Phylogenetic Characterization of *Orthohantavirus dobravaense* (Dobrava Virus)

Mert Erdin, Ceylan Polat, Teemu Smura, Sercan Irmak, Ortac Cetintas, Muhsin Cogal, Faruk Colak, Ahmet Karatas, Mustafa Sozen, Ferhat Matur, Olli Vapalahti, Tarja Sironen, Ibrahim Mehmet Ali Oktem

We report complete coding sequences of *Orthohantavirus dobravaense* (Dobrava virus) Igneada strains and phylogenetic characterization of all available complete coding sequences. Our analyses suggested separation of host-dependent lineages, followed by geographic clustering. Surveillance of orthohantaviruses using complete genomes would be useful for assessing public health threats from Dobrava virus.

Orthohantaviruses are globally distributed. Until now, they have been detected in rodents, insectivores, and bats. Rodentborne orthohantaviruses, which are associated with human diseases, are divided into 3 major groups, murid-borne, non–*Arvicolinae cricetidae*-borne, and *Arvicolinae*-borne viruses, according to their phylogeny and host species (*1*). Murid-borne orthohantavirus species, such as *Orthohantavirus dobravaense* (Dobrava virus; DOBV) and *O. hantanense* (Hantaan virus), which are associated with hemorrhagic fever with renal syndrome in humans, are distributed in the Old World (*1*,*2*). Non–*A. cricetidae*-borne orthohantaviruses, such as *O. bayoui* (Bayou virus) or *O. sinnombreense* (Sin Nombre virus), which cause hantavirus cardiopulmonary syndrome in human infections, are found in the Americas (*1*,*2*). *Arvicolinae*-borne orthohantaviruses, such as *O. puumalaense* (Puumala virus; PUUV) or *O. prospectense* (Prospect Hill virus), are either nonpathogenic or mildly pathogenic for humans (*1*,*2*) and are found in both the Old and New Worlds; *Arvicolinae*-borne strains are thought to serve as an evolutionary bridge between the other 2 groups.

Orthohantaviruses can be transmitted to humans through inhalation of virus-containing aerosols of rodent excreta or direct contact with reservoir hosts (*2*). In European Union/European Economic Area countries, the numbers of collective orthohantavirus case reports fluctuated between 1,647 and 4,249 cases during 2016–2020 (*3*). For instance, in 2020, PUUV virus caused 1,204 cases, Hantaan virus 14 cases, and DOBV 7 cases from the reports that confirmed laboratory information available for the causative viruses. The highest number of cases of hemorrhagic fever with renal syndrome have been detected in southeastern Europe, with 2,375 cases reported in the Balkan region during 1952–2012, most caused by PUUV or DOBV (*3*,*4*). DOBV-positive rodents have recently been found in northeastern Italy, suggesting potential geographic expansion of this clade (*5*).

Surveillance studies in rodent populations are essential for understanding the dynamics of fluctuations. Earlier studies have shown that geographic barriers might play a role in genetic diversity and clade separation among DOBV (*6*,*7*). Also, obtaining whole-genome sequences is a crucial step in understanding potential viral genetic determinants of phenotypic changes that might affect disease severity among these viruses. We report complete coding sequences of *O. dobravaense* Igneada strain and phylogenetic characterization of all available complete coding sequences of DOBV.

The Study

DOBV has caused human cases and outbreaks in the northern coastal region of Turkey (*8*–*12*). In a previous study, DOBV seropositivity and RNA positivity were discovered in rodents captured in Kirklareli Province in Eastern Thrace in Turkey, and phylogenetic analysis based on partial DOBV genomes

Author affiliations: University of Helsinki, Helsinki, Finland (M. Erdin, T. Smura, O. Vapalahti, T. Sironen); Hacettepe University, Ankara, Turkey (C. Polat); Balıkesir University, Balıkesir, Turkey (S. Irmak); Bulent Ecevit University, Zonguldak, Turkey (O. Cetintas, M. Cogal, F. Colak, M. Sozen); Ömer Halisdemir University, Niğde, Turkey (A. Karatas); Dokuz Eylul University, Izmir, Turkey (F. Matur, I.M.A. Oktem)

DOI: https://doi.org/10.3201/eid3004.230912

suggested that DOBV strains from Igneada, Turkey, are closely related to strains from Balkan countries (*13*). To understand the phyloepidemiologic distribution of DOBV, we sequenced complete coding regions of DOBV Igneada strains (GenBank accession nos. MW055917–9) from 1 archived sample that had been partially sequenced in a previous study (*13*); we compared results from the phylogenetic analyses with all available complete DOBV coding sequences in GenBank. Because of the limited number (n = 16) of complete DOBV coding sequences for all 3 segments currently available in GenBank, in addition to 55 complete small (S), 25 medium (M), and 16 large (L) sequences, we also analyzed a larger dataset of partial S-segment sequences (Appendix, https://wwwnc.cdc.gov/EID/article/30/4/23-0912-App1.pdf).

Phylogenetic analyses (Figure 1, panel A; Appendix Figure 1) and pairwise nucleotide identities (Appendix Figure 2) suggested 8 major clusters, designated by their main distribution ranges: Mediterranean, Sochi, Saaremaa, Central Europe, Germany, Rusne Island, Lithuania, and Russia. Consistent with a previous study (*13*), the DOBV Igneada strains sequenced in this study grouped together with strains from the cluster from the Mediterranean region. Of note, most human DOBV cases were from this region (*14*). The Mediterranean clade is further

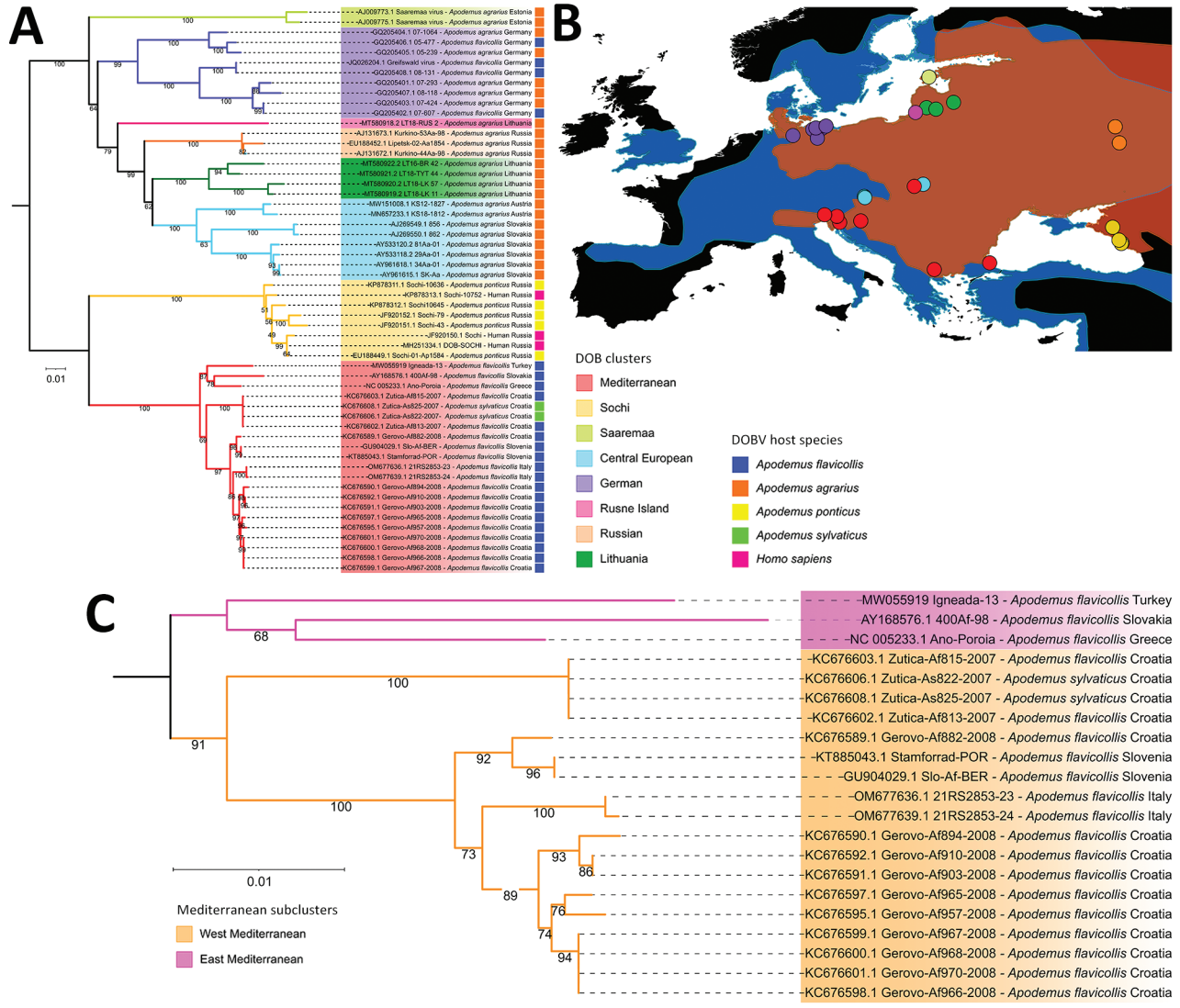


Figure 1. Phylogenetic characterization of DOBV combined with reservoir host and geographical distribution data. A) Maximum-likelihood tree based on all available complete DOBV sequences constructed using a transition plus empirical base frequencies plus gamma 4 substitution model. Colors indicate major clusters and hosts from which sequences were obtained. B) Distribution map of 2 major DOBV reservoir hosts, *Apodemus flavicollis* (blue) and *A. agrarius* (orange) mice, and their overlapping distribution zones. Solid circles indicate locations of complete sequences used in maximum-likelihood tree . C) Pruned version of the tree in panel A showing the division of the Mediterranean cluster into West and East Mediterranean subclusters. DOBV, Dobrava virus (*Orthohantavirus dobravaense*).

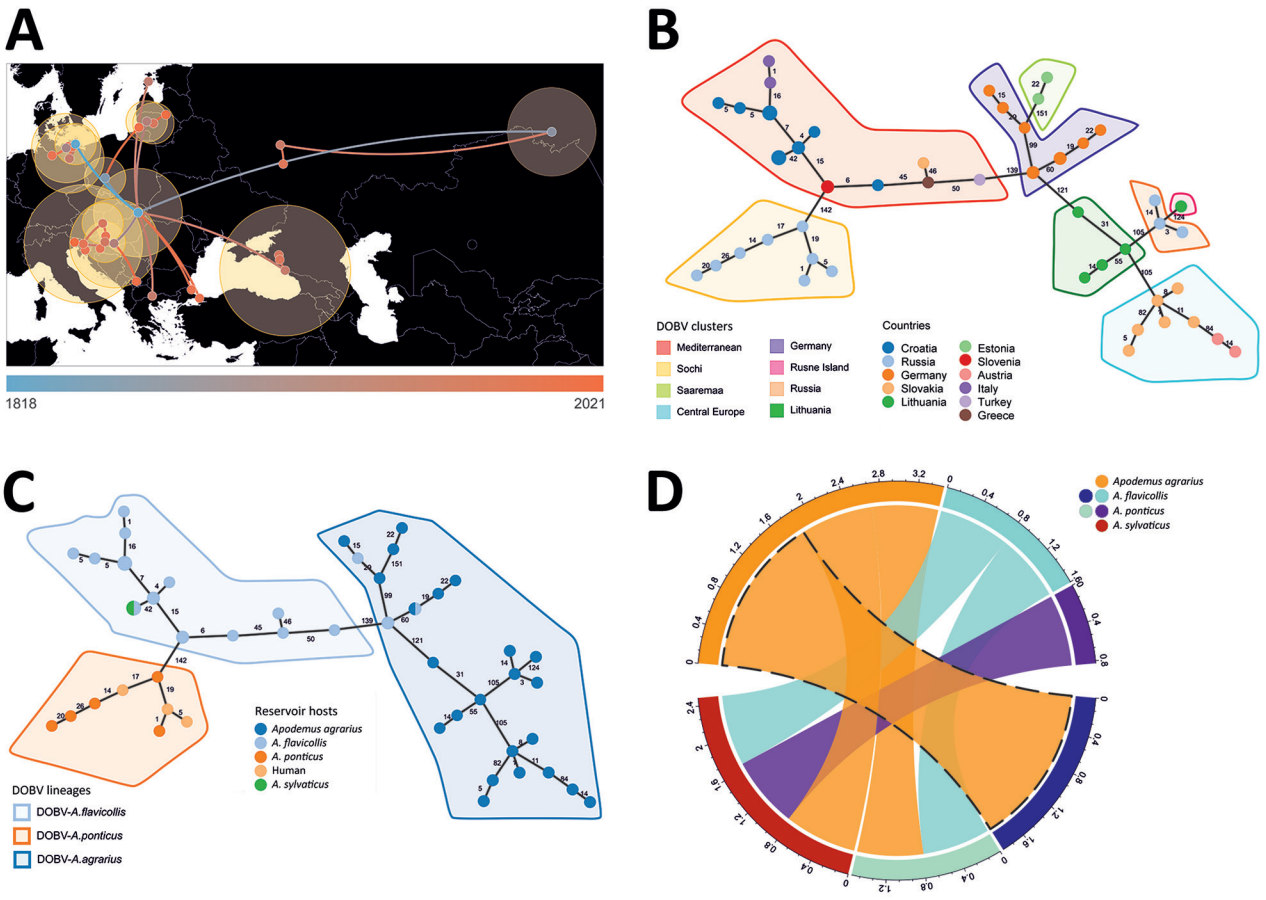


Figure 2. Host switching, phylogeographic reconstruction, and phylogenetic characterization of DOBV according to Bayesian analysis and minimum spanning tree constructions. A) Phylogeographic reconstruction of DOBV in discrete space. Each node is colored according to the estimated year of discovery, from the earliest (blue) to the latest (orange). Yellow shaded circles show the relative intensity of local viruses spread in the covered area. B) Minimum spanning tree showing the geographical cluster separation. C) Minimum spanning tree suggesting 3 major lineages according to reservoir host species. D) Chord diagram representing host switching rates of DOBV between 4 rodent species: *Apodemus flavicollis* (yellow-necked mouse), *A. agrarius* (striped field mouse), *A. ponticus* (Black Sea field mice), and *A. sylvaticus* (wood mice). DOBV, Dobrava virus (*Orthohantavirus dobravaense*).

regionally separated into West and East Mediterranean subclades (Figure 1, panel C). The West Mediterranean subclade consists of strains from Italy, Slovenia, Croatia, Hungary, and Kosovo; the East Mediterranean subclade comprises strains from Turkey, Greece, and eastern Slovakia (Appendix Figure 3).

Bayesian phylogeographic reconstruction based on all available complete and partial (>750 bases) S-segment sequences suggested that the estimated root location of DOBV is in Slovakia and Hungary in eastern Europe; from there, the virus has spread to other regions through multiple introductions, followed by local spreading (Figure 2, panel A). Minimum spanning tree phylogeny showing clear geographic clustering also supports that supposition (Figure 2, panel B). It should be noted, however, that sequence data are lacking for wide areas within the potential geographic distribution range of the main hosts of DOBV, and further studies are needed in those areas to confirm initial findings of clustering.

We derived DOBV sequences from 4 host species: *Apodemus flavicollis* (yellow-necked mice), *A. agrarius* (striped field mice), *A. sylvaticus* (wood mice), and *A. ponticus* (Black Sea field mice). Consistent with earlier studies (*15*), topology in the DOBV phylogenetic tree correlates with the geographic ranges of host species (Figure 1, panel A). Bayesian analysis suggested host-dependent lineage separation, followed by geographic clustering (Appendix Figure 3). The minimum spanning phylogenic tree correlated with the Bayesian analysis in showing clear host-dependent separation (Figure 2, panel C). In addition, our analysis suggested host-switching events between *A. flavicollis* and *A. agrarius* mice (Figure 2, panel D). The distribution ranges of *A. flavicollis* and *A. agrarius* mice overlap in eastern Europe and some parts of central Europe.

In northern Germany, there is a close phylogenetic relation of DOBV strains with those 2 reservoir hosts (Figure 1, panel B). Although probability estimates in our analysis did not support host-switching between the other host species, that lack of information might have resulted from lack of sufficient sequence data, especially on potential host-switching or spillover events between *A. flavicollis* and *A. sylvaticus* mice (Figure 1, panel A; Appendix Figure 3).

Conclusions

Tracking viral genetic changes using complete genome sequences to characterize viruses circulating in rodent populations is a crucial first step for understanding the spatiotemporal epidemiologic patterns of orthohantavirus-induced diseases and potential viral genetic determinants of virulence. Phylogenetic characterization of DOBV strains according to geographic regions within Europe and bordering countries suggests that more thorough genomic surveillance of orthohantaviruses, preferably using complete genomes, would be useful for assessing the DOBV-induced threat to public health.

This work was supported by the Federation of European Microbiological Societies, which provided a research and training grant to M.E., and by funding from Academy of Finland (grant number 339510).

About the Author

Mr. Erdin is a PhD student and doctoral researcher in University of Helsinki, Finland. His main research interests are discovery, characterization, evolution, and epidemiology of emerging and novel zoonotic viruses.

References

1. Mull N, Seifert SN, Forbes KM. A framework for understanding and predicting orthohantavirus functional traits. Trends Microbiol. 2023;31:1102–10. https://doi.org/10.1016/j.tim.2023.05.004
2. Vaheri A, Strandin T, Hepojoki J, Sironen T, Henttonen H, Mäkelä S, et al. Uncovering the mysteries of hantavirus infections. Nat Rev Microbiol. 2013;11:539–50. https://doi.org/10.1038/nrmicro3066
3. European Centre for Disease Prevention and Control. Hantavirus infection—annual epidemiological report for 2020. Stockholm: The Centre; 2023.
4. Avšič Županc T, Korva M, Markotić A. HFRS and hantaviruses in the Balkans/South-East Europe. Virus Res. 2014;187:27–33. https://doi.org/10.1016/j.virusres. 2013.12.042
5. Leopardi S, Drzewnioková P, Baggieri M, Marchi A, Bucci P, Bregoli M, et al. Identification of Dobrava-Belgrade virus in *Apodemus flavicollis* from North-Eastern Italy during enhanced mortality. Viruses. 2022;14:1241. https://doi.org/10.3390/v14061241
6. Korva M, Knap N, Rus KR, Fajs L, Grubelnik G, Bremec M, et al. Phylogeographic diversity of pathogenic and non-pathogenic hantaviruses in Slovenia. Viruses. 2013;5:3071–87. https://doi.org/10.3390/v5123071
7. Faber M, Krüger DH, Auste B, Stark K, Hofmann J, Weiss S. Molecular and epidemiological characteristics of human *Puumala* and *Dobrava-Belgrade hantavirus* infections, Germany, 2001 to 2017. Euro Surveill. 2019;24:1800675. https://doi.org/10.2807/1560-7917.ES.2019.24.32.1800675
8. Çelebi G, Öztoprak N, Öktem IMA, Heyman P, Lundkvist Å, Wahlström M, et al. Dynamics of *Puumala hantavirus* outbreak in Black Sea Region, Turkey. Zoonoses Public Health. 2019;66:783–97. https://doi.org/10.1111/zph.12625
9. Oncul O, Atalay Y, Onem Y, Turhan V, Acar A, Uyar Y, et al. Hantavirus infection in Istanbul, Turkey. Emerg Infect Dis. 2011;17:303–4. https://doi.org/10.3201/eid1702.100663
10. Kaya S, Yılmaz G, Erensoy S, Yağçı Çağlayık D, Uyar Y, Köksal I. Hantavirus infection: two case reports from a province in the Eastern Black Sea Region, Turkey [in Turkish]. Mikrobiyol Bul. 2010;44:479–87.
11. Ertek M, Buzgan T; Refik Saydam National Public Health Agency; Ministry of Health, Ankara, Turkey. An outbreak caused by hantavirus in the Black Sea region of Turkey, January–May 2009. Euro Surveill. 2009;14:19214. https://doi.org/10.2807/ese.14.20.19214-en
12. Oktem IM, Uyar Y, Dincer E, Gozalan A, Schlegel M, Babur C, et al. Dobrava-Belgrade virus in *Apodemus flavicollis* and *A. uralensis* mice, Turkey. Emerg Infect Dis. 2014;20:121–5. https://doi.org/10.3201/eid2001.121024
13. Polat C, Sironen T, Plyusnina A, Karatas A, Sozen M, Matur F, et al. *Dobrava hantavirus* variants found in *Apodemus flavicollis* mice in Kırklareli Province, Turkey. J Med Virol. 2018;90:810–8. https://doi.org/10.1002/jmv.25036
14. Avšič-Županc T, Saksida A, Korva M. Hantavirus infections. Clin Microbiol Infect. 2019;21S:e6–16. https://doi.org/10.1111/1469-0691.12291
15. Papa A. Dobrava-Belgrade virus: phylogeny, epidemiology, disease. Antiviral Res. 2012;95:104–17. https://doi.org/10.1016/j.antiviral.2012.05.011

Address for correspondence: Mert Erdin, University of Helsinki, Haartmaninkatu 3, 00290, Helsinki, Finland; email: mert.erdin@helsinki.fi

Article DOI: https://doi.org/10.3201/eid3004.230912

Phylogenetic Characterization of *Orthohantavirus dobravaense* (Dobrava Virus)

Appendix

Material and Methods

In earlier study, we obtained partial sequences of each segment of DOBV from Igneada region in Turkey (*1*). We used one of those sample in our complete genome sequencing. We used archived RNA extracted by using Invitrogen TRIzol (Thermo Fisher Scientific, https://www.thermofisher.com) following the manufacturer's guidelines. We used NEBNext rRNA depletion kit (human/mouse/rat) to remove host rRNA, and NEBNext Ultra II RNA library preparation kit (New England Biolabs, https://www.neb.com) to construct the sequencing library. We performed next-generation sequencing (NGS) using Illumina MiSeq system. We quality-filtered and de-novo assembled the raw data and annotated the contigs with LazyPipe (*2*). We filled the gaps in the sequences by designing primers to the genomic regions flanking the gaps (Appendix Table), performing polymerase chain reaction (PCR). These amplicons were sequenced by the Sanger method. We aligned our sequences for each segment separately with all available DOBV complete coding sequences of each encoded protein retrieved from the GenBank using ClustalW algorithm implemented in MegaX software and constructed maximum likelihood (ML) trees using IQ-TREE2 (http://www.iqtree.org) and ModelFinder for the best-fitted model for tree construction. We used PHYLOVIZ (https://www.phyloviz.net) for minimum spanning tree construction to support cluster hypothesis. We calculated pairwise identities from nucleotide sequences by using Sequence Demarcation Tool version 1.2 (University of Cape Town, http://web.cbio.uct.ac.za/~brejnev). For the phylogeographic reconstruction, host switching estimates, and Bayesian time tree construction of S segment, we used BEAST v1.10.4. In the BEAST analysis, the dataset included both complete coding

sequences and partial sequences which were equal to or longer than 750 nt and ended up total of 127 taxa for phylogeography and 107 taxa for host switching (DOBV sequences from human samples were excluded) analyses. Analysis parameters were as follows: tip dates enabled, host and geographic locations as discreet traits, uncorrelated relaxed clock for molecular clock, codon partition for each nucleotide separately to provide unique evolutionary rate for each position under Tamura-Nei 93 model with gamma categories as 5 and with invariant sites, population prior assumption to be constant, Markov chain Monte Carlo (MCMC) length to be 376.69×10^6 and echo sampling in every 10000 for phylogeographic reconstruction, and MCMC length to be 3.5×10^8 and echo sampling in every 1000 for host switching estimation. We used Tracer v1.7.2 to evaluate MCMC convergence for effective sample size to be >200 for each parameter. Bayesian trees were annotated to maximum clade credibility tree in TreeAnnotator v1.10.4. We used Spread3 v0.9.7.1 for the visualization of discreet phylogeographic reconstruction. Likelihood mapping assessment were done by IQ-TREE2 as 10000 quartets, and molecular saturation were extracted by DAMBE software with general time reversible distances. Data visualization was done in R v4.3.1/R studio. In earlier studies, it was hypothesized that recombination between DOBV strains may occur in nature (*3*). Thus, we tested our dataset for each segment with RDP version 5.30 (University of Cape Town, http://web.cbio.uct.ac.za/~darren/rdp.html) with all recombination testing methods implemented in this software package.

Appendix References

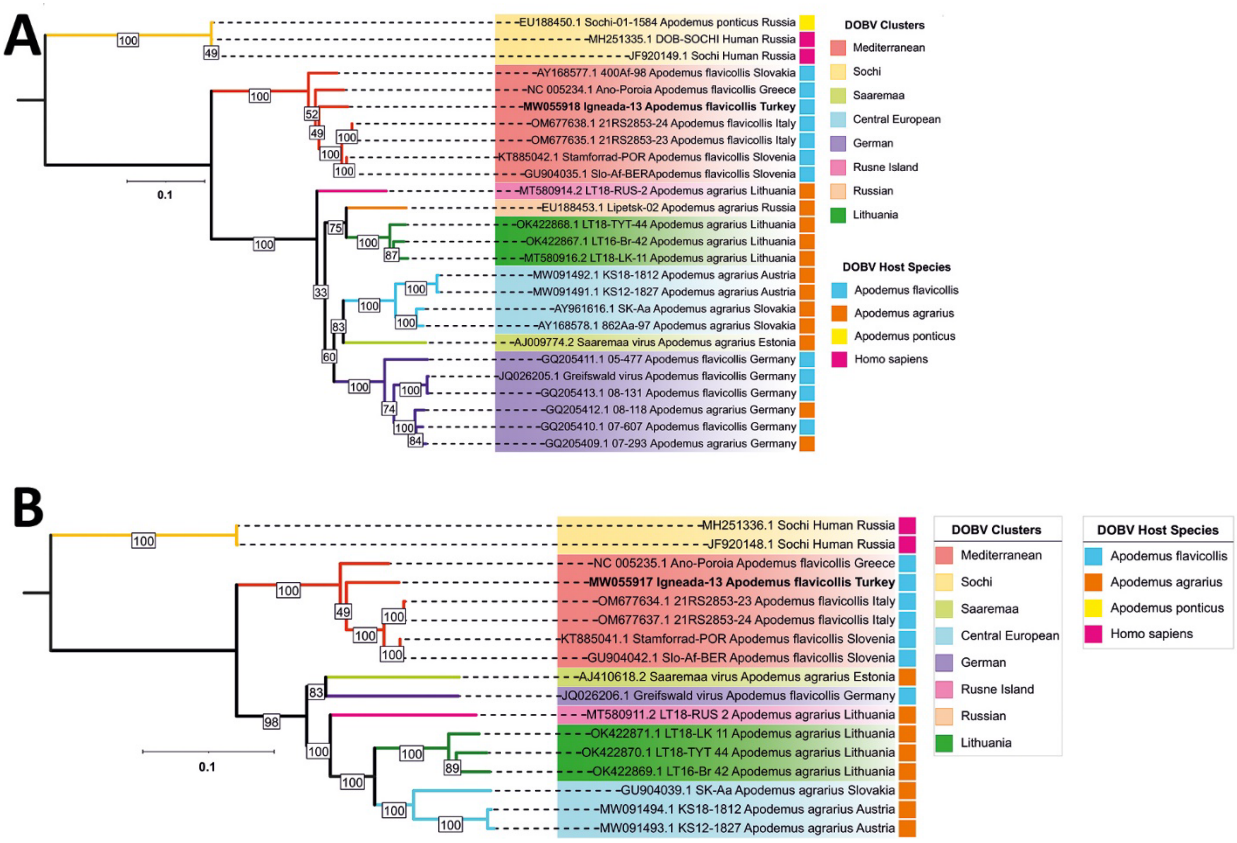
1. Polat C, Sironen T, Plyusnina A, Karatas A, Sozen M, Matur F, et al. *Dobrava hantavirus* variants found in *Apodemus flavicollis* mice in Kırklareli Province, Turkey. J Med Virol. 2018;90:810–8. PubMed https://doi.org/10.1002/jmv.25036

2. Plyusnin I, Kant R, Jaaskelainen AJ, Sironen T, Holm L, Vapalahti O, et al. Novel NGS pipeline for virus discovery from a wide spectrum of hosts and sample types. Virus Evol. 2020;6:veaa091.

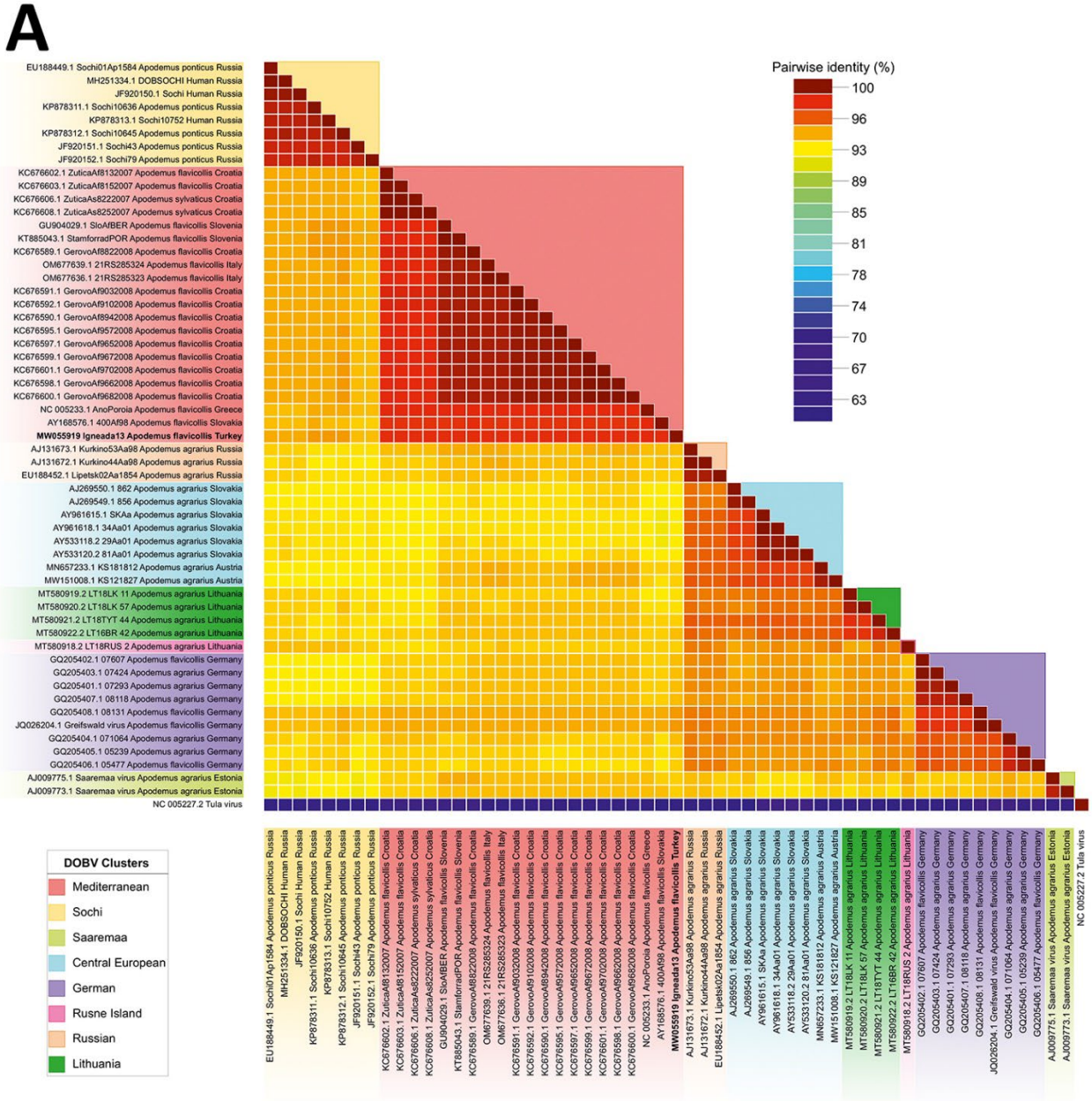
3. Klempa B, Schmidt HA, Ulrich R, Kaluz S, Labuda M, Meisel H, et al. Genetic interaction between distinct *Dobrava hantavirus* subtypes in *Apodemus agrarius* and *A. flavicollis* in nature. J Virol. 2003;77:804–9. PubMed https://doi.org/10.1128/JVI.77.1.804-809.2003

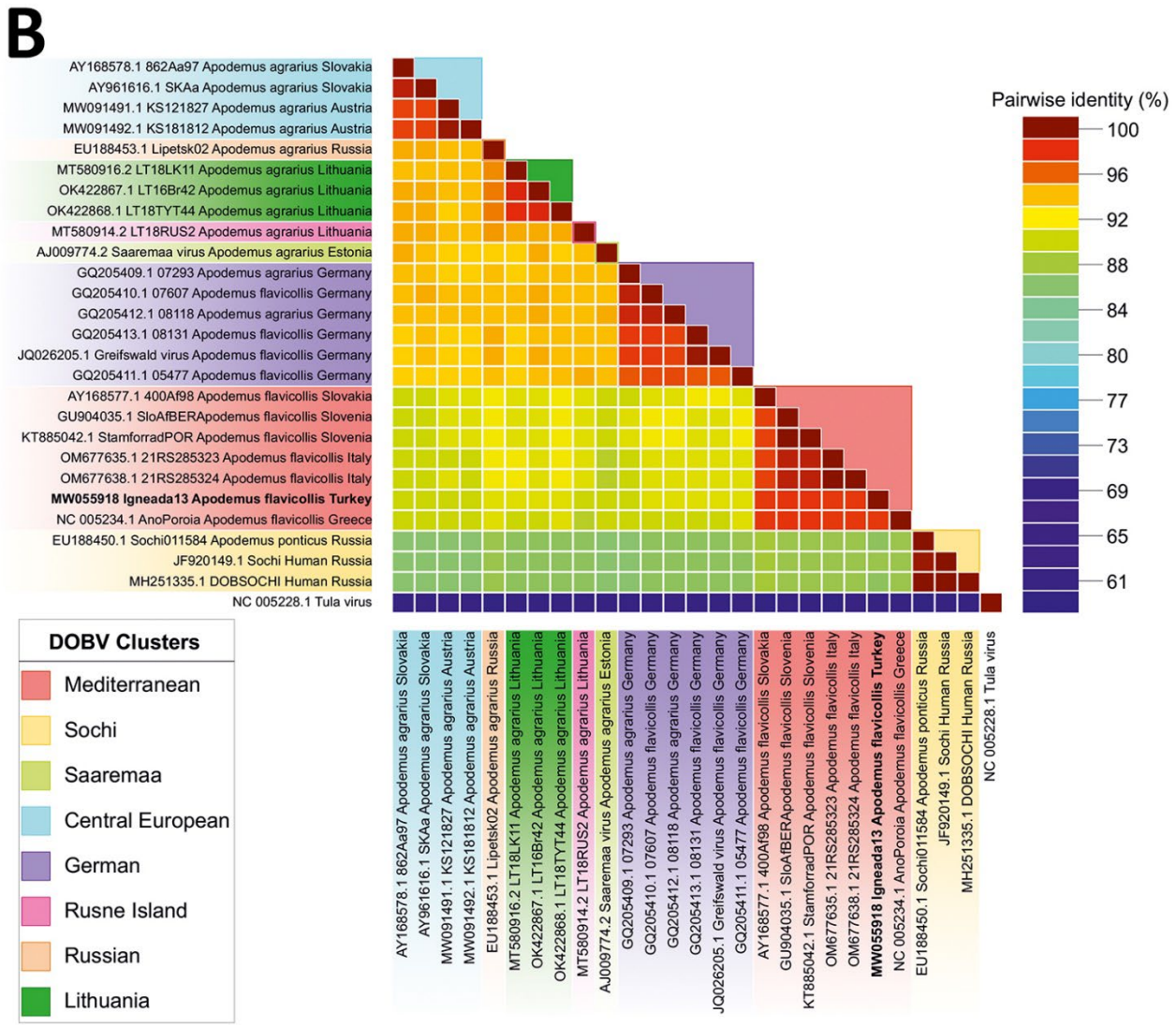
Appendix Table. List of designed primers to fill the gaps on the DOBV Igneada strain sequences by PCR and Sanger sequencing.

Primer	Sequence	Segment	Length
Forward_S_(G1)	ACAACCACGAAGGCCAACTG	S	20
Reverse_S_(G1)	TGTCCTGTAGTCTCATCAATGTC	S	23
Forward_S_(G2)	GATATGAGGAATACCATCATGGC	S	23
Reverse_S_(G2)	CCTAGTGCAAATACATCCACCAA	S	23
Forward_M_(G1)	GAGACAACATCAAGTGAGGTCAA	M	23
Reverse_M_(G1)	GAAACAATCCTGGGCTATAAACG	M	23
Forward_M_(G2)	GGTGTTACCGGACATTAAATCTC	M	23
Reverse_M_(G2)	CAGGATTACAGCCCCAACTG	M	20
Forward_L_(G1)	GAGGGATTGGTTATCAAAAAGCC	L	23
Reverse_L_(G1)	GTGGGTTCACTTATATTGAGCTC	L	23
Forward_L_(G2)	CGAAGTCTCAGGTTGTAGCTAA	L	22
Reverse_L_(G2)	GTTCAATAAAGCTCTCCCCAGA	L	22
Forward_L_(G3)	GAAGGCTGTGCTGTATCAATAC	L	22
Reverse_L_(G3)	TGCATGTAACCTAAAAGTGCC	L	21
Forward_L_(G4)	GAGGTAACTCAAGAAGATCTTG	L	22
Reverse_L_(G4)	GAAGGTCACCTTCATAGAGC	L	20
Forward_L_(G5)	CCCCTGCTGCATACTCATTAA	L	21
Reverse_L_(G5)	CCTTTTTGAGATACCAGAAGCT	L	22

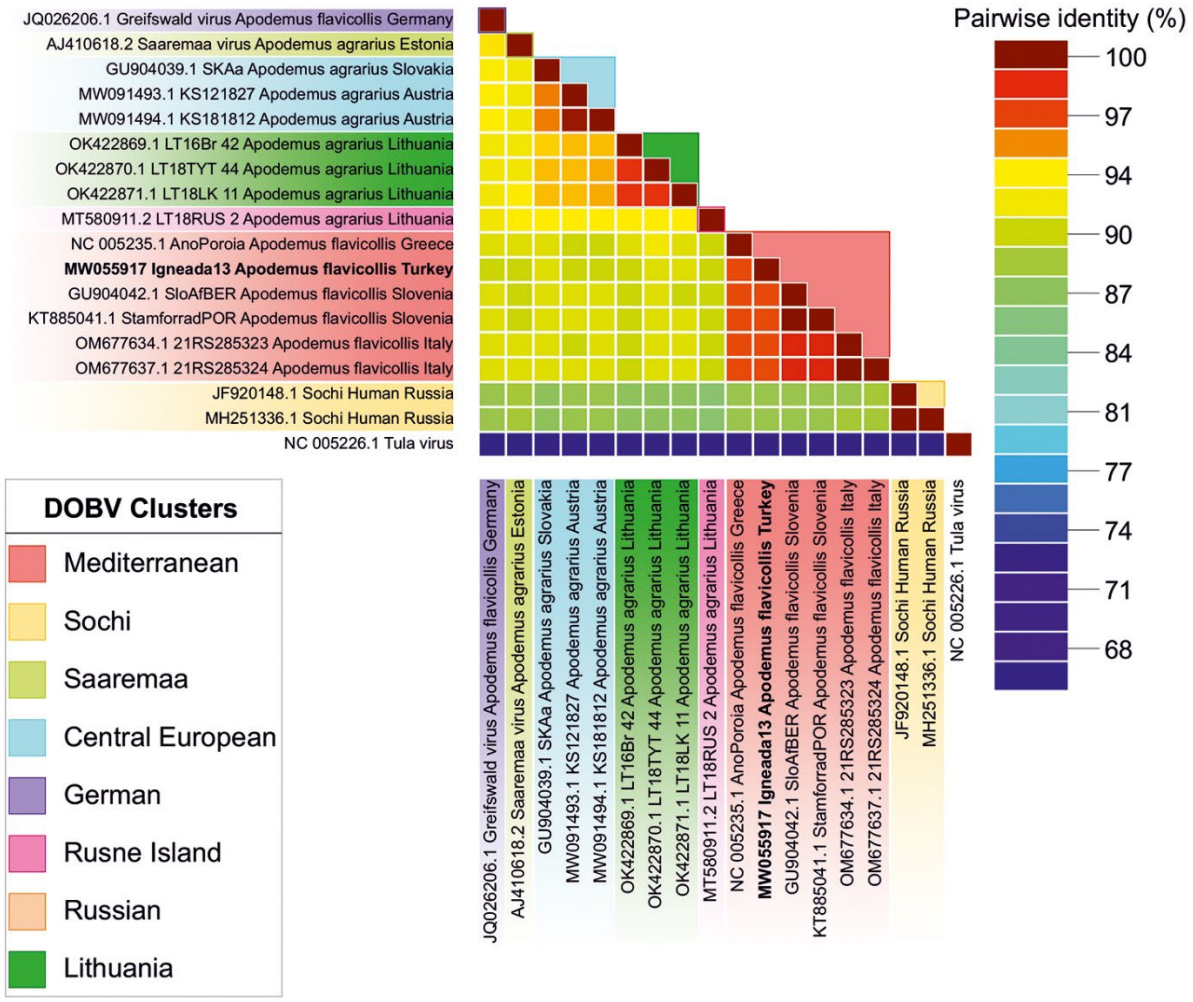


Appendix Figure 1. Phylogenetic characterization of DOBV combined with reservoir host and geographical distribution data. Maximum likelihood trees of A) M-segment based on all available complete DOBV sequences constructed using TIM2+F+I+R2 showed geographic clustering, and B) of L-segment based on all available complete DOBV sequences constructed using GTR+F+I+R2 showed geographical clustering in correlation with S and M segments.

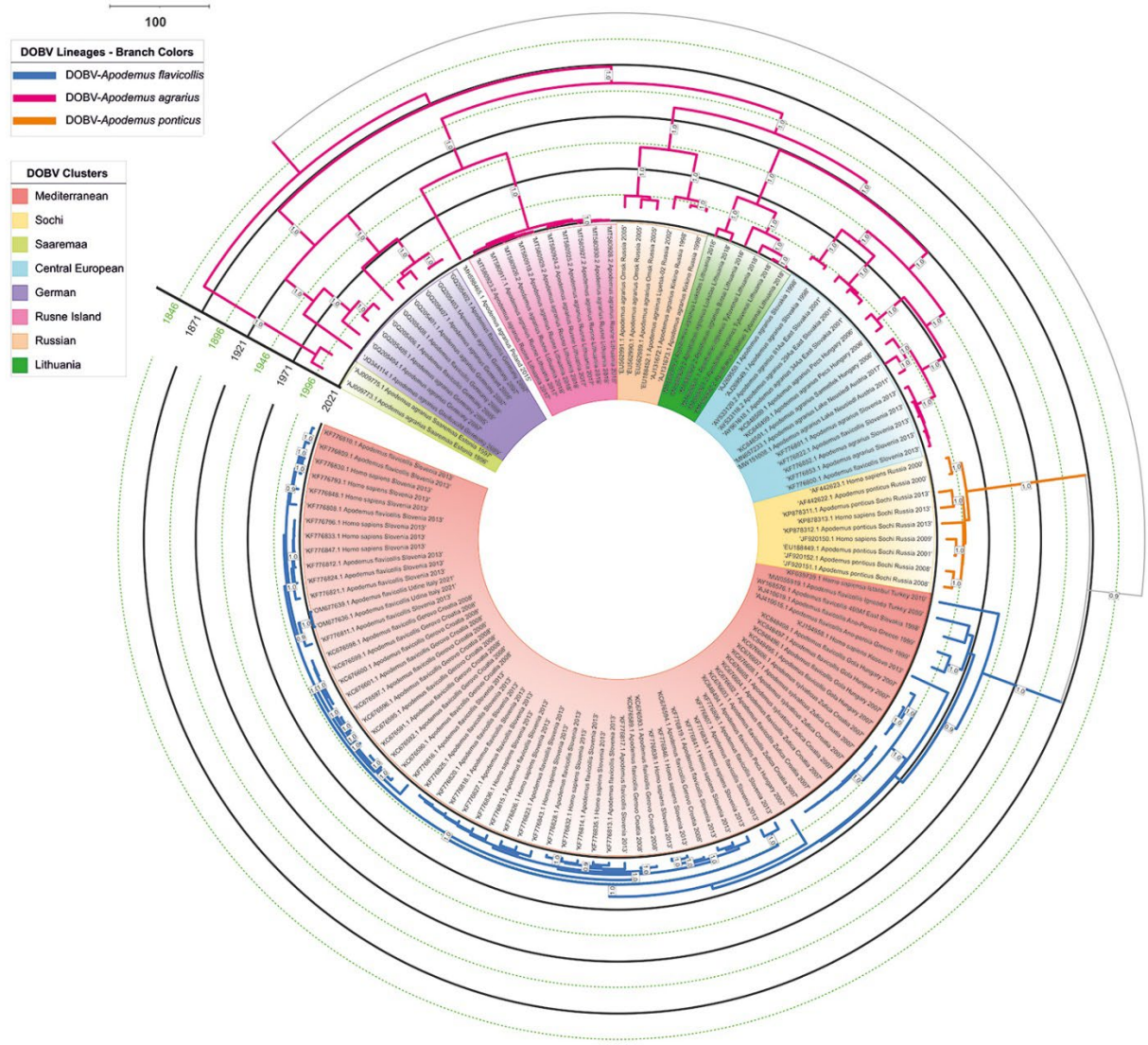

A
Pairwise identity (%)
100
96
93
89
85
81
78
74
70
67
63
EU188449.1 Sochi01Ap1584 Apodemus ponticus Russia
MH251334.1 DOBSOCHI Human Russia
JF920150.1 Sochi Human Russia
KP878311.1 Sochi10636 Apodemus ponticus Russia
KP878313.1 Sochi10752 Human Russia
KP878312.1 Sochi10645 Apodemus ponticus Russia
JF920151.1 Sochi43 Apodemus ponticus Russia
JF920152.1 Sochi79 Apodemus ponticus Russia
KC676602.1 ZuticaAf8132007 Apodemus flavicollis Croatia
KC676603.1 ZuticaAf8152007 Apodemus flavicollis Croatia
KC676606.1 ZuticaAs8222007 Apodemus sylvaticus Croatia
KC676608.1 ZuticaAs8252007 Apodemus sylvaticus Croatia
GU904029.1 SloAfBER Apodemus flavicollis Slovenia
KT885043.1 StamforradPOR Apodemus flavicollis Slovenia
KC676589.1 GerovoAf8822008 Apodemus flavicollis Croatia
OM677639.1 21RS285324 Apodemus flavicollis Italy
OM677636.1 21RS285323 Apodemus flavicollis Italy
KC676591.1 GerovoAf9032008 Apodemus flavicollis Croatia
KC676592.1 GerovoAf9102008 Apodemus flavicollis Croatia
KC676590.1 GerovoAf8942008 Apodemus flavicollis Croatia
KC676595.1 GerovoAf9572008 Apodemus flavicollis Croatia
KC676597.1 GerovoAf9652008 Apodemus flavicollis Croatia
KC676599.1 GerovoAf9672008 Apodemus flavicollis Croatia
KC676601.1 GerovoAf9702008 Apodemus flavicollis Croatia
KC676598.1 GerovoAf9662008 Apodemus flavicollis Croatia
KC676600.1 GerovoAf9682008 Apodemus flavicollis Croatia
NC 005233.1 AnoPoroia Apodemus flavicollis Greece
AY168576.1 400Af98 Apodemus flavicollis Slovakia
MW055919 Igneada13 Apodemus flavicollis Turkey
AJ131673.1 Kurkino53Aa98 Apodemus agrarius Russia
AJ131672.1 Kurkino44Aa98 Apodemus agrarius Russia
EU188452.1 Lipetsk02Aa1854 Apodemus agrarius Russia
AJ269550.1 862 Apodemus agrarius Slovakia
AJ269549.1 856 Apodemus agrarius Slovakia
AY961615.1 SKAa Apodemus agrarius Slovakia
AY961618.1 34Aa01 Apodemus agrarius Slovakia
AY533118.2 29Aa01 Apodemus agrarius Slovakia
AY533120.2 81Aa01 Apodemus agrarius Slovakia
MN657233.1 KS181812 Apodemus agrarius Austria
MW151008.1 KS121827 Apodemus agrarius Austria
MT580919.2 LT18LK 11 Apodemus agrarius Lithuania
MT580920.2 LT18LK 57 Apodemus agrarius Lithuania
MT580921.2 LT18TYT 44 Apodemus agrarius Lithuania
MT580922.2 LT16BR 42 Apodemus agrarius Lithuania
MT580918.2 LT18RUS 2 Apodemus agrarius Lithuania
GQ205402.1 07607 Apodemus flavicollis Germany
GQ205403.1 07424 Apodemus agrarius Germany
GQ205401.1 07293 Apodemus agrarius Germany
GQ205407.1 08118 Apodemus agrarius Germany
GQ205408.1 08131 Apodemus flavicollis Germany
JQ026204.1 Greifswald virus Apodemus flavicollis Germany
GQ205404.1 071064 Apodemus agrarius Germany
GQ205405.1 05239 Apodemus agrarius Germany
GQ205406.1 05477 Apodemus flavicollis Germany
AJ009775.1 Saaremaa virus Apodemus agrarius Estonia
AJ009773.1 Saaremaa virus Apodemus agrarius Estonia
NC 005227.2 Tula virus
DOBV Clusters
Mediterranean
Sochi
Saaremaa
Central European
German
Rusne Island
Russian
Lithuania


B
AY168578.1 862Aa97 Apodemus agrarius Slovakia
AY961616.1 SKAa Apodemus agrarius Slovakia
MW091491.1 KS121827 Apodemus agrarius Austria
MW091492.1 KS181812 Apodemus agrarius Austria
EU188453.1 Lipetsk02 Apodemus agrarius Russia
MT580916.2 LT18LK11 Apodemus agrarius Lithuania
OK422867.1 LT16Br42 Apodemus agrarius Lithuania
OK422868.1 LT18TYT44 Apodemus agrarius Lithuania
MT580914.2 LT18RUS2 Apodemus agrarius Lithuania
AJ009774.2 Saaremaa virus Apodemus agrarius Estonia
GQ205409.1 07293 Apodemus agrarius Germany
GQ205410.1 07607 Apodemus flavicollis Germany
GQ205412.1 08118 Apodemus agrarius Germany
GQ205413.1 08131 Apodemus flavicollis Germany
JQ026205.1 Greifswald virus Apodemus flavicollis Germany
GQ205411.1 05477 Apodemus flavicollis Germany
AY168577.1 400Af98 Apodemus flavicollis Slovakia
GU904035.1 SloAfBERApodemus flavicollis Slovenia
KT885042.1 StamforradPOR Apodemus flavicollis Slovenia
OM677635.1 21RS285323 Apodemus flavicollis Italy
OM677638.1 21RS285324 Apodemus flavicollis Italy
MW055918 Igneada13 Apodemus flavicollis Turkey
NC 005234.1 AnoPoroia Apodemus flavicollis Greece
EU188450.1 Sochi011584 Apodemus ponticus Russia
JF920149.1 Sochi Human Russia
MH251335.1 DOBSOCHI Human Russia
NC 005228.1 Tula virus
Pairwise identity (%)
100
96
92
88
84
80
77
73
69
65
61
DOBV Clusters
Mediterranean
Sochi
Saaremaa
Central European
German
Rusne Island
Russian
Lithuania

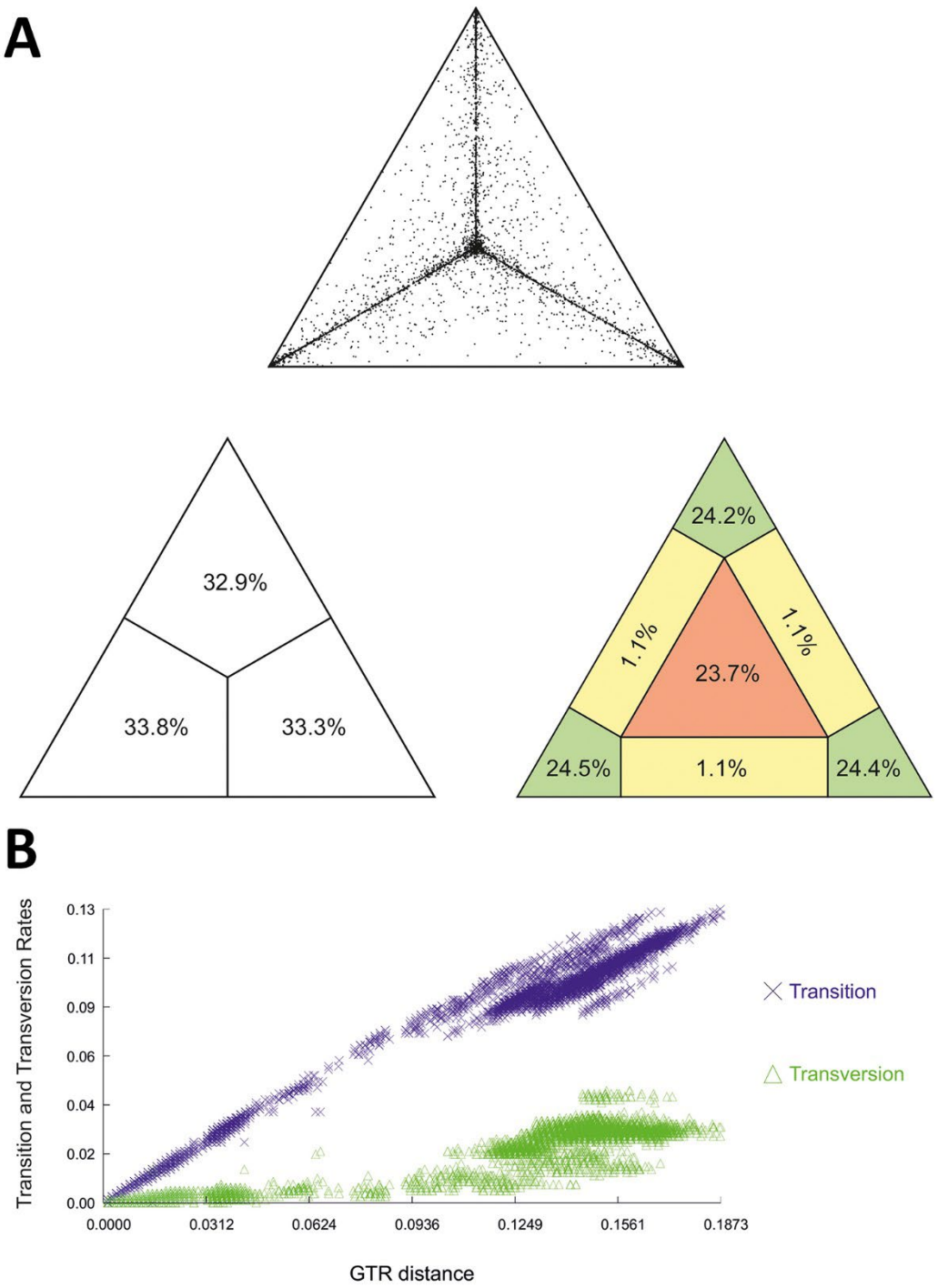
C



Appendix Figure 2. The pairwise identity matrices of A) S, B) M, and C) L segments showed correlating geographical clustering and different divergence among segments.



Appendix Figure 3. Bayesian maximum clade credibility (MCC) tree from phylogeographic reconstruction with total of 127 taxa. The MCC tree showed similar results as minimum spanning trees and ML trees by showing host-dependent lineage separation followed by geographic cluster separation. One sequence from Poland wasn't involved in any cluster under DOBV-*Apodemus agrarius* lineage due to insufficient data availability to make more detailed cluster hypothesis from that specific region.



Appendix Figure 4. The dataset phylogenetic information testing. (A) 73.1% of the quartets in the likelihood mapping placed at the corners of the triangle by being fully resolved, yet 23.7% of the quartets, as a big proportion, placed at the middle triangle and formed phylogenetically uninformative part of the assessment. (B) The molecular saturation of the dataset was none to low which provided the sight of some analysis estimates being underestimated.