Ticks and Tick-borne Diseases 10 (2019) 1168–1172

Contents lists available at ScienceDirect

Ticks and Tick-borne Diseases

journal homepage: www.elsevier.com/locate/ttbdis

Original article

"886-84-like" tick-borne encephalitis virus strains: Intraspecific status elucidated by comparative genomics

Renat V. Adelshin[a,b], Elena A. Sidorova[a], Artem N. Bondaryuk[a], Anna G. Trukhina[a], Dmitry Yu. Sherbakov[b,c], Richard Allen White III[d,e,f,g], Evgeny I. Andaev[a], Sergey V. Balakhonov[a]

[a] Irkutsk Anti-Plague Research Institute of Siberia and Far East, Trilisser 78, 664047, Irkutsk, Russia
[b] Irkutsk State University, Irkutsk, Russia
[c] Limnological Institute, Siberian Branch of the Russian Academy of Sciences, Irkutsk, Russia
[d] RAW Molecular Systems, Spokane, WA, United States
[e] Crop and Soil Sciences, Washington State University, Pullman, WA, United States
[f] Plant Pathology, Washington State University, Pullman, WA, United States
[g] Australian Centre for Astrobiology, University of New South Wales, Sydney, NSW, Australia

ARTICLE INFO

Keywords:
Tick-borne encephalitis virus (TBEV)
"886-84-like" strains
Phylogeny
Comparative genomics

ABSTRACT

Tick-borne encephalitis virus (TBEV) can cause severe meningitis, encephalitis, and meningoencephalitis. TBEV represents a pathogen of high zoonotic potential and an emerging global threat. There are three known subtypes of TBEV: Far-Eastern, Siberian and European. Since 2001 there have been suggestions that two new subtypes may be distinguished: "178-79" and "886-84". These assumptions are based on the results of the envelope gene fragment sequencing (Zlobin et al., 2001; Kovalev and Mukhacheva, 2017) and genotype-specific probes molecular hybridization (Demina et al., 2010). There is only one full-genome sequence of "178-79" strain and two identical ones of "886-84" strain can be found in GenBank. For clarification of the intraspecific position of the "886-84-like" strains group we completely sequenced six previously unknown "886-84-like" strains isolated in Eastern Siberia. As a result of applying different bioinformatics approaches, we can confirm that "886-84-like" strains group is a distinct subtype of TBEV.

1. Introduction

TBEV (tick-borne encephalitis virus) is transmitted by the bite of infected ticks or occasionally by ingestion of unpasteurized milk. Infection may induce a high fever and signs of central nervous involvement. Encephalitis developing during this second phase may result in paralysis, permanent sequelae or death (Bogovic and Strle, 2015).

TBEV belongs to the *Flavivirus* genus of the *Flaviviridae* (International Committee on Taxonomy of Viruses (ICTV) https://talk.ictvonline.org/taxonomy/). The genome of the virus is about 11,000 nucleotides (n.) which is a single open reading frame (ORF) encoding all structural and non-structural proteins at 10,245 n (Ecker et al., 1999). TBEV is divided into three subtypes: Far-Eastern, Siberian and European (Ecker et al., 1999). New TBEV subtypes (genotypes) may be distinguished: the "178-79 and the "886-84." This assumption is based both on the results of the envelope (E) gene fragment sequencing (Zlobin et al., 2001) and genotype-specific probes molecular hybridization (Demina et al., 2010). At the moment there is only one full-genome sequence of 178-79 strain (EF469661) available, and two identical ones of 886-84 strain (EF469662, KJ633033) are found in GenBank as prototype strains of these new subtypes. Also, eight fragments of the genome of the TBEV (from 255 to 2144 n.) of "886-84-like" strains isolated in the territory of Siberia and Mongolia are published. This work aimed to determine the intraspecific position of the "886-84-like strains group of TBEV based on analysis of full genome sequences.

2. Materials and methods

2.1. Ticks and small mammals

Six "886-84-like" strains of the TBEV have been isolated during 1999–2010 from taiga ticks (*Ixodes persulcatus*, n = 2) and small mammal brains (narrow-headed vole (*Microtus gregalis*, n = 1) and northern red-backed vole (*Myodes rutilus*, n = 3)) in the Trans-Baikal Territory of Eastern Siberia (Table1, Fig. 1). Ticks were sampled by flagging and collecting from the investigator's clothing. Rodents were collected by quick-killing traps.

E-mail address: adelshin@gmail.com (R.V. Adelshin).

https://doi.org/10.1016/j.ttbdis.2019.06.006
Received 3 August 2018; Received in revised form 7 June 2019; Accepted 9 June 2019
Available online 15 June 2019
1877-959X/

Table 1
Information about "886-84-like" strains isolated on the territory of Eastern Siberia (Trans-Baikal Territory). Alkhanay National Park: 50°44′49.31″N, 113°22′34.33″E; Chita city area: 52°05′52.78″N, 113°35′33.23″E.

№	Strain	Year	Source	Place	GenBank №
1	43/99	1999	*I. persulcatus* (Schulze, 1930)	Alkhanay National Park	MH490796
2	110/01	2001	*I. persulcatus* (Schulze, 1930)	Alkhanay National Park	MH481364
3	3033-1/10	2010	*Microtus gregalis* (Pallas, 1779)	Chita city area	MH481366
4	3094-9/10	2010	*Myodes rutilus* (Pallas, 1779)	Alkhanay National Park	MH481367
5	3094-18/10	2010	*Myodes rutilus* (Pallas, 1779)	Alkhanay National Park	MH490797
6	3094-29/10	2010	*Myodes rutilus* (Pallas, 1779)	Alkhanay National Park	MH481365

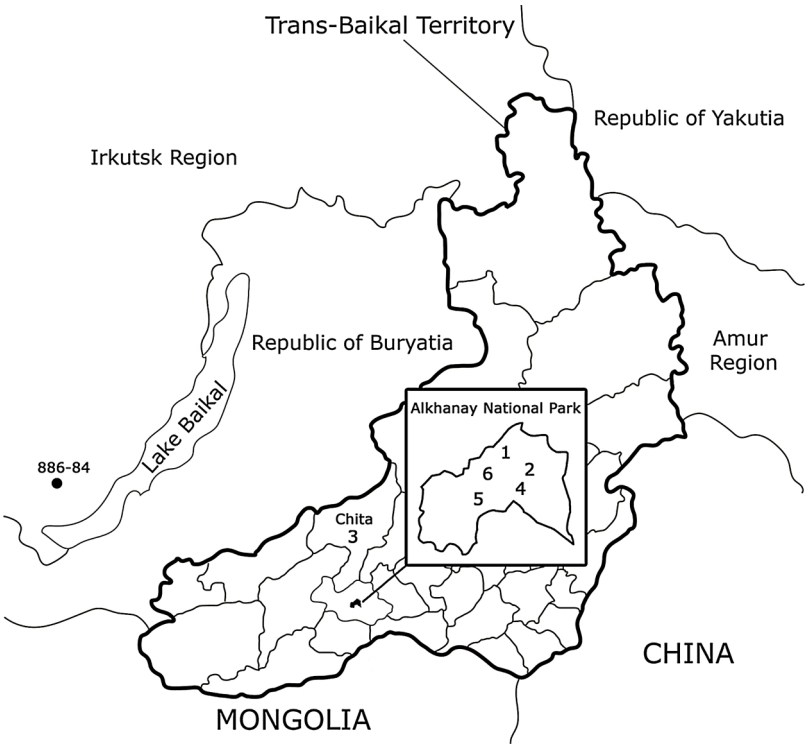

Fig. 1. Places of isolation of "886-84-like" strains of the TBEV in the Trans-Baikal Territory of Eastern Siberia. Notation (see Table 1): 1–43/99, 2–110/01, 3 – 3033-1/10, 4 – 3094-9/10, 5 – 3094-18/10, 6 – 3094-29/10. Isolation place of 886-84 prototype strain in Irkutsk Region is shown separately.

2.2. Virus isolation and RNA extraction

Each tick and small mammal's brain were homogenized in saline phosphate buffer (10% suspension) and then centrifuged. TBEV was isolated by intracerebral inoculation of newborn laboratory mice (Mahy, 1987) with 10% suspension. The animals were observed for 21 days. Total RNA was extracted from the brain tissue or tick (10% suspension in saline) using Riboprep kit ("NextBio," Russia).

2.3. Polymerase chain reaction (PCR) and sequencing

Amplification of the full genome of TBEV strains was performed via the reverse-transcription polymerase chain reaction (RT-PCR) using PCR kit ("Syntol," Russia) with specific original primers with the following cycling conditions: 95 °C–30 s; 95°C–10 s, 63°C–20 s, 72°C–30 s, 5 times; 95°C–10 s, 61°C–20 s, 72°C–30 s, 30 times. Primer sequences are listed in supplementary materials (supplemental Table 1). The PCR products have been sequenced with the ABI Prism Big Dye Terminator v.3.1 Cycle Sequencing Kit (Applied Biosystems, USA) and the Genetic Analyzer 3500xL (Applied Biosystems, Japan). Each PCR product (500–600 bp) was sequenced twice with forward and reverse primers. Primer sequences are listed in supplementary materials (supplemental Table 1). The PCR products have been sequenced with the ABI Prism Big Dye Terminator v.3.1 Cycle Sequencing Kit (Applied Biosystems, USA) and the Genetic Analyzer 3500xL (Applied Biosystems, Japan). Each PCR product (500–600 bp) was sequenced twice with forward and reverse primers.

2.4. Sequences and phylogenetic analysis

Analysis, alignment, and calculation of the identity of nucleotide sequences were carried out in the BioEdit program v. 7.0.5.3 (Hall, 1999). The model for phylogenetic analysis was selected by the jModelTest program2.1.10 (Darriba et al., 2012; Guindon and Gascuel, 2003) based on the minimum values of Akaike Information Criterion – AIC (Akaike, 1974) and Bayesian Information Criterion – BIC (Schwarz, 1978). Phylogenetic analysis was performed using BEAUti and BEAST 1.7 programs (Drummond et al., 2012) or PhyML program v. 3.3.20180129 (Guindon et al., 2010). The visualization of the analysis was carried out in the program FigTree v.1.4.2. (Rambaut, 2006-2014. Available from http://tree.bio.ed.ac.uk/software/figtree/). PhyML was also used for non-parametric bootstrap analysis. The topology was constrained to the best tree, 1000 bootstrap replicates of full TBEV genomes obtained and branch lengths estimated with PhyML. For each bootstrap reproduce the patristic distances were measured between the common ancestors of the four clades ("Far-Eastern," "Siberian," "European" and "886-84-like") with a custom script using ETE2 library for Python (Huerta- Cepas et al., 2010). The distances were collected, and

Table 2
The pairwise identity of nucleotide and amino acid sequences of TBEV groups (n. – nucleotides, a.a. – aminoacids).

	Sofjin (n./a.a.)	Vasilchenko (n./a.a.)	Neudoerfl (n./a.a.)	886-84 (n./a.a.)
Sofjin JX498940	*			
Vasilchenko AF069066	85.3/94.8	*		
Neudoerfl U27495	83.3/93.3	84.6/94.1	*	
886-84 EF469662	87.3/96.2	86.4/95.6	84.4/94.0	*

their density distributions were derived and plotted. All scripts used are available from http://github.com/dysh/886.

2.5. Divergence times estimation

To estimate divergence times for TBEV clades, the data were analyzed by the GTR nucleotide substitution model with a gamma distribution (4 categories) of among-site rate variation and a separate group of invariant sites. The relaxed clock model with an uncorrelated lognormal distribution of evolution rates was implemented since there was rate variation among branches (coefficient of variation is 0.27). A Bayesian skyline plot model was used as a population model. Statistical uncertainty in the tMRCA and evolution rates calculations was estimated as 95% highest probability density (HPD) intervals.

3. Results

The identity of nucleotide (10,245 n.) and amino acid (3415 a.a.) sequences of prototype strains of the TBEV was calculated pairwise (Table 2). By comparing the nucleotide and amino acid sequences of the prototype strains of the TBEV, strain 886-84 is identical to Neudoerfl (European subtype) by 84.4 n. / 94.0 a.a. %. The same identity value was found for the last one and the strain Vasilchenko of Siberian subtype: 84.6 n. / 94.1 a.a. %. The nucleotide and amino acid sequences of 886-84 strain is the closest to Sofjin (Far Eastern subtype) – 87.3 n./ 96.2 a.a.%.

The full-genome sequences of the six "886-84-like" strains form a homogeneous group on the phylogenetic tree (Fig. 2), the identity with the prototype strain (886-84 (EF469662, KJ633033)) is 99.4%. The number of variable sites within the group is 93 positions or 0.9% within the ORF.

*A comparison of the 1624 models (jModelTest 2.1.10) shows the minimum values of AIC (-lnL = 111,173.1247) and BIC (-lnL = 111,171.2409) corresponding to the GTR + G + I model (Lanave et al., 1984; Tavare, 1986; Gu et al., 1995; Waddell and Steel, 1997). The phylogenetic analysis was carried out with 164 full-genome sequences of the TBEV from GenBank and six sequences of "886-84-like" strains isolated in two regions of the Trans-Baikal Territory (Fig. 2).

The topology of the tree obtained in the BEAST 1.7 program (Drummond et al., 2012) is identical to the ML tree obtained with PhyML, which was used for phylogenetic distance calculation. After non-parametric bootstrap analysis in PhyML v. 3.3.20180129 (Guindon et al., 2010) the lengths of the branches were estimated. In the ETE2 package (standalone visualization script), the groups of sequences were selected according to the unconstrained tree topology except for the outgroup (AB507800, Omsk hemorrhagic fever virus, strain Guriev). A node (common ancestor) was found for each group. Internode patristic distances between common ancestors of the clades were calculated for each bootstrap replicate and used to obtain respective density distributions (Fig. 3).

Based on the results of the phylogenetic distance calculating between the common ancestors of the three subtypes and the group of "886-84-like" strains of the TBEV, a distance density graph has done (Fig. 4). The graph shows incomplete coincidence of the distances density between the Far Eastern and Siberian subtypes. The distances ratio between the European and Siberian subtypes is the same as between the Siberian subtype and the “886-84-like” strains group.

The mean evolution rate of the entire tree was 1.63×10^{-5} nucleotide substitutions/site/year^{-1} (95% HPD, 1.07×10^{-5}–2.27×10^{-5} nucleotide substitutions/site/year^{-1}). The mean evolution rate of the “886-84-like” strains clade was almost the same (1.53×10^{-5} nucleotide substitutions/site/year-1). The credible Bayesian interval of the "886-84-like" strains clade doesn't overlap with credible intervals of other subtypes. The estimated age of this group covers 212 years (122–333 years with 95% HPD) (Table 3).

4. Discussion

The incidence of tick-borne encephalitis in the period from 1996 to 2016 in the Trans-Baikal Territory ranged from 2.88 in 2004 to 11.62 in

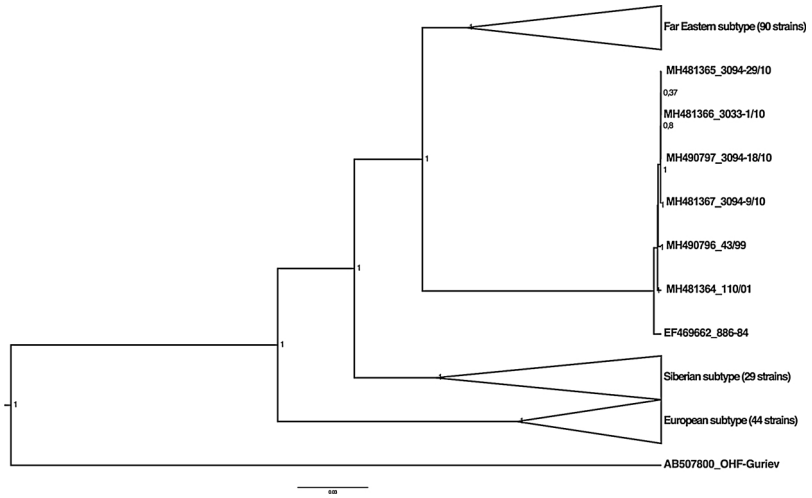

Fig. 2. The Bayesian phylogenetic tree of the TBEV full-genome sequences (10,245 n). The strict clock model of the evolution rate estimated with the GTR + G4 + I nucleotide substitution model and 100 million generations was used.

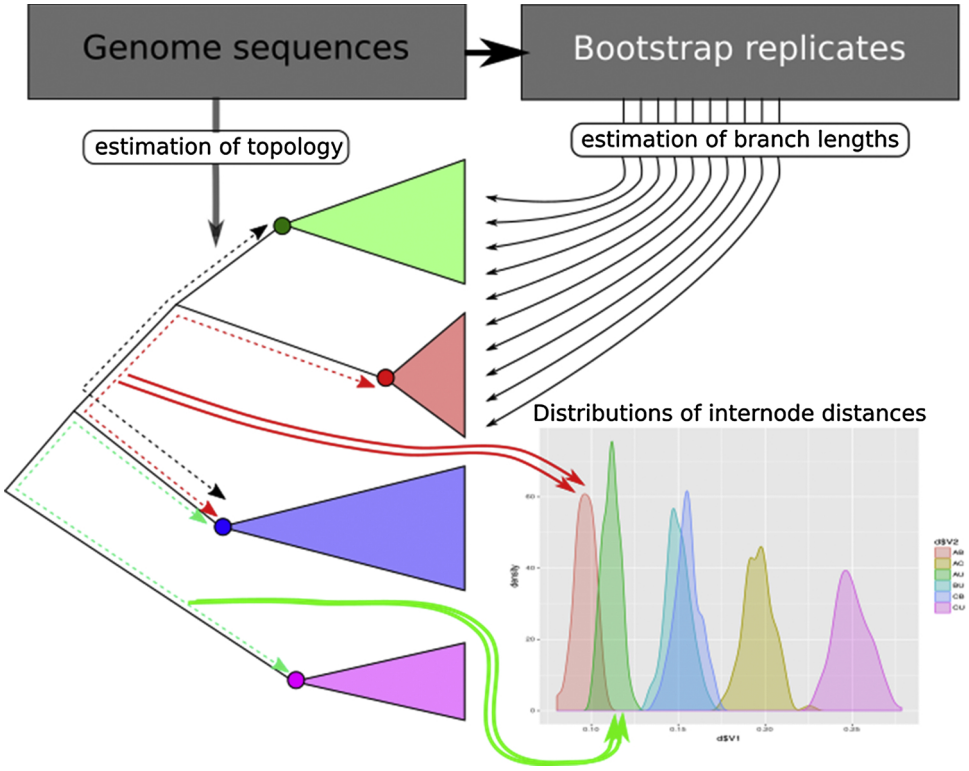

Fig. 3. Flowchart of the bootstrap statistical support for the phylogenetic separation between the common ancestors of the major TBEV groups.

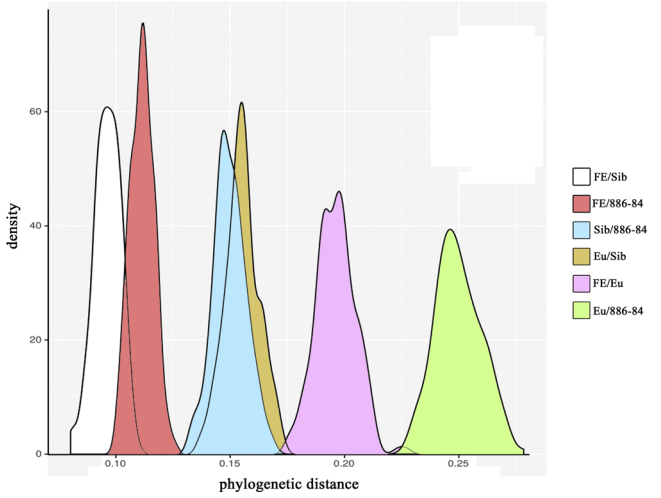

Fig. 4. The distribution of the distances values between common ancestors of the three subtypes and the group of "886-84-like" strains in the 1000 bootstrap replicates derived from the full TBEV genome sequences measured along the best topology inferred with PhyML and used afterwards as the topological constraint.

Table 3
The divergence times (shown in years) for TBEV clades.

TBEV clade	Date of the MRCA	(95% HPD)
"886-84-like" strains	212	122–333
Far-Eastern	2232	1467–3216
Siberian	2973	1932–4319
European	982	660–1412

2002 per hundred thousand of the population. Earlier we have studied molecular genetic characteristics of some TBEV strains isolated in the Alkhanay National Park territory. As a result of the study, representatives of the Siberian, Far Eastern subtypes and also "886-84-like" strains were found (Sidorova et al., 2012).

The degree of evolutionary separation of the 886-84-like clade from the three established significant subtypes of the TBEV as estimated from their complete genome sequences brings it to the same taxonomic level. Since the application of common Bayesian protocols used for estimation of a phylogeny statistical robustness in our case becomes too computationally intense due to genome size, the simplified procedure (Fig. 3) was implemented. It differs from the standard one only by applying strict topological constraint and estimating just branch lengths instead of topology + branch length. In the case of TBEV, the correctness of this protocol is justified by the strong robustness of the topology of the tree obtained from full-genome sequences.

The polyprotein analysis of the studied strains has revealed 23 variable sites, 16 of which are conserved and seven are radical compared to 886-84 strain. The identified seven mutations are located in the region of non-structural protein genes (NS1, NS2b, and NS5), one (S3014 → F) located in the NS5 region being a group mutation and there is in all six studied strains. A feature of mutations in the polyprotein is the replacement of six polar hydrophilic amino acids by non-polar hydrophobic ones. Only 3094-29/10 strain in NS5 protein has revealed a one-step replacement of positively charged arginine with uncharged hydrophilic serine (R3180 → S). All these mutations affect the nonstructural proteins NS1, NS2b, NS5 secondary and tertiary structures and, possibly, neurovirulence and replication of the TBEV.

In the phylogeography analysis of flaviviruses carried by ticks, Heinze et al. (2012) have named strains 886-84 and 178-79 – Irk-TBEV (Irkutsk-TBEV). The "886-84-like" strains were isolated not only around Lake Baikal territory in Eastern Siberia: Irkutsk Region (Trukhina et al., 2007; Paramonov et al., 2013), Republic of Buryatia (Demina et al., 2010) and Trans-Baikal Territory (Sidorova et al., 2012). The "886-84-like" strain was isolated from postmortal brain tissue sample of a patient

with meningoencephalitis in Mongolia (Khasnatinov et al., 2009, 2010). The "886-84-like" strain also was found in *Ixodes pavlovskyi* tick in Novosibirsk Region of Western Siberia territory (Rar et al., 2017).The article (Kovalev and Mukhacheva, 2017) presents a phylogenetic tree from a fragment of 1131 E gene sequences (454 n.) of the TBEV. Kovalev and Mukhacheva (2017) have suggested distinguishing the "886-84-like" strains group into a separate subtype named "Baikalian".

Based on the analysis of the seven full-genome sequences of the "886-84-like" strains, we can confirm the existence of a new “Baikalian” subtype. The genetic distances between the common root of the group of "886-84-like" strains full-genome sequences and some subtypes of the TBEV are the same as between the three known subtypes. The group of "886-84-like" strains of the TBEV is significantly monophyletic. The "886-84-like" strains group is "youngest" one comparing with three known subtypes. The divergence time of common ancestor of this group is 212 years (122–333 years with 95% HPD).

Funding sources

DS was partially supported by the topic of budget financing of Limnological Institute SB RUS 0345–2016–0004 (AAAA − A16 − 116,122,110,060 − 9).

Animal health and welfare compliance

All animal experiments complied with the ARRIVE guidelines and carried out in accordance with the U.K. Animals (Scientific Procedures) Act, 1986 and associated guidelines, EU Directive 2010/63/EU for animal experiments, or the National Institutes of Health guide for the care and use of Laboratory Animals (NIH Publications No. 8023, revised 1978) and the authors clearly indicated in the manuscript that such guidelines had followed.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.ttbdis.2019.06.006.

References

Akaike, H., 1974. A new look at the statistical model identification. IEEE Trans. Automat. Contr. 19, 716–723.
Bogovic, P., Strle, F., 2015. Tick-borne encephalitis: a review of epidemiology, clinical characteristics, and management. World J. Clin. Cas. 3 (5), 430–441.
Darriba, D., Taboada, G.L., Doallo, R., Posada, D., 2012. "jModelTest 2: more models, new heuristics and parallel computing". Nat. Meth. 9 (8), 772.
Demina, T.V., Dzhioev, Y.P., Verkhozina, M.M., Kozlova, I.V., Tkachev, S.E., Plyusnin, A., Doroshchenko, E.K., Lisak, O.V., Zlobin, V.I., 2010. Genotyping and characterization of the geographical distribution of tick-borne encephalitis virus variants with a set of molecular probes. J. Med. Virol. 82, 965–976.
Drummond, A.J., Suchard, M.A., Xie, D., Rambaut, A., 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol. Biol. Evol. 29 (8), 1969–1973. https://doi.org/10.1093/molbev/mss075.
Ecker, M., Allison, S.L., Meixner, T., Heinz, F.X., 1999. Sequence analysis and genetic classification of TBEV from Europe and Asia. J. Gen. Virol. 80, 179–185.
Gu, X., Fu, Y., Li, W., 1995. Maximum likelihood estimation of the heterogeneity of substitution rate among nucleotide sites. Mol. Biol. Evol. 12, 546–557.
Guindon, S., Gascuel, O., 2003. A simple, fast and accurate method to estimate large phylogenies by maximum-likelihood. Sys. Biol. 52, 696–704.
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. Sys. Biol. 59 (3), 307–321.
Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl. Acids Symp. Series 41, 95–98.
Heinze, D.M., Gould, E.A., Forrester, N.L., 2012. Revisiting the clinal concept of evolution and dispersal for the tick-borne encephalitis flaviviruses by using phylogenetic and bibliographic analysis. J. Virol. 86 (16), 8663–8671.
Huerta-Cepas, J., Dopazo, J., Gabaldon, T., 2010. ETE: a python environment for tree exploration. BMC Bioinformatics 11, 24.
Khasnatinov, M.A., Danchinova, G.A., Unursaikhan, U., Tyngalag, K.H., Tserennorov, D., Undraa, B., Baigalmaa, B., Abmed, D., Mironova, L.V., Otgonbaatar, D., 2009. Characterization of tick-borne encephalitis virus that caused the lethal meningoencephalitis human in Mongolia. The International Conference "Zoonotic Infectious Diseases and Tourism": Reports 30, 88–94.
Khasnatinov, M.A., Danchinova, G.A., Kulakova, N.V., Tungalag, K., Arbatskaia, E.V., Mironova, L.V., Tserennorov, D., Bolormaa, G., Otgonbaatar, D., Zlobin, V.I., 2010. Genetic characteristics of the causative agent of tick-borne encephalitis in Mongolia. Vopr. Virusol. 55 (3), 27–32 (In Russian).
Kovalev, S.Y., Mukhacheva, T.A., 2017. Reconsidering the classification of tick-borne encephalitis virus within the Siberian subtype gives new insights into its evolutionary history. Infect. Genet. Evol. 55, 159–165.
Lanave, C., Preparata, G., Saccone, C., Serio, G., 1984. A new method for calculating evolutionary substitution rates. J. Mol. Evol. 20, 86–93.
Mahy, B.W.J., 1987. Virology: a Practical Approach. IRL Press, Oxford, pp. 264.
Paramonov, A.I., Kozlova, I.V., Verkhozina, M.M., Dzhioev, Yu.P., Demina, T.V., Doroschenko, E.K., Lisak, O.V., Fedulina, O.O., 2013. Genetic and ecological characteristics an original variant borne encephalitis virus (the «group 886»), circulating in the Baikal region. Sib. Otd. Ross. Akad. Med. Nauk 2, 124–129 (In Russian).
Rar, V., Livanova, N., Tkachev, S., Kaverina, G., Tikunov, A., Sabitova, Y., Igolkina, Y., Panov, V., Livanov, S., Fomenko, N., Babkin, I., Tikunova, N., 2017. Detection and genetic characterization of a wide range of infectious agents in Ixodes pavlovskyi ticks in Western Siberia, Russia. Parasit. Vectors 10, 258. https://doi.org/10.1186/s13071-017-2186-5.
Schwarz, G., 1978. Estimating the dimension of a model. Ann. Stat. 6, 461–464.
Sidorova, E.A., Karan, L.S., Borisova, T.I., Adelshin, R.V., Andaev, E.I., Trukhina, A.G., Turanov, A.O., Nagibina, O.A., Pogodina, V.V., Lapa, S.E., Balakhonov, S.V., 2012. Genetic variety of tick-borne encephalitis virus population in Alkhanay National Park (Transbaikal region). Sib. Med. J. 4, 75–78 (in Russian).
Tavare, S., 1986. Some probabilistic and statistical problems on the analysis of DNA sequences. Lect. Math. Life Sci. 17, 57–86.
Trukhina, A.G., Andaev, E.I., Borisova, T.I., Pogodina, V.V., Karan’, L.S., Bochkova, N.G., 2007. Structure of the tick-borne encephalitis virus population in the Irkutsk region. Med. Parasitol. 24, 81–87 (In Russian).
Waddell, P., Steel, M., 1997. General time reversible distances with unequal rates across sites: mixing Gamma and inverse Gaussian distributions with invariant sites. Pers. - Int. Mycol. J. 8, 398–414.
Zlobin, V.I., Demina, T.V., Mamaev, L.V., Butina, T.V., Belikov, S.I., Gorin, O.Z., Dzhioev Iu, P., Verkhozina, M.M., Kozlova, I.V., Voronko, I.V., Adel’shin, R.V., Grachev, M.A., 2001. Analysis of genetic variability of strains of tick- borne encephalitis virus by primary structure of a fragment of the membrane protein E gene. Vopr. Virusol. 46, 12–16 (in Russian).