Arch Virol (1996) 141: 959–967

Archives of Virology

© Springer-Verlag 1996
Printed in Austria

Genetic and phylogenetic analyses of hantaviral sequences amplified from archival tissues of deer mice (*Peromyscus maniculatus nubiterrae*) captured in the eastern United States*

Brief Report

J.-W. Song[1], **L. J. Baek**[1], **J. W. Nagle**[2], **D. Schlitter**[3], and **R. Yanagihara**[1]

[1] Laboratory of Central Nervous System Studies and [2] Neurogenetics Section, National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, Maryland, [3] Section of Mammals, Carnegie Museum of Natural History Pittsburgh, Pennsylvania, U.S.A.

Accepted December 23, 1995

Summary. The S and M segments of a hantavirus, enzymatically amplified from tissues of Cloudland deer mice (*Peromyscus maniculatus nubiterrae*) captured during 1985 in West Virginia, diverged from strains of Four Corners virus from the southwestern United States by more than 16% and 6% at the nucleotide and amino acid levels, respectively. Phylogenetic analysis suggested that this virus strain (designated Monongahela) forms a possible evolutionary link between the Four Corners and New York hantaviruses.

*

An outbreak of a rapidly progressive respiratory illness with high mortality was reported in the spring of 1993 in the southwestern United States [3, 18]. Collective serologic and genetic data indicated that a newly identified hantavirus, known variously as Four Corners (FC) virus, Muerto Canyon virus, Convict Creek virus and Sin Nombre virus, caused this previously unrecognized clinical entity, now known as hantavirus pulmonary syndrome (HPS) [9, 14, 26]. The deer mouse (*Peromyscus maniculatus*) has been identified as the principal rodent reservoir of FC virus [8]. Like other members of the *Bunyaviridae* family, FC virus and other hantaviruses are enveloped viruses with a tripartite negative-

* This work was presented in part at the Third International Conference on HFRS and Hantaviruses, Helsinki, Finland in June 1995 and at the 14th Annual Meeting of the American Society for Virology in Austin, Texas in July 1995.

sense RNA genome. The segments, designated large (L), medium (M) and small (S), encode a putative viral transcriptase, the G1 and G2 envelope glycoproteins, and a nucleocapsid protein, respectively [10]. Each of the genetically and biologically distinct hantaviruses is harbored by a specific rodent or insectivore reservoir [39, 41].

Of the more than 100 HPS cases diagnosed in the United States and Canada since June 1993, the vast majority have occurred in the southwestern United States, particularly in the Four Corners region, where the borders of New Mexico, Arizona, Utah and Colorado meet. To date, only 7 HPS cases have been reported east of the Mississippi River. Of these, 4 cases have occurred in regions devoid of deer mice [4, 5, 21], and 3 cases have been reported from within the habitation range of the deer mouse (one case each in Indiana, West Virginia and Virginia). Not unexpectedly then, while a substantial genetic database has been amassed about the FC hantavirus harbored by deer mice in the western United States, little information is currently available for deer mouse-borne hantaviruses in the eastern United States.

To clarify this issue, we have retrieved from −80 °C storage lung tissues collected from 3 hantavirus-seropositive Cloudland deer mice (*Peromyscus maniculatus nubiterrae*), captured on July 25 and August 2, 1985 at two sites, approximately 7 mi apart, within the Monongahela National Forest in West Virginia (Pocahontas Co., 2.5 mi N, 3.75 mi W of Mill Point; and Webster Co., 4.75 mi S, 4.25 mi E of Bolair). As determined by the indirect immunofluorescent antibody technique, the 3 deer mice had IgG antibodies against Prospect Hill (PH) and Puumala (PUU) viruses (titer, 1:1024). Attempts to isolate hantaviruses from these tissues in 1985 were unsuccessful [1]. Total RNA, extracted from these archival tissues using RNAzol B (Tel-Test, Inc., Friendswood, TX), was examined for hantaviral sequences by reverse transcription-polymerase chain reaction, using oligonucleotide primers designed from sequences of FC virus [13, 19, 34], as well as a hantavirus (RI-1) amplified from postmortem tissues of an HPS case-patient, who was exposed in New York but died in Rhode Island [15, 16], and New York (NY) virus, a hantavirus isolated from a white-footed mouse (*Peromyscus leucopus*) captured on Shelter Island, New York [33]. Nucleotide sequences of 3 to 12 clones from each amplimer were determined in both directions using the Dye Primer Cycle Sequencing Ready Reaction Kit (Applied Biosystems Inc., Foster City, CA) on an automated sequencer (model 373A, Applied Biosystems Inc.). Hantaviral sequences from the 3 deer mice, designated Monongahela-1, -2 and -3, have been deposited in GenBank: S segment (U32646 to U32648, U32591); M segment (U32649 to U32656).

The complete S genomic segment of Monongahela-1, as amplified from RNA extracted from lung tissue of a Cloudland deer mouse captured near Mill Point, was 2082 nucleotides (nt) long. The virus complementary strand contained a large open reading frame (ORF) of 1284 nt beginning at position 43, capable of encoding a 428-amino acid nucleocapsid (N) protein, identical in length to that of other sigmodontine rodent-borne hantaviruses [34]. As determined by the Chou-Fasman algorithm, the secondary structures of the N protein of Monon-

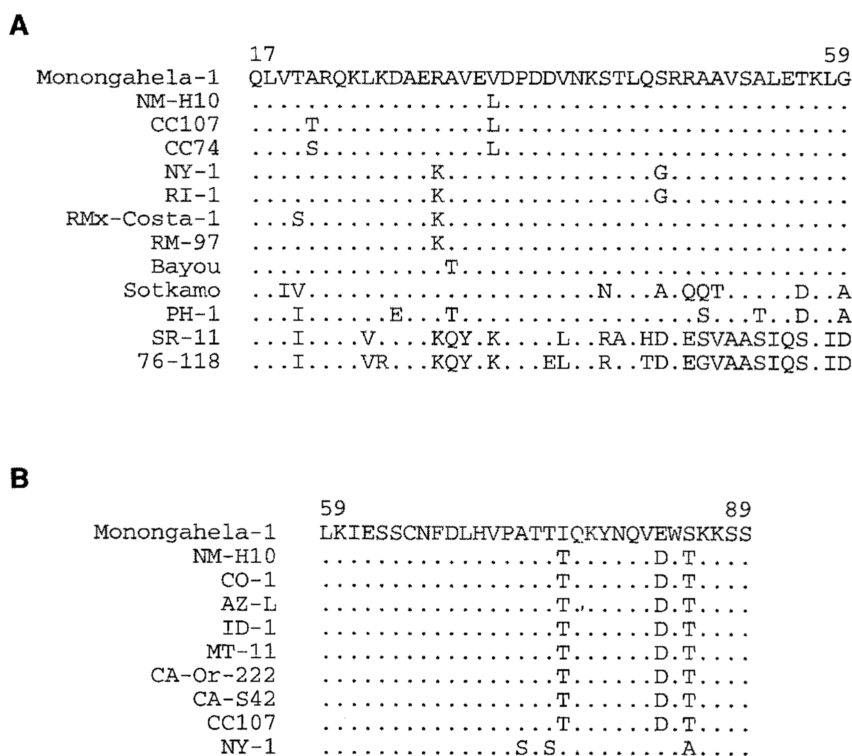

Fig. 1. Alignment and comparison of the deduced amino acid sequences of the immunodominant domains on **A** the nucleocapsid protein and **B** the G1 envelope glycoprotein of Monongahela-1, a hantavirus harbored by Cloudland deer mice (*Peromyscus maniculatus nubiterrae*), captured in West Virginia in 1985, as well as hantavirus strains amplified from postmortem tissues of HPS case-patients and from peromyscine rodents, including Four Corners virus strains CC74, CC107 and NM-H10, and New York virus strains NY-1 and RI-1. Comparison is also shown for other genetically distinct hantaviruses: El Moro Canyon virus strains RM-97, Rio Segundo virus strain RMx-Costa-1, Bayou virus, Prospect Hill virus strain Prospect Hill-1, Puumala virus strains Sotkamo, Seoul virus strain SR-11, and Hantaan virus strain 76–118. The amino acid coordinates of the first and last residue of each sequence are indicated. Identical residues are indicated by periods

gahela-1 and FC virus were nearly indistinguishable, and in the immunodominant domain on the N protein, defined by residues 17 to 59, Monongahela-1 differed from FC virus and other sigmodontine rodent-borne hantaviruses by only 1 or 2 amino acids (Fig. 1a).

Alignment and comparison of the entire N protein-encoding region of the S genomic segment revealed 83% and 94% nucleotide- and amino acid-sequence similarity, respectively, between Monongahela-1 and strains of FC virus amplified from postmortem tissues of HPS case-patients and from tissues of deer mice from the western United States. Monongahela-1, in turn, diverged from NY virus by 15.3% and 3.5% at the nucleotide and amino acid levels, respectively. Transitions, primarily thymine to cytosine or vice versa, accounted for most (62%) of the nucleotide substitutions between Monongahela-1 and FC virus.

As in FC and NY viruses [14, 19, 34, 36], a second, much smaller ORF of 189 nt (coordinates 122 to 310) was found in the S segment of Monongahela-1. The sequence of this 63-amino acid hypothetical nonstructural protein (NSx) was more similar to that of FC virus (89% identity) than to the NSx of other sigmodontine or arvicolid rodent-borne hantaviruses. The 756-nt 3′-untranslated or noncoding region of the Monongahela-1 S segment, which was comprised of numerous perfect and imperfect repeats, was similar in length to that of FC and NY viruses (728 to 757 nt) [19, 36]. Of 34 repeats having a minimum repeat length of 12 nt with 1 mismatch allowed, 33 repeats were located in the noncoding region and 1 was in the coding region.

Analysis of the 274-nucleotide, immunodominant epitope-spanning domain of the G1 envelope glycoprotein-encoding M genomic segment indicated that Monongahela-1, -2 and -3 diverged from FC virus by more than 15% and 6% at the nucleotide and amino acid levels, respectively. In this highly conserved immunodominant domain, defined by amino acids 59 to 89 [13], Monongahela virus strains differed by 3 residues (Fig. 1b) from NY virus and from multiple strains of FC virus from throughout the western United States and Canada [13, 16]. In a 1140-nt region of the G2-encoding M segment, Monongahela differed from FC and NY viruses by 2.6% and 2.9%, respectively, at the amino acid level, despite the 81.1% to 82.8% nucleotide sequence similarity. To what extent these differences in the G1 envelope glycoprotein sequences are important from the standpoint of virus neutralization is unclear.

Using the maximum parsimony and neighbor-joining methods [11, 37], as well as unweighted pair-group method of assortment, phylogenetic trees based on the nucleotide and deduced amino acid sequences of the coding region of the S genomic segment indicated that Monongahela shared a common ancestry with FC and NY viruses and with other hantaviruses harbored by sigmodontine rodents (Fig. 2). These topologies were strongly supported by bootstrap analysis of 1 000 iterations, but the phylogenetic position of Monongahela varied between deer mouse and white-footed mouse-borne hantaviruses, suggesting that Monongahela formed an evolutionary link between FC and NY hantaviruses. In trees based on the M genomic segment spanning the immunodominant region of the G1 glycoprotein, Monongahela was the most divergent of all deer mouse-borne hantaviruses and formed an outgroup for strains of FC virus from the western United States (data not shown).

Occasionally, the speciation of *Peromyscus* rodents can be problematic, particularly in geographic regions where *Peromyscus maniculatus* and *Peromyscus leucopus* are sympatric and synchronistic. In the present study, *Peromyscus maniculatus nubiterrae* were distinguished from *Peromyscus leucopus noveboracensis* by their long gray-colored dorsal pelage (versus short reddish-brown dorsal pelage, typically with a broad mid-dorsal band of dark hairs); their well-haired and bicolored long tails, measuring over 95 mm, usually exceeding 100 mm, and ending in a pencillated tip (versus short tails, covered with scattered hairs, and not effectively bicolored, measuring less than 95 mm, usually less than 90 mm, and not ending with a tufted tip); and their large and long ears, usually

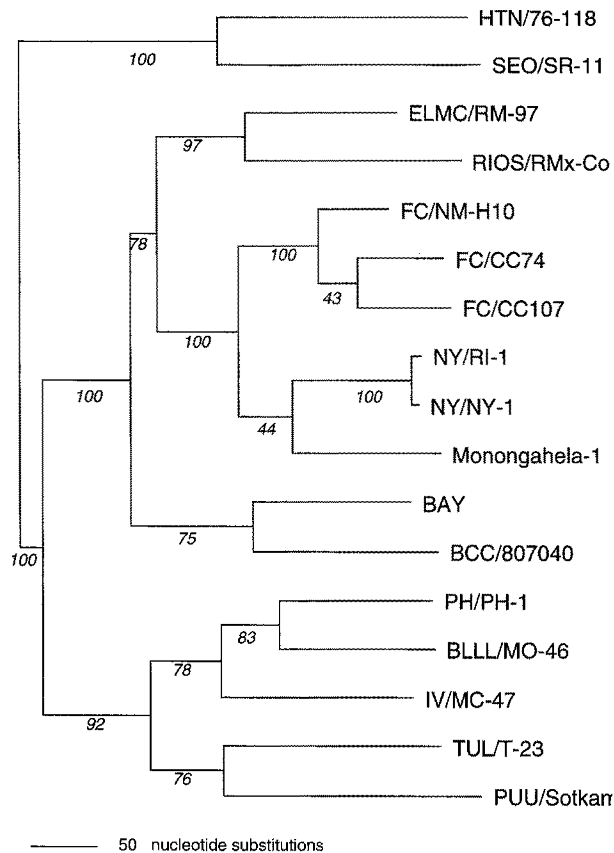

Fig. 2. Phylogenetic tree, based on the entire nucleocapsid-encoding S genomic segment (nucleotides 43 to 1 326) of Monongahela-1, constructed by the maximum parsimony method. A single parsimonious tree, rooted at the midpoint of the greatest patristic distance, was generated. The phylogenetic position of Monongahela-1 is shown in relation to a hantavirus (RI-1) amplified from postmortem tissues of an HPS case-patient, who was exposed in New York but died in Rhode Island [15, 16] and a hantavirus isolate (NY-1) from a white-footed mouse (*Peromyscus leucopus*) captured in New York [33], as well as other sigmodontine rodent-borne hantaviruses, including Four Corners (FC) virus strains CC74, CC107 and NMH-10; El Moro Canyon (ELMC) virus strain RM-97; Rio Segundo (RIOS) virus strain RMx-Costa-1; Bayou (BAY) virus; and Black Creek Canal (BCC) virus strain 807040. In addition, phylogenetic relationships are shown to Prospect Hill (PH) virus strain Prospect Hill-1, Bloodland Lake (BLLL) virus strain MO-46, Isla Vista (IV) virus strain MC-47, Tula (TUL) virus strain T23, Puumala (PUU) virus strain Sotkamo, Seoul (SEO) virus strain SR-11 and Hantaan (HTN) virus strain 76–118. Branch lengths are proportional to the number of nucleotide substitutions, while vertical distances are for clarity only. Bootstrap probabilities (expressed as percentages), as determined for 1 000 iterations by PAUP version 3.1.1 [37], are shown in italics for each adjacent node

more than 20 mm and often exceeding 22 mm (versus short ears, less than 20 mm). As determined by these morphological criteria, and as verified by mitochondrial DNA analysis, the 3 *Peromyscus* rodents in this study were identified as *Peromyscus maniculatus nubiterrae*. The geographic distribution

of this subspecies of deer mice extends from the northern tip of Georgia and South Carolina, through western North Carolina and eastern Tennessee, Virginia and West Virginia, and as far north as central Pennsylvania and the western half of New York [12]. Since there is considerable overlap with *Peromyscus leucopus noveboracensis*, it is uncertain whether the hantaviruses responsible for the retrospectively diagnosed, nonfatal cases of HPS occurring in Virginia and West Virginia were genetically more similar to Monongahela-1 than to NY virus.

Data presented in this report support our earlier demonstration that hantaviruses have long been enzootic in deer mouse populations [22, 24]. More importantly, these data indicate the existence of a highly divergent and phylogenetically distinct hantavirus in deer mice in the eastern United States. Recent comparisons of S and M segment sequences of PUU virus amplified from tissues of bank voles (*Clethrionomys glareolus*) trapped in Udmurtia and PUU virus isolates from Bashkiria and Sotkamo indicate that the degree of sequence similarity between PUU virus strains is dependent on geographic origin [27]. Similarly, sequence analyses of HTN and FC viruses indicate geographic dependence [13, 26, 35, 36].

Genotypic segregation of HTN, PUU and FC virus strains, according to their geographic origin, is consistent with the concept, originally proposed for St. Louis encephalitis virus [38], that viral genotypes are associated with specific geographic regions. Geographic-specific viral genotypes or topotypes have since been identified for dengue virus types 1 and 2 [2, 28], Japanese encephalitis virus [7], hepatitis C virus [31], poliovirus type 1 [29], hepatitis A virus [30], Ross River virus [20], rabies virus [32], vesicular stomatitis virus [25], human T-cell lymphotropic virus type I [23, 40, 42], and human papillomavirus type 16 [6, 17]. Further analyses of peromyscine rodent-borne hantaviruses from the eastern United States will be required to validate the apparent geographic-specific genetic and phylogenetic segregation of Monongahela-1 from strains of FC virus from the western United States and Canada. Such studies will also clarify if peromyscine rodent-borne hantaviruses like Monongahela-1 serve to link the evolutionary continuum between FC and NY viruses.

Acknowledgements

We thank Dr. Brian Hjelle for sharing unpublished sequences and Ms. Anne Book for technical assistance. Mitochondrial DNA sequences of deer mice were determined by Dr. Sergey Morzunov.

References

1. Baek LJ, Yanagihara R, Gibbs CJ Jr, Miyazaki M, Gajdusek DC (1988) Leakey virus: a new hantavirus isolated from *Mus musculus* in the United States. J Gen Virol 69: 3129–3132
2. Blok J, Gibbs AJ, McWilliam SM, Vitarana UT (1991) NS 1 gene sequences from eight dengue-2 viruses and their evolutionary relationship with other dengue-2 viruses. Arch Virol 118: 209–223

3. Centers for Disease Control and Prevention (1993) Outbreak of acute illness – southwestern United States, 1993. Morb Mortal Wkly Rep 42: 421–424
4. Centers for Disease Control and Prevention (1994) Hantavirus pulmonary syndrome – northeastern United States, 1994. Morb Mortal Wkly Rep 43: 548–549, 555–556
5. Centers for Disease Control and Prevention (1994) Newly identified hantavirus – Florida, 1994. Morb Mortal Wkly Rep 43: 99, 105
6. Chan S-Y, Ho L, Ong C-K, Chow V, Drescher B, Dürst M, ter Meulen J, Villa L, Luande J, Mgaya HN, Bernard H-U (1992) Molecular variants of human papillomavirus type 16 from four continents suggest ancient pandemic spread of the virus and its coevolution with humankind. J Virol 66: 2057–2066
7. Chen W-R, Rico-Hesse R, Tesh RB (1992) A new genotype of Japanese encephalitis virus from Indonesia. Am J Trop Med Hyg 47: 61–69
8. Childs JE, Ksiazek TG, Spiropoulou CF, Krebs JW, Morzunov S, Maupin GO, Gage KL, Rollin PE, Sarisky J, Enscore RE, Frey JK, Peters CJ, Nichol ST (1994) Serologic and genetic identification of *Peromyscus maniculatus* as the primary rodent reservoir for a new hantavirus in the southwestern United States. J Infect Dis 169: 1271–1280
9. Duchin JS, Koster FT, Peters CJ, Simpson GL, Tempest B, Zaki SR, Ksiazek TG, Rollin P, Nichol S, Umland E, Moolenaar RL, Reef SE, Nolte KB, Gallaher MM, Butler JC, Breiman RF, the Hantavirus Study Group (1994) Hantavirus pulmonary syndrome: a clinical description of 17 patients with a newly recognized disease. N Engl J Med 330: 949–955
10. Elliott RM, Schmaljohn CS, Collett MS (1991) Bunyavirus genome structure and gene expression. Curr Top Microbiol Immunol 169: 91–141
11. Felsenstein J (1993) PHYLIP (phylogeny inference package), version 3.5c. University of Washington, Seattle
12. Hall ER (1991) The Mammals of North America, vol 2, 2nd ed. New York, pp 777–838
13. Hjelle B, Chavez-Giles R, Torrez-Martinez N, Yamada T, Sarisky J, Ascher M, Jenison S (1994) Dominant glycoprotein epitope of Four Corners hantavirus is conserved across a wide geographic area. J Gen Virol 75: 2881–2888
14. Hjelle B, Jenison S, Torrez-Martinez N, Yamada T, Nolte K, Zumwalt R, MacInnes K, Myers G (1994) A novel hantavirus associated with an outbreak of fatal respiratory disease in the southwestern United States: evolutionary relationships to known hantaviruses. J Virol 68: 592–596
15. Hjelle B, Krolikowski J, Torrez-Martinez N, Chavez-Giles F, Vanner C, Laposata E (1995) Phylogenetically distinct hantavirus implicated in a case of hantavirus pulmonary syndrome in the northeastern United States. J Med Virol 46: 21–27
16. Hjelle B, Lee S-W, Song W, Torrez-Martinez N, Song J-W, Yanagihara R, Gavrilovskaya I, Mackow ER (1995) Molecular linkage of hantavirus pulmonary syndrome to the white-footed mouse, *Peromyscus leucopus*: genetic characterization of the M genome of New York virus. J Virol 69: 8137–8141
17. Ho L, Chan S-Y, Burk RD, Das BC, Fujinaga K, Icenogle JP, Kahn T, Kiviat N, Lancaster W, Mavromara-Nazos P, Labropoulou V, Mitrani-Rosenbaum S, Norrild B, Pillai MR, Stoeker J, Syrjaenen K, Syrjaenen S, Tay S-K, Villa LL, Wheeler CM, Williamson A-L, Bernard H-U (1993) The genetic drift of human papillomavirus type 16 is a means of reconstructing prehistoric viral spread and the movement of ancient human populations. J Virol 67: 6413–6423
18. Hughes JM, Peters CJ, Cohen ML, Mahy BWJ (1993) Hantavirus pulmonary syndrome: an emerging infectious disease. Science 262: 850–851

19. Li D, Schmaljohn AL, Anderson K, Schmaljohn CS (1995) Complete nucleotide sequences of the M and S segments of two hantavirus isolates from California: evidence for reassortment in nature among viruses related to hantavirus pulmonary syndrome. Virology 206: 973–983
20. Lindsay MDA, Coelen RJ, MacKenzie JS (1993) Genetic heterogeneity among isolates of Ross River virus from different geographical regions. J Virol 67: 3576–3585
21. Morzunov SP, Feldmann H, Spiropoulou C, Semenova VA, Rollin PE, Ksiazek TG, Peters CJ, Nichol ST (1995) A newly recognized virus associated with a fatal case of hantavirus pulmonary syndrome in Louisiana. J Virol 69: 1980–1983
22. Nerurkar VR, Song J-W, Song K-J, Nagle J-W, Hjelle B, Jenison S, Yanagihara R (1994) Genetic evidence for a hantavirus enzootic in deer mice (*Peromyscus maniculatus*) captured a decade before the recognition of hantavirus pulmonary syndrome. Virology 204: 563–568
23. Nerurkar VR, Song K-J, Bastian IB, Garin B, Franchini G, Yanagihara R (1994) Genotyping of human T-cell lymphotropic virus type I strains using Australo-Melanesian topotype-specific oligonucleotide primer-based polymerase chain reaction: insights into virus evolution and dissemination. J Infect Dis 170: 1353–1360
24. Nerurkar VR, Song K-J, Gajdusek DC, Yanagihara R (1993) Genetically distinct hantavirus in deer mice. Lancet 342: 1058–1059
25. Nichol ST, Rose JE, Fitch WM (1993) Punctuated equilibrium and positive Darwinian evolution in vesicular stomatitis virus. Proc Natl Acad Sci USA 90: 10424–10428
26. Nichol ST, Spiropoulou CF, Morzunov S, Rollin PE, Ksiazek TG, Feldmann H, Sanchez A, Childs J, Zaki S, Peters CJ (1993) Genetic identification of a hantavirus associated with an outbreak of acute respiratory illness. Science 262: 914–917
27. Plyusnin A, Vapalahti O, Ulfves K, Lehväslaiho H, Apekina N, Gavrilovskaya I, Blinov V, Vaheri A (1994) Sequences of wild Puumala virus genes show a correlation of genetic variation with geographic origin of the strains. J Gen Virol 75: 405–409
28. Rico-Hesse R (1990) Molecular evolution and distribution of dengue viruses type 1 and 2 in nature. Virology 174: 479–493
29. Rico-Hesse R, Pallansch M, Nottay BK, Kew OM (1987) Geographic distribution of wild poliovirus type 1 genotypes. Virology 160: 311–322
30. Robertson BH, Jansen RW, Khanna B, Totsuka A, Nainan OV, Siegl G, Widell A, Margolis HS, Isomura S, Ito K, Ishizu T, Moritsugu Y, Lemon SM (1992) Genetic relatedness of hepatitis A virus strains recovered from different geographical regions. J Gen Virol 73: 1365–1377
31. Simmonds P, Holmes EC, Cha T-A, Chan SW, McOmish F, Irvine B, Beall E, Yap PL, Kolberg A, Urdea MS (1993) Classification of hepatitis C virus into six major genotypes and a series of subtypes by phylogenetic analysis of the NS-5 region. J Gen Virol 74: 2391–2399
32. Smith JS, Orciari LA, Yager PA, Seidel HD, Warner CK (1992) Epidemiologic and historical relationships among 87 rabies virus isolates as determined by limited sequence analysis. J Infect Dis 166: 296–307
33. Song J-W, Baek LJ, Gajdusek DC, Yanagihara R, Gavrilovskaya I, Luft BJ, Mackow ER, Hjelle B (1994) Isolation of pathogenic hantavirus from white-footed mouse (*Peromyscus leucopus*). Lancet 344: 1637
34. Song J-W, Baek LJ, Gavrilovskaya IN, Mackow ER, Hjelle B, Yanagihara R (1995) Sequence analysis of the complete S genomic segment of a newly identified hantavirus isolated from the white-footed mouse (*Peromyscus leucopus*): phylogenetic relationship with other sigmodontine rodent-borne hantaviruses. Virus Genes (in press)

35. Song K-J, Kim SY, Choi YJ, Seong IW, Song J-W, Lee YJ, Yanagihara R (1995) Hantaan and Seoul virus infection: phylogenetic analysis of viral sequences amplified from sera of Korean patients with hemorrhagic fever with renal syndrome. Presented at the Third International Conference on HFRS and Hantaviruses, Helsinki, Finland, May 31–June 3
36. Spiropoulou CF, Morzunov S, Feldmann H, Sanchez A, Peters CJ, Nichol ST (1994) Genome structure and variability of a virus causing hantavirus pulmonary syndrome. Virology 200: 715–723
37. Swofford DL (1993) PAUP: phylogenetic analysis using parsimony, version 3.1. Illinois Natural History Survey, Champaign, Illinois
38. Trent DW, Grant JA, Vorndam AV, Monath TP (1981) Genetic heterogeneity among Saint Louis encephalitis virus isolates of different geographic origin. Virology 114: 319–332
39. Yanagihara R (1990) Hantavirus infection in the United States: epizootiology and epidemiology. Rev Infect Dis 12: 449–457
40. Yanagihara R (1994) Geographic-specific genotypes or topotypes of human T-cell lymphotropic virus type I as markers for early and recent migrations of human populations. Adv Virus Res 43: 147–186
41. Yanagihara R, Gajdusek DC (1988) Hemorrhagic fever with renal syndrome: a historical perspective and review of recent advances. In Gear JHS (ed) CRC handbook of viral and rickettsial hemorrhagic fevers. CRC Press, Boca Raton, pp 151–188
42. Yanagihara R, Saitou N, Nerurkar VR, Song K-J, Bastian I, Franchini G, Gajdusek DC (1995) Molecular phylogeny and dissemination of human T-cell lymphotropic virus type I viewed within the context of primate evolution and human migration. Cell Mol Biol 41: S145–S161

Authors' address: Dr. R. Yanagihara, Retrovirology Research Laboratory, Leahi Hospital, 3675 Kilauea Avenue, Honolulu, HI 96816, U.S.A.

Received October 20, 1995