<https://helda.helsinki.fi>

Estimation of main diversification time-points of hantaviruses using phylogenetic analyses of complete genomes

Castel, Guillaume

2017-04-02

Castel , G , Tordo , N & Plyusnin , A 2017 , ' Estimation of main diversification time-points of hantaviruses using phylogenetic analyses of complete genomes ' , Virus Research , vol. 233 , pp. 60-69 . <https://doi.org/10.1016/j.virusres.2017.03.011>

<http://hdl.handle.net/10138/235589>

<https://doi.org/10.1016/j.virusres.2017.03.011>

unspecified

publishedVersion

Downloaded from Helda, University of Helsinki institutional repository.

This is an electronic reprint of the original article.

This reprint may differ from the original in pagination and typographic detail.

Please cite the original version.

Estimation of main diversification time-points of hantaviruses using phylogenetic analyses of complete genomes

Guillaume Castel^{a,*}, Noël Tordo^{b,c}, Alexander Plyusnin^d

^a INRA-UMR 1062 CBGP, 755 Avenue Campus Agropolis, CS30016, 34988 Montpellier sur Lez, France

^b Unit Antiviral Strategies, Institut Pasteur, 25 Rue du Dr Roux, 75724 Paris Cedex 15, France

^c Institut Pasteur de Guinée, Gamal Abdel Nasser University, Conakry, Guinea

^d Department of Virology, University of Helsinki, Finland

ARTICLE INFO

Article history:

Received 7 November 2016

Received in revised form 8 March 2017

Accepted 9 March 2017

Available online 16 March 2017

Keywords:

Hantavirus
Diversification
Co-speciation
Phylogeny

ABSTRACT

Because of the great variability of their reservoir hosts, hantaviruses are excellent models to evaluate the dynamics of virus-host co-evolution. Intriguing questions remain about the timescale of the diversification events that influenced this evolution. In this paper we attempted to estimate the first ever timing of hantavirus diversification based on thirty five available complete genomes representing five major groups of hantaviruses and the assumption of co-speciation of hantaviruses with their respective mammal hosts. Phylogenetic analyses were used to estimate the main diversification points during hantavirus evolution in mammals while host diversification was mostly estimated from independent calibrators taken from fossil records. Our results support an earlier developed hypothesis of co-speciation of known hantaviruses with their respective mammal hosts and hence a common ancestor for all hantaviruses carried by placental mammals.

© 2017 Elsevier B.V. All rights reserved.

1. Introduction

Hantaviruses constitute the genus *Hantavirus* in the family *Bunyaviridae* (Plyusnin et al., 2012). The family includes three other genera of mammalian viruses: *Orthobunyavirus*, *Phlebovirus*, and *Nairovirus*, while the genus *Tospovirus* consists of plants viruses. Hantaviruses are an exception in the family *Bunyaviridae* since they are not transmitted by bites of arthropod vectors, but directly by aerosols from one infected mammal (rodent, bat or insectivore) to another, including human where hantaviruses can provoke Haemorrhagic Fever with Renal Syndrome (HFRS) or Hantavirus Cardio Pulmonary Syndrome (HCPS). However, mosquito-born (MOSBO) hantaviruses have been recently reported (Li et al., 2015), a discovery that might have important consequences and change our view on the hantavirus evolution in general.

The prototype *Hantaan virus* (HTNV), that has given the name to the genus, has been discovered back in the 1980-ies in Korea as a causative agent of a disease previously known as Korean Haemor-

rhagic Fever, then HFRS [for a review, see (Lee et al., 2014)]. Later, more hantaviruses provoking HFRS in humans have been found in Asia (*Seoul virus*, SEOV (Lee et al., 1982) and others) and Europe (*Puumala virus*, PUUV (Brummer-Korvenkontio et al., 1980), and others). In the Americas hantaviruses were mostly discovered later (LeDuc et al., 1984) and often associated with HCPS (Nichol et al., 1993; Padula et al., 1998). Hantaviruses were thought for several decades to be associated exclusively with rodents (order *Rodentia*) and although the first hantavirus isolated in cell culture, *Thottapalayam virus* (TPMV), originated from a shrew (Song et al., 2007), it was initially believed to represent a spill-over event from a rodent. Consequently, the collective term “ROBO-viruses” (from ROdent-BORne) has been initially coined. The recent outburst of novel hantavirus genotypes (likely to represent novel species as well) in shrews and moles (order *Eulipotyphla*) and later in bats (order *Chiroptera*) has proven the variety of mammal hosts being much wider than initially anticipated [for a review, see (Yanagihara et al., 2014; Zhang, 2014)]. These groups of hantaviruses could be called INBO- (from Insectivore-BORne) and BABO (from Bat-BORne) hantaviruses, respectively. Most recently, hantavirus partial genome sequences (encoding the L protein) have been recovered from arthropods: *Culex* and *Armigeres* mosquitoes (Li et al., 2015). This finding was in line with other publications on arthropod-specific

* Corresponding author.

E-mail addresses: guillaume.castel@supagro.inra.fr (G. Castel), ntordo@pasteur.fr (N. Tordo), Alexander.Plyusnin@Helsinki.Fi (A. Plyusnin).

bunya- and togaviruses that are thought to maintain in insects only (reviewed in (Junglen and Drosten, 2013)). These viruses would be logically called “MOSBO-hantaviruses” (from MOSquito-BORne) (this term is used in the present manuscript awaiting for a possible more appropriate one). In addition to their discovery in new hosts, hantaviruses have continued to expand geographically and plenty of novel genotypes/species were found in Africa, thus enlarging “the hantavirus world” to this continent (for a review, see (Witkowski et al., 2014)).

Although their medical importance was the initial reason to focus on hantaviruses (Jonsson et al., 2010), they rapidly appeared as excellent models to study virus evolution through phylogenetic analyses (Antic et al., 1992; Nichol et al., 1993; Plyusnin et al., 1996). Their genome evolves with relatively low speed and strong stabilizing selection (for a review, see (Holmes and Zhang, 2015; Sironen and Plyusnin, 2011)). The principal mechanism generating genetic diversity is genetic drift, i.e. gradual accumulation of point mutations, mostly neutral or quasi-neutral, and small deletions/insertions in the non-coding regions. Reassortment of genome RNA segments (Bennett et al., 2014) and, to a lesser extent, recombination may have also contributed (Plyusnin et al., 2002).

Although some examples of host switching have been described (Guo et al., 2013; Holmes and Zhang, 2015; Lin et al., 2012; Vapalahti et al., 1996), the main evolutionary forces, natural selection and genetic drift, have been shaping hantavirus diversification in their respective hosts and/or geographic location (Bennett et al., 2014). As hantaviruses have a tri-segmented RNA genome, they also have been subjected to reassortment events. During the two last decades, geographic clustering of genetic variants has been demonstrated, supporting the hypothesis of a close association and hence long-term co-speciation with natural hosts on a timescale of millions of years (MY) (Plyusnin and Morzunov, 2001; Plyusnin and Sironen, 2014; Sironen et al., 2001). This hypothesis has been challenged by several studies evidencing different timescales between viral and host evolution (Ramsden et al., 2009; Souza et al., 2014). However, these studies used short-term rates of evolution not well adapted to estimate very ancient divergence events in phylogenies (Castel et al., 2014; Sharp and Simmonds, 2011). Moreover, the use of common models (Wertheim and Kosakovsky Pond, 2011) with mutation rates calculated on recent viral genes evolving under strong purifying selection like hantavirus (Castel et al., 2014; Hjelte et al., 1995) can lead to severe underestimation of divergences for viral ancestors (Taylor et al., 2014; Wertheim and Kosakovsky Pond, 2011). Today, the gradual accumulation of complete genome sequences in the databases helps to precise the initial estimates based on partial sequences (see (Sironen et al., 2001) for example) and thus to improve our understanding of hantavirus genetics and speed of evolution.

One of the most intriguing question remains the timescale of diversification events in hantavirus evolution (Bennett et al., 2014; Holmes and Zhang, 2015). This point is difficult to address due to the lack of fossils of viruses. To overcome this point, an alternative consists at using fossil records of the host to calibrate the phylogenetic trees (Gustafsson et al., 2010; Ho et al., 2011). Of course, the origin of animals (Metazoa) itself might be viewed as controversial: e.g. the dating based on molecular data suggests around 800 million years, a projection that is older than the fossil evidence of the “Cambrian Explosion” estimated from 550 million years to 100 million years for the most recent estimations (Goswami, 2012). There are other limitations as well. But, those do not seem to make the approach less fruitful.

Based on the assumption that ancient hantaviruses were already present at the points of diversification of major placental clades (Plyusnin and Sironen, 2014) and would have co-evolved with their hosts, we attempt in this paper to estimate the main diversification time-points during hantavirus evolution for the last 100 MY

using the accumulated wealth of 35 hantaviruses with completely sequenced genomes. We confront them to the current knowledge of host diversification based on independent calibrators taken from fossil records.

2. Materials and methods

2.1. Sequence datasets

Table 1 displays the list of the selected the 35 complete coding genome sequences of hantaviruses, with the three segments S-M-L that were used either separately or concatenated for phylogenetic studies. They encompass *Murinae*-, *Arvicolinae*-, *Sigmodontinae*-, and *Neotominae*- ROBO- as well as INBO- and BABO- hantaviruses. In addition, the complete L protein-encoding sequence of the recently described MOSBO-hantavirus *Jianxia Mosquito Virus 2* (JMV-2) (Li et al., 2015) was compared to the 35 L protein-encoding sequences mentioned above.

2.2. Phylogenetic analyses

Multiple sequence alignments were prepared with the MUSCLE alignment program (Edgar, 2004) implemented in SEAVIEW v4.5.4 (Gouy et al., 2010). The GBLOCK program (Talavera and Castresana, 2007) was then used to remove gaps and poorly aligned positions to improve the pertinence of alignments for phylogenetic analyses. Phylogenies were inferred using the Maximum Likelihood (ML) approach in PhyML v3.0 (Guindon et al., 2010) (implemented in SEAVIEW v4.5.4.), with a statistical approximate likelihood ratio test (aLRT) of branch support. The Model Test function implemented in MEGA6 (Tamura et al., 2013) allowed to identify the GTR+G+I model (General Time Reversible, with the Gamma-distribution allowing some sites to be evolutionary invariable) as the optimal substitution model. The transition/transversion ratio was fixed to the value 4.0 and nucleotide (nt) frequencies were calculated from the dataset. Rate heterogeneity was applied using the discrete gamma distribution with four rate categories, and the shape parameter alpha was estimated from the dataset.

Phylogenetic trees were compared by calculating the Robinson-Foulds (RF) distance measuring the topological distance between unrooted phylogenetic trees (Robinson and Foulds, 1981) with the CompPhy program (Fiorini et al., 2014) available online on the ATGC bioinformatics platform at ATGC (2017) <http://www.atgc-montpellier.fr/compphy/>.

2.3. Evolutionary distances

Genetic distances between Hantavirus groups were calculated using functions implemented in the MEGA6 program. Analyses were conducted using the Maximum Composite Likelihood substitution model (for nucleotide—nt sequences) or the Poisson substitution model (for amino acid—aa sequences). The rate variation among sites was modeled with the gamma-distribution (shape parameter alpha = 1). All the other parameters were set to their default values.

2.4. Estimation of the divergence time points

Divergence times for all branching points in the phylogenetic tree were calculated with the RelTime method described by Tamura et al. (2012) and implemented in MEGA6. This method supports multiple user-defined calibration constraints. The Maximum Likelihood method based on the GTR-model was used and relative times were optimized and converted to absolute divergence times

Table 1

Complete hantavirus genomes selected for analysis (groups I, IIa, IIb, IIIa, and IIIb). The colours refer to those used in the phylogenetic trees in Figs. 1–3.

Isolate Name	Species Name	Abbreviation	GenBank Accession	Sequence Length	Collection Year	Sample source	GenBank Host	Country	Segment	Group
BT20	Laibin virus	LBNV	KM102247	1935	2012	Bat	Taphozous melanopogon	China	S	I
			KM102248	3908					M	
			KM102249	6531					L	
Cixi-Cl-23	Imjin virus	IMJV	KJ420559	1567	2013	Shrew	Crocidura lasiura	China	S	
			KJ420541	3649					M	
			KJ420567	6581					L	
BE/Namur/TE/2013/1	Nova virus	NVAV	KT004445	1826	2013	Mole	Talpa europaea	Belgium	S	
			KT004446	3590					M	
			KT004447	6563					L	
Wencheng-Sm-305	Thottapalayam virus	TPMV	KJ420560	1535	2013	Shrew	Suncus murinus	China	S	
			KJ420542	3635					M	
			KJ420568	6579					L	
76-118	Hantaan virus	HTNV	M14626	1696	1976	Mouse	Apodemus agrarius	South Korea	S	
			M14627	3616					M	
			X55901	6533					L	
SAA/Saaremaa/160V	Saaremaa virus	SAAV	AJ009773	1696	1996	Mouse	Apodemus agrarius	Estonia	S	
			AJ009774	3645					M	
			AJ410618	6532					L	
Slo/Af-BER	Dobrava-Belgrade virus	DOBV	GU904029	1672	2009	Mouse	Apodemus flavicollis	Slovenia	S	
			GU904035	3647					M	
			GU904042	6456					L	
Anjozorobe/Rr/MDG/2009/A TD56	Anjozorobe virus	ANZV	KC490916	1816	2008	Rat	Rattus rattus	Madagascar	S	
			KC490921	3586					M	
			KC490923	6496					L	
SA14	Sangassou virus	SANGV	JO082300	1746	2004	Mouse	Hylomyscus simus	Guinea	S	
			JO082301	3651					M	
			JO082302	6531					L	
80-39	Seoul virus	SEOV	NC005236	1769	N/A	Rat	Rattus norvegicus	South Korea	S	
			NC005237	3651					M	
			NC005238	6530					L	
ApJLCB2011-99	Amur-Soochong virus	ASV	JX473004	1695	2009	Mouse	Apodemus peninsulae	China	S	
			JX473003	3595					M	
			JX473002	6477					L	
"10-11"	Jeju virus	JJUV	HQ834695	1754	2010	Asian White-Toothed	Crocidura shantungensis	South Korea	S	
			HQ834696	3570					M	
			HQ834697	6588					L	
VN1512	Bowe virus	BOWV	KC631782	1731	2012	Shrew	Crocidura douceti	Guinea	S	
			KC631783	3597					M	
			KC631784	6551					L	
Fuyuan-Sr-326	Kenkeme virus	KENV	KJ857341	1597	2011	Shrew	Sorex roboratus	China	S	
			KJ857337	3635					M	
			KJ857320	6499					L	
3	Cao Bang virus	CBNV	EF543524	1833	2006	Shrew	Anourosorex squamipes	Viet Nam	S	
			EF543526	3637					M	
			EF543525	6533					L	
Tula/Moravia/5302v/95)	Tula virus	TULV	Z69991	1831	1995	Vole	Microtus arvalis	Czech Republic	S	
			Z69993	3694					M	
			AJ005637	6541					L	
Sotkamo	Puumala virus	PUUV	X61035	1830	1980	Vole	Myodes glareolus	Finland	S	
			X61034	3682					M	
			Z66548	6550					L	
Fuyuan-Mm-217	Khabarovsk virus	KHAV	KJ857342	1845	2011	Vole	Microtus maximowiczii	China	S	
			KJ857338	3706					M	
			KJ857321	6581					L	
Kitahiyama128/2008	Hokkaido virus	HOKV	AB675463	1832	2008	Vole	Myodes rufocanus	Japan	S	
			AB676848	3682					M	
			AB712372	6547					L	
isolate 11-1	Muju virus	MUJV	JX028273	1831	2011	Vole	Myodes regulus	South Korea	S	
			JX028272	3652					M	
			JX028271	6544					L	
LX309	Luxi virus	LXV	HM756286	2107	2009	Vole	Eothenomys miletus	China	S	
			HM756287	3663					M	
			HQ404253	6538					L	
Yuanjiang-Mf-13	Yuanjiang virus	YNJV	FJ170795	1861	2007	Vole	Microtus fortis	China	S	
			KJ857333	3653					M	
			KJ857316	6550					L	
Chile/123v/97	Andes virus	ANDV	AF291702	1871	1997	Rat	Oligoryzomys longicaudatus	Chile	S	
			AF291703	3671					M	
			AF291704	6562					L	
HV F0260003	Bayou virus	BAYV	GQ200820	1957	1996	Rice Rat	Oryzomys palustris (marsh rice rat)	USA	S	
			GQ244521	3664					M	
			GQ244526	6554					L	
HV C1280001	Catacamas virus	CATV	DQ256126	1943	2001	Rice Rat	Oryzomys couesi	Honduras	S	
			DQ177347	3616					M	
			FJ858378	6560					L	
200/2006	Huitzilac virus	HUIV	AB620106	1869	2006	Mouse	Reithrodontomys megalotis	Mexico	S	
			AB620107	3655					M	
			AB620108	6540					L	
BOR	Maripa hantavirus	MRPV	JO611712	1946	2009	Human	Human	French Guiana	S	
			JO611714	3684					M	
			JO611713	6562					L	
AN717313/BRA300	CASV-2 Hantavirus	CASV	JX443692	1876	2005	Rice Rat	Oligoryzomys utiariensis	Brazil	S	
			JX443702	3635					M	
			JX443698	6562					L	
H731172/BRA259	LANV-2	LANV	JX443704	1853	2007	Mouse	Calomys laucha	Brazil	S	
			JX443703	3665					M	
			JX443696	6527					L	
104/2006	Montano virus	MTNV	AB620100	2006	2006	Orizaba Deermouse	Peromyscus beatae	Mexico	S	
			AB620101	3636					M	
			AB620102	6529					L	
26/2006	Carrizal virus	CRZV	AB620103	1980	2006	Mouse	Reithrodontomys sumichrasti	Mexico	S	
			AB620104	3675					M	
			AB620105	6517					L	
NM R11	Sin Nombre virus	SNV	L37904	2060	1993	Mouse	Peromyscus maniculatus	USA	S	
			L37903	3696					M	
			L37902	6562					L	
AN693292/BRA293	Rio Mamore virus	RIOMV	JX443679	1939	2005	Rice Rat	Oligoryzomys sp.	Brazil	S	
			JX443701	3647					M	
			JX443697	6528					L	
VHV-574	Cano Delgadito virus	CADV	DQ285566	1986	1996	Cotton Rat	Sigmodon alstoni	Venezuela	S	
			DQ284451	3675					M	
			GQ200821	6578					L	
MSB57412	Rockport virus	RKPV	HM015223	1830	1986	Mole	Scalopus aquaticus	USA	S	
			HM015219	3647					M	
			HM015221	6558					L	
JX1-14	Jianxia Mosquito Virus	JMV-2	KM817671	5599	2013	Mosquito	Culex tritaeniorhynchus	China	L	

(shown next to branching points) based on user-supplied calibration constraints (defined below). A discrete Gamma distribution was used to model evolutionary rate differences among sites with the default parameters: 5 categories of substitutions, Gamma-distribution, the shape parameter $\alpha = 0.5030$.

The rate variation model allowed some sites to be evolutionarily invariable.

The inputs for the analyses were the calculated ML-trees (see “Phylogenetic analyses” above) and the corresponding MUSCLE nt sequence alignments (hantaviruses alone) or aa sequence alignment (hantaviruses and JMV-2).

The following calibration points were used (unit = million years ago (MYA)):

- 1) The split point (the Most Recent Common Ancestor, MRCA) of placental super-orders *Laurasiatheria* (including the orders *Eulipotyphla* and *Chiroptera*) from *Euarchontoglires* (including *Rodentia*): Time = 90 to 100 MYA (Bininda-Emonds et al., 2007; Meredith et al., 2011);
- 2) The order *Rodentia* starting to diversify: Time = 69.0 to 85.0 MYA (Bininda-Emonds et al., 2007; Meredith et al., 2011);
- 3) MRCA of all *Murinae* species: Time = 12.0 MYA; this value was assigned by Stepan et al. (2004), in contrast to Jacobs and Downs, (Jacobs and Downs, 1994) who assigned it earlier for MRCA of the families *Mus* and *Rattus* only;
- 4) First appearance of *Arvicolinae* rodents in fossil records: Time = 8.0 MYA (Engel et al., 1998; Jacobs and Lindsay, 1984);
- 5) First appearance of *Neotominae* rodents in fossil records: Time = 7.0 MYA (Leite et al., 2014; Slaughter and Ubelaker, 1984);

- 6) First appearance of *Sigmodontinae* rodents in fossil records: Time = 9.5 MYA (Engel et al., 1998; Webb et al., 1981);
- 7) The split point (MRCA) of *Myodes glareolus* from *Myodes rufocanus*: Time = 1.0 to 3.0 MYA (Kaneko et al., 1998; Sironen et al., 2001).

For the analyses of ROBO-, INBO- and BABO-hantaviruses, tree was rooted according to the phylogeny with JMV-2 and the age for the root of the tree (calibration point no. 1) was fixed between 90 and 100 MYA. For the analysis that included the sequence of MOSBO-hantavirus JMV-2, the age for the root of the tree was fixed at 900 MYA. This value corresponds to the mean of the divergence time estimates in different studies (560–1197 MYA) as determined by the Timetree webserver (Hedges et al., 2006) available at <http://timetree.org/search/pairwise/arthropods/Placentalia> <http://timetree.org/search/pairwise/arthropods/Placentalia>.

3. Results

3.1. Phylogenetic analyses using complete hantavirus genomes

Thirty-five complete hantavirus genome sequences have been selected for the analysis (Table 1). They include representatives of all three previously established groups of hantaviruses associated with mammalian hosts (Hartenberger, 2001): I, II and III (both represented by two subgroups). However, the diversity within and between each group shown in Fig. 1 is clearly different. Actually, one could upgrade at least some subgroups: for example, the patristic distance (the sum of length of all branch connecting the considered species in a phylogenetic tree) within group IIb is higher

Fig. 1. Global picture of hanta-evolution. Hantavirus phylogenetic unrooted tree(s) calculated on (A) S segment ORF, (B) M segment ORF, (C) L segment ORF, and (D) concatenated ORFs: Groups IIb, IIIa and IIIb correspond to ROBO viruses; Groups I and IIb to INBO and BABO viruses (listed in Table 1). Phylogenies were inferred using the Maximum Likelihood (ML) approach (in PhyML v3.0 implemented in SEAVIEW v4.5.4.), with a GTR + G + I substitution model and a statistical approximate likelihood ratio test (aLRT) of branch support. Only branching supported by bootstrap values of ≥ 0.7 are considered as statistically significant. Virus names at the tips of the trees are followed by a color point indicating the animal where the virus was isolated; brown from Sigmodontinae, grey from Neotominae, green from Murinae, blue from Arvicolinae, red from insectivore and orange from bat.

Table 2

Robinson and Foulds (RF) distances measuring the topological difference between the S, M, L and concatenated S-M-L phylogenetic trees. RF distance is a metric which transform one tree into another, enabling the distance between any pair of trees to be calculated efficiently and providing an objective procedure for comparing different phylogenetic trees (Robinson and Foulds, 1981). RF distances were calculated using the CompPhy program (Fiorini et al., 2014).

S tree	M tree	L tree	Concatenated S-M-L tree
10			
10	6		
10	4	2	

than that within group III (a + b combined); however for more clarity, we have kept the established nomenclature that is based on host association.

The group I includes hantaviruses carried by members of the order *Eulipotyphla* (insectivore-borne, INBO) and *Chiroptera* (bat-borne, BABO): *Thottapalayam virus* (TPMV), *Imjin virus* (IMJV), *Nova virus* (NVAV), and *Laibin virus* (LBNV) (Plyusnin and Sironen, 2014; Yanagihara et al., 2014; Zhang, 2014).

The group II and III are more complex and deserve analysis by sub-grouping. Subgroups IIa, IIIa and IIIb correspond to hantaviruses hosted by members of the order *Rodentia* (rodent-borne, ROBO). They are associated, with families/sub-families *Muridae/Murinae*, *Cricetidae/Arvicolinae* and *Cricetidae/Sigmodontinae/Neotominae* hosts, respectively. The group IIa includes: *Hantaan virus* (HTNV), *Amur-Soochong virus* (ASV), *Anjzorobe virus* (ANZV), *Dobrava virus* (DOBV), *Saaremaa virus* (SAAV), *Sangassou virus* (SANGV), and *Seoul virus* (SEOV). The group IIIa includes: *Hokkaido virus* (HOKV), *Khabarovsk virus* (KHAV), *Luxi virus* (LXV), *Muju virus* (MUJV), *Puumala virus* (PUUV), *Tula virus* (TULV), and *Juangjiang virus* (YNJV). The group IIIb includes: *Andes virus* (ANDV), *Bayou virus* (BAYV), *Catacamas virus* (CATV), *Huitzilac virus* (HUIV), *Maripa virus* (MRPV), *CASV-2 virus* (CASV), *Laguna Negra virus* (LANV), *Montano virus* (MTNV), *Carrizal virus* (CRZV), *Sin Nombre virus* (SNV), *Rio Mamore virus* (RIOMV), and *Cano Delgado virus* (CADV). Interestingly, group IIb consists of INBO-viruses like group I, but both groups (IIb and I) are not monophyletic, group IIb sharing a common ancestor with ROBO-viruses from the group IIa. This group IIb includes: *Bowe virus* (BOWV), *Cao Bang virus* (CBNV), *Jeju virus* (JJUV), and *Kenkeme virus* (KENV).

Phylogenetic trees based on the complete coding regions (ORF) of either S-, M-, and L segments or concatenated segments (Fig. 1A–D, respectively) were globally congruent suggesting a similar evolutionary history for the different viral genes. Some exceptions (listed below) indicated that reassortment event(s) might have been involved. When the RF distances were estimated (Table 2) the topology of the S-trees appeared more divergent than the topologies of other trees. This leads to the conclusion that the S segment, encoding the N nucleoprotein which covers the RNA genome within the Ribonucleoprotein (and that is most frequently used in phylogenetic studies) might not always be the best choice. Even if the use of S segment sequences gives generally good enough phylogenies, the use of longer sequences such as the concatenated full length genome provides more reliable results, attenuate the effect potentially linked to recombination and reassortment events and finally help to avoid misinterpretations. For example, some incongruities exist between phylogenetic positions of RKPV and LXV in the S, M and L segments phylogenetic trees suggesting that their segments might have different evolutionary histories (Bennett et al., 2014).

Phylogenetic clustering of hantaviruses resembles that of their hosts, with some important exceptions. INBO- and BABO-viruses from group I are clustered together in all phylogenies and form two pairs: (1) shrew-borne TPMV and IMJV and (2) bat-borne LBNV and mole-borne NVAV. This tendency of NVAV not being monophyletic with other INBO-viruses (TPMV, IMJV as well as Uluguru and Kilimanjaro viruses, not shown) but instead to branch together with BABO-viruses (LBNV as well as Xuan Son and Makokou viruses, not shown) has been recently evidenced in phylogenetic analysis using partial sequences (Witkowski et al., 2016). This may be explained by a spill-over from a BABO-virus to a mole. Alternatively, it may be that hantaviruses hosted by *Talpidae* animals from the order *Eulipotyphla* would be distinct from hantaviruses hosted by *Soricidae* mammals from the same order. More complete sequences of viruses isolated from moles and bats have to be analyzed to clarify this point. Anyway, the group I seems to be linked to other hantaviruses by the most ancient common ancestors (see later Fig. 2) and may have resulted from spill-over and co-diversification within the super-order *Laurasiatheria* between the orders *Chiroptera* and *Eulipotyphla*.

In all four phylogenies, *Murinae*-borne viruses (group IIa) form a cluster supported by high bootstrap values. INBO viruses of Group IIb are also branching together in all trees. The lack of monophyly of Old-World INBO-viruses of the groups IIb and I suggests different evolution histories. While the Group I virus NVAV clusters with bat-borne viruses such as LBNV (see above), Group IIb viruses might have resulted from a horizontal transmission of the ROBO-viruses associated with *Muridae*. Within group IIb, CBNV seems closer to KENV except in the S-tree where it was found more related to JJUV and BOWV, themselves always paired with a very significant bootstrap value. This atypical picture delivered by S segment-based phylogeny, when compared to the three other trees, is again observed for the *Arvicolinae*-borne viruses (group IIIa): LXV is more distant from the rest of the group and branches together with *Rockport virus* (RKPV) found in the mole *Scalopus aquaticus* (*Eulipotyphla* order, family *Soricidae*). RKPV itself is hesitating between the *Arvicolinae*-borne (group IIIa) and the *Sigmodontinae/Neotominae*-borne (group IIIb) hantaviruses and swings depending on the genome segment used to design the tree: from being closer to LXV in the S-tree, it becomes more associated to *Arvicolinae*-borne viruses in the M-tree and finally monophyletic and separated from both groups in the L- and concatenated-trees. One possible explanation for such atypical behavior might be a spill-over (or a host-switch) of a ROBO-hantavirus to moles. Interestingly, this observation further insists on the global receptivity of the insectivore reservoir for hantaviruses (Group 1, Group IIb INBO-viruses are spread across the tree) and more particularly the mole-reservoir with RKPV and NVAV which seem to have been acquired from rodent and bat, respectively.

The group IIIb includes all *Sigmodontinae/Neotominae*-borne hantaviruses. Positions of individual viruses may fluctuate between *Sigmodontinae* and *Neotominae* hosts, depending of the S-, M-, and L-trees. For example, the *Sigmodontinae*-borne HUIV is closer to the *Neotominae*-borne viruses, particularly to CRZV, in all trees, whereas, at the opposite, the *Neotominae*-borne RIOMV sits closer to *Sigmodontinae*-borne MPRV and LANV. The *Neotominae*-borne CADV and SNV are behaving more independently, sometimes monophyletic (S-tree) sometimes closer to the *Sigmodontinae*-borne (i.e. CATV in M- and L- trees).

In terms of evolutionary distances (Table 3), INBO- and BABO-members of the group I are the most distant from the others groups (up to 0.683 nt – 0.645 aa substitutions per site) which are themselves ranging from 0.443 nt to 0.325 aa subs/site (between *Arvicolinae*-borne viruses of group III-a and *Sigmodontinae/Neotominae* viruses of group III-b) to 0.559 nt – 0.472 aa

Fig. 2. Hantavirus complete coding genome timetree analysis by Maximum Likelihood method. The timetree shown was generated using the RelTime method. Divergence times for all branching points in the supplied topology were calculated using the Maximum Likelihood method based on the General Time Reversible model (GTR + G + I). Relative times were optimized and converted to absolute divergence times (shown in black next to branching points) based on supplied calibration constraints 1–7. Bars around each node represent 95% confidence intervals, which were computed using the method described in (Tamura et al., 2013). The tree is drawn to scale, with branch lengths measured in the relative number of substitutions per site. The analysis involved 35 nucleotide sequences. There were a total of 11364 positions in the final dataset. Evolutionary analyses were conducted in MEGA6 (Tamura et al., 2013). Time estimates are in Million of Years Ago (MYA) and the corresponding geological periods are indicated below. All pictures are taken from Wikimedia Commons. Photographers are Catania K (*S. aquaticus*), Josti A (*T. melanopogon*), Descouens D (*T. europaea*), Lindsey J (*A. flavicollis*), Salix S (*S. minutus*), Ritter S (*Crosidura Sp.*), Haye S (*M. glareolus*), Dieter TD (*M. arvalis*), Hogan CM (*R. megalotis*), Dias C (*C. laucha*), Hussein E Y (*O. longicaudatus*), unknown member of U.S federal government (*P. maniculatus* and *R. norvegicus*). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 3

Estimates of evolutionary distance over sequence pairs between groups. The number of base (under diagonal) and of amino acid (above diagonal) substitutions per site was calculated on averaging all sequence pairs between groups. Brackets indicate Standard error estimate(s). Analyses were conducted using the Maximum Composite Likelihood model (nt) or the Poisson correction model (aa). The rate variation among sites was modeled with a gamma distribution (shape parameter = 4). Ambiguous positions were removed for each sequence pair resulting in a total data set of 11364 (nt) and 3780 (aa) positions. Evolutionary analyses were conducted in MEGA6 (Tamura et al., 2013).

	I	II-a	II-b	III-a	III-b
I		0.634 [0.013]	0.645 [0.013]	0.631 [0.014]	0.627 [0.015]
II-a	0.676 [0.015]		0.420 [0.010]	0.472 [0.013]	0.459 [0.011]
II-b	0.683 [0.015]	0.512 [0.010]		0.460 [0.011]	0.420 [0.010]
III-a	0.671 [0.016]	0.559 [0.014]	0.541 [0.012]		0.325 [0.009]
III-b	0.673 [0.015]	0.554 [0.013]	0.514 [0.011]	0.443 [0.012]	

subs/site (between the Arvicolinae-borne viruses of group III-a and the Murinae-borne viruses of group II-a).

This supports the notion that the group I viruses are descendants of most ancient placental hantaviruses as already suggested (Witkowski et al., 2016).

3.2. Estimation of the divergence time points

In order to confront the divergent time points predicted by the phylogenetic trees based on hantavirus sequences to those estimated of the host species, we have searched the available literature for information about timing of the placental mammal radiation.

The split point (the Most Recent Common Ancestor, MRCA) of the placental super-order *Laurasiatheria* (including the orders *Eulipotyphla* and *Chiroptera*) from the super-order *Eumarchontoglires* (including the order *Rodentia*) has been evaluated at about 90–100 MYA (Bininda-Emonds et al., 2007; Meredith et al., 2011). Based on this assumption, we estimated the divergence dates in hantavirus phylogeny by fixing several published calibration points (see Material and Methods) to impose some temporal constraints to specific nodes in the tree. In order to validate this method we compared the results obtained by using two different options: (1) by fixing the 7 estimated calibration points (values in black at the nodes); (2) by omitting the calibration point n°1, i.e. the timing for splitting between super-orders *Laurasiatheria* and *Eumarchontoglires*. Both options resulted in very coherent estimates for the MRCA of ROBO-, INBO- and BABO-hantaviruses, within 71–109 MYA (medium fixed at 90) when using all seven calibration points and from 61 to 94 MYA (medium 78 MYA) when the first point is neglected (Supplementary Table 3).

Assuming this result, the phylogeny presented in Fig. 2 was inferred by fixing the 7 estimated calibration points.

The more interesting results were obtained concerning the points that were not fixed but deduced from the other calibrations. For example, the estimation of the divergence between *Sigmodontinae*–*Neotominae*-borne (group IIIa) from *Arvicolinae*-borne (group IIIb) hantaviruses was 30.7 MYA [24–37 MYA], i.e. in the late Oligocene (Supplementary Table 1). This finding is consistent with Fabre et al. (2012), who showed that divergence of rodent families might have occurred before the end of the Oligocene (around 31 MYA) and that the majority of radiations expanding rodent diversity occurred during the Neogene (around 22 MYA). Indeed, the beginning of the Oligocene was marked by a general cooling, increasing aridity and the development of open savannah habitats at the expense of forest environments. These environmental changes directly affected the fauna, especially in the Northern hemisphere. Europe experienced a mass extinction while Asian animals declined. In South America, isolated from other continents, humid forests became increasingly common, and allowed the development of a very typical wildlife with unique mammals (Hartenberger, 2001; Prothero, 1994; Weaver, 2002).

Another interesting result is the estimation of the different MRCA timings for the two groups of INBO-viruses: 70.4 [55–86 MYA] for the group I, and 58.3 [46–71 MYA], for the group IIb (Supplementary Table 1). According to a previous hypothesis (Plyusnin and Sironen, 2014), the first point might reflect the ancient split of INBO/BABO- from ROBO-hantaviruses while the second point is connected to the beginning of diversification of those INBO-hantaviruses, which would have originated through a host-switch from ROBO-viruses of the group IIa. This hypothetical host-switch might have happened between 64.9 MYA [51–79 MYA] (the MRCA of the whole group II viruses) and 58.3 MYA [46–71 MYA]. If true, this should affect all the later diversification point estimates such as for example MRCA of *Crocidura*-borne JUV and BOWV 39.1 MYA [30–48 MYA].

Finally, the phylogenetic tree in Fig. 2 predicts that the split of BABO- and INBO-viruses was around 52.1 [41–63 MYA] and that a common ancestor of *Cricetidae*-borne hantaviruses and the only currently known representative of yet another group of hantaviruses, RKPV, is dated to approximately same period: 50.2 [39–61 MYA].

4. Discussion

In this paper we attempted to estimate the first ever timing of diversification during hantavirus evolution based on thirty-five available complete genomes representing major groups of han-

taviruses and the assumption of co-speciation of hantaviruses with their respective mammal hosts. We agree that the method suffers several limitations such as 1) a limited and biased panel since focusing only on complete sequences, 2) the use of fossil calibrations and 3) the limit of the divergence dates estimation method itself. Nevertheless, the phylogenies presented globally support a hypothesis of co-evolution/co-speciation/co-divergence of hantaviruses with their respective mammal hosts and hence a common ancestor of for all hantaviruses carried by placental mammals. It is in particular suggested that INBO-hantaviruses of the group I split first while ROBO-viruses started to diversify later (Fig. 2), a conclusion in agreement with data of other authors (Witkowski et al., 2016). One possible scenario would include only one major host-switch between *Eulipotyphla* hosts and *Muridae* rodents (Plyusnin and Sironen, 2014). Alternative scenarios have been proposed as well (Bennett et al., 2014): they involve several major host-switches such as, for example, transmission of an ancient hantavirus from a shrew (order *Eulipotyphla*) ancestor into bats (order *Chiroptera*) and subsequently into rodents (order *Rodentia*) followed by even more transmissions within this order as well as secondary spill-overs to *Soricidae*.

As divergence dating by phylogeny is challenged by the lack of an independent method of time calibration for viruses (Bennett et al., 2014), the maximum likelihood approach was used in combination with the outside calibration points taken from the host fossil records. Indeed, the alternative of calculating molecular evolutionary rates between successive generations (microevolution), are expected to result in much higher rates than rates inferred from the fossil record (macroevolution). This molecular calibration is more convenient to be applied to the tips of the branches (“leaves”) rather than to the internal nodes (Ho et al., 2011). The diversification points-dating analysis revealed several important figures such as for example 41–63 MYA for a hypothetical split of BABO-viruses and INBO-viruses of the group I, and generally supports the notion that hantaviruses are ancient. This view received a strong support from the recent publication by Li and co-authors who described the first arthropod-associated hantavirus, *Jiangxia Mosquito Virus* type 2 (JMV-2), in *Culex* mosquitoes (Li et al., 2015). JMV-2 is found to be closer to *Hantavirus* genus than to any other virus clade/genus (see Figure Sup 2 of Li et al., 2015 for the phylogenetic analysis), even if JMV-2 showed low sequence identity to ROBO-, INBO- and BABO-hantaviruses: approximately 19%, at the amino acid sequence level. Most notably, when this virus (only a partial sequence of the L segment) was added to the dataset for the TimeTree analysis by fixing the root of the tree at 900 MYA (mean of the estimated split between arthropods and *Placentalia*), it formed a branch distinct from the rest of hantaviruses and occupied the most ancestral node on the tree (Fig. 3), in good agreement with the conclusions made by Li and co-authors that arthropods contain viruses falling basal to most groups of negative strand RNA viruses.

Thus hantaviruses might well be even more ancient than anticipated so far and their ancestry could be placed prior the diversification time-point of all placental animals.

It has been suggested that the “more ancient” hantavirus(es) may have been inhabiting arthropods (Plyusnin and Sironen, 2014; Yanagihara et al., 2014). One possible scenario would include transmission to *Placentalia* in which the virus(es) were first adapted to early *Eulipotyphla* or *Chiroptera* ancestor (Guo et al., 2013; Witkowski et al., 2016) before emerging in rodent species. This scenario is implying an ancient host switching of a hantavirus presumably occurring many millions of years ago (remnants in group I) and also more recent host-switches from rodents into *Eulipotyphla* members (viruses of the group II-b) and more recently to host(s) of RKPV and alike.

Fig. 3. Hantavirus and JMV-2L segment timetree analysis by Maximum Likelihood method. The timetree shown was generated using the RelTime method. Divergence times for all branching points in the user-supplied topology were calculated using the Maximum Likelihood method based on the LG model (Le and Gascuel, 2008). Relative times were optimized and converted to absolute divergence times (shown next to branching points) based on user-supplied calibration constraints. Bars around each node represent 95% confidence intervals, which were computed using the method described in (Tamura et al., 2013). A discrete Gamma distribution was used to model evolutionary rate differences among sites (5 categories (+G, parameter=0.4478)). The rate variation model allowed for some sites to be evolutionarily invariable (+I). The tree is drawn to scale, with branch lengths measured in the relative number of substitutions per site. The analysis involved amino acid sequences of 36 hantaviruses. There were a total of 2502 positions in the final dataset. Evolutionary analyses were conducted in MEGA6 (Tamura et al., 2013). All pictures are taken from Wikimedia Commons. Photographers are Gathary J (arthropods), Coin P (shrew), CSIRO (rat) and an unknown member of U.S federal government (bat). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Alternatively, one could imagine proper diversification of primordial hantavirus(es), existing already in a predecessor of arthropods and placentalia, following the split of hosts into different orders. It is tempting to speculate that the whole family *Bunyaviridae* might have been originally infecting arthropods exclusively and only later some “pioneers” were transmitted (horizontally) to plants and placental animals. This would suggest an early split between arthropod-associated hantaviruses and ROBO-, INBO- and BABO-hantaviruses.

Author contributions

GC and AP conceived the experiments; GC performed the phylogenetic analyses; GC, AP and NT analysed the data and wrote the paper.

Conflicts of interest

The authors declare no conflict of interest.

Acknowledgments

This work was supported by the European FP7 ANTIGONE Program no. 278976. The contents of this publication are the sole responsibility of the authors and don't necessarily reflect the views of the European Commission.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.virusres.2017.03.011>.

References

- ATGC 2017 South of France bioinformatics platform. French National Institute of Bioinformatics (IFB).
- Antic, D., Kang, C.Y., Spik, K., Schmaljohn, C., Vapalahti, O., Vaheri, A., 1992. Comparison of the deduced gene products of the L, M and S genome segments of hantaviruses. *Virus Res.* 24 (1), 35–46.
- Bennett, S.N., Gu, S.H., Kang, H.J., Arai, S., Yanagihara, R., 2014. Reconstructing the evolutionary origins and phylogeography of hantaviruses. *Trends Microbiol.* 22 (8), 473–482.
- Bininda-Emonds, O.R., Cardillo, M., Jones, K.E., MacPhee, R.D., Beck, R.M., Grenyer, R., Price, S.A., Vos, R.A., Gittleman, J.L., Purvis, A., 2007. The delayed rise of present-day mammals. *Nature* 446 (7135), 507–512.
- Brummer-Korvenkontio, M., Vaheri, A., Hovi, T., von Bonsdorff, C.H., Vuorimies, J., Manni, T., Penttinen, K., Oker-Blom, N., Lahdevirta, J., 1980. Nephropathia epidemica: detection of antigen in bank voles and serologic diagnosis of human infection. *J. Infect. Dis.* 141 (2), 131–134.
- Castel, G., Razzauti, M., Jousset, E., Kergoat, G.J., Cosson, J.F., 2014. Changes in diversification patterns and signatures of selection during the evolution of murinae-associated hantaviruses. *Viruses* 6 (3), 1112–1134.
- Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32 (5), 1792–1797.
- Engel, S.R., Hogan, K.M., Taylor, J.F., Davis, S.K., 1998. Molecular systematics and paleobiogeography of the South American sigmodontine rodents (vol 15, pg 35, 1998). *Mol. Biol. Evol.* 15 (2), 224–224.
- Fabre, P.H., Hautier, L., Dimitrov, D., Douzery, E.J., 2012. A glimpse on the pattern of rodent diversification: a phylogenetic approach. *BMC Evol. Biol.* 12, 88.
- Fiorini, N., Lefort, V., Chevenet, F., Berry, V., Chifolleau, A.M., 2014. CompPhy: a web-based collaborative platform for comparing phylogenies. *BMC Evol. Biol.* 14, 253.
- Goswami, A., 2012. A dating success story: genomes and fossils converge on placental mammal origins. *Evodevo* 3 (1), 18.
- Gouy, M., Guindon, S., Gascuel, O., 2010. SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol. Biol. Evol.* 27 (2), 221–224.
- Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., Gascuel, O., 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biol.* 59 (3), 307–321.
- Guo, W.P., Lin, X.D., Wang, W., Tian, J.H., Cong, M.L., Zhang, H.L., Wang, M.R., Zhou, R.H., Wang, J.B., Li, M.H., Xu, J., Holmes, E.C., Zhang, Y.Z., 2013. Phylogeny and origins of hantaviruses harbored by bats, insectivores, and rodents. *PLoS Pathog.* 9 (2), e1003159.
- Gustafsson, A.L., Verola, C.F., Antonelli, A., 2010. Reassessing the temporal evolution of orchids with new fossils and a Bayesian relaxed clock, with implications for the diversification of the rare South American genus *Hoffmannseggella* (Orchidaceae: Epidendroideae). *BMC Evol. Biol.* 10, 177.
- Hartenberger, J.-L., 2001. Une brève histoire des mammifères: bréviaire de mammalogie. Paris.
- Hedges, S.B., Dudley, J., Kumar, S., 2006. TimeTree: a public knowledge-base of divergence times among organisms. *Bioinformatics* 22 (23), 2971–2972.
- Hjelle, B., Krolkowski, 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 (1), 21–27.
- Ho, S.Y., Lanfear, R., Bromham, L., Phillips, M.J., Soubrier, J., Rodrigo, A.G., Cooper, A., 2011. Time-dependent rates of molecular evolution. *Mol. Ecol.* 20 (15), 3087–3101.
- Holmes, E.C., Zhang, Y.Z., 2015. The evolution and emergence of hantaviruses. *Curr. Opin. Virol.* 10, 27–33.
- Jacobs, L., Downs, W., 1994. The evolution of murine rodents in Asia. In: Tomida, Y., Li, C., Setoguchi, T. (Eds.), *Rodent and Lagomorph Families of Asian Origins and Diversification*. National Science Museum Monographs, Tokyo, pp. 149–156.
- Jacobs, L.L., Lindsay, E.H., 1984. Holarctic radiation of Neogene murid rodents and the origin of South American cricetids. *J. Vertebr. Paleontol.* 4 (2), 265–272.
- Jonsson, C.B., Figueiredo, L.T., Vapalahti, O., 2010. A global perspective on hantavirus ecology, epidemiology, and disease. *Clin. Microbiol. Rev.* 23 (2), 412–441.
- Junglen, S., Drosten, C., 2013. Virus discovery and recent insights into virus diversity in arthropods. *Curr. Opin. Microbiol.* 16 (4), 507–513.
- Kaneko, Y., Nakata, K., Saitoh, T., Stenseth, N.C., Bjørnstad, O.N., 1998. The biology of the vole *Clethrionomys rufocanus*: a review. *Researches Popul. Ecol.* 40 (1), 21–37.
- Le, S.Q., Gascuel, O., 2008. An improved general amino acid replacement matrix. *Mol. Biol. Evol.* 25 (7), 1307–1320.
- LeDuc, J.W., Smith, G.A., Johnson, K.M., 1984. Hantaan-like viruses from domestic rats captured in the United States. *Am. J. Trop. Med. Hyg.* 33 (5), 992–998.
- Lee, H.W., Baek, L.J., Johnson, K.M., 1982. Isolation of Hantaan virus, the etiologic agent of Korean hemorrhagic fever, from wild urban rats. *J. Infect. Dis.* 146 (5), 638–644.
- Lee, H.W., Vaheri, A., Schmaljohn, C.S., 2014. Discovery of hantaviruses and of the Hantavirus genus: personal and historical perspectives of the Presidents of the International Society of Hantaviruses. *Virus Res.* 187, 2–5.
- Leite, R.N., Kolokotronis, S.O., Almeida, F.C., Werneck, F.P., Rogers, D.S., Weksler, M., 2014. In the wake of invasion: tracing the historical biogeography of the South American cricetid radiation (Rodentia, Sigmodontinae). *PLoS One* 9 (6), e100687.
- Li, C.X., Shi, M., Tian, J.H., Lin, X.D., Kang, Y.J., Chen, L.J., Qin, X.C., Xu, J., Holmes, E.C., Zhang, Y.Z., 2015. Unprecedented genomic diversity of RNA viruses in arthropods reveals the ancestry of negative-sense RNA viruses. *Elife* 4.
- Lin, X.D., Wang, W., Guo, W.P., Zhang, X.H., Xing, J.G., Chen, S.Z., Li, M.H., Chen, Y., Xu, J., Plyusnin, A., Zhang, Y.Z., 2012. Cross-species transmission in the speciation of the currently known murinae-associated hantaviruses. *J. Virol.* 86 (20), 11171–11182.
- Meredith, R.W., Janecka, J.E., Gatesy, J., Ryder, O.A., Fisher, C.A., Teeling, E.C., Goodbla, A., Eizirik, E., Simao, T.L., Stadler, T., Rabosky, D.L., Honeycutt, R.L., Flynn, J.J., Ingram, C.M., Steiner, C., Williams, T.L., Robinson, T.J., Burk-Herrick, A., Westerman, M., Ayoub, N.A., Springer, M.S., Murphy, W.J., 2011. Impacts of the Cretaceous Terrestrial Revolution and KPg extinction on mammal diversification. *Science* 334 (6055), 521–524.
- Nichol, S.T., Spiropoulou, C.F., Morzunov, S., Rollin, P.E., Ksiazek, T.G., Feldmann, 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 (5135), 914–917.
- 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 (2), 323–330.
- Plyusnin, A., Morzunov, S.P., 2001. Virus evolution and genetic diversity of hantaviruses and their rodent hosts. *Curr. Top. Microbiol. Immunol.* 256, 47–75.
- Plyusnin, A., Sironen, T., 2014. Evolution of hantaviruses: co-speciation with reservoir hosts for more than 100 MYR. *Virus Res.* 187, 22–26.
- Plyusnin, A., Cheng, Y., Lehtvaslahti, H., Vaheri, A., 1996. Unique and conserved features of Tula hantavirus M gene encoding envelope glycoproteins G1 and G2. *Virus Genes* 12 (3), 257–263.
- Plyusnin, A., Kukkonen, S.K.J., Plyusnina, A., Vapalahti, O., Vaheri, A., 2002. Transfection-mediated generation of functionally competent Tula hantavirus with recombinant S RNA segment. *EMBO J.* 21 (6), 1497–1503.
- Plyusnin, A., Beaty, B., Elliott, R., Goldbach, R., Kormelink, R., Lundkvist, A., Schmaljohn, C., Tesh, R., 2012. *Bunyaviridae*. Elsevier.
- Prothero, D.R., 1994. The Eocene-Oligocene transition: paradise lost. Columbia University Press.
- Ramsden, C., Holmes, E.C., Charleston, M.A., 2009. Hantavirus evolution in relation to its rodent and insectivore hosts: no evidence for codivergence. *Mol. Biol. Evol.* 26 (1), 143–153.
- Robinson, D.F., Foulds, L.R., 1981. Comparison of phylogenetic trees. *Math. Biosci.* 53 (1), 131–147.
- Sharp, P.M., Simmonds, P., 2011. Evaluating the evidence for virus/host co-evolution. *Curr. Opin. Virol.* 1 (5), 436–441.
- Sironen, T., Plyusnin, A., 2011. Genetics and evolution of hantaviruses. *Bunyaviridae*, 61–94.
- Sironen, T., Vaheri, A., Plyusnin, A., 2001. Molecular evolution of Puumala hantavirus. *J. Virol.* 75 (23), 11803–11810.
- Slaughter, B.H., Ubelaker, J.E., 1984. Relationship of South American cricetines to rodents of north america and the old world. *J. Vertebr. Paleontol.* 4 (2), 255–264.
- Song, J.W., Baek, L.J., Schmaljohn, C.S., Yanagihara, R., 2007. Thottapalayam virus, a prototype shrewborne hantavirus. *Emerg. Infect. Dis.* 13 (7), 980–985.
- Souza, W.M., Bello, G., Amarilla, A.A., Alfonso, H.L., Aquino, V.H., Figueiredo, L.T., 2014. Phylogeography and evolutionary history of rodent-borne hantaviruses. *Infect. Genet. Evol.* 21, 198–204.
- Stephan, S., Adkins, R., Anderson, J., 2004. Phylogeny and divergence-date estimates of rapid radiations in murid rodents based on multiple nuclear genes. *Syst. Biol.* 53 (4), 533–553.
- TimeTree: the timescale of life.
- Talavera, G., Castresana, J., 2007. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Syst. Biol.* 56 (4), 564–577.
- Tamura, K., Battistuzzi, F.U., Billings, P., Murillo, O., Filipski, A., Kumar, S., 2012. Estimating divergence times in large molecular phylogenies. *Proc. Natl. Acad. Sci. U. S. A.* 109 (47), 19333–19338.
- Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* 30 (12), 2725–2729.

- Taylor, D.J., Ballinger, M.J., Zhan, J.J., Hanzly, L.E., Bruenn, J.A., 2014. [Evidence that ebolaviruses and cuevaviruses have been diverging from marburgviruses since the Miocene](#). *PeerJ* 2, e556.
- Vapalahti, O., Lundkvist, A., Kukkonen, S.K., Cheng, Y., Gilljam, M., Kanerva, M., Manni, T., Pejcoch, M., Niemimaa, J., Kaikusalo, A., Henttonen, H., Vaheiri, A., Plyusnin, A., 1996. Isolation and characterization of Tula virus, a distinct serotype in the genus Hantavirus, family Bunyaviridae. *J. Gen. Virol.* 77 (Pt. 12), 3063–3067.
- Weaver, N., 2002. *Oligocene*. *Anim. Sci. (Encyclopedia.com)*.
- Webb, S.D., MacFadden, B.J., Baskin, J.A., 1981. Geology and paleontology of the Love Bone Bed from the late Miocene of Florida. *Am. J. Sci.* 281 (5), 513–544.
- Wertheim, J.O., Kosakovsky Pond, S.L., 2011. Purifying selection can obscure the ancient age of viral lineages. *Mol. Biol. Evol.* 28 (12), 3355–3365.
- Witkowski, P.T., Klempa, B., Ithete, N.L., Auste, B., Mfuni, J.K., Hoveka, J., Matthee, S., Preiser, W., Kruger, D.H., 2014. [Hantaviruses in Africa](#). *Virus Res.* 187, 34–42.
- Witkowski, P.T., Drexler, J.F., Kallies, R., Lickova, M., Bokorova, S., Mananga, G.D., Szemes, T., Leroy, E.M., Kruger, D.H., Drosten, C., Klempa, B., 2016. [Phylogenetic analysis of a newfound bat-borne hantavirus supports a laurasiatherian host association for ancestral mammalian hantaviruses](#). *Infect. Genet. Evol.* 41, 113–119.
- Yanagihara, R., Gu, S.H., Arai, S., Kang, H.J., Song, J.W., 2014. [Hantaviruses: rediscovery and new beginnings](#). *Virus Res.* 187, 6–14.
- Zhang, Y.Z., 2014. [Discovery of hantaviruses in bats and insectivores and the evolution of the genus Hantavirus](#). *Virus Res.* 187, 15–21.