Fig. 10), which belongs to the family Cricetidae, subfamily Arvicolinae [Wilson & Reeder, 2005]. The bank vole has a wide distribution from the British Isles through continental Europe and Russia to Lake Baikal. In the north, its range extends beyond the Arctic Circle (up to latitude 68N), and in the south it reaches northern parts of Spain, Greece, Turkey and Kazakhstan (down to latitude 40N). It is widespread in Europe, although it is not found in southern Iberia or the Mediterranean islands, and it occurs from sea level to 2400 m [IUCN, 2011].

The bank vole has a small (12-40 g), stocky body (8-12 cm) and a blunt, rounded snout. The fur is reddish-brown above and creamy-grey below, the flanks are greyish and the rump is whitish-grey, and ears covered with fur. The tail (3-6 cm) is bicoloured (black above, white below) usually slightly bushy at the tip [Spitzenberger, 1999] (Fig. 10). The bank vole typically inhabits all types of forest, densely vegetated clearings, hedgerows, and the banks of woodland rivers and creeks [IUCN, 2011]. Burrows are lined with moss and vegetable fibres and are used for nesting and food storage. The species is omnivorous, eating insects, leaves, seeds, nuts, berries and other fruits. It is predominately diurnal with activity peaks at dawn and dusk [Greenwood, 1978]. In Finland, the breeding season occurs between April and September in most years. Females can rear up to four litters per year, each with three to six pups. Early offspring may mature and breed in the first year, whereas those born later overwinter as sub-adults and begin breeding the next year. Bank voles do not hibernate but replace a dense winter fur with a lighter coat in the spring. Maximum life span is over a year but the average individual lives only a few months. Thus, population turnover is rapid.
Bank vole population density fluctuates seasonally and over the course of several years. Patterns of multiannual variations are biome-specific and modulated by a combination of endo- and exogenous factors such as high reproductive potential, annual weather conditions, masting, density-dependent intraspecific competition, community diversity, and predation pressure. Voles in northern Europe exhibit cyclical population dynamics with a latitudinal gradient in cycle length, amplitude and interspecific synchrony [Hansson & Henttonen, 1985, 1988; Hanski et al., 1991; Sundell et al., 2004]. The length of the bank vole population cycle in Fennoscandia is 3-4 years and its period varies from north to south [Hansson & Henttonen, 1985; Hanski et al., 1991]. In temperate Europe, masting generally coincides with bank vole peaks and NE epidemics [Tersago et al., 2009].

Occurrence of NE in humans depends strongly on local bank vole population dynamics and outbreaks typically follow the population peaks [Brummer-Korvenkontio et al., 1982; Olsson, 2003; Kallio et al., 2009; Tersago et al., 2009] (Fig. 5). Persistence of PUUV within a bank vole population is influenced by the local environment [Kallio et al., 2009; Olsson, 2003], and age structure and stage of the breeding season (maternal immunity may modulate PUUV infection dynamics) [Kallio et al., 2010].

Genetic factors may regulate the chronic infections caused by PUUV in bank voles. Tumor necrosis factor (TNF) can limit parasite infection by regulating host tolerance or resistance. A recent study suggested that independent evolutionary histories of PUUV and its hosts account for different tolerance of or resistance to PUUV, some of which is mediated by the variable production of TNF [Guivier et al., 2010a]. Furthermore, several alleles of the class II major histocompatibility complex (MHC) gene may play a role in resistance to PUUV [Guivier et al., 2010b].
Geographic distribution of Puumala virus and the post-glacial history of bank voles

Genetic variation of PUUV exhibits some geographical structure within the bank vole distribution [Plyusnin et al., 1994b; Hörling et al., 1996; Heiske et al., 1999; Asikainen et al., 2000; Escutenaire et al., 2001; Sironen et al., 2001; Plyusnina et al., 2006; Johansson et al., 2008; Razzauti, et al. 2009; Nemirov, et al. 2010; Razzauti et al., 2012]. To date, eight PUUV lineages have been described in Eurasia: (1) the Central European (CE) lineage, which includes strains from France, Belgium, Germany, and Slovakia; (2) the Alpe-Adrian (ALAD) lineage that occurs in Austria, Slovenia, Croatia and Hungary; (3) the Latvian (LAT) lineage observed in the vicinity of Jelgava town; (4) the Danish (DAN) lineage from the island of Fyn in Denmark; (5) the South-Scandinavian (S-SCA) lineage including strains from Norway and central and southern Sweden; (6) the North-Scandinavian (N-SCA) lineage in northern Sweden and parts of northwestern Finnish Lapland; (7) the Finnish (FIN) lineage found in Finland, Russian Karelia and western Siberia (Omsk strains); and (8) the Russian (RUS) lineage with strains from pre-Ural Russia, Estonia and Latvia. It is believed that the current distribution of distinct PUUV lineages is based on the isolation of rodent populations during the Last Glacial Maximum (LGM) and subsequent recolonization of Eurasia (Fig. 11).

**Figure 11.** Distribution of the eight described Puumala virus lineages: Finnish (FIN), North-Scandinavian (N-SCA), South-Scandinavian (S-SCA), Russian (RUS), Latvian (LAT), Danish (DAN), European (CE) and Alpe-Adrian (ALAD).
During the LGM (ca. 0.25-0.013 Ma), ice sheets covered much of North America, northern Europe and Asia. This period was followed by substantial climate warming that brought about their accelerating retreat north (ca. 13-10 ka). During this period, animal populations previously confined within glacial refugia began to repopulate the newly deglacied landscape [Grosswald, 1993, 1998; Taberlet et al., 1998].

Phylogeographic patterns of voles are consistent with colonization of boreal regions from southern refugia following the LGM [Runck & Cook, 2005]. Yet, locations of all refugia that contributed to the modern vole population remain uncertain. In Eurasia, possible refugia sites for the bank vole include the Iberian Peninsula, Apennines, Balkans, Carpathians, Ukraine and the Ural mountains (Fig. 12) [Taberlet et al., 1998; Hewitt, 1999; Kotlík et al., 2006; Wójcik et al., 2010; Jaarola & Searle, 2002]. After the late Weichselian (ca. 25000-15000 years ago) [Svendsen et al., 2004], bank vole populations expanded from their refugia throughout Europe. Phylogeographic analyses using bank vole mitochondrial DNA (mtDNA) indicate that eight mt clades currently occur, referred to as Basque, Spanish, Italian, Balkan, Carpathian, Western, Eastern and Ural [Deffontaine et al., 2005; Kotlík et al., 2006; Deffontaine et al., 2009; Wójcik et al., 2010] (Fig. 12).

Figure 12. Distribution of mitochondrial DNA (mtDNA) clades throughout the bank vole range (orange shadow). The approximate ice sheet limit during the Last Glacial Maximum is delineated in light blue, and assumed glacial refugia for small mammals are shown as dashed circles. Modified from Wójcik et al. [2010].
Range variation during Pleistocene glacial cycles explains patterns of intraspecific genetic variation among many European species [Hewitt, 1996, 1999]. Molecular clock analyses estimate that the major European bank vole phylogroups differentiated during the late Pleistocene and thus preceded the LGM [Deffontaine et al., 2005], and palaeontological data confirm that bank voles were present in Europe as early as the middle Pleistocene (1.2 Ma) [Bauchau & Chaline, 1987]. Fragmentation of bank vole populations during LGM likely spurred differentiation of the phylogroups [Deffontaine et al., 2005]. High endemism of the Mediterranean phylogroups suggests that these did not contribute to the postglacial recolonization of central and northern Europe [Bilton et al., 1998], rather, this region was recolonized by phylogroups originating in Carpathian, Ukraine and Ural refugia. The western phylogroup likely arose via expansion from the central European refugia. There is clear evidence that the Ural phylogroup originated from the introgression of red vole (Myodes rutilus) mtDNA into bank voles during their postglacial recolonization of Fennoscandia (ca. 30-60 ka) [Tegelström, 1987; Deffontaine-Deurbroeck, 2008; Abramson et al., 2009].

Several contact zones of bank vole phylogroups have been described. One is located in northern Sweden and Norway, which is a well-known suture zone for many other species. The width of this suture zone appears to be ca. 30-60 km [Tegelström & Jaarola, 1998]. The other contact zone in Fennoscandia is ca. 50 km wide and extends across north-central Finland between the towns of Oulainen and Kuhmo [Deffontaine-Deurbroeck, 2008; Boratynski et al., 2011]. Another suture zone of several plant and animal species has been described in Lithuania and western Poland, where distinct mtDNA lineages of the field vole (Microtus agrestis) [Jaarola et al., 2002], the sibling vole (Microtus levis, formerly rossiaemeridionalis) [Jaarola & Searle, 2002], the root vole (Microtus oeconomus) [Brunhoff et al., 2003], and the bank vole (Myodes glareolus) [Deffontaine et al., 2005] can all be detected. The Carpathian phylogroup occurs in western Lithuania and the Eastern phylogroup in eastern Lithuania [Deffontaine et al., 2005]. In contrast to Fennoscandia, the contact zone for the Western and the Eastern phylogroups is very wide, ranging up to 400 km [Deffontaine et al., 2005].

Known PUUV lineages show some geographical clustering according to the distribution of each bank vole phylogroup (Fig. 11). Yet, it seems that during recolonization some PUUV lineages transgressed from one host clade to another, and nowadays two PUUV lineages are hosted by the same phylogroup; e.g., the FIN and N-SCA PUUV lineages co-circulate in the Ural bank vole phylogroup at Pallasjärvi in northern Finland [Razzauti et al., 2009], and the RUS and the LAT lineages have been found in the Carpathian phylogroup in Jelgava, western Latvia [Razzauti et al., 2012]. Details concerning these results are provided later.
AIMS OF THE STUDY

Zoonotic viruses are major contributors to emerging diseases in humans. Hantaviruses are the etiological agents of two significant human diseases around the globe, HCPS and HFRS. Understanding the mechanisms and dynamics of human hantavirus epidemics is crucial to control and respond to outbreaks of these diseases.

Since the discovery of Hantaan virus, studies have focused on the ecology of host species, which affects virus transmission, structure, life cycle, infection process, disease course, genetics and evolution of hantaviruses. Results have led to the development of diagnostics and the first vaccine formulation for these viruses.

This study aimed to improve our understanding on microevolution of Puumala virus (PUUV) in relation to bank vole population dynamics.

Specific aims were the following:

- elucidation of the mechanism(s) of PUUV microevolution
- analysis of PUUV genetic diversity in its host population by examining nucleotide mutations and amino acid substitutions
- analysis of viral genes for reassortment and recombination events
- monitoring of PUUV variants throughout the density cycle of a bank vole population
MATERIALS AND METHODS

Rodent trapping

Bank voles reported in article I and manuscript IV were trapped at Konnevesi, Central Finland (62°34'N, 26°24'E) during May and October of each year from 2005 to 2009. Trapping was performed at 38 sites within 120 km² of typical taiga forest, dominated by Norway spruce (*Picea abies*), Scots pine (*Pinus sylvestris*), downy birch (*Betula pubescens*) and silver birch (*Betula pendula*). Of the 38 trapping sites, 14 (alphabetically named in Fig. 13) were sampled using live-trapping techniques throughout the entire study period. Live trapping consisted of grids of 3x3 Ugglan Special live traps (Grahnab, Hillerstorp, Sweden) set 15 m apart. Traps were baited with oat seeds for three nights and checked twice per day to minimize stress to the animals. Occupied traps were replaced with empty ones and immediately taken to the lab. From May 2007 to October 2009, 24 additional trapping sites (numerically named in Fig. 13) were sampled using snap traps. Snap trapping consisted of 4 transects of 15 standard snap traps set at 15 m intervals over two nights. Sites were selected to cover bank vole habitats and traps baited with rye bread were placed in areas with fresh rodent signs (runways, burrows, fresh pellets or grass clippings). Traps were checked each morning for two days. All trapping sites were situated 500 to 1000 m apart from each other. Live-trapped bank voles were bled, then sacrificed via cervical dislocation and frozen immediately on dry ice. Similarly, snap-trapped bank voles were flash frozen.

Rodents reported in article II were trapped at Pallasjärvi in the Pallas-Ylläs National Park in western Finnish Lapland (68°30'N, 24°09'E) during September 1998. Rodents were live-trapped, and euthanized sacrificed via cervical dislocation and necropsy was performed immediately. Tissue samples were stored in liquid nitrogen until subsequent analysis.

Rodents reported in article III were snap trapped in 33 sites around the Latvian towns of Jelgava (56°34'N, 26°31'E) and Madona (56°48'N, 26°42'E). Trapping sessions took place during the spring and autumn of 2008 and early winter of 2009. Animals were flash frozen in dry ice and later processed in the laboratory.
**Figure 13.** Map of Konnevesi region. Trapping sites are shown in boxes; alphabetically-named trapping sites (Kv-A to Kv-P) were sampled throughout the study period from May 2005 to October 2009. Numerically-named trapping sites (010 to 036) were only sampled from May 2007 to October 2009. The red square on the insert shows Konnevesi location (62°34’N, 26°24’E) within Finland.
Materials and methods

Processing of rodent samples

In the laboratory, voles were weighed and their sex and age was determined. Blood samples of approx. 200 µl were taken from the retro-orbital sinus of live-trapped voles using ammonium heparinised glass capillary tubes (Hirschmann Laborgeräte, Eberstadt, Germany) of 75 µl capacity. Animals were then euthanized using cervical dislocation and immediately flash-frozen on dry ice prior to storage at −80 °C. Snap-trapped bank voles were similar flash-frozen and stored until necropsy could be performed in full.

Frozen voles were later dissected in the rodent laboratory at the Finnish Forest Research Institute. At dissection, animal annotation was completed (trapping site, habitat, species, weight, sex, maturity, age) and lung, heart, kidney, liver, spleen, brain, tongue together with salivary glands and tail were sampled individually and placed into deep-frozen storage. Tissue samples were further processed in a class II laminar flow biosafety hood in a biosafety level 3 laboratory at the Department of Virology, Haartman Institute, University of Helsinki.

The age of voles was estimated on the basis of fur and molt patterns, or root length was determined of the lower first molar (M\textsubscript{1}) for a more precise age. For Myodes glareolus, M\textsubscript{1} roots do not appear until they are two months old after which they grow at a rate of 0.18 mm per month [Lowe, 1971; Abe, 1973]. This is a reliable method to distinguish between breeding voles that are old (i.e., overwintered) from voles born in the spring and late summer (i.e., summer-borne).

Lung samples from each animal were divided into two subsamples, one lung sample (approx. 20 mg) was placed in 500 µl of Laemmli sample buffer and the second (approx. 200 mg) in 1000 µl Tri-Pure Isolation reagent (Roche Diagnostics, Espoo, Finland) to be used in immunoblotting and reverse transcription – polymerase chain reaction (RT-PCR), respectively. Both lung samples were stored at −80 °C prior to analysis.

Immunoblotting

Rodent lung tissue samples were screened for the presence of PUUV N-Ag using western blot (WB) as described earlier [Plyusnin et al., 1995] with some modifications. Briefly, lung tissue samples (approx. 20 mg) in 500 µl of Laemmli buffer were incubated at room temperature overnight, then homogenized by sonication and heated to 80-100 °C for 5 minutes. Aliquots of 25 µl were separated by electrophoresis in 10% sodium dodecyl sulphate polyacrylamide gels and proteins were subsequently transferred to nitrocellulose membranes. To verify the blotting efficacy, membranes were stained with 1x Ponceau S staining solution for one minute and destained in distilled water until the background was clean. Membranes were blocked via overnight immersion in 1% bovine serum albumin solution at +4 °C. Membranes were incubated with rabbit polyclonal antiseraum made
Materials and methods

against rN-Ag for two hours at room temperature, washed with 0.1% Ten-Tween20 solution, and incubated for one hour at room temperature with Odyssey IRDye 800CW goat anti-rabbit secondary Ab (LI-COR), diluted at 1:10000 in PBS. The Odyssey Infrared Imaging system was used to visualize the proteins.

RNA/DNA extraction

Lung tissues used in article I and II were homogenized by hand in sterile mortars under liquid nitrogen. Tissue samples used in article III and manuscript IV were automatically homogenized with the MagNA Lyser Instrument (Roche Applied Science). Viral RNA and rodent DNA used in article III were obtained by a phenol-chloroform method using TriPure Isolation reagent (Roche Diagnostics, Espoo, Finland) or TRIsure reagent (Bioline, UK) according to the manufacturer's instructions.

Recovery of viral genome sequences

RT was performed with the SuperScript II reverse transcriptase (Invitrogen/Gibco-BRL, Burlington, Canada) and RevertAid™ H Minus M-MuLV Reverse Transcriptase (Fermentas, Lithuania) as specified in the manufacturers instructions. AmpliTaq® DNA polymerase (Applied Biosystems, Foster City, CA, USA) was used to amplify viral cDNA via PCR. PUUV S-, M- and L-segment sequences were amplified using different combinations of specific oligonucleotide primers.

Complete S segment PCR products were TA-ligated into the pGEM®-T Vector System (Promega, USA) overnight at +4 ºC. A ligation cocktail was used to transform JM-109 Escherichia coli competent cells (Promega, USA), and cell transformation was provoked by water-bath sonication of the previously cooled competent cells with the ligation product. pGEM®-T vector uses the activity of β-galactosidase recovered with lacZα peptide of the vector DNA and lacZΔM15 to select successfully transformed cells. Therefore, transformation products were sown on bacterial plates containing x-gal and IPTG for screening. Plasmids were extracted from white transformed cells and exposed to restriction digest. Plasmids that contained the insert of appropriate length were mixed with the vector-based commercial primers M13F and M13R (Promega, USA) and sequenced by ABI PRISM™ Dye Terminator sequencing kit (Perkin Elmer/ABI).

Multi-band PCR products were separated by electrophoresis in a low-melting-point agarose gel (peqGOLD Low Melt Agarose, peqLab, Erlangen, Germany), excised and purified with the QIAquick Gel Extraction kit (Qiagen, Helsinki, Finland). Alternatively, PCR amplicons were purified with ExoSAP-IT™ PCR clean-up reagent (USB Corporation, Miles Road, Cleveland, USA). Automated sequencing was performed using the ABI
PRISM™ Dye Terminator sequencing kit (Perkin Elmer/ABI). Newly recovered S-, M- and L-segment sequences were deposited in GenBank, accession numbers are detailed in each article.

Mitochondrial DNA (mtDNA) amplification

Total DNA from rodent tissue samples (spleen or piece of tail) was purified with TriPure Isolation reagent (Roche Diagnostics, Espoo, Finland) according to the manufacturers recommendations. Alternatively, the DNeasy Blood & Tissue Kit (Qiagen, Helsinki, Finland) was used. The cytochrome b (cyt b) gene and the D-loop region of mtDNA were amplified as described in Michaux et al. [2003] and Morzunov et al. [1998], respectively.

Genetic variation and phylogenetic analyses

Genetic distances were estimated by calculating pairwise sequence identities using BioEdit v7.0.9 [Hall, 1999]. To estimate DNA sequence variation within and between viral populations, DnaSP [Librado et al., 2009] was used.

Nucleotide sequence alignments were generated using ClustalX [Thompson et al., 1997] and/or Muscle [Edgar, 2010] and inspected visually. Sequence alignments were further analyzed using BioEdit v7.0.9 [Hall, 1999] and MEGA 5 [Tamura et al., 2011].

The most appropriate model of sequence evolution for each alignment was determined according to the Akaike Information Criterion in jModelTest [Posada, 2008]. Phylogenies were constructed with neighbour-joining (NJ) or Fitch-Margoliash (FM) distant methods in PHYLIP [Felsenstein, 1993] and distances were calculated with dnadist according to the F84 substitution model. Maximum-likelihood (ML) phylogenies were estimated with PHYLIP, RAxML III [Stamatakis et al., 2008] and PhyML [Dereeper et al., 2010] with default heuristic search algorithms. Maximum-Parsimony (MP) trees were inferred with PHYLIP. Bayesian inference of phylogeny was performed with MrBayes [Ronquist & Huelsenbeck, 2003] and BEAST [Drummond & Rambaut, 2007] to estimate lineage divergence times assuming relaxed molecular clock. Phylogenies were visualized with TreeView v3.4.0 [Page, 1996.] and FigTree v1.3.1 [Rambaut, 2010].

Sequence alignments were submitted to phylogenetic Network 4.600 (Fluxus-Engineering) to generate genetic and evolutionary relationships among the segment genotypes using the Median-Joining (MJ) algorithm [Bandelt et al., 1999].

Finally, recombination of sequences was inferred with the recombination detection program (RDP) [Martin & Rybicki, 2000] and SimPlot v3.5.1 [Lole et al., 1999].
RESULTS AND DISCUSSION

Microevolution of PUUV in relation to host population dynamics

The primary objective of this study was to gain further insight into the mechanisms of PUUV microevolution in relation to host population dynamics. Genetic variants of the virus were studied in a local bank vole population at Konnevesi in Central Finland. This region is highly endemic for PUUV. Long-term studies of PUUV in rodents have been performed at the Konnevesi Research Station of the University of Jyväskylä since the 1990’s. Thus, this study continued this program with biannual (May and October) rodent sampling during 2005-2009.

The first part of the study [article I] included a genetic analysis of PUUV circulating in the bank vole population during a peak phase in 2005 at Konnevesi. A total of 147 voles were captured and 44 of those contained the PUUV N protein according to immunoblotting. Partial sequences of S, M and L viral genome segments were recovered. As expected, PUUV variants (strains) detected at Konnevesi were found to be most closely related to the Finnish lineage including strains collected in Finland and Russian Karelia. A pairwise sequence analysis revealed that the S segments of Konnevesi strains were most similar to Puumala strain (95-96% identical residues), the M segments were most closely related to Kolodozero and Gomselga strains (92-94%), and L segments most closely related to Sotkamo strain (90-93%). The most divergent strains among the Finnish lineage were the Karhumäki strain for S segments (89-90%), Längelmäki strain for M segments (88-90% identity), and Pallasjärvi strain for L segments (89-93%) (Fig. 14). It should be stressed that the number of L segment sequences available for comparison in GenBank is still very limited, which may affect the interpretation of the results.

Partial S- (nt 640 to 1082), M- (nt 2180 to 2610) and L- (nt 502 to 1036) segment amplicons were recovered from 40 PUUV-infected bank voles. Pairwise genetic comparison revealed that genetic diversity (nucleotide polymorphism within a population) among Konnevesi strains was up to 4.9% for S segment sequences, 4.8% for M segment sequences and 9.7% for L segment sequences. Twenty-eight, 29 and 67 point mutations were observed for the S-, M-, and L-sequences, respectively (Fig. 2 in article I). This surprisingly high genetic diversity was due to the presence of two different groups of PUUV genetic variants circulating in the bank vole population at Konnevesi. Numerous intrinsic substitutions (13, 9 and 35 for the S, M and L segments, respectively) could be used as genetic markers to separate the groups, from hereon referred to as “A” and “B”. The A genogroup included two S-segment genotypes, six M-segment genotypes, and eight L-segment genotypes. Corresponding numbers for B genogroup were five, six and nine (Fig. 3 in article I). Notably, groupings based on genetic markers were confirmed by phylogenetic analysis (Fig. 16 herein and Fig. 4 in article I).
Results and discussion: Microevolution of PUUV related to host population dynamics

Figure 14. Geographical locations of PUUV strains from the Finnish genetic lineage. PUUV was discovered at Puumala [Brummer-Korvenkontio et al., 1982] and the prototype Sotkamo strain was sequenced by Vapalahti et al. [1992]. Subsequently, genetic studies have been conducted using samples collected at Evo [Plyusnin et al., 1995], Virrat [Plyusnin et al., 1997], Aitoo and Längelmäki [Plyusnin et al., 1999], Karhumäki, Gomselga and Kolodozero [Asikainen et al., 2000], Konnevesi [Razzauti et al., 2008 and manuscript IV]; Pallasjärvi [Razzauti et al., 2009], Kuhmo [Sironen, Henttonen & Plyusnin, unpublished data], Pieksämäki, Taivalkoski, Kolari and Saariselkä [Razzauti, Henttonen & Plyusnin, unpublished data].

The easy discrimination between the two genogroups allowed the recognition of reassortants. A substantial proportion of the analyzed PUUV genomes (8 of 40, or 20%) possessed segments originating from different ancestral genogroups (Table 2 and Fig. 3 in article I). Of six possible S/M/L segment combinations, only three were observed in the reassortant PUUV strains from Konnevesi in 2005: B_S A_M A_L (in four variants), A_S B_M B_L (in three variants), and B_S A_M B_L (in one variant). The relatively high numbers of PUUV S-, M- and L-segment sequences and their genetic diversity within a small sampling area allowed, for the first time, an estimate to be made of the natural rate of PUUV segment reassortment.
Both the accumulation of point mutations and reassortment of the genomic RNA segments contributed to the generation of genetic diversity in the PUUV population. A total of 124 point mutations, mostly transitions, were detected along 1409-nt amplicon. Although the high diversity of L segment genotypes was remarkable, the majority of mutations were at the 3rd codon's position and only six substitutions were non-synonymous, suggesting strong selection at the aa level. It should be noted that increasing evidence suggests that silent mutations may not always be neutral [Novella et al., 2004; Hamano et al., 2007], especially for RNA viruses since the genome itself can potentially be a target of selection [Domingo et al., 2001]. High genetic diversity measured at a given point in time should not necessarily be considered typical over longer periods. With this in mind, the unexpectedly high level of diversity observed among L-segment genotypes becomes especially intriguing. Perhaps greater heterogeneity confers PUUV with more plasticity or potential for adaptation to difficult situations. Whether this bias is a long-term survival strategy for the virus remains to be investigated.

Interestingly, the proportion of reassortants was slightly higher in October than in May, but total numbers were too small for any definite conclusions to be made with respect to seasonal dynamics. These data showed that reassortants emerge and co-circulate with their parental strains. Competition with parental variants and the survival of reassortants through successive viral generations remains to be studied in greater detail. This issue is addressed in manuscript IV.

Direct sequencing of S amplicons recovered from two bank voles trapped in October revealed double peaks at several positions of the chromatogram (Fig. 15). Interestingly, the nt at these positions is diagnostic for genogroups A and B. Direct sequencing of the corresponding M- and L-amplicons revealed only one genotype. These data suggest that both of these bank voles were infected with two PUUV variants, each possessing a distinct S segment genotype but identical M and L segments.

![Figure 15](image.png)

**Figure 15.** Sequence chromatogram of part of the S segment(s) amplified from the same vole. Yellow circles mark the positions of double peaks, which coincide with discriminating nucleotide positions for A and B genogroups.
Genetic analysis of PUUV strains from Konnevesi was continued (with biannual trapping) until October 2009, i.e., through one complete bank vole population cycle. The study period covered two peak phases in 2005 and 2008, and two population declines in 2006 and 2009 (i.e., presumable viral bottlenecks). Altogether, from May 2005 to October 2009, 1369 bank voles were captured. Of those, 360 voles (26.3%) were found to be PUUV N-Ag-positive and the infecting PUUV genomes were subsequently amplified. A 1443-nt long sequence, or 12% of the genome, was recovered from 356 infections (of which we failed to recover three of the M segments); 455 nt of the S segment (631-1085nt), 452 nt of the M segment (2162-2613nt), and 536 nt of the L segment (505-1040nt).

Table 2. Puumala virus prevalence during the bank vole density cycle

<table>
<thead>
<tr>
<th>Year</th>
<th>Season</th>
<th>no. of trapped bank voles</th>
<th>no. of PUUV N-Ag positives</th>
<th>PUUV Ag-prevalence (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2005</td>
<td>Spring</td>
<td>47</td>
<td>22</td>
<td>46.8</td>
</tr>
<tr>
<td></td>
<td>Autumn</td>
<td>100</td>
<td>22</td>
<td>22</td>
</tr>
<tr>
<td>2006</td>
<td>Spring</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Autumn</td>
<td>6</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2007</td>
<td>Spring</td>
<td>54</td>
<td>7</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>Autumn</td>
<td>132</td>
<td>15</td>
<td>11.4</td>
</tr>
<tr>
<td>2008</td>
<td>Spring</td>
<td>237</td>
<td>106</td>
<td>44.7</td>
</tr>
<tr>
<td></td>
<td>Autumn</td>
<td>625</td>
<td>155</td>
<td>24.8</td>
</tr>
<tr>
<td>2009</td>
<td>Spring</td>
<td>78</td>
<td>28</td>
<td>35.9</td>
</tr>
<tr>
<td></td>
<td>Autumn</td>
<td>88</td>
<td>5</td>
<td>5.7</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td><strong>1369</strong></td>
<td><strong>360</strong></td>
<td><strong>26.3</strong></td>
</tr>
</tbody>
</table>

PUUV prevalence in bank voles was higher in spring, when the majority of rodents in the population were born the previous year and have over-wintered, in contrast to the autumn when young animals predominate (Table 2). Maturation of voles during the spring contributes to PUUV transmission in that sexually mature males disperse longer distances and those infected spread PUUV over a wider area. However, in summer and early autumn, a significant proportion of young bank voles may be protected with maternal Abs. When these animals seroconvert (i.e., cease breast-feeding), their interaction rate typically increases and transmission of PUUV continues over the winter [Kallio et al., 2010]. It is also possible that recent infections with fewer viral particles, mostly occurring in autumn, may go undetected if sensitive methods such as RT-PCR are not used [our unpublished data].
The pairwise sequence comparison of 356 PUUV genomes (totally, 1443 nt analyzed for each genome; with exception of the M segments failed to recover) revealed 182 mutations; 53 in the S segment, 50 in the M segment, and 79 in the L segment. Two genogroups (A and B), described in article I, were co-circulating throughout the population cycle. Genetic diversity within A and B genogroups was 1.5-3.2% and between the genogroups was 4.8-10.1%. The highest diversity was observed among L segments. One non-synonymous mutation that led to an aa substitution (Ile83Val) could be used as a molecular marker to discriminate L segment genotypes belonging to A or B genogroups (Table 2 in manuscript IV).

The number of segment genotypes (SA1-11 and SB1-34, MA1-18 and MB1-20, LA1-26 and LB1-35; see Fig. 1 in manuscript IV) observed over five years was larger than in the time-point analysis [article I]. Newly emerged segment genotypes were found every season except in the decline year of 2006, when no PUUV-infected voles were trapped. The A genogroup included 11 S segment genotypes, 18 M segment genotypes, and 26 L segment genotypes. The corresponding numbers for genogroup B were 34, 20 and 35. Overall diversity was always higher for genogroup B. For both genogroups, the L segment presented a larger number of genotypes than S or M segments. Remarkably, L segment diversity between genogroups was higher than that of the other two segments (Table 2 and Fig. 1 in manuscript IV), and the groupings based on genetic markers were further confirmed by phylogenetic analyses (Fig. 16 a-c).

The 356 PUUV sequences mentioned above represented 184 distinct genetic strains (variants). Of those, 46 belonged to the A genogroup \((A_S A_M A_L)\), 82 were of the B genogroup \((B_S B_M B_L)\), and 56 were reassortants of A and B genogroups (all six possible types of reassortants were registered, Fig. 7). Genogroup B was more diverse and PUUV variants of this genogroup were always more abundant than variants of A (Table 1 in manuscript IV). Nevertheless, the proportion of both genogroups was stable during the study period and no signs of competition between the two genogroups were apparent. Interestingly, only 39 of these 184 distinct variants were observed during two or more seasons, and the remaining variants (nearly the 80%) were registered only once (Table 3 in manuscript IV).
Results and discussion: Microevolution of PUUV related to host population dynamics

**Figure 16a.** Phylogenetic tree of the coding region of PUUV S segment. The tree topology, its support and branch lengths were estimated using neighbour-joining (NJ) with the F84 substitution model in PHYLIP [Felsenstein, 1993]. A maximum clade credibility tree with an arbitrary root is shown with mean branch lengths (substitutions per site), and non-parametric bootstrap percentages are shown for each node. TPMV, HTNV, RIOMV, SNV and TULV were used as representatives of hantaviruses from each host clade and omitted for the graphical representation.
Figure 16b. Phylogenetic tree of partial M segment (nt 2180-2610) of PUUV. The tree topology, its support and branch lengths were estimated using neighbour-joining (NJ) with the F84 substitution model in PHYLIP [Felsenstein, 1993]. A maximum clade credibility tree with an arbitrary root is shown with mean branch lengths (substitutions per site), and non-parametric bootstrap percentages are shown for each node. TPMV, HTNV, RIOMV, SNV and TULV were used as a representatives of hantaviruses from each host clade and omitted for the graphical representation.
Results and discussion: Microevolution of PUUV related to host population dynamics

Figure 16c. Phylogenetic trees of partial L segment (nt 502-1036) of PUUV. The tree topology, its support and branch lengths were estimated using neighbour-joining (NJ) with the F84 substitution model in PHYLIP [Felsenstein, 1993]. A maximum clade credibility tree with an arbitrary root is shown with mean branch lengths (substitutions per site), and non-parametric bootstrap percentages are shown for each node. TPMV, HTNV, RIOMV, SNV and TULV were used as a representatives of hantaviruses from each host clade and omitted for the graphical representation.
Most abundant genotypes ("copious") could be distinguished for each of the genome segments: 6A, 7A, 8A, 5B, 20B, 21B, 22B and 28B for the S segment; 1A, 8A, 14A, 15A, 3B, 8B and 15B for the M segment, and 3A, 9A, 14A, 4B, 8B, 11B, 19B, 21B and 23B for the L segment. Figure 1 [manuscript IV] shows the frequency of occurrence of all segment genotypes and their relationships. For both A and B genogroups, several clusters can be distinguished. These clusters are composed of at least one copious segment genotype and several other genotypes rarely observed ("sporadic") (e.g., the S segment copious genotype 28B and sporadic genotypes 27B, 30B, 31B, 32B, 33B, 34B form a cluster). Sporadic genotypes are probably derived from a closely-related copious genotype by genetic drift. Clusters of genotypes can be easily distinguished for the S- and M-segments. However, the relationships of L segment genotypes in both genogroups show a net-like pattern (Fig. 1 in manuscript IV), probably stemming from the accumulation of many point mutations in that segment.

The analysis of segment genotype frequencies revealed that the copious ones (names for those shown in color in Fig. 2 of manuscript IV) were prevalent throughout the bank vole cycle whereas sporadic genotypes appeared and then disappeared during the observation window. Only a few of the sporadic segment genotypes (names for those are given in squares in Fig. 2 of manuscript IV) were observed at different times, reflecting that sporadic genotypes cannot outcompete the copious ones. Newly detected segment genotypes were observed at all times except in 2006 when all trapped voles were PUUV free.

For the genetic analysis, circulating PUUV variants were classified into four categories depending on their occurrence: (i) the most frequently occurring variants (9), detected ≥5 times; (ii) repeatedly observed variants (27), detected 2-4 times; (iii) transient variants (145), detected once; and (iv) reassortant variants (56), which are analyzed separately due to their peculiar genetic nature. The most frequently occurring variants (≥5 times) included only copious segment genotypes. S and M segments of the repeatedly observed variants (2-4 times) were composed of approx. 70% copious genotypes, while the L segment was represented less often (~30%) by copious genotypes. The three segments of the transient variants (1 time) were equally represented by copious and sporadic segment genotypes. Approximately 70% of reassortant variants were composed of copious S and M segment genotypes, whereas the L genome segment was equally represented by copious and sporadic genotypes (Table 4 in manuscript IV). On this basis, reassortants were similar to the repeatedly occurring variants. S and M segments were, for the most part, composed of copious genotypes although they were not observed to be circulating throughout the bank vole cycle (with the exception of three variants detailed in Table 5 in manuscript IV) in the way repeatedly occurring variants did.

A total of 68 (19.1%) reassortant PUUV genomes were found in 2005-2009, similar proportion (20%) of reassortants detected in our first study [article I]. The reassortants comprised 58 distinct variants (i.e., some variants were observed more than once). Reassortants were mainly composed of copious segment genotypes (Table 5 in
Results and discussion: Microevolution of PUUV related to host population dynamics

As one might expect since copious (or most abundant) genotypes are more likely to co-infect the same host and undergo reassortment. Alternatively, dominant (or more frequently occurring) variants may have a somewhat higher ability to counterpart in co-infection, while transient variants could be outcompeted.

All six possible segment combinations ($A_S B_M B_L$, $B_S A_M A_L$, $A_S A_M B_L$, $B_S B_M A_L$, $A_S B_M A_L$, $B_S A_M B_L$) were observed among reassortant PUUV strains. Although this study ascertained that all kinds of reassortant are viable, their relative fitness appears to be lower than that of their parental strains. Nearly all (95%) of reassortant variants were transient, with only three circulating repeatedly in the population. Examination of reassortants revealed a preferred pattern of segment combination, i.e., a non-random exchange of segments. The L genome segment tended to pair with S or M segments of the same genogroup (reassortants 1-4 in Fig. 7). Hence, S and M segments are extensively exchanged (95.6% of cases) during reassortment (Table 6 in manuscript IV). Earlier studies on hantavirus reassortment reported an exchange of the M segment on seven occasions [Henderson et al., 1995; Rodriguez et al., 1998; Rizvanov et al., 2004; Handke et al., 2010; Kivsanot et al., 2010], of the S segment on three [Li et al., 1995; Rodriguez et al., 1998; Zou et al., 2008], and the L segment was observed to be exchanged in only two studies, and always with lower frequency than other segments [Henderson et al., 1995; Rodriguez et al., 1998]. This suggests that different inter-segmental compatibilities may exist and influence viral replication.

Analysis of emerging variants revealed that the number of reassortants varied from one season to another and inversely to PUUV prevalence. Reassortants were less abundant in spring when PUUV prevalence is high, but more abundant in autumn when prevalence is low. Such a cyclicity was not observed for non-reassortants of either A or B genogroups (Table 1 in manuscript IV). The seasonality of reassortment could be due to a higher proportion of young voles during the autumn. These naive animals may become suddenly susceptible to PUUV when they stop breast-feeding. Thus, infection dynamics shift in favour of the virus, co-infections become more likely and the chance for reassortment increases. An alternative explanation could be that at the beginning of a co-infection, PUUV variants replicate at similar rates but then one of the variants could outcompete the others and becomes dominant.

In nature, the intra-genogroup reassortants are difficult to detect and only the most apparent reassortants are easily recognized. These are inter-species, inter-lineage or inter-genogroup reassortants. In nature, the intra-genogroup reassortants are very difficult to detect. Imperceptible reassortment between close genetic variants could be a
useful mechanism for PUUV to maintain a steady-state viral population; it could counteract the effects of Muller’s ratchet [Muller, 1932, 1964]. Experiments with the segmented RNA Φ6 virus support that reassortment can reduce an excessive mutational load in a population and hence help it to escape from accumulated deleterious effects [Chao et al., 1992, 1997].

Visual inspection of chromatograms revealed 12 partial sequences (six of the S, two of the M, and four of the L segment) that showed double peaks at diagnostic positions (Fig. 15), suggesting the presence of multiple variants in the host. Sequencing of individual cDNA clones confirmed the presence of two sequences and hence a two-strain co-infection. Interestingly, 11 of such co-infection cases were recovered from voles trapped in October, suggesting them to be recently acquired during the autumnal peak in infection rate. Perhaps, co-infections are detectable in their early stages but as time passes either variant is replaced by the dominant. This would explain the seasonal occurrence of co-infections in October.

To find traces of genome recombination in PUUV, two variants of each genogroup where thoroughly investigated. Near-complete genomes (complete S and M segments and approx. 80% of the L segment) of these four variants were recovered but no clear signs of recombination were observed.

**PUUV lineages and their post-glacial history**

The Konnevesi study focused on genetic variants of PUUV circulating in a host population, the dynamics of which strongly fluctuated over the 5-year study period. It was important to compare those with genetic variants of PUUV circulating in a host population with only seasonal cycles where declines are less apparent. Population cycles of bank voles in Lapland stabilized in the early 1980’s and cycling continued seasonally until 1998 and thereafter [Henttonen, 2000].

In September 1998, 171 bank voles were trapped at Pallasjärvi, in the Pallas-Yllästunturi National Park in northern Finland (Fig. 14). Of them, 25 were PUUV N-Ag-positive (14.6% viral prevalence). Partial sequences of S segment (nt 640-1082), M segment (nt 2180-2632) and L segment (nt 577-1036) were analysed. Mutation characteristics are shown in Table 1 of article II. Similar to PUUV at Konnevesi, S, M and L genotypes observed at Pallasjärvi could be divided into two groups. Within these groups, genetic diversity of PUUV variants was modest (up to 1.3%) but between groups it was surprisingly high (up to 18.9%) (Table 2 in article II). Analysis of deduced protein sequences showed similar results. Only one aa substitution was detected within each group; one in the N protein of the first group (Val260Ille), and one in the L protein of the second group (Asp245Asn). However, aa
diversity between groups was as high as 7.5%. Three of the 11 aa residues observed in the N protein sequence were earlier described as lineage-specific residues [Sironen et al., 2001]; Met262 and Asp304 from FIN and Asp272 from N-SCA. These observations suggested that the PUUV variant groups at Pallasjärvi belonged to different phylogroups, and further phylogenetic analysis (Fig. 3a in article II) confirmed the FIN and N-SCA lineages (Fig. 16-c). This situation granted us an opportunity to study inter-lineage interaction(s) in the field.

A total of 13 PUUV genetic variants were observed in the bank vole population at Pallasjärvi. These resulted from a combination of five, seven and seven different genotypes for the S, M and L segments, respectively (Fig. 2 in article II). Genotyping and phylogenetic analyses revealed that reassortment of genome segments belonging to different PUUV lineages may have occurred. Eight of 25 infected bank voles (32%) carried reassortant viruses. Of the six possible reassortant S/M/L combinations, only two were found and for all reassortants both S and L segments derived from the same lineage. One variant possessed a genome of the S\textsubscript{FIN}M\textsubscript{N-SCA}L\textsubscript{FIN} type and five of the S\textsubscript{N-SCA}M\textsubscript{FIN}L\textsubscript{N-SCA} type (Table 3 in article II). Although few in number, these observations suggest a non-random reassortment pattern resembling that seen in the Konnevesi study.

To test the hypothesis of co-evolution of bank vole phylogroups and PUUV lineages, genetic analyses of rodent mtDNA were conducted. Results showed that the D-loop region in bank voles from Pallasjärvi belonged to the red vole (Myodes rutilus). As such, the Ural phylogroup (Fig. 12) is represented in northernmost Fennoscandia, in support of conclusions drawn by Deffontaine-Deurbroeck [2008].

This study identified a contact zone for FIN and N-SCA PUUV lineages and detected co-circulation of two distinct lineages within a host population of monomorphic mtDNA. Furthermore, inter-lineage reassortment was demonstrated in PUUV for the first time. Earlier studies showed that the contact zone of the Ural and Western European bank vole phylogroups in North-Central Sweden is congruent with the contact zone for the S-SCA and N-SCA PUUV lineages [Hörling et al., 1996; Johansson et al., 2008; Nemirov & Lundkvist, personal communication]. However, recent observations demonstrate that PUUV lineages are indifferent to bank vole phylogroups (Fig. 11 and 12 herein; and Fig. 1 in article II). It has also been shown that PUUV strains of the FIN lineage recovered from bank voles of both Eastern and Ural clades occur at a contact zone in Kuhmo, eastern Finland, [Sironen, Henttonen & Plyusnin, personal communication], supporting our observations at Pallasjärvi that PUUV lineages are not tied to bank vole phylogroups.

It seems likely that the two PUUV lineages have been sharing their host population for some time and neither has achieved dominance, perhaps due to similar fitness. Boldin and Diekmann [2008] developed a mathematical model for superinfections of pathogens that fits this hypothesis. Intra-group genetic diversity of PUUV variants found at Pallasjärvi was relatively low (up to 1.7%). Earlier descriptions of genetic diversity of the FIN and N-
SCA lineages reported higher values of up to 3.2% for FIN [manuscript IV] or up to 7% for N-SCA variants [Johansson et al., 2008]. One can speculate that the presence of two lineages exerts stronger selection pressure and, therefore, less fit variants are removed from the population leading to a reduction in genetic diversity.

Genetic variation of PUUV exhibits some geographical structure within the host distribution. The isolation of rodent populations during the LGM (Weichselian, ca. 10–13 ka) and subsequent recolonization of Eurasia left signatures in the current distribution of PUUV lineages. Our findings at the northern contact zones stimulated our interest in other PUUV lineages and their post-glacial history. NE has been reported over the last 10–15 years in the Baltic countries [Golovljova et al., 2000; Lundkvist et al., 2002; Sandmann et al., 2005]. Although Estonian PUUV strains clustered within the RUS lineage [Golovljova et al., 2002], Sironen and colleagues [2001] suggested a recombinant origin. Based on these findings, we began a study of PUUV in the Baltic region.

Rodents in Latvia were trapped in 2008 and 2009. Of the 257 bank voles analyzed, only six were PUUV N-Ag positive. These animals were trapped at three separate locations: two in the vicinity of Jelgava in western Latvia (#149 and #778, location 1: 56°59’N 24°10’E; and #136 and #140, location 2: 56°74’N 23°53’E) and one close to Madona in eastern Latvia (#99 and #233, location 3: 56°81’N 26°71’E) (Fig. 1 in article III). Complete S-segment sequences (1835 nt) and partial M- and L-segment sequences (391 nt and 413 nt, respectively) of PUUV were recovered from all N-Ag+ samples. Sequence comparison revealed that PUUV strains from Latvia formed two groups; one included strains Madona99, Madona233, Jelgava136 and Jelgava140, and the other contained Jelgava149 and Jelgava778. The four strains of the first group were closely related, and S, M and L segments of Madona99 and Madona233 were 99.9, 100, and 98.3% similar, respectively. S segments of Jelgava136 and Jelgava149 were 99.8% similar and partial M- and L-segment sequences were identical. For the whole group, divergence of the S-, M- and L-segment sequences was up to 4.8, 7.8 and 10.7%, respectively. Recovered sequences of Jelgava149 and Jelgava778 were identical. Genetic divergence between the two groups was up to 17.1% for the S segment, 21% for the M segment, and 22.1% for the L segment. Moreover, comparison of protein sequences revealed several synapomorphic aa markers that discriminate the two groups; 11 in the complete N protein sequence (A5T, R30K, V34M, Y61F, P233A, D234E, K237R, D250E, R265K, N272Q, and D302N), nine in the partial Gc protein sequence (A775G, Y769F, T771S, V803A, V807I, R810K, V815A, T827S, and I/L835V), and eight in the partial L protein sequence (R200K, Q234P, V239A, N242S, D245S, Q265H, D273E and R315Q). Such levels of diversity are typical between distinct PUUV lineages. Further phylogenetic analyses (Fig. 2 in article III) confirmed the patterns observed by aa markers; strains Madona99, Jelgava136, Jelgava140 and Madona233 formed a well-supported group within the RUS lineage in phylogenetic trees of all segments; and Jelgava149/778 formed an exclusive clade designated the Latvian lineage (LAT) (Fig. 2 in article III).
The LAT lineage was at least 14.7% divergent from the other seven lineages described so far. The S-segment sequence was equally distant to all other PUUV lineages. Interestingly, the LAT M-segment sequence was most similar to ALAD and CE lineages, while the L segment sequence appeared closer to FIN lineage (Table I in article III). This was also reflected in the close kinship of the M- and L-segment phylogenies (Fig. 2b and c in article III). Such discrepancies suggest the LAT lineage is a chimeric reassortant. Unfortunately, M- and L-segment sequences of the S-SCA lineage and L-segment sequences of the CE lineage were unavailable for comparison. An expanded PUUV database might help to resolve this question.

Phylogenetic analysis of the cyt b and D-loop mtDNA regions of PUUV-infected Latvian bank voles revealed that all six animals belonged to the Carpathian clade [Wójcik et al., 2010]. Thus, genetic distinction observed between the RUS and the LAT lineages occurred within one mtDNA phylogroup of the host species. This again supports the idea that PUUV lineages are not tightly associated with bank vole phylogroups. Due to low PUUV prevalence among trapped rodents, we did not detect many variants of the LAT lineage and, therefore, were not able to investigate reassortment between the RUS and LAT lineages.

The distribution of the LAT lineage remains uncertain, although it likely includes neighbouring countries. Since no apparent geographical barriers constrain the lineage, LAT may have been brought from more southern or eastern glacial refugia, and exchange could have occurred with other lineages during recolonization. It should be mentioned that a suture zone has been reported in Lithuania among phylogroups of several vole species, as well as plants and other animals. Interestingly, this postglacial contact zone also coincides with the distributional limits of two bank vole mtDNA phylogroups; the Carpathian clade in west Lithuania and the Eastern clade in east Lithuania [Deffontaine et al., 2005]. It would appear that, while the distribution of bank vole phylogroups has been established since postglacial recolonization, PUUV lineages have crossed these borders indiscriminately.
CONCLUDING REMARKS AND FUTURE PROSPECTS

This study focused on the mechanisms of PUUV microevolution. The five-year monitoring of PUUV variants in a local bank vole population unveiled distinct strategies that PUUV uses to survive in a host population of fluctuating density. High genetic diversity seems essential for survival, and both the accumulation of point mutations and reassortment of genomic RNA segments generate genetic diversity in the PUUV population. The contribution of recombination to PUUV evolution is rather slight, if any. Point mutations, of which the majority was silent, occurred with relatively high frequency. The dn/ds ratios indicate the operation of negative (or purifying) selection at the protein level. Notably, the highest genetic diversity was observed among L-segment genotypes, probably due to a different evolutionary origin of that genome segment.

Co-circulation of two PUUV genogroups in the same bank vole population did not promote competition between groups (subpopulations of viral variants). Exchange of genome segments between the co-circulating PUUV was extensively observed (in 19.1% of the PUUV population), and it was suspected that many reassortants remained undetected. Although all six possible segment combinations were found in reassortant PUUV strains, a preferred pattern of reassortment was observed with the L segment being more likely to combine with S or M homologous segments. Thus, when reassortment occurs, S or M segments are the most commonly exchanged. The L segment was swapped in only 4.4% of cases.

The occurrence of reassortment was inversely related to seasonal prevalence of PUUV in the bank vole population. Viral prevalence is always higher during the spring due to the age structure of the host population, but the number of reassortments was lower during this period. In contrast, PUUV prevalence is lower during the autumn but PUUV reassortants were more numerous. This can be explained by an increase in infection rate during the autumn when newly weaned voles become susceptible and suffer co-infection by two or more PUUV variants that undergo reassortment in their common host.

A huge number of PUUV variants was observed co-circulating throughout the bank vole population cycle. Variants of one genogroup were always more diverse and abundant in the population, but no clear signs of competition among genogroups were observed. In spite of the significant number of variants, only a few were detected more than once during the five-year study. These variants were mainly composed of the most abundant genotypes (copious) of each segment. At each point in time, new segment genotypes (sporadic) were detected, and those appeared to be mutated versions of the copious genotypes. Sporadic segment genotypes did not outcompete copious genotypes. The majority (78.8%) of observed PUUV variants were transient, and comprised approx 50% of the cases by sporadic segment genotypes. Repeatedly observed variants as well as S and M segments of reassortants were approx. 70% composed of copious genotypes. Notably, the L segment was equally represented by copious and sporadic segment genotypes. An
imperceptible intra-genogroup reassortment could explain such differences in the composition of L-segment genotypes. PUUV may rely on intra-genogroup reassortment to counteract the effects of Muller's ratchet and maintain a steady-state viral population.

The co-circulation of and interaction between two distinct PUUV lineages (FIN and N-SCA) were observed in the bank vole population at Pallasjärvi in northern Finland. Genetic diversity within each of the circulating lineages was modest and substitutions were mostly synonymous, probably reflecting stabilizing selection at the aa level. In contrast, genetic differences between lineages was high. These distinct lineages were shown to reassort their RNA genome segments with a frequency comparable to that of PUUV genogroups circulating at Konnevesi in central Finland. In this case, a preferred pattern of segment combination was also found, where only the M segment was exchanged between the two lineages. The swapped M segment always belonged to the FIN lineage, suggesting unequal fitness.

Swedish and Finnish research groups have been looking for the contact zone of the FIN and N-SCA lineages. This study has now determined a location where these two lineages converge. Whereas the contact zone for the Ural and Western European bank vole phylogroups in Central Sweden is congruent with the contact zone for the S-SCA and N-SCA lineages [Hörling et al., 1996; Johansson et al., 2008; Nemirov et al., 2010], no such congruence was observed for bank vole phylogroups and FIN or N-SCA PUUV lineages.

When in co-circulation, neither the FIN nor N-SCA lineage showed dominance over the other, suggesting that their fitness was similar. In the future, it would be interesting to monitor the survival of inter-lineage reassortants. Rodent research in this area has been conducted from 1970 to the present, and tissue samples are available for much of this period. Hence, a continuation of this research is feasible and a 30-year study of PUUV microevolution in the bank vole population at Pallasjärvi is in our plans.

Co-circulation of two distinct PUUV lineages was observed in Latvia and, again, in a single bank vole phylogroup. Although trapping sites were close to the suture zone of East European and Carpathian phylogroups in Lithuania, bank voles from the study area belonged to the Carpathian clade. One of these PUUV lineages was identified for the first time and named as the Latvian lineage (LAT). Distribution of the LAT lineage out of the Latvian political borders remains to be determined.

These two recent observations, in addition to unpublished data of Sironen and colleagues, demonstrate that PUUV lineages do not assort according to bank vole phylogroups. Moreover, genetic variation across the host distribution suggests that PUUV lineages were isolated during the LGM, recolonized Europe with their hosts, and were capable of spreading through the bank vole phylogroups even in their contact zones. Future plans include a study of PUUV genetics in Central and Eastern Europe in relation to host phylogeography.
ACKNOWLEDGEMENTS

The work described in this thesis was carried out at the Department of Virology, Haartman Institute of the University of Helsinki, at the Finnish Forest Research Institute (Metla), and at the Konnevesi Research Station of the University of Jyväskylä.

I wish to express my sincere gratitude to my thesis supervisors Professor Alexander Plyusnin and Professor Heikki Henttonen for guiding me through my research. I thank them for the dedicated support they granted me during the development of the thesis. I enjoyed all our stimulating scientific discussions. It has been an honour to work with them.

I thank the head of the Virology Department at the Haartman Institute, Professor Kalle Saksela for providing excellent research facilities, Professor Heikki Henttonen, for supporting the rodent laboratory at Metla, and Professor Hannu Ylönen for accommodating the rodent trapping team at the Konnevesi Research station.

I sincerely appreciate the constructive comments from thesis reviewers Docent Petri Susi and Professor Dennis Bamford. I am glad that my opponent, Professor Esteban Domingo, accepted the invitation to audit the public defence of this doctoral thesis.

I want to acknowledge Angelina Plyusnina for the invaluable help in the lab during these past years. I am indebted to Jukka Niemimaa and Liina Vuotilainen for their great help in the field. They were always ready at 5 a.m. to dig meter-deep holes in the snow at -25 °C. I sincerely appreciate their dedication.

I thank Cédric Moisy for his work designing the cover of the thesis paperback.

I am grateful for the charming company of past and present members of the zoonoses research group, particularly the “scientific discussions” that took place every now and then Friday afternoon. I thank the ever-smiling laboratory technicians Leena Kostamovaara, Tytti Manni, Irina Suomalainen and Kirsi Aaltonen. My office colleagues deserve a special mention. Anna Katz always provided a helping hand, and Angė Alminaitė for some unforgettable conference experiences.

Above all, I am grateful for my “Finnish” friends outside the lab for the great moments we have spent together during these last years. They have smoothed the path of this long trip. I also want to thank my overseas friends, who have supported me remotely at the most difficult times, and especially to Pilar Escandell for her infinite friendship.

There are no words to express my affection for my family members. They have provided unconditional support at all stages of my life. I thank them from the bottom of my heart for their commitment and love.

This work was financially supported from Funcació-Universitat Empresa de les Illes Balears, the Centre for International Mobility in Finland (CIMO), the Sigrid Jusélius
Acknowledgements

Foundation, the Academy of Finland, the EU-FP6 grant GOCE-2003-010284 EDEN (http://www.eden-fp6project.net) and EU grant FP7-261504 EDENext (http://www.edenext.eu) and is catalogued by the EDENext Steering Committee as EDENext 0009. The contents of this publication are the sole responsibility of the authors and do not necessarily reflect the views of the European Commission.

Maria Razzauti Sanfeliu
REFERENCES


References


Li D. 2010. Trends of HFRS epidemiology and the expanded program on immunization with hantavirus vaccines in China. Personal communication at the VIII International Conference on HFRS, HPS & Hantaviruses, Athens, Greece 2010.


References


