Recent Discoveries of New Hantaviruses Widen Their Range and Question Their Origins

Heikki Henttonen,[a] Philippe Buchy,[b] Yupin Suputtamongkol,[c] Sathaporn Jittapalapong,[d] Vincent Herbreteau,[e] Juha Laakkonen,[a] Yannick Chaval,[f] Maxime Galan,[f] Gauthier Dobigny,[f] Nathalie Charbonnel,[f] Johan Michaux,[f] Jean-François Cosson,[f] Serge Morand,[g] and Jean-Pierre Hugot[h]

[a]*Finnish Forest Research Institute, Vantaa Research Centre, Vantaa, Finland*

[b]*Institut Pasteur in Cambodia, Virology Unit, Phnom Penh, Cambodia*

[c]*Department of Medicine, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand*

[d]*Kasetsart University, Department of Parasitology, Bangkok, Thailand*

[e]*Laboratoire Espace, Santé, Territoire,Université Paris X-Nanterre, Nanterre, France*

[f]*CBGP, Campus International de Baillarguet, Montferrier-Sur-Lez, France*

[g]*ISEM, Université-Montpellier-Ii, Montpellier, France*

[h]*Muséum National d'Histoire Naturelle, UMR CNRS 5202, Origine, Structure et Évolution de la Biodiversité, Paris, France*

Hantaviruses belong to the Bunyaviridae family. While usually hosted by wild mammals, they are potentially pathogenic for humans, and several serologically distinct groups associated with different syndromes have been identified. Yet, investigations have mostly been conducted where human infections by hantaviruses constitute a real and well-identified public health problem, i.e., the holarctic and neotropical areas. Some hantaviruses have also been described from a *Suncus murinus* in India and a *Bandicota indica* in Thailand. In addition, recent investigations in Cambodia revealed new *Hantavirus* types. More recently, two new *Hantavirus* species were described: *Sangassou* from a *Hylomyscus simus*, and *Tanganya* from a *Crocidura theresae*, both from Africa (Guinea), thus strongly questioning the current views about geographic range, evolution, and epidemiology of hantaviruses. In such a framework, we have conducted a survey of *Hantavirus* diversity in Southeast Asia which allows us to isolate the Thailand virus and address questions about the taxonomy of their rodent hosts. Here we present a molecular analysis of representatives of all currently known *Hantavirus* species, thus allowing the comparison between the newly described ones with a large range sample of rodent hantaviruses. Our results clearly point to the presence of a particular lineage of hantaviruses in Southeast Asia. It also strongly suggests that new viruses, additional mammalian hosts and different related syndromes in humans are likely to be discovered in the near future, particularly in Southeast Asia and in Africa, where Muridae rodents are highly diversified. Furthermore, additional work is also urgently needed to investigate the hantaviruses associated with Crociduridae and Soricidae.

Key words: **rodent-borne; shrew-borne; hantaviruses; *Thottapalayam*; S gene; phylogeny; Bayesian analysis; biogeography; coevolution**

Address for correspondence: Jean-Pierre Hugot, Doctor in Veterinary Medicine, Doctor es Sciences, Research Director CNRS, Muséum National d'Histoire Naturelle, Département Systématique et Évolution, UMR 5202, Origine, Structure et Évolution de la Biodiversité, 75231 Paris Cedex 05, France. hugot@mnhn.fr

Animal Biodiversity and Emerging Diseases: Ann. N.Y. Acad. Sci. 1149: 84–89 (2008).
doi: 10.1196/annals.1428.064

Introduction

The genus *Hantavirus* is the member of Bunyaviridae family, which contains more than 350 species. Most of them are arboviruses that are vectored by mosquitoes, ticks, and sand flies. Within this family, only the genus *Tospovirus*, is associated with plants. Most of the Bunyaviridae may cause human diseases. For instance, *Bunyavirus* is the agent of La Crosse and California encephalites, *Phlebovirus* is responsible for the Rift Valley fever as well as sand fly fever, and *Nairovirus* causes Crimean–Congo hemorrhagic fever.

Hantaviruses, which are usually hosted by wild mammals, such as rodents and shrews, are potentially pathogenic for humans. Several serologically distinct virus species, associated with different syndromes, have been recognized. In Eurasia, *Hantaan*, *Dobrava*, *Seoul*, and *Puumala* cause the clinical forms of hemorrhagic fever with renal syndrome (HFRS).[1] In North and South America, *Sin Nombre* and *Andes* are responsible for the hantavirus pulmonary syndrome (HPS).[2] A last group, *Tula*, widely distributed in Russia and Eastern Europe, has never been associated with any human disease.

Because murid rodents are the most frequently recorded hosts and because each virus group seems to be associated with a particular rodent group, it was suggested that all hantaviruses may have a common origin and were coevolving with the Muridae.[1,3,5] However, some recent studies have revealed new hantaviruses which were hosted by different rodent species, as well as by shrews.[6,9] Moreover, these new data were all recorded in Southeast Asia and in Africa, that is, far away from the geographic range where hantaviruses are traditionally investigated, and where most of the human cases are detected. This clearly raises questions about the extent of the range of the hantaviruses, and suggests that their origins as well as their evolution in relation to their hosts urgently need to be readdressed.

In order to address these questions, we here present a molecular Bayesian analysis of the S sequence from representatives of the main known *Hantavirus* lineages, including all the recently discovered species. These questions are then discussed.

Materials and Methods

Sequences

The data set includes 100 *Hantavirus* S sequences. Most of them, downloaded from GenBank, were isolated from rodent hosts. The only exceptions are *Thottapalayam*, isolated from an Indian shrew (*Suncus murinus*)[10,11] and recently deposited in GenBank by Schmaljohn and Toney as a direct submission in 2004, and *Tanganya* isolated from the African shrew *Crocidura theresae*.[7] Our data set also includes the virus sequences retrieved from *Rattus rattus* and *R. norvegicus* in Cambodia,[9] the Sangassou sequence isolated by Klempa *et al*.[6] from *Hylomyscus simus* in Guinea, as well as the sequences that we recently isolated from *Bandicota indica* in Thailand.[5,12] Although sequence alignment, gap coding methods dealing with insertion–deletions (indels), choice of outgroup, methods used for analyses, and evaluation of node supports were extensively described in a preceding work,[3] the most pivotal methodologic points are briefly recalled here.

Coding Indels

To express potential phylogenetic information contained in zones with inter-nested insertions/deletions and substitutions, eight characters coding the presence/absence of deletions between nucleotides 766 and 813 were added. Finally, the matrix includes 1323 RNA characters and 8 presence/absence characters.

Choice of an Outgroup

In our data set, two sequences, each one hosted by a different shrew species, may be used as outgroups: *Thottapalayam,* collected from *S. murinus* in India, and *Tanganya*, collected from

C. theresae in Guinea. Finally, we retained *Thottapalayam* as an outgroup because its sequence includes 1530 pairs of bases (pb), while the original *Tanganya* sequence only includes 442 pb.

Sequence Analyses

A Bayesian analysis was performed using MrBayes v3.0B4.[13] Two partitions were distinguished in our original data set: partition 1 = nucleotide (characters 1–1323) for which the likelihood model chosen was the GTR + I + G; partition 2 = indels (characters 1324–1331) treated as presence/absence. Analyses were conducted with four independent Markov chains, run for 5,000,000 metropolis-coupled MCMC generations, with tree sampling every 10 generations and burn-in after 3300 trees. Consensus tree was computed using the "halfcompat" option, equivalent to the 50% majority rule. Proportion values of posterior probability of bipartition were used for evaluation of robustness of the nodes.

Results

The cladogram was rooted between a basal branch corresponding with *Thottapalayam* and a monophyletic group including all the rodent-borne parasites. The latter split into four main lineages (Fig. 1): Clade-1 groups—all the viruses hosted by Murinae rodents; this includes *Seoul, Hantaan, and Dobrava*. Clade-2 groups—all the viruses hosted by Sigmodontinae rodents including *Bayou, Sinnombre,* and *Andes*. Clade-3 groups—all the viruses hosted by Arvicolinae rodents including *Islavista, Tula,* and *Puumala*. Each clade and the sister-grouping of Clade-2 and Clade-3, have a support superior or equal to 78%. All recently discovered hantaviruses are included in, or close to Clade-1 (Fig. 2). *Sangassou* is the sister taxon of the *Dobrava/Saarema* clade. It is important to note that Thailand viruses are closely related to the *Hantavirus* found by Reynes *et al*.[9] in *R. rattus* from Cambodia. Together Cambodian and

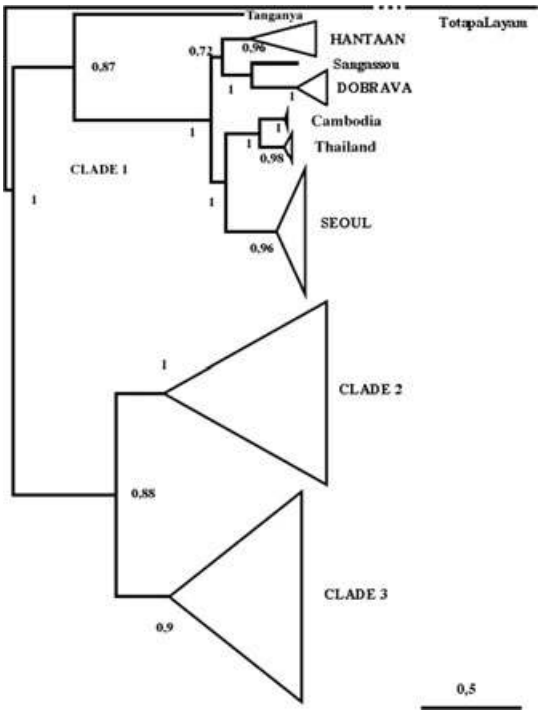

Figure 1. Simplified cladogram resulting of Bayesian analysis of 100 *Hantavirus* strains using GTR + I + G model. Number at nodes are the posterior probabilities.

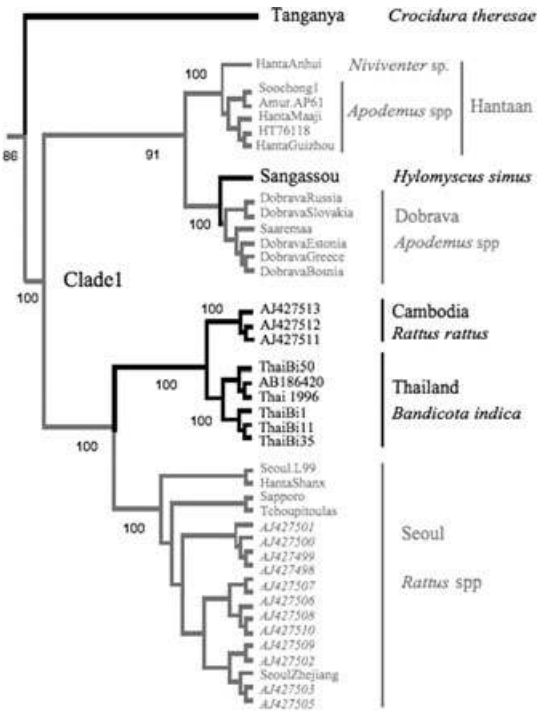

Figure 2. Detail of Clade-1 of Figure 1.

Thailand viruses are the sister group of Seoul. The *Seoul* clade includes all the hantaviruses described by Reynes *et al*.[9] in *R. norvegicus* from Cambodia. Note that relationships for and within this Southeast Asian group is strongly supported. Clade-4 group—the *Tanganya* virus, which appears as the sister group of Clade-1 ($Pp = 0.87$).

Discussion

The topology of the three main clades matches the phylogeny of the three host subfamilies from which they are respectively devolved and confirm previous results, thus still supporting the hypothesis of very old coevolution, between the *Hantavirus* and the Muridae rodents. In particular, the addition of the recently discovered *Hantavirus* species in our present analysis strongly reinforces this view (Fig. 2). Tanganya sister grouping with the Dobrava group matches with its host classification: *H. simus*, a Praomiyini, is considered close to the *Apodemus* group. The virus strains isolated by Reynes *et al*.,[9] from specimens of *R. rattus* in Cambodia, all fall together and are closely related to the different strains of Thailand hantavirus which were all isolated from *B. indica* in different parts of Thailand.[4,5] This group is close to, although different from, the cosmopolitan Seoul strains hosted by different *Rattus* species, including all Cambodian strains recently isolated from *R. norvegicus*. Of importance, our own investigations in Thailand seriously question the taxonomic status of the different components of the *R. rattus* complex in Southeast Asia (unpublished data). Indeed, several clearly distinct clades appear in our first analyses, with some of them being clearly divergent from the *R. rattus* individuals sampled from Asia. This may lead to subdivision of *R. rattus* into different clades, showing different species or subspecies endemic to Southeast Asia. These taxonomic studies on rodents strongly suggest that a particular *Hantavirus* group, hosted by endemic Muridae rodents, might exist in Southeast Asia, and that the Cambodian viruses previously associated with *R. rattus* may in fact be borne by other *Rattus* taxa.

Most *Hantavirus* species are known from North Eurasia, North America, and the neotropics, where different groups are associated with different syndromes. In Eurasia, *Hantaan, Dobrava, Seoul*, and *Puumala* cause the clinical forms of HFRS. In South America, Sin Nombre and Andes are responsible for HPS. A last group, *Tula*, widely distributed in Russia and Eastern Europe, has never been associated with any human disease. In spite of their importance for public health, hantaviruses have been only very rarely investigated outside of the regions where infected humans have been detected.

Yet, comparison of the respective geographic distributions of the Muridae and their associated *Hantavirus* (Table 1) clearly demonstrates that the currently known distribution of the *Hantavirus* does not match the distribution of the Muridae.

In a previous work, we suggested that further investigations were urgently needed to provide a better understanding of Hantavirus distribution, especially in South Asia and in Africa, where murid rodents are present and highly diversified.[3,5] The results presented here reinforced our previous assertions.

In a recent work, Song *et al*.,[8] questioned the significance for understanding *Hantavirus* evolution of *Thottapalayam* isolated in India from a *S. murinus*. Since the publication of this work, *Tanganya* isolated in Guinea from another shrew, *C. theresae*, was discovered (Table 2), allowing us to compare the patristic distances between shrew and/or rodent hantaviruses. *Thottapalayam* and *Tanganya* appear to be highly divergent from the rodent hantaviruses. The distance between *Thottapalayam* and *Tanganya* is 0.56, equal to the highest distance value between *Tanganya* and the rodent-borne hantaviruses. Yanagihara's group found hantaviruses in several insectivore species.[8] However, shrew hantaviruses were so different that none of rodent PCR isolates might help to

TABLE 1. Comparison of Geographic Distributions of Muridae Rodents and Respective Diversities of *Hantavirus* Strains

	1 Number of Muridae species	2 Percentage	3 Number of *Hantavirus* strains	4 Percentage
Neotropics	305	23.05	35	15.84
Ethiopian	280	21.16	2	0.90
Oriental	222	16.78	22	9.95
Palearctic	220	16.63	133	60.18
Nearctic	158	11.94	29	13.12
Australian	138	10.43	0	0.00

Note: Number of Muridae species (column 1) and percentage (column 2) of 1323 species recorded in mammal species of the world. Number of *Hantavirus* strains (column 3) and percentage (column 4) of 221 strains recorded in GenBank.

TABLE 2. Distances calculated using PAUP[14]

	1 (Totapalayam)	2 (Tanganya)	3 (within rodent strains)
Maximum	0.50	0.56	0.51
Minimum	0.42	0.48	0.08
Average	0.46	0.53	0.38
Median	0.46	0.53	0.41
DEVSQ	0.05	0.04	1.94

Note: The differences between Totapalayam and Tanganya and the rodent strains included in this study are shown in column 1, those between Tanganya and the rodent strains included in this study are shown in column 2, and those within rodent strains are seen in column 3. The distance between Totapalayam and Tanganya is 0.56.

identify the shrew viruses. Thus, if shrews carried hantaviruses that are very distant from ones carried by rodents, they also are very distant from each other. This of course does not support the hypothesis of a "shrew-borne" *Hantavirus* lineage, but also seems to exclude that the shrew as an incidental host of *Hantavirus* might be result of recent host-switching between rodents and sympatric shrews. The sequence data so far do not allow too-far-reaching conclusions: The *Tanganya* S sequence is incompletely described and its position in our analysis (Fig. 2) is weakly supported; the presence of different hantaviruses in different insectivores may be suspected, but must be confirmed by virus isolations. However, because Bunyaviridae viruses are usually insect-borne, these new discoveries ask questions about *Hantavirus* origins: did hantaviruses originate from insects to insectivores and later to rodents? Can we hypothesize a shrew–hantavirus coevolution older than in rodents?

All this suggests that if rodents are most probably a pivotal reservoir, other mammals may be involved in the circulation of hantaviruses in the wild. New viruses, associated with various hosts and leading to new human syndromes, may also be expected in the near future. For this reason extensive additional work is urgently needed, especially in the areas where hantaviruses have been traditionally recorded. In such a context, Southeast Asia and Africa, where Muridae rodents are present and highly diversified, should be of particular interest. Finally, additional efforts are also necessary to investigate the relationships between hantaviruses and the shrews.

Acknowledgments

This study was supported by the French GIP-ANR, Programme Santé Environnement–Santé Travail (Program 00121 05), by the PHC Franco-Thai Cooperation in Higher Education and Research (Program No. 16601PK).

Conflicts of Interest

The authors declare no conflicts of interest.

References

1. Schmaljohn, C.A., A. Schmaljohn & J. Dalrymple. 1987. Hantaan virus M RNA: coding strategy, nucleotide sequence, and gene order. *Virology* **157:** 31–39.
2. Childs, J.E. *et al*. 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.
3. Herbreteau, V., J.P. Gonzalez & J.P. Hugot. 2006. Implication of phylogenetic systematics of rodent-borne hantaviruses on their distribution. *Ann. N. Y. Acad. Sci.* **1081:** 39–56.
4. Dekonenko, A. *et al*. 2003. Genetic similarity of Puumala viruses found in Finland and western Siberia and of the mitochondrial DNA of their rodent hosts suggests a common evolutionary origin. *Infect. Gen. Evol.* **3:** 245–257.
5. Hugot, J.P. *et al*. 2006. Genetic analysis of Thailand hantavirus in *Bandicota indica* trapped in Thailand. *Virology J.* **3:** 72–81.
6. Klempa, B. *et al*. 2006. Hantavirus in African wood mouse, Guinea. *Emerg. Infect. Dis.* **12:** 838–840.
7. Klempa, B. *et al*. 2007. Novel Hantavirus sequences in shrew, Guinea. *Emerg. Infect. Dis.* **13:** 520–522.
8. Song, J.-W. *et al*. 2007. Thottapalayam virus, a prototype shrewborne hantavirus. *Emerg. Infect. Dis.* **13:** 1–15.
9. Reynes, J.M. *et al*. 2003. Evidence of the presence of Seoul virus in Cambodia. *Microbes Infect.* **5:** 769–773.
10. Carey, D. *et al*. 1971. Thottapalayam virus: a presumptive arbovirus isolated from a shrew in India. *Indian J. Med. Res.* **59:** 1758–1760.
11. Xiao, S. Y. *et al*. 1994. Phylogenetic analyses of virus isolates in the genus Hantavirus, family Bunyaviridae. *Virology* **198:** 205–217.
12. Pattamadilok, S. *et al*. 2006. Geographical distribution of hantaviruses in Thailand and potential human health significance of *Thailand* virus. *Amer. J. Trop. Med. Hyg.* **75:** 994–1002.
13. Huelsenbeck, J.P. & F. Ronquist. 2001. MRBAYES: Bayesian inference of phylogeny. *Bioinformatics* **17:** 754–755.
14. Swofford, D.L. *PAUP*. Phylogenetic Analysis Using Parsimony (and Other Methods)*. Version 4. 0b10 2001 Sinauer. Sunderland, MA.