Hantaviruses in Central South America: Phylogenetic Analysis of the S Segment from HPS Cases in Paraná, Brazil

Sonia M. Raboni,*‡ Christian M..Probst,* Juliano Bordignon,* Aurélio Zeferino* and Claudia N. Duarte dos Santos*#.

*Instituto de Biologia Molecular do Paraná, IBMP /Fiocruz/Brazil

‡ Universidade Federal do Paraná, Brazil

#Corresponding author: Instituto de Biologia Molecular do Paraná, IBMP, Rua Prof. Algacyr Munhoz Mader 3775, CIC, 81350 010, Curitiba, PR, Brazil

Phone # 55 41 316 3230

Fax # 55 41 316 3267

Email: clsantos@tecpar.br

Word counts:

Abstract – 145

Text – 5802

Footnote: Sonia M. Raboni and Christiam M. Probst contributed equally to this study

Running Head: Phylogenetic Analyses of Hantavirus in Southern Brazil

Keywords:

Hantavirus pulmonary syndrome; S segment genomic sequence; Hantavirus phylogeny, Nucleocapsid.

Abstract

We sequenced the complete S segments of hantaviruses detected from 12 HPS patients living in southern of Brazil. Samples were obtained from patients diagnosed in different years, in distinct areas and with a broad spectrum of clinical signs. Despite these differences, all the Paraná's hantavirus S proteins were identical, except for one amino acid substitution. Phylogenetic analyses of the complete S segment nucleotide and amino acid sequences indicated that Paraná's hantaviruses form a distinct clade from those circulating in South and North America. Other hantaviruses from Brazil were not placed in the same clade. The *Oligoryzomys nigripes*-associated strains ITA37 and ITA38 from Paraguay were found to belong to the same clade as the Paraná's hantaviruses. Paraná state and Paraguay are located at the same latitude and some ecosystems are similar in both places. The geographic position and common rodent hosts could explain this phylogenetic relationship.

Introduction

Hemorrhagic fever with renal syndrome (HFRS) and hantavirus pulmonary syndrome (HPS) are rodent-borne viral diseases caused by members of the genus *Hantavirus*, family *Bunyaviridae* (Beaty and Calisher, 1991), which includes approximately 30 distinct virus serotypes/genotypes (Monroe *et al*., 1999; Plyusnin and Morzunov, 2001). Hantaviruses are mainly transmitted to humans through the inhalation of contaminated aerosols of rodent excreta (Zu *et al*., 1985; Schmaljohn and Hooper, 2001; Lednicky, 2003), but human-to-human transmission has also been described (Padula *et al*., 1998).

The hantavirus genome consists of three negative single-stranded RNA segments: Large (L), medium (M), and small (S). These segments encode the viral RNA polymerase, a precursor glycoprotein that is processed into two separate envelope glycoproteins (G1 and G2), and a nucleocapsid protein (N), respectively (Schmaljohn and Hooper, 2001).

Each hantavirus is predominantly associated with a specific rodent host in a certain geographic region, but occasional spillover infection in related rodents occurs (Vicent *et al*., 2000; Wang *et al*., 2000; Plyusnin and Morzunov, 2001). The close association between each hantavirus and a particular rodent species and the phylogenetic analysis of hantaviruses suggest that hantaviruses coevolved with their hosts (Plyusnin *et al*., 1996; Vapalahti *et al*., 1999). The geographic and ecologic restrictions of the rodents separate these viruses into two phylogenetically distinct groups, one in Eurasia (Old World) associated with HFRS and the other in the Americas (New World) associated with HPS (Nichol *et al*., 1993; Nichol, 2001; Plyusnin and Morzunov, 2001; Lednicky, 2003). HPS was only described recently and is more severe than HFRS, with mortality rates of up to 40% (Nichol et al., 1993; Hjelle, et al., 1994; Schmaljohn and Hjelle, 1997).

In the Old World, the Hantaan, Seoul, Dobrava and Puumala viruses are known to be associated with HFRS of varying degrees of severity (Schmaljohn and Hjelle, 1997). The reservoirs of these viruses are rodents of the Murinae and Arvicolinae subfamilies (Plyusnin and Morzunov, 2001). Since HPS was first recognized in the United States of America (Nichol *et al*., 1993), several new *Sigmodontinae*-borne hantaviruses have been described in North, Central and South America (Peters and Khan, 2002). These HPS-associated viruses included several distinct genotypes in North and South America (Levis *et al*., 1998; Padula *et al*., 2000; Bohlman *et al*., 2002; Galeno *et al*., 2002; Peters and Khan, 2002; Chu *et al*., 2003; Lednicky, 2003; Pini *et al*., 2003).

In Brazil, the first evidence that these viruses were circulating came from an investigation of three individuals living in a rural area of Juquitiba, São Paulo, in 1993. They all had an acute disease characterized by fever, headaches, prostration, nausea and vomiting. Two of them died from acute respiratory failure. Serological tests revealed hantavirus infection in all cases (Iversson *et al*., 1994). Between 1993 and July 2004, 407 confirmed HPS cases had been detected in Brazil and three novel HPS-associated hantaviruses had been sequenced: Juquitiba (JUQ), Araraquara (ARA) and Castelo dos Sonhos (CAS) (Johnson *et al*., 1999).

Bolomys lasiurus and perhaps *Akodon* spp. are putative vectors of ARA (Katz *et al.*, 2001). Seropositive *Akodon* spp. and *Oligoryzomys* spp. have also been detected in Paraná state (Suzuki *et al.* 2004). However, until now, no hantaviruses have been isolated from rodents in Brazil.

Since 1998, Paraná has been the Brazilian state with the highest number of reported HPS cases (98 cases, 33%). Young adult males (mean age, 22 years) are most likely to develop HPS. Most HPS patients work in reforestation areas and live in

precarious housing conditions, with poor sanitation and a plentiful food supply, creating the ideal conditions to attract rodents.

No studies of hantavirus genome diversity using complete S segment have been carried out in the South region of Brazil. Aiming to adress this issue, we determined the complete S segment nucleotide sequence of hantaviruses detected from HPS patients in Paraná state to gain new insights into the genetic diversity of HPS-associated hantaviruses in Brazil and their phylogenetic relationships with other hantaviruses described elsewhere.

Methods

Collection of hantavirus samples

Samples from 12 HPS IgM-positive subjects living in Paraná State, listed in Table 1, were diagnosed for HPS as described (Raboni *et al*., 2005).

All patients were young males who contracted HPS whilst working in a pine tree reforestation region or during their rural working activities in Paraná State (Figure 1).

Extraction of total RNA, nested RT-PCR

Viral RNA was extracted using the Qiamp Viral RNA kit (Qiagen Inc, Ontario, CA) or the High Pure Viral RNA kit (Roche Inc, Mannhein, GE), according to the manufacturers' protocols.

Specific primers were used to synthesize overlapping cDNAs corresponding to the complete S segment using Expand Reverse Transcriptase (Roche Molecular Biochemicals, Mannhein, GE) and ImProm II RT (Promega Inc, USA), essentially as described in the manufacturers' protocols. The resulting cDNA was subjected to nested PCR using primers designed to detect the S segment of hantaviruses associated with sigmodontine rodents (Johnson *et al*., 1999, Raboni *et al*., 2005).

Sequencing and Assembling

The PCR products were purified (High Pure PCR kit, Roche Inc, Mannhein, GE) and both strands were directly sequenced with Thermo Sequenase kit (USB Inc, Ohio, USA) and on an ABI 3100 device using the BigDye® Terminator method (Applied Biosystems Inc, USA).

The Phred/Phrap/Consed system package (Ewing *et al*., 1998; Ewing and Green, 1998; Gordon *et al*., 1998) was used to assemble the fragments into the most likely S segment sequence

Multiple sequence alignment, nucleotide sequence comparison and phylogenetic analysis

A set of S segment sequence was retrieved from Genbank (http://www.ncbi.nlm.nih.gov/), comprising representative hantaviruses genotypes from Old World and North America, along with all complete sequence from South American hantaviruses until August 2004 (Table 2).

The CLUSTALx software (Thompson *et al*., 1994 Thompson *et al*., 1997) was used to align the hantavirus S sequences. Nucleotide differences were quantified using the MegaAlign software (DNASTAR, Inc, USA).

The PAUP v. 4.0b10 software (Swofford, 2003) was used to phylogenetic analyses with the entire N gene CDS (coding sequence) without start and stop codons (1,281 nucleotides in length), the incomplete N gene CDSs, and the complete N protein without the first methionine (427 amino acids). The Kimura two-parameter distance model (Kimura, 1980) and the Maximum Parsimony (MP) method were used to calculate evolutionary relationships between nucleotide sequences. The MP method was used to analyze complete N protein sequences. Both methods were used with default parameters.

The neighbor-joining and heuristic search methods were used to create phylogenetic trees for Kimura and MP analysis, respectively. Bootstrap analysis was performed with 1,000 repetitions.

Results

Sequence of the S segment

We sequenced the complete S segments from hantaviruses detected from 12 HPS patients living in Paraná State, southern Brazil (GenBank accession nos. AY 740622 to AY 740633). Nine were efficiently amplified by RT-PCR in two segments encompassing the complete (~1.9 kb) S segment (HPR/01-55, HPR/02-71, HPR/02-72, HPR/02-73, HPR/02-85, HPR/03-97, HPR/03-99 and HPR/04-102). For the other three, one RT-PCR segment, covering a 1.6-kb fragment starting at position 135 of the South American hantavirus CDS, was amplified.

The 5' and 3' segment termini from the cRNA, comprising 27 and 30 nucleotides respectively, are derived from the primers designed from the Laguna Negra (LN) virus sequence (Johnson *et al*., 1997). All viruses for which the S segment was completely sequenced had a sequence spanning 1,904 nucleotides, with the exception of HPR/01-55 (1,901 nt) and HPR/02-85 (1,903 nt). All deletions were located in the 3´non-coding region (NCR) of the S segment cRNA.

Coding regions

All complete sequenced S segments contained one CDS spanning 1,287 nt, from position 43 to 1,329 in the cRNA, which could be translated into a 428 amino acid nucleocapsid protein. The three partial sequences contained an abrogated CDS comprising 1,153 nt coding for the 383 carboxyl-terminal amino acids of the nucleocapsid protein.

An overlapping CDS, from position 122 to 313 in the cRNA and encoding a 63 amino acid protein, was identified. There are no experimental data supporting the existence of this protein, but there is an abnormally low third-base substitution frequency in this region (Bowen *et al*., 1995; Ravkov *et al*., 1995; Parrington and Kang, 1990). A similar CDS has also been identified in PUU, PH, SN, BAY, TUL and BCC viruses, and is

predicted to encode a NSs-like protein, similar to that seen in many other members of the Bunyaviridae family (Ravkov *et al*., 1995).

Non-coding region

The 42 nucleotides composing the 5´NCR of the S segment cRNA, including 27 nucleotides from the primer, are identical in all Paraná's hantaviruses. The S segment of the viruses described here also possessed a long 3'NCR, as is the case for other hantaviruses. This region occupies nearly one third of the segment (nucleotide positions 1329 to 1904). The 3´NCR sequences of Paraná's hantaviruses are very similar to each other, with a similar degree of divergence as that seen in the CDS (data not shown). However, this region is poorly conserved when compared to distantly related hantaviruses but contains several relatively well-conserved repeats. The CTACCTCA sequence motif was found in three copies, at positions 1813, 1841 and 1855 of Paraná's hantavirus cRNA. The fact that these sequences are located close to the highly conserved 3'-terminus of the S segment cRNA may have some functional significance that is important for virus replication (Ravkov *et al*., 1995; Lopez *et al*., 1997). The GGGT sequence motif was found in six copies in the Paraná hantaviruses. One of these copies is located in the CDS, at position 956 of the cRNA. The other five repeats are located in the 3´NCR, at positions 1422, 1436, 1450, 1469 and 1506. All repeats are separated by fourteen nucleotides, with the exception of the fourth. This motif is suspected to be a transcription termination motif (Hutchinson *et al*., 1996). In other South American hantaviruses, the number of repeats is different and the repeats are not located at regular intervals. The biological meaning of the intervals between these motifs in the Paraná's hantaviruses remains to be determined.

Sequence Comparisons

We aligned the nucleotide and amino acid sequences of the S segment from the Paraná's hantaviruses (Table 3). The CDSs of the nine completely sequenced S segments

displayed seven different nucleotide sequences, as HPR/02-72 was identical to HPR/04-102 and HPR/02-71 was identical to HPR/02-73. When the entire amplified sequence was considered, only HPR/02-71 and HPR/02-73 were identical. HPR/02-72 and HPR/04-102 have distinct 3´NCR regions. The most divergent nucleotide sequences were HPR/03-97 in comparison to HPR/02-72 and HPR/04-102, with 97.0% similarity. A similar level of divergence was observed when we considered the 3´NCR region of the S segment cRNA (data not shown).

All nucleotide differences were synonymous, with the exception of one mutation in HPR/02-85 (CTA_{1153}→GTA), which led to the replacement of a leucine by a valine at codon 385. This amino acid substitution is also present in DOB, HTN, SEO and TOP.

We compared the nucleotide and amino acid sequences of the two most divergent Paraná's hantaviruses and of other known hantaviruses (Table 4). The percentage nucleotide similarity ranged from 63% (HTN versus NY) to 97% (HPR/02-72 versus HPR/03-97) if we included all Paraná's hantaviruses, and to 87.8% (Hu39694 and ORN) otherwise. The percentage amino acid similarity ranged from 61% (HTN and PUU) to 100% (HPR/02-72 versus HPR/03-97).

Phylogenetic relationship

NJ dendrograms were constructed using the Kimura two-parameter distance method to show the relationships between selected hantavirus genotypes (Figures 2 to 4). The MP trees showed a very similar pattern and are not shown here.

The NJ tree shown in figure 2 is based on the nucleotide sequences of the S segment CDS. The Paraná's samples formed a well-defined group (100% bootstrap value) within the South American hantavirus clade, which was also supported by a high bootstrap value. The South American viruses could be divided into two well-supported clades, one

composed of Argentinean and Chilean viruses, containing the Paraná's samples, and the other composed of Laguna Negra and Rio Mamoré viruses.

The Argentinean-Chilean clade could be divided into four distinct subclades: Andes viruses (9717869, 9718133, CHI-7913 and AH-1); Lechiguanas-like viruses (BMJ, LEC, Hu39694, AND-Nort and ORN); Paraná's hantaviruses and Maciel-Pergamino viruses. AND-NORT is very similar to ORN and was located in a distinct subclade from other Andes viruses. The phylogenetic relationship between Paraná's, Andes and Lechiguanas-like viruses is unclear, as the bootstrap value was below 60%.

When the Paraná's hantaviruses with partially sequenced S segments were included, the phylogenetic tree remained very similar to that depicted in figure 2 (data not shown).

The NJ tree constructed using the amino acid sequences of the S segments showed branching pattern and bootstrap values very similar to those obtained with the nucleotide sequence-based phylogenetic analysis (data not shown).

All completely sequenced S segment CDSs from South American hantaviruses were included in phylogenetic analyses depicted in figures 2. To obtain further insight into the relationships between Paraná's hantaviruses and other less well-characterized South American hantaviruses, we performed a phylogenetic analysis using smaller fragments of the S sequence.

Figure 3 shows the dendrogram obtained by comparing the nucleotide sequences of the partial S segments (650 nucleotides), including one Peruvian sequence (HTN-007) and the two available Brazilian S sequences (Johnson *et al*., 1999). The CAS and ARA viruses were placed in distinct subclades from the Paraná's hantaviruses, although with low confidence for the CAS virus.

We next constructed a tree based on the nucleotide sequences of the partial S segment CDSs, from position 669 to 999, including five distinct Paraguayan viruses (Figure 4). The Paraguayan samples did not cluster together and were distributed into three distinct South America subclades: Lechiguanas-like (NEM), Paraná (ITA37 and ITA38) and Central South America (ALT and ITA-16). The small fragment used does not provide enough evolutionary information. Consequently, the branching order of subclades is not well supported, except for the Andes viruses and, more remarkably, for the Paraná subclade, which remains as a well-supported group, including the related Paraguayan samples, ITA-37 and ITA-38.

Discussion

We conducted the first genomic characterization of the complete S segment from hantaviruses in Paraná State, Brazil. The location of Paraná State is strategic because it is close to the South America countries with the highest incidence of HPS: Argentina, Chile and Paraguay. Besides that, the hantaviruses characterized in Paraguay are heterogeneous and this diversity may have epidemiological significance (Williams et al., 1997; Chu et al., 2003). Hence, it is of utmost importance to characterize the hantaviruses circulating in Paraná.

We obtained hantaviruses RNA from 12 IgM-positive HPS patients and sequenced their S segment. We have chosen this nucleoprotein-encoding genome region due to the large number of complete South American S segment sequences available.

We found that the nucleotide sequences of the S segment of the Paraná's hantaviruses were extremely similar, both in the nucleocapsid CDS (Table 3, 97% to 100% similarity) and in the less conserved 3´ NCR of the cRNA (97% to 100% similarity, data not shown).

The degree of conservation was even more striking when we considered the putative protein sequences. All variable nucleotide positions except one out of 64 (1.5%) were synonymous. We found five distinct amino acids at position 385 among the viruses analyzed, suggesting that this position is, at least, not subjected to negative selective pressure.

The phylogenetic analyses of the complete S segment nucleotide and amino acid sequences indicated that Paraná's hantaviruses form a distinct clade. This clade is placed in the South American hantavirus branch, together with other groups of South American viruses, such as Andes, Lechiguanas-like, Akodontine-borne (MAC and PRG) and Central South America (LN and OM-556). The relation between Paraná's hantaviruses and

Akodontine-borne and Central South-America hantaviruses is clear, but the same is not true for Andes and Argentinean viruses; in these cases the sequence region analyzed does not provide enough bootstrap support. Further studies, using the other genome segments, may provide the information required to resolve this issue.

To date, the criteria used to designate novel hantaviruses are controversial and it seems premature to state that novel South American hantavirus lineages are distinct virus species. Some authors have suggested that there are not enough sequences or identified rodent hosts available to allow us to consider the *Oligoryzomys*-borne hantavirus lineages as independent viruses or to classify them as Andes viruses (Levis *et al*., 1998; Padula *et al*., 2000). Meissner *et al*. (2002) suggests that whole genome sequencing will ultimately be required to verify preliminary designations and that in the meantime the genotype names should be considered provisory. While sequence data from complete genome segments are not available, we named the hantaviruses described here as "Araucária hantavirus".

We also assessed the phylogenetic relationships between Araucária and other Brazilian hantaviruses by using a small section of the S segment sequence and all conclusions must be interpreted with caution. None of these Brazilian viruses were placed in the same clade as the Araucária hantaviruses. ARA was placed in the same clade as the Maciel and Pergamino hantaviruses, with good bootstrap support. Despite the geographic distance between the Araraquara and Maciel-Pergamino isolation sites, the host seems to be a major determinant in the phylogenetic similarity, as the host of ARA is presumed to be *Bolomys lasiurus*, an akodontine rodent. The relationship between CAS and the other South American hantaviruses is less clear. A more detailed study is required to establish the relationship of all Brazilian hantaviruses.

Paraná and Paraguay are located at the same latitude and some ecosystems present in Paraná are continuous with their Paraguayan counterparts. When we included the available sequences from Paraguayan hantaviruses, an interesting picture emerged, despite the small size of the sequences used.

The samples from western Paraguay are most closely related to LN, OM-556 and HTN-007 hantaviruses. The hantaviruses from southern Paraguay are clustered in three distinct groups. The NEM virus has the same geographic position and host as other viruses included in the Lechiguanas-like clade. The *Oligoryzomys nigripes*-associated viruses ITA37 and ITA38 are in the same clade as the Paraná hantaviruses. The ITA16 virus, although detected in the same region, has a different host and was placed in a distinct clade, together with Central South America viruses.

The recent findings of hantavirus-seropositive *Akodon* spp. and *Oligoryzomys* spp. in the eastern Paraguay (Chu *et al*., 2003) and in the Southern region of Paraná (Suzuki *et al*. 2004) are compatible with a close relationship between the hantaviruses of both countries. Further studies are needed to confirm the phylogenetic relationships between the *Oligoryzomys* spp-borne hantaviruses found in these regions.

The association between Araucária hantaviruses and severe cases of HPS suggests that this virus poses a public health threat throughout the distribution range of the rodent host. This is reinforced by the results of a human serological survey carried out in both countries. Interestingly, unlike in Brazil (Paraná) where the case fatality rate is usually high (39%) during outbreaks, no cases of HPS have been reported among residents of these Paraguayan communities, despite the similar seroprevalence in the two populations: 17.9% in Paraguay (Chu *et al*., 2003) and 19.7% in Brazil (www.saude.parana.gov.br). This high seroprevalence was not reported for the classical Sin Nombre- or Andes-related syndrome in North and South America, respectively (Nichol, 2001, Peters and Khan,

2002; Pini *et al*., 2003). This observation suggests that these differences are due to other factors, such as the risk of contact and the behavior of the rodent reservoir species, and/or differences in viral virulence. Ongoing investigations that include genetic characterization of the viruses associated with different clinical forms and the rodent hosts and also "healthcare programs" should provide better insight into the dynamics of this disease.

Acknowledgments

The authors thank Irene Skrabas from Lacen-Curitiba for providing HPS samples and Gisélia Rubio from Secretaria Estadual de Saude do Paraná for helping with the rodent serological data.

This work received financial support from Fiocruz and CNPq.

References

1. Beaty, B. J. & Calisher, C. H. (1991). Bunyaviridae: natural history. Curr. Top. Microbiol. Immunol. 169, 27-78.

2. Bohlman, M. C., Morzunov, S. P., Meissner, J., Taylor, M. B., Ishibashi, K., Rowe, J., Levis, S., Enria, D. & St Jeor, S. C. (2002). Analysis of hantavirus genetic diversity in Argentina: S segment-derived phylogeny. J. Virol. 76, 3765-3773.

3. Bowen, M. D., Kariwa, H., Rollin, P. E., Peters, C. J. & Nichol, S. T. (1995). Genetic characterization of a human isolate of Puumala hantavirus from France. Virus Res. 38, 279-89.

4. Chu, Y. K, Owen, R. D, Gonzalez, L. M. & Johnsson C. B. (2003). The complex ecology of hantavirus in Paraguay. Am. J. Trop. Med. Hyg. 69, 263-68.

5. Ewing, B., Hillier, L., Wendl, M. & Green, P. (1998). Basecalling of automated sequencer traces using phred. I. Accuracy assessment. Genome Research. 8, 175-185.

6. Ewing, B. & Green, P. (1998). Base calling of automated sequencer traces using phred. II. Error probabilities. Genome Research 8, 186-194.

7. Galeno, H., Mora, J., Villagra, E., Fernandez, J., Hernandez, J., Mertz, G. J. & Ramirez, E. (2002). First human isolate of hantavirus (Andes virus) in the Americas. Emerg. Infect. Dis. 8, 657-60.

8. Gordon, D., Abajian, C. & Green, P. (1998). Consed: A graphical tool for sequence finishing. Genome Research. 8, 195-202.

9. Hjelle, B., Jenison, S., Torrez-Martinez, N., Yamada, T., Nolte, K., Zumwalt, R., M_{AC}Innes, 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.

10. Hutchinson, K. L., Peters, C. J. & Nichol, S. T. (1996). Sin Nombre vírus mRNA synthesis. Virology. 224, 139-149.

11. Iversson, L. B., Travassos da Rosa, A. P. A., Rosa, M. D. B., Lomar, A. V., Sasaki, M. G. M. & LeDuc, J. W. (1994). Infecção humana por hantavírus no Sul e Sudeste do Brasil. Rev. Ass. Med. Brasil. 40, 85-92.

12. Johnson, A. M., Souza, L. T. M., Ferreira, I. B., Pereira, L. E., Ksiazek, T. G., Rollin, P. E., Peters, C. J. & Nichol, S. T. (1999). Genetic investigation of novel hantaviruses causing fatal HPS in Brazil. J. Med. Virol. 59, 527-535.

13. Johnson, A. M., Bowen, M. D., Ksiazek, T. G., Willians, R. J., Bryan, R. T., Mills, J. N., Peters, C. J. & Nichol, S. T. (1997). Laguna Negra virus associated with HPS in western Paraguay and Bolivia. Virology 238, 115-127.

14. Katz, G., Williams, R. J., Burt, M. S., de Souza, L. T., Pereira, L. E., Mills, J. N., Suzuki, A., Ferreira, I. B., Souza, R. P., Alves, V. A., Bravo, J. S., Yates, T. L., Meyer, R., Shieh, W., Ksiazek, T.G., Khan, A. S. & Peters, C. J. (2001). Hantavirus pulmonary syndrome in the state of Sao Paulo, Brazil, 1993-1998. Vector Borne Zoonotic Dis. 1, 181-90.

15. Kimura, M. (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. 16, 111-20.

16. Levis, S., Morzunov, S. P. & Rowe, J. E. (1998). Genetic diversity and epidemiology of hantaviruses in Argentina. J. Infect. Dis. 177, 529-38.

17. Lednicky, J. A. (2003). Hantaviruses a short review. Arch. Pathol. Lab. Med. 127, 30-35.

18. Lopez, N., Padula, P., Rossi, C., Miguel, S., Edelstein, A., Ramirez, E. & Franze-Fernandez, M. T. (1997). Genetic characterization and phylogeny of Andes virus and variants from Argentina and Chile. Virus Res. 50, 77-84.

19. Meissner, J. D., Rowe, J. E., Borucki, M. K. & St Jeor, S. C. (2002). Complete nucleotide sequence of a Chilean hantavirus. Virus Res. 89, 131-143.

20. Monroe, M. C., Morzunov, S. P., Johnson, A. M., Bowen, M. D., Artsob, H., Yates, T., Peters, C. J., Rollin, P. E., Ksiazek, T. G. & Nichol, S. T. (1999). Genetic diversity and distribution of *Peromyscus*-borne hantaviruses in North America. Emerg. Infect. Dis. 5, 75-86.

21. Nichol, S. T. (2001). Bunyaviruses. *In* Fields Virology. Edited Fields, B.N., Knipe, D.M. & Howley, Lippincott Willians & Wilkins, Philadelphia. p 1603-1633.

22. Nichol, S. T., Spirolou, C., Morzunov, S., Rollin, P., Ksiazek, T. G., Feldman, H., Sanchez, A., Childs, J., Zaki, S. & Peters, C. J. (1993). Genetic identification of a hantavirus associated with an outbreak of acute respiratory illness. Science. 262, 914-917.

23. Padula, P. J., Colavecchia, S. B., Martinez, V. P., Della-Valle, M. O. G., Edelstein, A., Miguel, S. D. L., Russi, J., Riquelme, M., Colucci, N., Almirón, M. & Rabinovich R. D. (2000). Genetic diversity, distribution, and serological features of hantavirus infection in five countries in South America. J. Clin. Microbiol. 38(8), 3029-3035.

24. Padula, P. J., Edelstein, A., Miguel, S. D., Lopez, N. M., Rossi, C. M. & Rabinovich, R. D. (1998). Hantavirus pulmonary syndrome outbreak in Argentina: molecular evidence for person-to-person transmission of Andes virus. Virology. 241, 323-30.

25. Parrington, M. A. & Kang, C. Y. (1990). Nucleotide sequence analysis of the S genomic segment of Prospect Hill virus: comparison with the prototype Hantavirus. Virology. 175, 167-75.

26. Peters, C. J. & Khan, A. S. (2002). Hantavirus pulmonary syndrome: The new American hemorrhagic fever. Emerg. Infect. Dis. 34, 1224-31.

27. Pini, N., Levis, S., Calderón, G., Ramirez, J., Bravo, D., Lozano, E., Ripoll, C., St Jeor, S., Ksiazek, T. G., Barquez, R. M. & Enria, D. (2003). Hantavirus infection in humans and rodents, Northwestern Argentina. Emerg. Infect. Dis. [Online] http://www.cdc.gov/ncidod/EID/vol9no9/02-0768-G2.htm

28. Plyusnin, A., Vapalahti, O. &Vaheri, A. (1996). Hantaviruses; genome structure, expression and evolution. J.Gen.Virol. 77, 2677-2687.

29. Plyusnin, A. & Morzunov, S. P. (2001). Virus evolution and genetic diversity of hantaviruses and their rodent hosts, p.47-75. *In* Schmaljohn, C.S., Nichol, S.T., editors. Hantaviruses. Springer-Verlag, Belin.

30. Raboni, S. M., Rubio, G., Borba, L., Zeferino, A., Skraba, I., Goldenberg, S. & Duarte dos Santos, C. N. Am. J. Trop. Med. Hyg., in press.

31. Ravkov, E. V., Rollin, P. E., Ksiazek, T. G., Peters, C. J. & Nichol, S. T. (1995). Genetic and serologic analysis of Black Creek Canal virus and its association with human disease and *Sigmodon hispidus* infection. Virology. 210, 482-489.

32. Schmaljohn, C. & Hjelle, B. (1997). Hantaviruses: A global disease problem. Emerg. Infect. Dis. 3, 95-104.

33. Schmaljohn, C. S. & Hooper, J. W. (2001). Bunyaviridae: the viruses and their replication, p.1581-1633. *In* Fields BN, Knipe DM, Howley editors. Fields Virology. Lippincott Willians & Wilkins, Philadelphia.

34. Suzuki, A., Bisordi, I., Levis, S., Garcia, J., Pereira, L. E., Souza, R. P., Sugahara, T. K. N., Pini, N., Enria, D. and Souza, L. T. M. (2004). Identifying rodent hantavirus reservoirs, Brazil. Emerg. Infect. Dis. 10, 2127-2134.

35. Swofford, D. L. 2003. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.

36. Thompson, J. D, Higgins, D. G. & Gibson, T. J. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. Nucleic Acids Research. 22, 4673-80.

37. Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. & Higgins, D. G. 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Research 24, 4876-4882.

38. Vapalahti, O., Lundkvist, Å., Fedorov, V., Conroy, C. J, Hirvonen, S., Plyusnina, A., Nemirov, K., Fredga, K., Cook, J. Á., Niemimaa, J., Kaikusalu, A., Henttonen, H., Vaheri, A. & Plyusnin, A. (1999). Isolation and characterization of a hantavirus from *Lemmus sibiricus*: Evidence for host switch during hantavirus evolution. J. Virol. 73, 5586-5592.

39. Vicent, M. J., Quiroz, E., Gracia, F., Sanchez, A. J., Ksiazek, T. G., Kitsutani, P. T., Ruedas, L. A., Tinnin, D. S., Caceres, L., Garcia, A., Rollin, P. E., Mills, J. N., Peters, C. J. & Nichol, S. T. (2000). Hantavirus pulmonary syndrome in Panama: identification of novel hantaviruses and their likely reservoirs. Virology. 277, 14-19.

40. Wang, H., Yoshimatsu, K., Ebihara, H., Ogino, M., Araki, K., Kariwa, H., Wang, Z., Luo, Z., Hang, C. & Arikawa, J. (2000). Genetic diversity of hantaviruses isolated in China and characterization of novel hantaviruses isolated from *Niviventer confucionatus* and *Rattus rattus*. Virology. 278, 332-45.

41. Williams, R. J., Bryan, R. T., Mills, J. N., Palma, R. E., Vera, I., De Velasquez, F., Baez, E., Schimidt, W. E., Figueroa, R. E., Peters, C. J., Zaki, S. R., Khan, A. S. & Ksiazek, T. G. (1997). An outbreak of hantavirus pulmonary syndrome in Western Paraguai. Am. J. Trop. Med. Hyg. 57, 274-282.

42. Zu, Z. Y., Guo, C. S., Wu, Y. L., Zhang, X. W. & Liu, K. (1985). Epidemiological studies of hemorrhagic fever with renal syndrome: analysis of risk factors and mode of transmission. J. Infect. Dis. 152, 137-43.

Table 1. Patient codes, clinical evolution, place of origin and date

Patient*	Clinical evolution	Place of origin	Month/year
HPR/01-50	Recovered	Pinhão	10/2001
HPR/01-55	Recovered	Guarapuava	11/2001
HPR/02-71	Recovered	General Carneiro	08/2002
HPR/02-72	Fatal	General Carneiro	08/2002
HPR/02-73	Recovered	General Carneiro	08/2002
HPR/02-85	Fatal	Porto Vitória	11/2002
HPR/03-95	Fatal	Pinhão	05/2003
HPR/03-97	Fatal	São Mateus do Sul	11/2003
HPR/03-99	Fatal	Guarapuava	12/2003
HPR/03-100	Fatal	União da Vitoria	12/2003
HPR/04-101	Fatal	Pato Branco	01/2004
HPR/04-102	Recovered	Barbosa Ferraz	12/2003

* nomenclature: H = human PR = Parana State/ year of HPS diagnosis-patient registration number

Table 2. Representative hantaviruses genotypes from Old World and North, Central and South America

Sample	Description	CDS range	Location	Accession
Old World hantaviruses				
DOB	Dobrava virus	Complete	Europe	NC_005233
HTN	Hantaan virus	Complete	Asia	NC_005218
KHA	Khabarovsk virus	Complete	Asia	U35255
PUU	Puumala virus	Complete	Europe	NC_005224
SEO	Seoul virus	Complete	Asia	NC_005236
TOP	Topografov virus	Complete	Asia	AJ011646
TUL	Tula virus	Complete	Europe	NC_005227
North and Central American hantaviruses				
BAY	Bayou virus	Complete	Louisiana	L36929
BCC	Black Creek Canyon virus	Complete	Florida	L39949
CAL	Calabazo virus	271-648	Panamá	AF395443
CC074	Sin Nombre virus	Complete	California	L33816
CC107	Sin Nombre virus	Complete	California	L33683
CHO	Choclo virus	264-656	Panamá	AF395442
IV1	Isla Vista virus	Complete	California	U31534
MO46	Prairie Vole virus	Complete	North America	U19303
MON	Monongahela virus	Complete	West Virginia	U32591
MUL	Muleshoe virus	Complete	Texas	U54575
NM-H10	Sin Nombre virus	Complete	Four Corners	NC_005216
NM-R11	Sin Nombre virus	Complete	Four Corners	L37904
Nya	New York virus	Complete	New York	U29210
NYb	New York virus	Complete	New York	U47135
PHV	Prospect Hill virus	Complete	United States	M34011
RI-1	New York virus	Complete	New York	U09488
RM-97	El Moro Canyon virus	Complete	California	U11427
RMx-1	Rio Segundo virus	Complete	Costa Rica	U18100
South American hantaviruses				
9717869	Andes virus	Complete	Chile	NC_003466
9718133	Andes virus	Complete	Chile	AF482712
AND-NORT	Andes virus	Complete	Argentina	AF325966
CHI-7913	Andes virus	Complete	Chile	AY228237
AH-1	Andes virus	Complete	Argentina	AF324902
BMJ	Bermejo virus	Complete	Argentina	AF482713
Hu39694	Strain Hu39694 hantavirus	Complete	Argentina	AF482711
LEC	Lechiguanas virus	Complete	Argentina	AF482714
LN	Laguna Negra virus	Complete	Paraguay	AF005727
MAC	Maciel virus	Complete	Argentina	AF482716
OM-556	Rio Mamoré virus	Complete	Bolivia	U52136
ORN	Oran virus	Complete	Argentina	AF482715
PRG	Pergamino virus	Complete	Argentina	AF482717
ARA	Araraquara virus	15-657	Brazil	AF307325
CAS	Castelo dos Sonhos virus	15-657	Brazil	AF30732
ALT	Alto Paraguay hantavirus	669-999	Paraguay	AY515592
ITA16	Itapua hantavirus strain 16	669-999	Paraguay	AY515594
ITA37	Itapua hantavirus strain 37	669-999	Paraguay	AY515595
ITA38	Itapua hantavirus strain 38	669-999	Paraguay	AY515596
NEM	Neembucu hantavirus	669-999	Paraguay	AY515593
HTN-007	HTN-007 hantavirus	1-1020	Peru	AF133254

From Genbank (http://www.ncbi.nlm.nih.gov/)

Table 3. Comparison of the nucleotide (top) and amino acid (bottom) sequences of the S segment CDS between Paraná's hantaviruses

Virus	% Similarity								
	03-99	01-55	03-95	02-72	04-102	02-71	02-73	03-97	02-85
03-99		98.6	99.6	99.3	99.3	99.8	99.8	97.4	98.9
01-55	100.0		98.5	98.1	98.1	98.6	98.6	97.4	98.2
03-95	100.0	100.0		99.5	99.5	99.5	99.5	97.1	98.5
02-72	100.0	100.0	100.0		100.0	99.3	99.3	97.0	98.2
04-102	100.0	100.0	100.0	100.0		99.3	99.3	97.0	98.2
02-71	100.0	100.0	100.0	100.0	100.0		100.0	97.5	98.8
02-73	100.0	100.0	100.0	100.0	100.0	100.0		97.5	98.8
03-97	100.0	100.0	100.0	100.0	100.0	100.0	100.0		97.3
02-85	99.8	99.8	99.8	99.8	99.8	99.8	99.8	99.8	

Table 4. Comparison of the amino acid (top) and nucleotide (bottom) sequences of the S segment CDS among several hantaviruses.

Virus	% Similarity																		
	02-72	03-97	AH_1	BAY	BCC	BER	HTN	Hu	LEC	LN	MAC	NM	NY	OM	ORN	PRG	PUU	SEO	TUL
02-72		97.0	83.1	77.9	76.5	83.4	66.7	82.4	83.1	79.4	80.7	76.8	78.5	79.3	82.9	82.0	70.5	65.6	69.9
03-97	100.0		83.1	78.0	76.4	83.1	66.7	82.5	82.8	80.2	80.2	77.2	78.2	79.2	83.6	81.7	70.6	65.1	69.8
AH_1	95.3	95.3		77.2	76.8	84.1	66.8	84.5	84.5	79.1	81.2	76.5	77.5	79.8	83.8	81.5	69.5	65.9	70.2
BAY	87.4	87.4	88.3		81.0	77.1	66.2	77.0	77.2	77.5	76.1	76.8	77.9	77.6	78.6	76.9	71.4	63.6	71.5
BCC	85.2	85.2	86.4	92.3		78.5	65.7	76.5	78.3	76.8	76.8	75.7	76.2	76.8	76.6	75.8	69.1	65.5	70.6
BER	95.6	95.6	96.7	88.8	86.2		66.9	87.6	91.5	79.3	82.0	77.8	77.5	79.9	87.3	82.0	68.9	64.7	71.7
HTN	63.7	63.7	65.1	64.2	64.4	65.8		66.4	66.0	66.8	66.9	64.4	63.0	67.8	66.8	66.6	64.3	75.0	64.8
Hu	95.8	95.8	97.0	89.0	86.4	99.8	65.1		87.4	79.6	82.1	76.1	77.3	79.8	87.8	82.4	68.5	66.2	70.8
LEC	95.3	95.3	96.5	88.5	85.9	99.8	65.6	99.5		80.0	82.2	77.7	78.2	80.5	87.2	83.1	70.5	65.8	72.3
LN	90.2	90.2	90.4	87.1	85.5	89.9	64.6	90.2	89.7		78.2	76.7	77.0	82.4	79.3	78.5	71.9	66.3	71.5
MAC	93.0	93.0	93.9	87.6	86.4	94.4	64.2	94.6	94.1	88.3		78.1	78.4	78.5	82.3	82.9	71.2	66.2	70.9
NM	85.9	85.9	85.9	86.9	83.8	87.4	62.8	87.1	87.1	85.2	85.2		83.6	76.1	77.5	76.2	69.8	65.3	72.2
NY	86.7	86.7	87.6	87.1	85.0	87.6	63.5	87.8	87.4	86.9	87.1	93.0		77.9	77.2	77.9	70.7	65.7	71.4
OM	90.2	90.2	91.1	88.1	86.7	90.4	64.4	90.6	90.6	93.2	89.5	84.5	86.4		81.0	79.4	70.4	64.8	72.0
ORN	96.5	96.5	96.7	89.0	86.4	98.8	65.3	99.1	98.6	90.2	94.6	86.9	87.6	90.6		83.1	69.4	65.9	71.8
PRG	93.0	93.0	95.1	88.5	85.9	95.8	64.2	96.0	95.6	89.7	96.0	85.9	87.6	89.9	96.0		70.6	64.4	69.9
PUU	72.8	72.8	73.1	73.3	74.2	72.8	61.0	72.8	72.6	72.6	73.8	71.0	71.2	72.6	72.8	72.8		64.0	74.3
SEO	63.7	63.7	64.9	63.5	64.2	64.4	83.2	64.4	64.6	63.7	63.5	62.3	63.2	64.6	64.4	63.9	62.4		66.0
TUL	74.5	74.5	75.2	76.1	75.9	75.6	63.1	75.4	75.9	75.2	74.2	73.8	72.6	75.4	75.2	74.9	79.4	62.9	

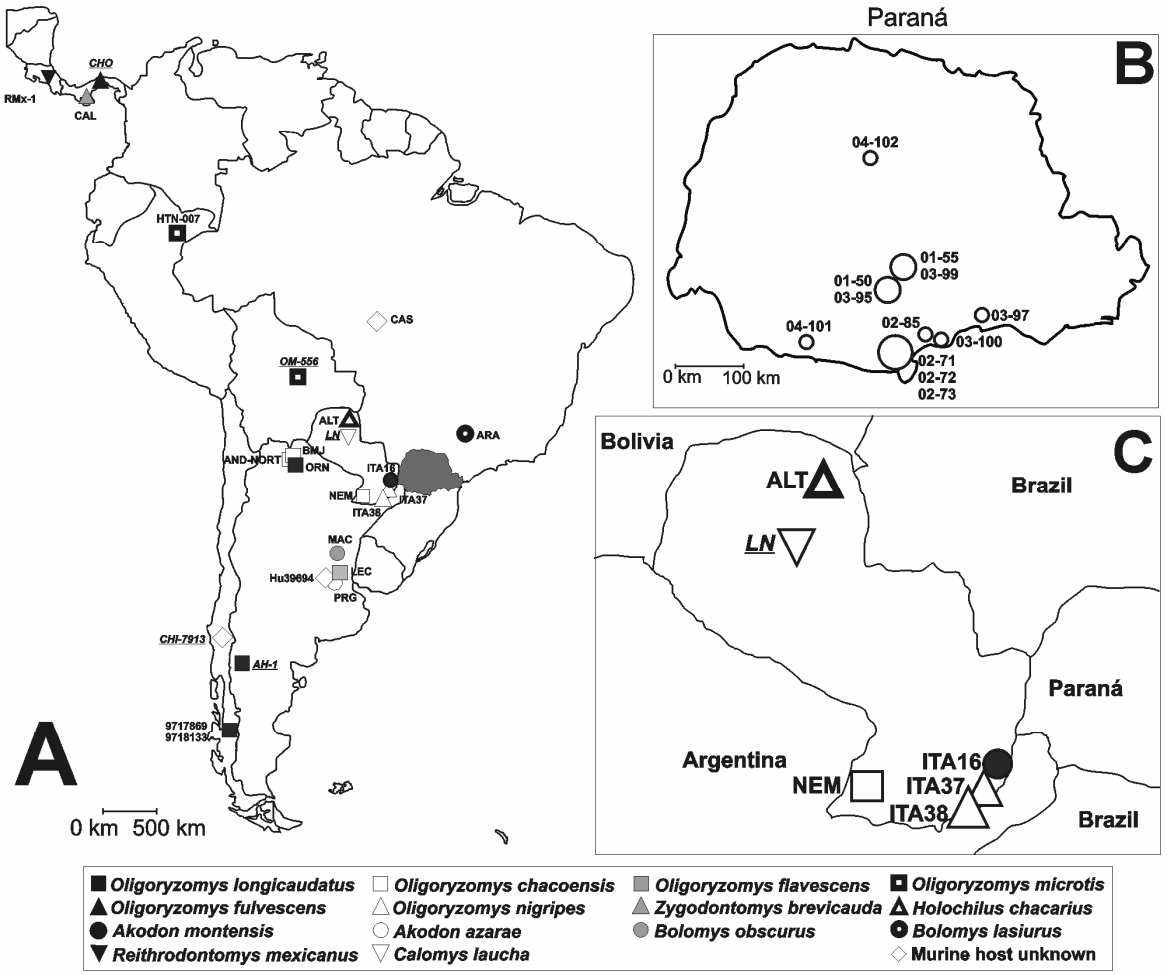

Figure 1. Map showing position of Paraná State (shaded area). A: the locations of the South American hantaviruses analyzed in this study are shown, together with their codename and rodent host. B: the locations of the Paraná's hantaviruses are shown, together with their names. The circle size represents the number of cases analyzed in each location. C: a closer view of Paraguayan hantaviruses showing their proximity to Paraná State.

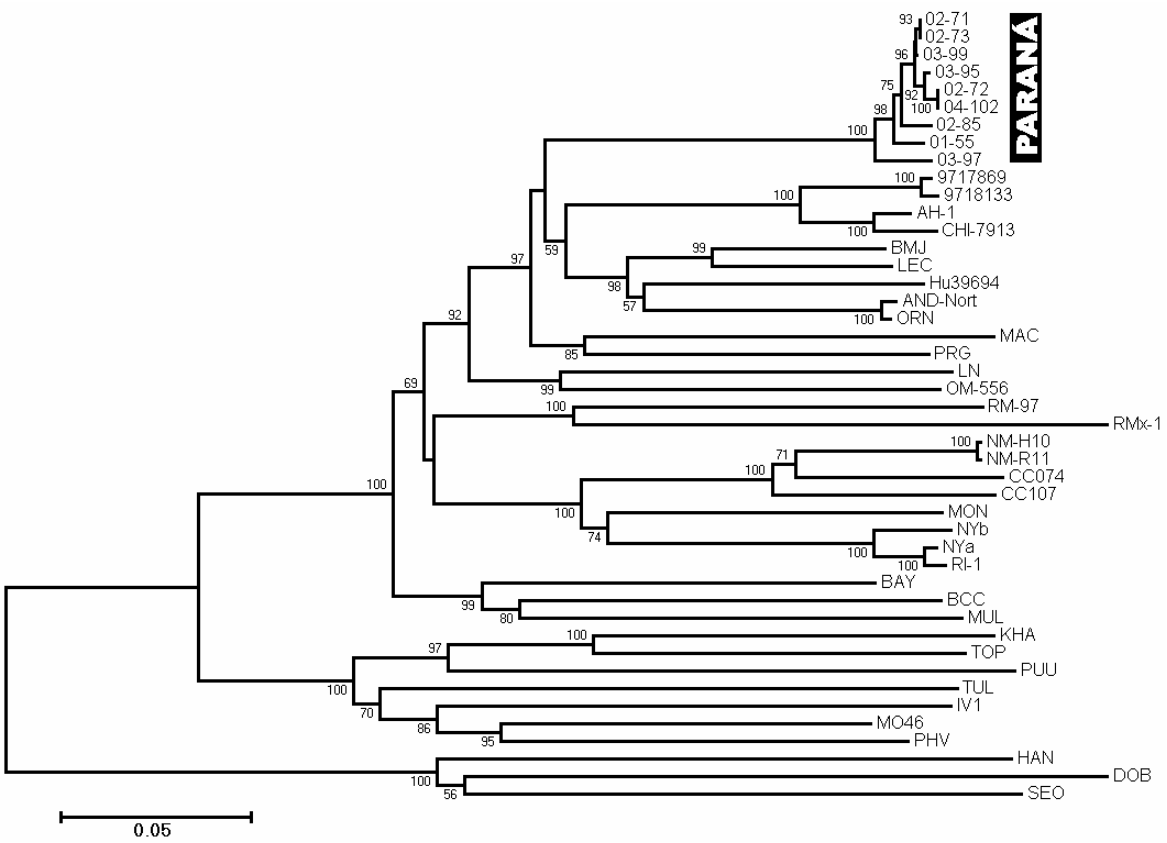

Figure 2. Dendrogram based on the complete nucleotide sequences of the S segment CDS (Hantaviruses included in the analysis are listed in Table 2).

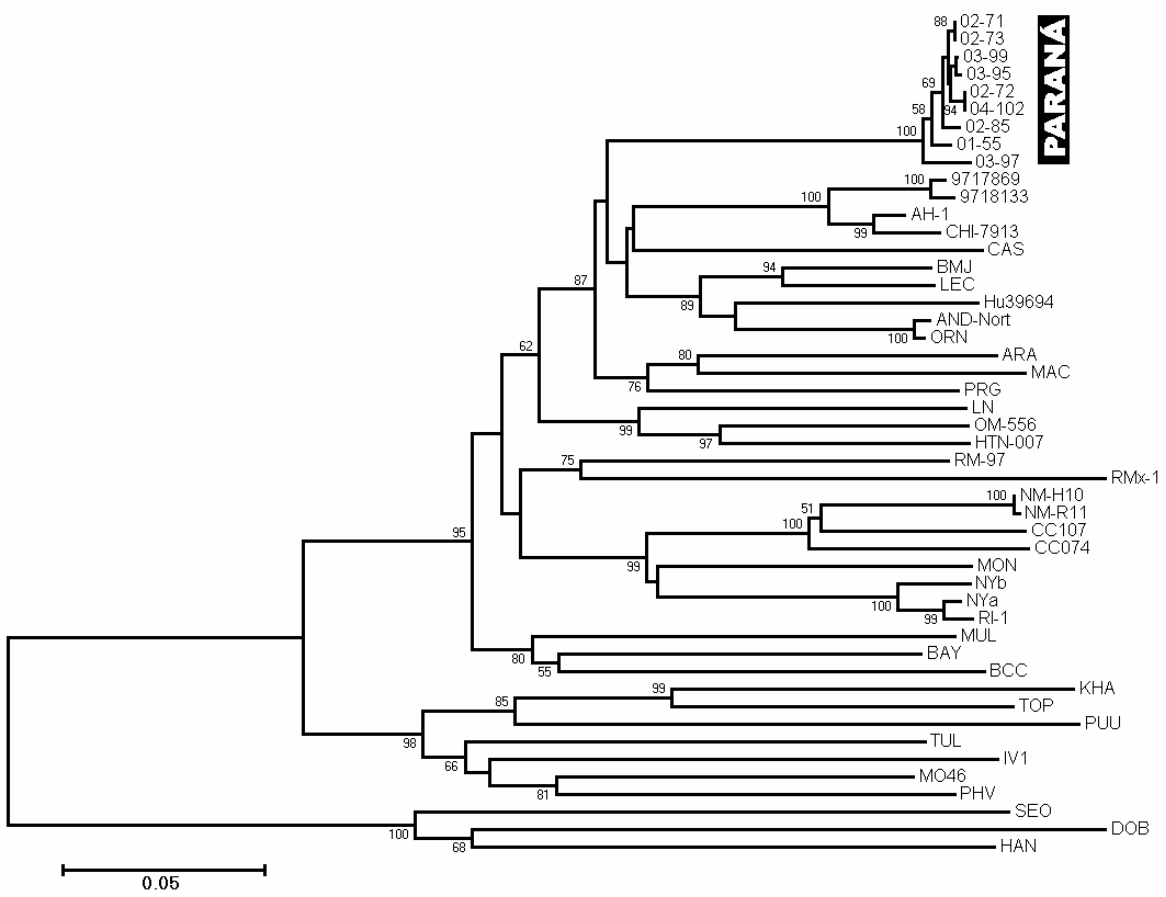

Figure 3. Dendrogram based on the partial nucleotide sequences of the S segment, including all available Brazilian sequences.

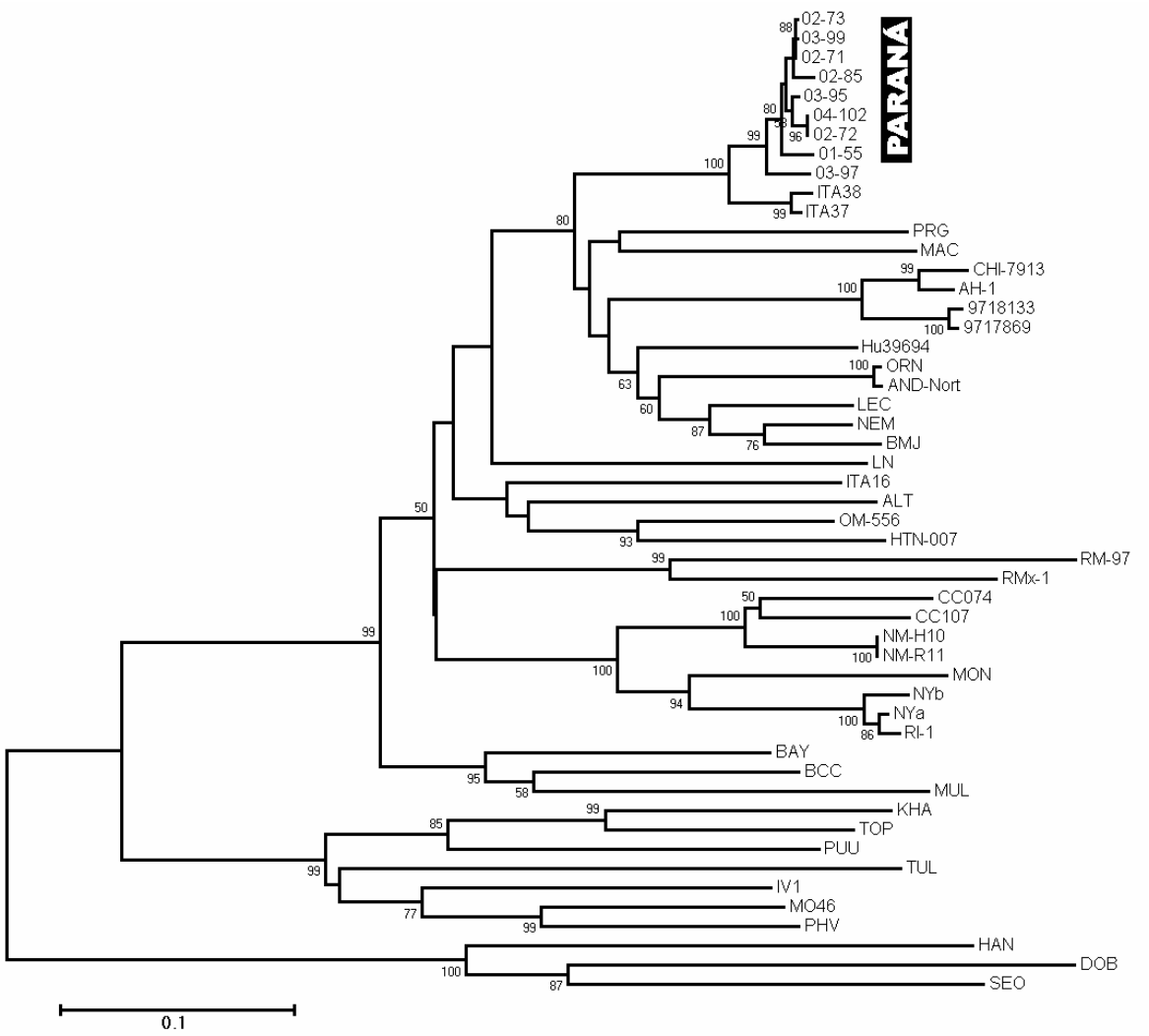

Figure 4. Dendrogram based on the partial nucleotide sequences of the S segment, including several hantaviruses from Paraguay.