

Research

Open Access

Phylogenetic evidence for the distinction of Saaremaa and Dobrava hantaviruses

Tarja Sironen, Antti Vaheri and Alexander Plyusnin*

Address: Department of Virology, Haartman Institute, Haartmaninkatu 3, FIN-00014 University of Helsinki, Finland

Email: Tarja Sironen - Tarja.Sironen@helsinki.fi; Antti Vaheri - Antti.Vaheri@helsinki.fi; Alexander Plyusnin* - Alexander.Plyusnin@helsinki.fi

* Corresponding author

Published: 08 December 2005

Received: 27 June 2005

Virology Journal 2005, 2:90 doi:10.1186/1743-422X-2-90

Accepted: 08 December 2005

This article is available from: <http://www.virologyj.com/content/2/1/90>

© 2005 Sironen et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Dobrava virus (DOBV) and Saaremaa virus (SAAV) are two closely related hantaviruses carried by different rodent species. The distinction of these two viruses has been a matter of debate. While the phylogenies based on the viral M segment sequences were repeatedly showing monophyly of SAAV strains, some trees based on the S segment sequences were not, thus causing questions on the demarcation between these two viruses. In order to clarify this issue, the current collection of the virus S segment sequences was subjected to extensive phylogenetic analysis using maximum likelihood, maximum parsimony and distant matrix methods. In all inferred phylogenies, the SAAV sequences were monophyletic and separated from DOBV sequences, thus supporting the view that SAAV and DOBV are distinct hantavirus species. Since collection of the S segment sequences used in this study "obeyed" the molecular clock, calculations of the split of DOBV and SAAV were now repeated resulting in an estimation of 3.0–3.7 MYA that is very close to the values obtained earlier.

Background

Hantaviruses (genus *Hantavirus*, family *Bunyaviridae*) are enveloped viruses with a segmented, single-stranded RNA genome of negative polarity [1]. The large (L) segment encodes the viral RNA polymerase, the medium (M) segment the two surface glycoproteins, and the small (S) segment the nucleocapsid protein (N). Hantaviruses cause two human zoonoses, hemorrhagic fever with renal syndrome (HFRS) in Eurasia and hantavirus pulmonary syndrome (HPS) in the Americas [reviewed in [2]]. DOBV is carried by yellow-necked mouse (*Apodemus flavicollis*) and is associated with severe HFRS in the Balkans (Slovenia, Albania and Greece). SAAV is carried by striped field mouse (*A. agrarius*) [3]. So far, the virus has been found in Estonia, the European part of Russia, Slovakia, Slovenia, Hungary, Denmark and Germany [2].

SAAV was initially called an *A. agrarius*-carried variant of Dobrava virus [3], but the accumulating data suggest that the virus should be regarded as a distinct hantavirus species. It is carried by a specific rodent host [3], there is a four-fold difference in two-way cross-neutralization tests [4], and the coexistence of SAAV and DOBV in the same geographic region [5,6] indicates reproductive isolation. They also exhibit 6.1–6.3% difference in the glycoprotein precursor amino acid sequences. This level is a fraction lower than the officially accepted 7% cut-off value [1]. It should be mentioned that some of the officially approved, distinct hantavirus species show lower than 7% diversity in their N or GnGc-sequences: Sin Nombre and New York viruses, Topografov and Khabarovsk viruses, Rio Mamore and Laguna Negra viruses, and Blood Land Lake and Prospect Hill viruses [7].

Table 1: Sequences used in the analysis

	Strain	Accession number
Saaremaa virus (SAAV)	Saaremaa/160 V	AJ009773
	90Aa/97	AJ009775
	Lolland/Aa1403/2000	AJ616854
	Kurkino/44Aa/98	AJ131672
	Kurkino/53Aa/98	AJ131673
	East Slovakia/856/Aa	AJ269549
Dobrava virus (DOBV)	East Slovakia/862/Aa	AJ269550
	Slovenia	L41916
	East Slovakia/400Af/98	AY168576
	Ano-Poroia/9Af/1999	AJ410615
	Ano-Poroia/13Af/99	AJ410619
	As-1/Goryachiy Klyuch-2000	AF442622
Seoul virus (SEOV)	P-s1223/Krasnodar-2000	AF442623
	Gou3	AB027522
	L99	AF288299
	Z37	AF187082
Hantaan virus (HTNV)	SR11	M34881
	Ah09	AF285264
	84FlI	AY017064
	76-118	M14626
Andes virus (ANDV)	Lr1	AF288294
	AH-1	AF324902
Topografov virus (TOPV)	Ls136V	AJ011646
Sin Nombre virus (SNV)	NM H10	L25784
El Moro Canyon virus (ELMCV)	RM-97	U11427
Puumala virus (PUUV)	Sotkamo	X61035
Tula virus (TULV)	Moravia/5302v/95	Z69991

SAAV and DOBV also exhibit only 3% diversity on their N protein sequences. This unusually low level of diversity is most probably a reflection of host switching in their evolution [8,9]; this event seems to be historically recent

(2.7–3.4 MYA) and these two viruses are still diverging [8]. There is another important feature differentiating DOBV and SAAV, and that is the apparently different pathogenicity in humans: while DOBV causes severe

Table 2: Bootstrap and puzzle support values for DOBV and SAAV clades in phylogenetic trees calculated using different methods.

method	outgroup	support for: DOBV	support for: SAAV
maximum likelihood	SEOV	100	70
	collection*	100	49
	no outgroup	100	100
maximum parsimony	SEOV	100	75
	collection*	100	75
distance matrix: Neighbor-joining	SEOV	100	84
	collection*	100	91
distance matrix: Fitch-Margoliash	SEOV	79	58
	collection*	100	79
	no outgroup	100	99
TreePuzzle**	SEOV	99	87
	collection*	99	75

*A collection of hantavirus sequences including SNV, ANDV, ELMCV, TULV, TOPV, PUUV, SEOV strains SR11 and Gou3, HTNV strains 76-118 and 84FlI **Tamura-Nei was used as the nucleotide (nt) substitution model in TreePuzzle, as suggested by Modeltest.

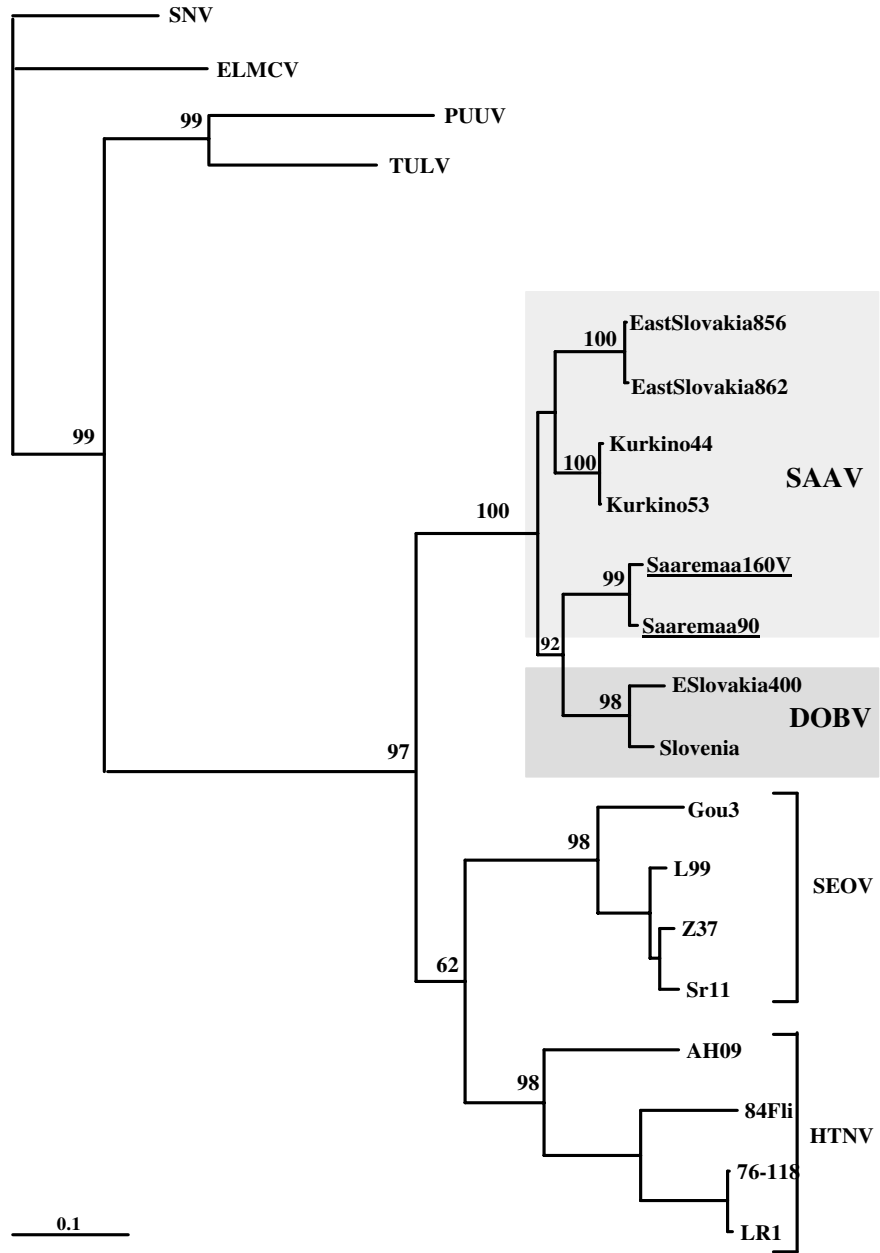


Figure 1

Phylogenetic tree created with TreePuzzle for a smaller data set. The tree is based on the nt 37–1232 of the S segment sequences.

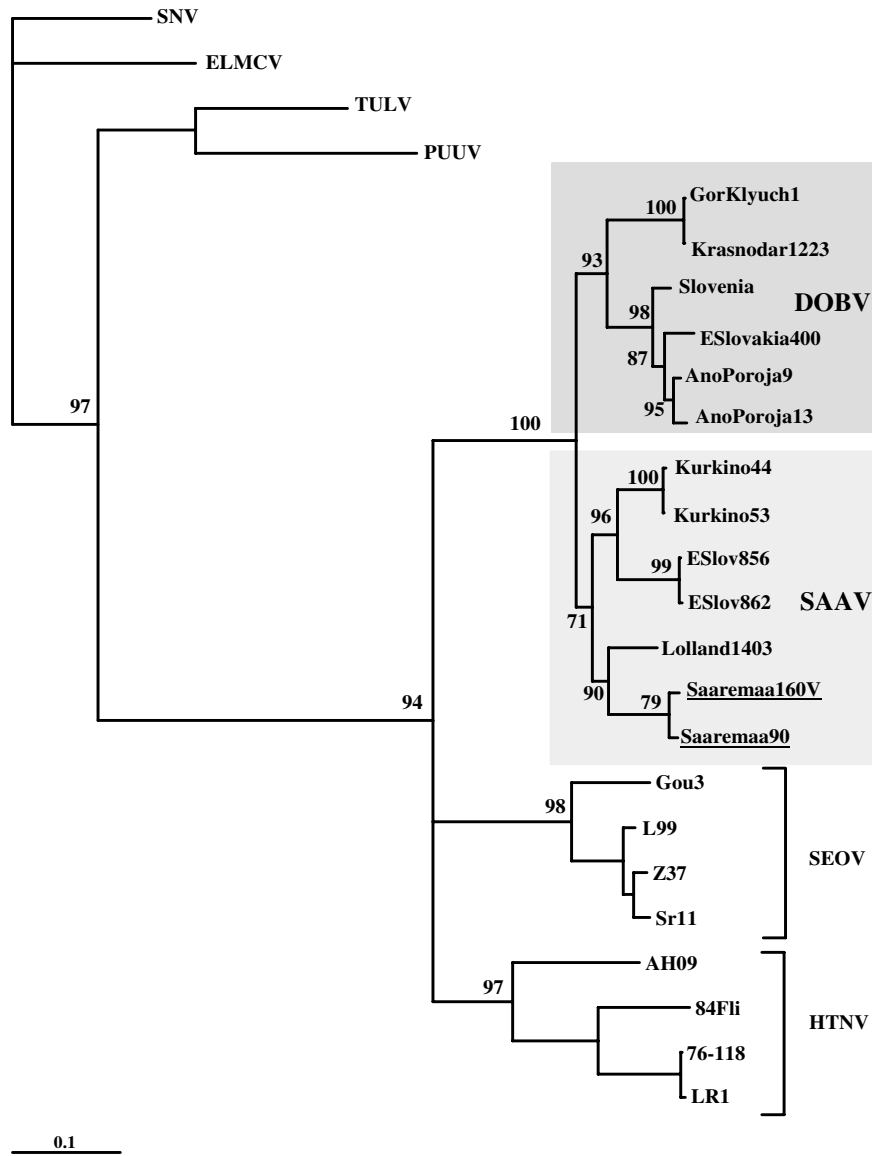


Figure 2

Phylogenetic tree created with TreePuzzle for a more representative data set. The tree is based on the nt 37–1232 of the S segment sequences. Two SAAV sequences that are placed differently on the trees shown on Fig. 1 and Fig 2 are underlined.

HFRS in humans, SAAV causes a milder form of the disease, similar to nephropathia epidemica [2]. This difference is also reflected in different pathogenicity in suckling mice: DOBV is lethal to suckling mice, while SAAV is not [10].

The phylogenetic distinction of SAAV and DOBV was recently a matter of debate [11,12]. While the phylogenies based on the M segment/GnGc protein sequences were repeatedly showing monophyly of SAAV strains, some trees based on the S segment/N protein sequences were not [[11,13], and our unpublished observations], thus causing questions on the demarcation between these two viruses. In order to clarify this issue, the current collection of DOBV and SAAV S segment sequences was subjected to extensive phylogenetic analysis. Especially important additions to the dataset include an *A. agrarius*-derived SAAV strain from Denmark, Saaremaa/Lolland/Aa1403/2000 [AJ616854], and two DOBV sequences from southern Russia, P-s1223/Krasnodar-2000 (AF442623) and As-1/Goryachiy Klyuch-2000 (AF442622). Our earlier data indicated that these sequences could be helpful for resolving the S phylogeny [14].

Results and discussion

Our analysis was restricted to nt 37–1232 of the S segment available for all the strains. This part of the S segment includes almost complete coding region for the N protein. Accession numbers for the sequences are given in Table 1.

Since recombinant sequences might influence phylogenetic reconstructions (e.g. by "breaking" the molecular clock [15]), we wanted to check whether the sequences used in this study included any recombinants ones. A similarity plot (Stuart Ray's SIMPLOT2.5) was created in order to visualize the pattern of similarity between the DOBV and SAAV S segment nucleotide sequences, and phylogenetic trees were created on partial sequences, that were possibly of recombinant origin. Although we have found some indications on a recombinant origin of the strain Lolland (in particular, nt 200–460 were most similar to the Estonian SAAV strains, while other regions, especially nt 1150–1450, were more similar to SAAV strains from Russia and Slovakia), they were not unequivocal. For instance, the SIMPLOT data were not mirrored by a mosaic-like pattern of the N protein sequence of Lolland strain. Moreover, the presence of this sequence did not "break" the molecular clock (see below). The Lolland sequence was, therefore, not excluded from our data set.

Next, we wanted to study whether the new additional sequences would have any effect on the clustering of DOBV and SAAV. A phylogenetic tree was re-calculated with the same collection of sequences and same parameters as has been done by Klempa et al. [11] (Fig. 1). The

additional DOBV and SAAV sequences were then included to this set, a new phylogenetic tree was created, and indeed, a change in the topology was seen. The SAAV sequences turned monophyletic with a puzzle support of 71% (Fig. 2).

In order to confirm the phylogeny, trees were calculated using different algorithms listed earlier (Table 2). All methods agreed on placing DOBV and SAAV sequences into their own clusters. Placing of the two above mentioned DOBV sequences derived from southern Russia was more variable, but in most cases they were sharing a common ancestor with the other DOBV strains. The puzzle support values and bootstrap support for the DOBV cluster were in most cases very high (79–100%). For SAAV, the support was more variable, but only in two out of 12 phylogenies below the widely accepted confidentiality limit (70%) [16]. The support values were also varying depending on the phylogenetic algorithm, on the parameters used, and on the sequences chosen as outgroup. In the case of maximum likelihood trees, the use of additional hantavirus sequences as outgroup resulted in a lower bootstrap support for SAAV. In fact, a 100% support for SAAV monophyly was reached, when no outgroup sequences were used at all. This algorithm goes through an exhaustive search of all the possible trees, and it is possible that additional information creates an interfering noise to the phylogenetic signal. The opposite was happening with Fitch-Margoliash distance-matrix method. As more sequences were added, the bootstrap support for SAAV was increasing, most probably due to more accurate distance estimations. Nevertheless, in every tree, all the SAAV sequences were monophyletic and separated from DOBV. It should be stressed that bootstrap or puzzle support values do not estimate accuracy of a tree (i.e. right topology), but precision (how many trees had to be rejected) [17]. Phylogenies inferred here with different algorithms, and by varying the parameters used in the analyses (Table 2), gave a consensus answer on the monophyly of all SAAV strains, thus suggesting that this tree topology is most accurate.

Earlier it has been estimated, that the split of DOBV and SAAV happened 2,7–3.4 million years ago (MYA) (10). Since the larger collection of the S segment sequences used in this study "obeyed" the molecular clock, these calculations were now repeated resulting in an estimation of 3.0–3.7 MYA.

Conclusion

In all phylogenies inferred in this study using different approaches such as maximum likelihood, maximum parsimony and distant matrices, the SAAV sequences were monophyletic and separated from DOBV sequences, thus

supporting the view that SAAV and DOBV are distinct hantavirus species.

Methods

Sequences were handled with BIOEDIT [18], and alignments were created using CLUSTALX [19]. The various methods used for phylogenetic analysis included maximum likelihood ("classic" maximum likelihood from PHYLIP [20] and TreePuzzle [21], maximum parsimony (PHYLIP) and distance matrix methods Neighbor joining and Fitch-Margoliash (PHYLIP). 500 bootstrap replicates were used in PHYLIP programs and 10000 puzzling steps in TreePuzzle. MODELTEST and PAUP were used to check, which DNA substitution model would fit best to this data set [22,23]. The test for molecular clock and estimation of the time of split of these two viruses was done with TreePuzzle [21].

Competing interests

The author(s) declare that they have no competing interests.

Authors' contributions

TS carried out experiments, participated in the analysis of the results and drafted the manuscript. AV participated in the analysis of the results and helped to draft the manuscript. AP designed the study, participated in the analysis of the results and helped to draft the manuscript.

References

- Elliott RM, Bouloy M, Calisher CH, Goldbach R, Moyer JT, Nichol ST, Pettersson R, Plyusnin A, Schmaljohn CS: **Family Bunyaviridae**. In *Virus taxonomy. 11th report of the International Committee on Taxonomy of Viruses* Edited by: van Regenmortel MHV, Fauquet CM, Bishop DHL, Carsten EB, Estes MK, Lemon SM, Maniloff J, Mayo MA, McGeoch DJ, Pringle CR, Wickner RB. San Diego: Academic Press; 2000:599-621.
- Vapalahti O, Mustonen J, Lundkvist Å, Henttonen H, Plyusnin A, Vaheri A: **Hantavirus infections in Europe**. *Lancet* 2003, **361**:653-661.
- Nemirov K, Vapalahti O, Lundkvist Å, Vasilenko V, Golovljova I, Plyusnina A, Niemimäa J, Laakkonen J, Vaheri A, Plyusnin A: **Isolation and characterization of Dobrava hantavirus carried by the striped field mouse (*Apodemus agrarius*) in Estonia**. *J Gen Virol* 1999, **80**:371-379.
- Brus-Sjölander K, Golovljova I, Plyusnin A, Lundkvist Å: **Serological divergence of Dobrava and Saaremaa hantaviruses: evidence for two distinct serotypes**. *J Epidemiol Infect* 2002, **128**:99-103.
- Avsic-Zupanc T, Nemirov K, Petrovec M, Trilar T, Poljak M, Vaheri A, Plyusnin A: **Genetic analysis of wild-type Dobrava hantavirus in Slovenia: co-existence of two distinct genetic lineages within the same natural focus**. *J Gen Virol* 2002, **81**:1747-1755.
- Sibold C, Ulrich R, Labuda M, Lundkvist Å, Martens H, Schutt M, Gerke P, Leitmeyer K, Meisel H, Krüger DH: **Dobrava hantavirus causes hemorrhagic fever with renal syndrome in central Europe and is carried by two different *Apodemus* mice species**. *J Med Virol* 2001, **63**:158-167.
- Plyusnin A: **Genetics of hantaviruses: implications to taxonomy (review)**. *Arch Virol* 2002, **147**:665-682.
- Nemirov K, Henttonen H, Vaheri A, Plyusnin A: **Phylogenetic evidence for host switching in the evolution of hantaviruses carried by *Apodemus* mice**. *Virus Res* 2002, **90**:207-215. Erratum 2003, **92**:125-126
- Wang H, Yoshimatsu K, Ebihara H, Ogino M, Araki K, Kariwa H, Wang Z, Luo Z, Li D, Hang C, Arikawa J: **Genetic diversity of hantaviruses isolated in china and characterization of novel hantaviruses isolated from *Niviventer confucianus* and *Rattus rattus***. *Virology* 2000, **278**:332-345.
- Klingström J, Hardestam J, Lundkvist Å: **Dobrava, but not Saaremaa, hantavirus is lethal and induces nitric oxide production in suckling mice**. *Microbes and Infection* 2005 in press.
- Klempa B, Schmidt HA, Ulrich R, Kaluz S, Labuda M, Meisel H, Hjelte B, Krüger DH: **Genetic interaction between distinct Dobrava hantavirus subtypes in *Apodemus agrarius* and *A. flavicollis* in nature**. *J Virol* 2003, **77**:804-809.
- Plyusnin A, Vaheri A, Lundkvist Å: **Genetic interaction between Dobrava and Saaremaa hantaviruses: now or millions of years ago?** *J Virol* 2003, **77**:7156-7157.
- Plyusnin A, Krüger DH, Lundkvist Å: **Hantavirus infections in Europe. (Review)**. *Adv Vir Res* 2001, **57**:105-136.
- Nemirov K, Andersen HK, Leirs H, Henttonen H, Vaheri A, Lundkvist Å, Plyusnin A: **Saaremaa hantavirus in Denmark**. *J Clin Virol* 2004, **30**:254-257.
- Schierup MH, Hein J: **Recombination and the molecular clock**. *Mol Biol Evol* 2000, **17**:1578-1579.
- Hillis DM, Bull JJ: **An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis**. *Syst Biol* 1993, **42**:182-192.
- Page RDM, Holmes EC: **Inferring molecular phylogeny**. In *Molecular evolution: a phylogenetic approach* UK: Blackwell Science Ltd; 1998:216-225.
- Hall T: **BioEdit. Biological sequence alignment editor for Windows**. 1998 [<http://www.mbio.ncsu.edu/BioEdit/bioedit.html>]. North Carolina State University, NC, USA
- Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG: **The CLUSTAL X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools**. *Nucl Acids Res* 1997, **25**:4876-4882.
- Felsenstein J: **PHYLIP – Phylogeny Inference Package (Version 3.2)**. 1989.
- Strimmer K, von Haeseler A: **Quartet puzzling: A quartet maximum likelihood method for reconstructing tree topologies**. *Mol Biol Evol* 1996, **13**:964-969.
- Posada D, Crandall KA: **MODELTEST: testing the model of DNA substitution**. *Bioinformatics* 1998, **14**:817-818.
- Swofford DL: **PAUP*: Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4**. Sinauer Associates, Sunderland, Massachusetts; 2003.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
http://www.biomedcentral.com/info/publishing_adv.asp

