Hantaviruses: genome structure, expression and evolution

Alexander Plyusnin, Olli Vapalahti and Antti Vaheri

Haartman Institute, Department of Virology, POB21, FIN-00014 University of Helsinki, Finland

Introduction

The isolation by Ho Wang Lee and collaborators of the virus causing Korean haemorrhagic fever, now called Hantaan virus (HTN), from the lungs of striped field mice (Apodemus agrarius) in 1976 (Lee & Lee, 1976) launched a new era in the study of haemorrhagic fever with renal syndrome (HFRS) throughout the world. This was soon followed by the discovery of the causative agent of the European form of HFRS, nephropathia epidemica, now known as Puumala virus (PUU) (Brummer-Korvenkontio et al., 1980), and of the urban rat virus, Seoul virus (SEO) (Lee et al., 1980). The identification in 1993 of Sin Nombre virus (SN) as the causative agent of the North American form of HFRS, now known as Hantavirus pulmonary syndrome (HPS) (Nichol et al., 1993a) led to intensive search for further hantaviruses and as a result today a total of as many as 16 well-established sero/genotypes may be listed.

It was largely the pioneering molecular virology studies by Connie Schmaljohn and the late Joel Dalrymple in the 1980s (Schmaljohn & Dalrymple, 1983; Schmaljohn et al., 1985, 1986b, 1987) that gave hantaviruses the status of a separate genus, Hantavirus, in the family Bunyaviridae (Murphy et al., 1995). Hantaviruses, HTN being the prototype, are today recognized to have their own distinguishing features, both ecologically and in molecular properties. In contrast to viruses in other Bunyaviridae genera, hantaviruses are not transmitted by arthropods. The reservoir hosts of hantaviruses are specific rodent/insectivore species; on occasion they infect humans but do not spread from human-to-human. These features make it of special interest to study the genetic evolution of hantaviruses. Hantaviruses are serologically related to one another but their proteins do not exhibit extensive sequence similarities or antigenic relationships to proteins of viruses representing other genera. On the one hand, hantaviruses give rise to severe clinical manifestations, HFRS and HPS. On the other hand, they cause no detectable cytopathology in cell cultures and produce persistent non-pathogenic infections in rodents. In this review, we summarize the current knowledge on genetic structure, protein expression and evolution of hantaviruses.

Background

Currently, the genus Hantavirus of the family Bunyaviridae consists of at least 16 sero/genotypes (Table 1). The list of hantaviruses is increasing rapidly and several more types are being characterized at the moment: for instance, Tobetsu virus from Japan (Kariwa et al., 1995; K. Yoshimatsu, personal communication), Rio Mamore virus from Bolivia (Fijelle et al., 1996, and personal communication) and Topograpoval virus from Arctic Siberia (Plyusnin et al., 1996b). Each hantavirus is carried predominantly by its own natural host, rodent or insectivore. Taking into consideration that only about 5% of rodent species have been tested for the presence of hantaviruses (Henttonen et al., 1996), it is to be expected that many more new members of the genus will be encountered in the near future.

Some hantaviruses are causative agents of HFRS or HPS; the pathogenicity and infectivity to humans of others still needs to be clarified (Lundkvist & Niklasson, 1994). HFRS presents as a generalized infection with fever, haemorrhages and acute renal insufficiency. The three major pathogens, HTN, PUU and SEO, cause annually about 200 000 HFRS cases of varying severity in China, Korea, Russia and Europe (Lee et al., 1990). The mortality in HFRS ranges from 3–7% for HTN infection to 0.1–0.2% for PUU (Lee et al., 1990; Mustonen et al., 1994). Dobrava virus (DOB) has been demonstrated to cause HFRS in the Balkans (T. Avsic-Zupanc & Å. Lundkvist, personal communications). HPS, first detected in 1993 in the Four Corners region of the southwestern United States (Nichol et al., 1993a) is an acute febrile illness, followed by development of respiratory failure and, in about 50% of cases, death. Altogether, 128 cases of HPS (mostly in the United States and Canada) were reported for February 1996 (Khan et al., 1996; Nichol et al., 1996).

The demarcation between HPS and HFRS is not as absolute as perhaps thought initially. This is because Bayou virus (BAY)- and Black Creek Canal virus (BCC)-associated cases of HPS exhibited renal involvement (Morzuov et al., 1995a; Ravkov et al., 1995) and HFRS cases have frequently had pulmonary manifestations (Alexeev & Baranov, 1993; Kanerva et al., 1996). However, being clinically distinct, HFRS and HPS have a common pathogenetic feature, capillary injury that can lead to haemorrhages and shock in HFRS or pulmonary oedema and suffocation in HPS.
Table 1. Hantaviruses, their natural hosts and pathogenicity for humans

<table>
<thead>
<tr>
<th>Hantavirus geno/serotype</th>
<th>Host rodent or insectivore</th>
<th>Areas where virus has been found</th>
<th>Disease</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Black Creek Canal (BCC)</td>
<td>Cotton rat (Sigmodon hispidus)</td>
<td>North America (USA)</td>
<td>HPS</td>
<td>Ravkiov et al. (1995)</td>
</tr>
<tr>
<td>Dobrava (DOB)</td>
<td>Yellow-necked field mouse (Apodemus flavicollis)</td>
<td>Balkans (Slovenia)</td>
<td>HFRS</td>
<td>Avisc-Zupanc et al. (1995)</td>
</tr>
<tr>
<td>El Moro Canyon (ELMO)*</td>
<td>American harvest mouse (Rethrodontomys megalotis)</td>
<td>North America (USA)</td>
<td>?</td>
<td>Hjelle et al. (1994)</td>
</tr>
<tr>
<td>Hantaan (HTN)</td>
<td>Field mouse (Apodemus agrarius)</td>
<td>Asia (China, Japan, Korea), Balkans</td>
<td>HFRS</td>
<td>Lee &amp; Lee (1976)</td>
</tr>
<tr>
<td>Isla Vista (ILV)*</td>
<td>California vole (Microtus californicus)</td>
<td>North America (USA)</td>
<td>?</td>
<td>Song et al. (1995)</td>
</tr>
<tr>
<td>Khabarovsk (KBR)</td>
<td>Reed vole (Microtus fortis)</td>
<td>Asia (Far East Russia)</td>
<td>?</td>
<td>Höfling et al. (1996a)</td>
</tr>
<tr>
<td>New York (NY)*</td>
<td>White-footed mouse (Peromyscus leucopus)</td>
<td>North America (USA)</td>
<td>HPS</td>
<td>Hjelle et al. (1995 b)</td>
</tr>
<tr>
<td>Puumaia (PUU)</td>
<td>Bank vole (Clethrionomys glareolus)</td>
<td>Northern and Central Europe, Russia, Balkans</td>
<td>HFRS</td>
<td>Brummer-Korvenkontio et al. (1980)</td>
</tr>
<tr>
<td>Rio Segundo (ROIS)*</td>
<td>Harvest mouse (Rethrodontomys mexicanus)</td>
<td>North America (USA)</td>
<td>?</td>
<td>Hjelle et al. (1995 a)</td>
</tr>
<tr>
<td>Seoul (SEO)</td>
<td>Rats (Rattus norvegicus, Rattus rattus)</td>
<td>Asia (China, Korea, Japan); harbours worldwide</td>
<td>HFRS</td>
<td>Lee et al. (1980)</td>
</tr>
<tr>
<td>Sin Nombre (SN)</td>
<td>Deer mouse (Peromyscus maniculatus)</td>
<td>North America (USA, Canada)</td>
<td>HPS</td>
<td>Nichol et al. (1993a)</td>
</tr>
<tr>
<td>Thailand (THAI)</td>
<td>Bandicoot (Bandicota indica)</td>
<td>Asia (Thailand)</td>
<td>Apathogenic?</td>
<td>Xiao et al. (1994)</td>
</tr>
<tr>
<td>Thottapalayam (TPM)</td>
<td>Shrew (Suncus murinus)</td>
<td>Asia (India)</td>
<td>Apathogenic?</td>
<td>Xiao et al. (1994)</td>
</tr>
<tr>
<td>Tula (TUL)</td>
<td>European common vole (Microtus arvalis, Microtus rassiasemidionalinis)</td>
<td>Europe (Russia, Czech Republic, Slovakia)</td>
<td>?</td>
<td>Plyusnin et al. (1994 a), Vapalahti et al. (1996 b)</td>
</tr>
</tbody>
</table>

* Isolation in tissue culture has not been reported. Each of the 16 hantaviruses in Table 1 fulfils the criteria for a distinct genotype (specific natural host and a unique genetic structure). Hantaviruses that have been isolated in Vero E6 cells also fulfill criteria for a distinct serotype (at least 4-fold difference between homologous and heterologous titres in cross-focus/plaque reduction neutralization assay).

Factors which determine and influence the pathogenicity of hantaviruses are still poorly understood. However, the nature of the rodent host should be considered as one of the most important factors: SN-like hantaviruses carried by Sigmodontinae rodents, cause, in general, more severe illnesses in humans than HTN-like hantaviruses and PUU carried, respectively, by Murinae and Arvicolinae rodents. For hantaviruses belonging to the same type, reassortant variants (Schmaljohn et al., 1995; Plyusnin et al., 1996 c) as well as strains with an altered pattern of protein glycosylation (Pilaski et al., 1994; Bowen et al., 1995 b) or palmitoylation (Isegawa et al., 1994) have been proposed to possess a higher virulence. On the other hand, genetic susceptibility of patients can also contribute substantially to the outcome of the hantavirus-human interaction: the severity of PUU infection (Mustonen et al., 1996) and the level of virus expression in the patient (Plyusnin et al., 1996 c) are associated with HLA B8 and DRB1*0301 haplotypes.

**Genome structure**

Hantaviruses are negative-strand RNA viruses with a tripartite genome (Fig. 1) which consists of a large (L), medium (M) and small (S) segment encoding, respectively, an L protein (which acts at least as a replicase, transcriptase and endonuclease), a glycoprotein precursor (GPC), which is processed to yield two surface glycoproteins G1 and G2, and a nucleocapsid protein (N) (Elliott et al., 1991; Antic et al., 1992; Elliott, 1996). An open reading frame (ORF) for a putative nonstructural protein NSs (7–10 kDa) was found in BAY,
Fig. 1. Genome structure of hantaviruses. Inset: panhandle-forming 3'- and 5'-terminal sequences of the genomic S, IV] and L
RNA segments of HTN. The genus-specific terminal sequences are tinted.

BCC, ELMC, ILV, KBR, NY, PH, PUU, RIOS, SN and TUL
hantavirus S segments but not in HNT, SEO and DOB (data on
THAI and TPM are not available). The M and/or the S
segments of other members of the family Bunyaviridae encode
nonstructural proteins; bunyaviruses, for instance, contain a
functional ORF of similar size and in about the same region of
the S segment (Elliott, 1989). So far, NS₃ has not been found in
hantavirus-infected cells, but the decreased frequency of
nucleotide substitutions in the region with double coding
capacity suggests that the ORF may be functional in at least
SN, PUU and TUL (Plyusnin et al., 1994a; Spiropoulou et al.,
1994; Bowen et al., 1995 b). The absence of a corresponding
ORF in HTN-like viruses remains to be explained.

The 3' and 5' termini of the hantavirus RNA genome are
highly conserved and complementary and thus capable of
forming panhandle structures (Fig. 1), a hallmark of the
Bunyaviridae (Elliott et al., 1991). Panhandles in hantaviruses are
at least 17 bp long; 14 of the 17 bases from each end are genus-
specific and, in all currently known hantavirus sequences, the
complementarity of the ends is incomplete, with a mismatch at
position 9 (Yoo & Kang, 1987; Parrington & Kang, 1990;
Bowen et al., 1995 b; Chizhikov et al., 1995). A noncanonical
U-G pair in position 10 occurs in HTN, SEO, PUU, SN and
BCC viruses, but not in BAY and PH, where a G → A change
leads to restoration of the canonical pair. The predicted
panhandle-like structures of hantaviruses are thought to serve
a role in the regulation of viral transcription and replication,
similarly to what has been demonstrated for vesicular stomatitis (Wertz et al., 1994) and influenza viruses (Tiley et al., 1994;
Lee & Seong, 1996). The ends of the genome contain three
trinucleotide repeats (5' pUAGUAGUAG) which are sug-
gested to be involved in the proposed prime-and-realign
mechanism of replication resulting in 5' ends containing only
monophosphate (see below).

A complete genome sequence has been so far determined
for five hantavirus genotypes: HTN, SEO, PUU, SN and TUL.
For others only partial sequence information is available:
complete S and M sequences and partial L sequence for BCC
and PH, complete M and S for BAY, DOB, ELMC and NY, and
partial sequences for the others. While the genomic RNAs of
hantaviruses show relatively few differences in the length of
the L and the M segment, the length of the S segment varies
significantly, mainly in its 3' noncoding region (3' NCR)
(Antic et al., 1992; Hjelle et al., 1994; Plyusnin et al., 1994a,
1995 a; Spiropoulou et al., 1994; Morzunov et al., 1995 a;
Ravkov et al., 1995; Song et al., 1995). [Antigenomic sense.
The abbreviation '5' NCR' is used for the noncoding region
located upstream from the ORF and the '3' NCR' for the one
located downstream.) The lengths of the deduced protein
products (L, G1, G2 and N) of the different hantaviruses are

Panhandle-forming terminal nucleotides of the genomic RNA segments

<table>
<thead>
<tr>
<th>Segment</th>
<th>3' Genomic Sense (-)</th>
<th>Antigenomic Sense (+)</th>
</tr>
</thead>
<tbody>
<tr>
<td>S</td>
<td>AUCAUCAUCAUCAUGGC</td>
<td>[...]</td>
</tr>
<tr>
<td>M</td>
<td>AUCAUCAUCAUCAUGGC</td>
<td>[...]</td>
</tr>
<tr>
<td>L</td>
<td>AUCAUCAUCAUCAUGGC</td>
<td>[...]</td>
</tr>
</tbody>
</table>

Hantavirus genome

S

3' Genomic sense (-) 1.8-2.1 kb
Antigenomic sense (+) N protein (~ 430 aa)
5' NCR 3' NCR

M

3' Genomic sense (-) GPC (~ 1140 aa) G1 and G2 glycoproteins
Antigenomic sense (+)
5' NCR 3' NCR

L

3' Genomic sense (-) Viral polymerase (~ 2150 aa)
Antigenomic sense (+)
5' NCR 3' NCR
almost identical. Sequence comparisons of hantaviruses belonging to different sero/genotypes show 60–70% identity at the nucleotide level for all three RNA segments. For the deduced protein products the corresponding values vary from 70–90% for the L to 60–85% for the N and to 50–80% for G1–G2 proteins. Different levels of variability between hantavirus strains belonging to a single serotype have been reported, ranging from 6% for the M segment nucleotides of HTN (Schmaljohn et al., 1988) and 7% of SEO (Kariwa et al., 1994) to 13–14% and 20% for the S segment nucleotides of SN (Spiropoulou et al., 1994; Li et al., 1995) and PUU (Bowen et al., 1995b; Plyusnin et al., 1995a), respectively. The corresponding values for the viral G1–G2 and N proteins are up to 5%.

The S segment 3' NCR represents perhaps the most puzzling part of the hantavirus genome. Within a certain hantavirus type the length and sometimes even the sequence of this region does not undergo dramatic changes suggesting that it has a functional role. In contrast, between different hantavirus types, the S segment 3' NCR varies widely both in length (from 229 nt in PH to 728 nt in SN) and in its nucleotide sequence, except for the terminal nucleotides forming the panhandle structures. Most hantaviruses carry within this region motifs(s) that resemble the sequence 3' CCCACCCAGUCA 5' found at the proposed mRNA termination site in HTN (Dobbs & Kang, 1994) and in the corresponding regions of other bunyaviruses (Dunn et al., 1994; Bowen et al., 1995a; Vapalahti et al., 1996a). Another motif, 3' GAUGAGU 5', with a still unclear function, can be found in single or multiple copies in all hantaviruses close to the highly conserved 5' terminus of the S segment (Ravkov et al., 1995). Also, numerous precise and imprecise repeats with no overall similarity in their structure and pattern may be detected in the 3' NCR. Supposing that the 3' NCR participates in such steps of viral reproduction as packaging, there could be at least two possible explanations for the above-mentioned differences in its primary structure: (i) molecular mechanisms operating at these steps differ from one host to another; or (ii) the secondary rather than the primary structure of the 3' NCR is crucial for its proper activity.

**Genome replication and protein expression**

The hantavirus particle is generally spherical, somewhat pleomorphic and approximately 100–120 nm in diameter, and consists of three circular nucleocapsids each containing one RNA segment and nucleocapsid and polymerase proteins. The two glycoproteins, which are embedded in a lipid envelope, form a grid-like pattern. The budding of the virus takes place in the Golgi and endoplasmic reticulum (Hung et al., 1988). However, detailed morphogenesis and the mechanism of release of virions are not known. The cellular receptor(s) has not been identified and details of the events during virus entry remain unsolved.

The negative-sense genomes serve as templates for the L protein to produce positive-strand genomes and mRNA. The 5' ends of HTN genome segments were shown to contain only monophosphate by Garcin et al. (1995), who suggested a 'prime-and-realign' mechanism for the initiation of both mRNA and genomic RNA synthesis. In the latter, GTP aligns with the third residue (C) at the 3' end of the molecule, priming or fewer nucleotides take place and afterwards the primer 'slides' back one trinucleotide repeat, leaving a protruding GTP non-aligned with the 3' end. This is subsequently cleaved off by the polymerase leaving a pU residue at the 5' end. This system would allow restoration of exact genome ends if the template were to be degraded for some nucleotides, but awaits experimental verification.

The mRNAs of hantaviruses, like those of all members in the family *Bunyaviridae* studied, have 5'-terminal capped, heterogeneous extensions of cellular origin (less than 20 nucleotides in length), which are cleaved by the L protein preferably after a G residue (Fin & Elliott, 1991; Garcin et al., 1995). As proposed for genome RNA synthesis, the G residue at the end of the capped host cell primor probably aligns with the 3' end, and priming, sliding and realigning would take place as for genomic RNA three nucleotides upstream (but without cleavage) before further elongation of the mRNA (Garcin et al., 1995). Alignment may also occasionally take place in the next trinucleotide repeat. Hantavirus mRNAs are not polyadenylated and the exact length of the mRNAs remains to be clarified, although transcription termination motifs have been suggested at least for HTN (see above).

The hantaviral 49–51 kDa N protein consists of 428–433 amino acids (aa) with four conserved cysteines and seems to be expressed in excess in infected cells, and has been reported to form large granular to filamentous inclusion bodies (Hung, 1988). The N-terminal 100 aa are mainly hydrophilic and highly antigenic - most PUU and TUL monoclonal antibodies (MAbs) and several HTN MAbs bind to this region and practically all human sera containing hantavirus antibodies recognize this region. The middle part of the protein, aa 210–310, is a mainly hydrophilic region, highly variable within different hantaviruses, and also contains some MAb epitopes (Jenison et al., 1994; Vapalahti et al., 1995; Elgh et al., 1996; Lundkvist et al., 1995, 1996a, b; Yoshimatsu et al., 1996). The C terminus of the protein is highly conserved and has been shown to bind RNA, preferentially double-stranded, although not specifically viral sequences (Gött et al., 1993). Recombinant hantavirus N proteins are expressed efficiently in bacterial, baculovirus or vaccinia virus systems for use as specific diagnostic antigens, but can also elicit protective immunity in animal models (Schmaljohn et al., 1990; Yoshimatsu et al., 1993) and natural host rodents (Lundkvist et al., 1996a). Recombinant HTN N expressed in insect or mammalian cells has been shown also to form nucleocapsid-like structures similar to authentic nucleocapsids (Bettenbauch et al., 1995).
The glycoprotein precursor of 1132–1148 aa encoded by the M segments cannot be found in infected cells, suggesting that it is cotranslationally cleaved into G1 and G2 proteins, which are thought to form a heterodimer. The G2s of some hantaviruses have an extra ATG preceding or in the G2 signal sequence, but although it can be used to initiate G2 if the normal initiation codon of G1 is experimentally deleted, separate translation of G2 seems to be insignificant (Kamrud & Schmaljohn, 1994). The glycoproteins of hantaviruses have a high content of conserved cysteine residues (5-4%) and virtually identical hydropathy profiles (Antic et al., 1992). The G1 glycoprotein has a predicted signal peptide of 16–23 aa; however, the exact N terminus of the proteins has been determined only for SEO and HTN (Schmaljohn et al., 1987; Arikawa et al., 1990). The external part of G1, 450 aa in length, is followed by an approximately 75 aa hydrophobic membrane associated domain, in which the exact orientation to the lipid bilayer is not known, and a cytoplasmic loop followed by a second, ~ 20 aa domain, which serves as a signal peptide for G2 (Pensiero & Hay, 1992). In all hantaviruses the G1 protein contains an identical motif, WAASA, at the C terminus, after which cleavage has been determined to occur for HTN and SEO (Arikawa et al., 1990; Schmaljohn et al., 1987). In the C terminus of G2 there is a conserved, ca. 30 aa transmembrane domain. G1 contains three N-linked glycosylation sites and G2 one, conserved in all hantaviruses; many hantaviruses have additional sites. Digestion of HTN G1 and G2 with endoglycosidases H and F has shown that both glycoproteins, glycosylation suggesting that the virus is not completely retained in the endoplasmic reticulum. Ruusala et al. (1986a) found that the glycoproteins have to be coexpressed for Golgi localization of either protein while Pensiero & Hay (1992) suggested that G1 alone is sufficient to determine the cellular targeting. However, in the former study, misfolding of a separately expressed G1 was detected by immunofluorescence which may explain the retention in the endoplasmic reticulum. The role of G1 as the protein responsible for Golgi localization is supported also by analogy to other Bunyaviridae (Matsuoka et al., 1994) and by the presence of a conserved motif, CKLTXRFGENLKKSL, in the cytoplasmic part of G1 of the PUU- PH-like and SN-like hantaviruses (Plyusnin et al., 1996a) which has 40% similarity to the proposed Golgi retention signal of Uukuniemi virus, family Bunyaviridae (Pettersson et al., 1995).

The three-dimensional structure of the glycoprotein is complex and most MAbs epitopes seem to be conformational. Using truncated recombinant antigens, epitopes of non-neutralizing MAbs have been mapped to the N-terminal one-third of HTN G1 (Wang et al., 1993); the N terminus of SN G1 has been shown to contain also a linear human B cell epitope (Jenison et al., 1994). For G2, however, truncation to less than 80% of the protein results in loss of reactivity with MAbs. MAbs escape mutants have been genetically characterized for HTN and PUU glycoproteins showing, in each case, single amino acid substitutions. Peptides from the altered regions in PUU M do not react with the MAbs and peptide-antisera do not recognize the native protein, suggesting that these epitopes are non-linear, complicated by glycosylation or binding to regions distant from the mutations (J. Hörling, personal communication). Interestingly, HTN G1 MAbs which did not recognize their 'own' escape mutant failed in addition to recognize a G2 MAbs escape mutant and vice versa, suggesting a close structural interplay between the two glycoproteins (Wang et al., 1993).

Recombinant hantavirus glycoproteins have been produced in baculovirus and with somewhat better expression levels in vaccinia virus systems (Schmaljohn et al., 1990; Xu et al., 1991; Kamrud & Schmaljohn, 1994). However, constructs expressing both glycoproteins result in higher antibody levels and better protection from infection in animal experiments (Schmaljohn et al., 1990). Coexpression of HTN S and M segments generated virus-like particles in the vaccinia virus system, but not efficiently in the baculovirus-insect cell system, relating probably to differences in processing of the two glycoproteins in insect cells as compared with mammalian cells (Betenbauch et al., 1995).

The 2150 aa L protein acts at least as a replicase, transcriptase and endonuclease; however, the function of no hantaviral polymerase has been thoroughly studied. In their N-terminal one-third, hantavirus L proteins have two regions with motifs similar to other segmented negative-strand RNA virus polymerases (Müller et al., 1994) and in the middle one-third there are polymerase motifs, described by Poch et al. (1989), conserved in all RNA-dependent RNA polymerases. In addition, hantaviruses L proteins have also an acidic amino acid-rich C terminus (Stohwasser et al., 1991; Chizhikov et al., 1995).

**Evolution**

In contrast to other members of the family Bunyaviridae, which are arthropod-borne viruses? (Beatty & Calisher, 1991), members of the genus Hantavirus are rodent- (or insectivore- ) borne, as indicated by the fact that rodent-to-rodent transmission is sufficient for maintenance of the virus within laboratory rodent colonies (Lee et al., 1981; Yanagihara et al., 1985). (†Tospoviruses are not transmitted to a vertebrate host and, therefore are not true arboviruses.) This difference is mirrored in the phylogenetic tree based on the L protein sequences (Chizhikov et al., 1995): while bunya-, phlebo- and tospoviruses seem to share a common ancient ancestor, hantaviruses form an independent group suggesting their distinct evolutionary origin.

The phylogenetic relatedness between hantaviruses resembles that of their natural hosts (Wilson & Reeder, 1993).
supporting the hypothesis of hantavirus–host coevolution (Antic et al., 1992; Plyusnin et al., 1994a; Spiropoulou et al., 1994; Xiao et al., 1994; Morzunov et al., 1995b). On the phylogenetic tree calculated for the S segment nucleotide sequences (Fig. 2) hantaviruses are grouped in three separate clades — (i) HTN-like (HTN, SEO and DOB), (ii) PUU–PH-like (PUU, PH, TUL, KBR and ISL) and (iii) SN-like (SN, NY, BAY, BCC, ELMC and RIOS) — carried by Murinae, Arvicolinae and Sigmodontinae rodents, respectively. Comparison of the limited nucleotide sequence of TPM, the only known hantavirus carried by an insectivore, shows that this virus, while diverged from all other rodent-borne hantaviruses, never-
theless shares a common ancestor with them (Xiao et al., 1994). Assuming that there was no re-introduction of hantaviruses into evolving Rodentia or Insectivora, it seems possible that an ancient hantavirus coexisted with their ancient predecessors, i.e. in the Cretaceous period, about 100 million years before the present (Eisenberg, 1981).

Phylogenetic trees calculated for the smaller numbers of hantavirus M and L sequences show branching orders similar to that of the S-gene tree suggesting similar evolutionary pathways for all three genome segments. Indeed, thus far no genome reassortment has been found between hantaviruses of different genotypes (Xiao et al., 1994; Nichol et al., 1996). Within individual hantavirus types, genetic variants show geographical clustering (Fig. 2): thus PUU strains from Finland (PUU/Sotk), Sweden (PUU/Vind) and Russia (PUU/Bash), SN strains from the Four Corners area (SN/H10) and California (SN/CC), and TUL strains from Russia (TUL/Tula) and Central Europe (TUL/Mal and TUL/Mor) are situated on the tree according to their geographical origins. Clustering of genetic variants has, in addition, been demonstrated for ELMC, PH and SEO (Hjelle et al., 1994; Kariwa et al., 1994; Rowe et al., 1995), and therefore might be regarded as a general feature of hantaviruses. The time of hantavirus isolation (or collection of virus-containing specimen) seems to play a minor role (Kariwa et al., 1994; Plyusnin et al., 1995 a) and the evolutionary pattern of hantaviruses resembles that of vesicular stomatitis virus (Nichol et al., 1993 b) with a correlation to the geographical distribution of the agent rather than to the time dimension.

Since the evolution of hantaviruses follows that of their natural carriers and human infection is merely an accidental dead-end event, the primary animal host should be regarded as the evolutionary scene for these agents. Persistent infections of the natural hosts provide the opportunity for both genetic drift (accumulation of base substitutions and deletions/insertions) and genetic shift (reassortment of genome segments), the two main mechanisms driving RNA virus evolution (Elliot, 1996). Genetic drift seems to be the main source of hantavirus genetic variation (Schmaljohn et al., 1986; Kariwa et al., 1994; Plyusnin et al., 1994 a, b, 1995 a, b; Spirigouloa et al., 1994; Ravkov et al., 1995; Rowe et al., 1995). While heterologous (inter-type) reassortment has not been observed for hantaviruses, there is evidence for reassortment of genomic RNA segments between strains belonging to the same type (Henderson et al., 1995; Li et al., 1995). Increased virulence for humans has been proposed for such reassortant variants of pathogenic hantaviruses (Schmaljohn et al., 1995; Plyusnin et al., 1996 c), but more sequences from human cases of hantavirus infection, both HFRS and HPS, are needed to clarify this point.

As shown for PUU and TUL, populations of wild-type hantaviruses within an individual rodent are represented by complex mixtures of different but closely related variants – quasispecies (Plyusnin et al., 1994 a, 1995 a, 1996 c). Thus, hantaviruses resemble other RNA viruses existing in the form of heterogeneous populations. The likely basic reason for this is base misincorporation by viral RNA-dependent RNA polymerases in the absence of proof-reading and repair mechanisms. In general, the quasispecies nature of viral RNA populations allows for rapid evolution, via selection of pre-existing variants, resulting in the establishment of altered mutant spectra with a higher fitness to a new environment. Being highly adaptable and rapidly changeable, quasispecies populations might, nevertheless, exhibit long periods of stasis in a certain environment, if the master geno/phenotype has high fitness to it (Holland et al., 1992; Domingo & Holland, 1994). Most probably, the latter is the case for hantaviruses, well adapted to their natural hosts long ago. It does not look as if there would be strong immunological pressures operating in the persistently infected natural hosts in which both viral antigen and neutralizing antibodies are commonly detected. The current hantaviruses should be regarded as rather stable agents. This point is illustrated, for instance, by the coexistence of two SN sublineages in deer mice in Nevada and by the cocirculation of several different hantavirus genotypes in the same territory (Rowe et al., 1995).

Data on TUL S/N quasispecies (Plyusnin et al., 1996 d) suggest that the random fixation of quasi-neutral mutations may play a role in hantavirus evolution. There may be more freedom for virus variability within an individual rodent; in fact, a larger proportion of the nucleotide substitutions in quasispecies result in amino acid changes. On the other hand, as with other RNA viruses, there is no evidence to suggest that this mechanism contributes disproportionally more than selection. Selection might be involved at the stage of virus transmission from one rodent to another, when randomly sampled aliquots of virus population are distributed into another environment, and might operate preferentially against changes in master geno- and phenotypes. Indeed, the resulting nucleotide substitutions between master genotypes of different hantavirus strains are mostly silent and deduced amino acid substitutions are mostly synonymous.

Geographical clustering of genetic variants is without doubt connected to the fact that small rodents such as mice and voles are very local; usually, they spend their life-time within an area of a few square kilometres. In the past, however, species have been engaged in intensive and complicated migrations. The current populations of mice and voles in Europe and North America were established in the process of recolonizing of these territories during the postglacial period, about 10 000 years before the present, and from that time are thought to have been quite stable in location. Thus, geographically isolated rodent populations allow comparison of genetic properties of hantavirus strains that were separated long ago and have been evolving independently since then. From this point of view, Fennoscandia may constitute one of the best regions to study both faunal history (Siivonen, 1982) and coevolution of hantaviruses together with their natural hosts. For example, the contact zone between two major phylo-
geographical groups of the bank vole, which recolonized the territory of current central Sweden from two directions (the north and south) is still only 50 km wide (Jaarola, 1995), and, most notably, PUU strains circulating on both sides of the population border in Sweden belong to two different genetic sublineages (Höring et al., 1996b). Similarly, geographically separated PUU strains from Finland and from pre-Ural Russia form two sublineages (Plyusnin et al., 1995a) which most probably originated from two different, more-recent ancestors present in separate glacial refugia of bank voles. In North America, on the eastern slopes of the Sierra Nevada range, SN sublineages of Type 1 and Type 2 are present, while on the western slope only Type 1 is found, suggesting the importance of geographical location for the distribution of genetic variants (Rowe et al., 1995).

**Concluding remarks**

We know today that many species and populations of rodents are infected by hantaviruses. These viruses have adapted to their natural hosts during coevolution, and produce chronic infection with no apparent harm. As is often the case, transmission of a virus to a new host may result in severe disease, such as HFRS and HPS in human hantavirus infection. Human epidemics are associated with fluctuations of rodent populations, caused by climatic, ecological and environmental changes, or changes in human activities associated with nature or agriculture. These changes make it possible for known hantaviruses to re-emerge and maybe also for new infections to emerge from the plethora of rodent species. The hantaviruses themselves, or their virulence to man, have not changed and there is no reason to expect abrupt changes in the future.

On the other hand hantaviruses, being confined to their specific host rodents or insectivores and not spread by mobile species such as insects, birds, or humans, provide a good model for evolutionary studies ranging from quasispecies in an individual host rodent to coevolution with their natural host in the different continents of the world.

From a taxonomic point of view, it seems that the most important factor determining the classification of hantaviruses is a separate main rodent/insectivore reservoir characteristic of each of the 16 hantaviruses. This classification is supported by genetic comparisons and phylogeny and, for most hantaviruses, by cross-neutralization assays also.

We have focused in this review on the genetic properties of the hantaviruses and therefore, a large body of important work concerning immunology, pathogenesis, diagnostics and vaccine development is not covered. The global map of hantaviruses has become more accurate, but many geographical areas and rodent species are yet to be explored as well as the association of hantaviruses with human infection and disease. Much work also still remains to be carried out to delineate the replication cycle, genetic determinants that influence virus-host cell interaction, tissue tropism and host specificity of hantaviruses.

We thank Tatiana Avisic-Zupanc, Brian Hjelle, Jan Hörling, Åke Lundkvist, Mani Kanerva, Kimiko Yoshimatsu, Stephen C. St. Jeor and Stuart T. Nichol for sharing with us unpublished data. Helpful comments from Heikki Lehväslaiho, Åke Lundkvist and Heikki Henkitonen are greatly appreciated.

**References**


