Ticks and Tick-borne Diseases 11 (2020) 101496

Contents lists available at ScienceDirect

Ticks and Tick-borne Diseases

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

Original article

Reporting of New tick-borne encephalitis virus strains isolated in Eastern Siberia (Russia) in 1960-2011 and explaining them in an evolutionary context using Bayesian phylogenetic inference

Artem N. Bondaryuk[a,*], Elena A. Sidorova[a], Renat V. Adelshin[a,b], Evgeny I. Andaev[a], Sergey V. Balakhonov[a]

[a] Irkutsk Antiplague Research Institute of Siberia and Far East, Trilisser 78, 664047, Irkutsk, Russia
[b] Irkutsk State University, Irkutsk, Russia

ARTICLE INFO

Keywords:
TBEV
Baikalian subtype
Phylogenetics
Population dynamics

ABSTRACT

Tick-borne encephalitis virus (TBEV) is one of the main tick-borne viral pathogens of humans. Infection may induce signs of meningitis, encephalitis, paralysis and high fever. TBEV is well studied by molecular phylogenetic methods. The present-day implementation of Bayesian phylogenetic models allows population dynamics to be tracked, providing changes in population size that were not directly observed. However, the description of the past population dynamics of TBEV is rare in the literature. In our investigation, we provide data on the dynamics of viral genetic diversity of TBEV in Zabaikalsky Krai (Eastern Siberia, Russia) revealed by the Bayesian coalescent inference in a BEAST program. As a data set, we used the envelope (E) protein partial gene sequences (1308 nt) of 38 TBEV strains (including six "886 – 84-like" or Baikalian subtype strains (TBEV-B)), isolated in Zabaikalsky Krai (Eastern Siberia, Russia) in 1960–1963 and 1995–2011. To increase estimations reliability, we compared 9 model combinations by Path sampling and Stepping-stone sampling methods. It has been shown that the genetic diversity decline in the population history of TBEV in the 1950s coincides with the date of the beginning of wide dichlorodiphenyltrichloroethane forest dusting in Siberia. We assumed that the TBEV population on the territory of Siberia went through a genetic bottleneck. Also, we provide data estimating the divergence time of TBEV-B strains and indicate the specific evolution rate of an ancestor lineage of the Baikalian subtype, illustrated on a phylogenetic tree, and reconstructed under a relaxed clock model.

1. Introduction

Molecular phylogenetics has had a profound impact on the study of infectious diseases, especially rapidly evolving infectious agents such as RNA viruses (Grenfell et al., 2004). Modern bioinformatics approaches allow the reconstruction of evolutionary processes of microorganisms, which were previously unavailable for consideration by applying classical biological methods. In particular, phylogenetic analysis using nucleotide or amino acid sequences of pathogens permits the reconstruction of their genealogies and reveals that population dynamics left a signature in their genomes (Ho and Shapiro, 2011). Today, several works revealed a correlation between the past population dynamics of viruses and their actual epidemic development (e.g. hepatitis C virus (Drummond et al., 2005), Ebola virus (Dellicour et al., 2018), dengue virus (Wei and Li, 2017), and Coxsackievirus B5 (Henquell et al., 2013) etc.).

Tick-borne encephalitis virus (TBEV) is the most notorious tick-borne RNA virus from the family *Flaviviridae*, genus *Flavivirus*. It is the aetiological agent of a severe human neurological infectious disease (Shi et al., 2018). During the study of TBEV, researchers obtained sufficient information allowing to use TBEV as an object for efficient bioinformatics analysis. At present, there are descriptions of a TBEV evolutionary rate, recombination events, spatial distribution, and interspecies genetic diversity (Adelshin et al., 2019; Bertrand et al., 2016; Heinze et al., 2012; Uzcátegui et al., 2012). However, studies of the past population dynamics of TBEV are rare in the literature (Deviatkin et al., 2020).

Zabaikalsky Krai is a unique location concerning TBEV genetic diversity, where we previously exhibited circulation of two major TBEV subtypes (Far-Eastern (TBEV-FE) and Siberian (TBEV-Sib) subtypes)

* Corresponding author.
E-mail addresses: ui.artem.ui@gmail.com (A.N. Bondaryuk), sidorovavirusolog@yandex.ru (E.A. Sidorova), adelshin@gmail.com (R.V. Adelshin), e.andaev@gmail.com (E.I. Andaev), adm@chumin.irkutsk.ru (S.V. Balakhonov).

https://doi.org/10.1016/j.ttbdis.2020.101496
Received 13 October 2019; Received in revised form 3 June 2020; Accepted 12 June 2020
Available online 15 June 2020
1877-959X/

and a relatively new “886 − 84-like” genetic variant (Sidorova et al., 2012) (also known as the “Baikalian subtype” (Adelshin et al., 2019; Kovalev and Mukhacheva, 2017)).

The present contribution aims to report new envelope (E) protein partial gene sequences (1308 nt) of 38 TBEV strains, isolated in Zabaikalsky Krai (Eastern Siberia, Russia) in 1960–1963 and 1995–2011 and interpreting them in an evolutionary context using modern bioinformatics approaches.

2. Materials and methods

2.1. Virus isolation and RNA extraction

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

2.2. 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 Reverta-L (“Central institute of epidemiology”, Russia) and PCR kit ("Syntol", Russia). 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 was sequenced twice (with forward and reverse primers).

2.3. Nucleotide sequence data sets

The sequence data set was generated by E protein partial gene sequences (1308 nt) of 38 TBEV strains, isolated in Zabaikalsky Krai territory (Eastern Siberia, Russia) in 1960–1963 and 1995–2011 by “Irkutsk Antiplague Research Institute of Siberia and the Far East” and “Chita Antiplague Station” (Fig. 1, Table 1).

To perform the extra reconstruction of the past population dynamics, we also used an extended data set (n_{seq} = 69) containing nucleotide sequences of TBEV strains from the territory of Siberia (Novosibirsk region, Irkutsk region, and Zabaikalsky Krai) and the Far East. Additional sequences were downloaded from ViPR (https://www.ViPRbrc.org) (Pickett et al., 2012).

BEAST xml-projects and the sets of nucleotide sequences (Fasta format) used in current work are available from https://github.com/trilisser/TBEV_Zabaikalye.

2.4. Sequence alignment

Nucleotide sequence alignment was performed in AliView v.1.26 (Larsson, 2014) by MAFFT v.7.394 algorithm (Katoh et al., 2017; Kuraku et al., 2013).

2.5. Phylogenetic analysis, model comparison, and past population dynamics reconstruction

Phylogenetic analysis was performed by BEAST v.1.10.4 (Suchard et al., 2018). As the substitution models, we employed three main models (GTR (Tavaré, 1986), HKY (Hasegawa et al., 1985), SRD06 (Shapiro et al., 2006)). The selection of the most appropriate combination of evolutionary models described above and other additional parameters (rate heterogeneity (+G), proportion of invariant sites (+I), nucleotide frequencies) was based on a preliminary test of models in IQTREE v.1.6.12 (Nguyen et al., 2015).

A molecular clock model selection was based on the analysis of temporal structure of heterochronous sequences performed in TempEst v.1.5.3 (Rambaut et al., 2016). Nucleotide data informativeness was assessed by comparing prior and posterior distributions in BEAST.

To infer past population dynamics, a constant size model (CS), a coalescent exponential growth (EG) model, and a non-parametric Bayesian skyline population (BSP) model were applied (Drummond et al., 2005). The best-fit model combination was determined by the Path sampling and Stepping-stone sampling (PS/SS) methods (Baele et al., 2012, 2013). Additional statistics, such as Tajima’s D and Fu and Li’s tests with an outgroup, were applied using DnaSP 6 (Rozas et al., 2017).

Markov chain Monte Carlo (MCMC) analyses for 100 million generations, with a tree sampled every 2000 steps (total amount of trees for each run = 50,000). The reproducibility of each analysis was tested by three independent BEAST runs (see Supplemental Fig. S3-S11). MCMC convergence of multiple runs and effective sample sizes (ESS) were assessed using a Tracer v.1.7.1 program (Rambaut et al., 2018). Statistical uncertainty in the time to the most recent common ancestor (tMRCA) and substitution rate calculations was estimated as the 95 % highest probability density (HPD) intervals.

3. Results

3.1. Model comparison

A comprehensive model comparison performed by PS/SS showed that the best-fit model combination was the relaxed clock model with an uncorrelated lognormal distributions (UCLD) of substitution rates with the SRD06 substitution model and the BSP model ($M_{UCLD+SRD06+BSP}$). Marginal likelihood for each model combination showed in Table 2.

3.2. Phylogenetic analysis

A phylogenetic tree based on partial E gene sequences was reconstructed (Fig. 2). The phylogeny was assessed using the model with the highest marginal likelihood ($M_{UCLD+SRD06+BSP}$). The age of the most recent sample (2011) was taken as a zero in timescale (the following tMRCA estimates were reported in the years before 2011).

Phylogenetic analyses revealed that TBEV strains separated on three main clusters: Baikalian subtype strains (n = 6; tMRCA = 36; 95 % HPD interval, 16–62) TBEV-FE strains (n = 12; tMRCA = 265; 95 % HPD interval, 110–475), TBEV-Sib strains (n = 20; tMRCA = 695; 95 % HPD interval, 320–1200).

Notably, TBEV-FE includes cluster H consisting of strains isolated from only patients with an encephalitic form of the disease (tMRCA = 49; 95 % HPD, 20–62).

TBEV-Sib was divided into two distinct groups, including “Vasilchenko” and “Zausaev” strains (clusters V and Z respectively). The tMRCA of the cluster V is 348 years (95 % HPD, 165–600). Strains isolated in the 1960s form a monophyletic group V60 (tMRCA = 61; 95 % HPD, 51–76). Isolates “Cht-653″ and “11 − 99” are most closely related to the cluster V60. The other five TBEV-Sib strains (“68B-00″, “17 − 11”, “Cht-22″, “1 − 09”, “8 − 98”) isolated in 1998–2000 form a monophyletic group with long internode patristic distances between common ancestors. The tMRCA of the group Z is 224 years (95 % HPD, 105–398). Six TBEV-Sib strains (“516 − 60,” “569 − 60”, “562 − 60”, “260 − 63”, “253 − 63”, “312 − 60”) isolated in the 1960s fall into the generic group Z60 (tMRCA = 88; 95 % HPD, 55–140). Isolate “46 − 99” is more closely related to the Z60 (tMRCA = 225).

The remaining six strains belonging to TBEV-B were isolated from the Zabaikalsky Krai territory during 1999–2010 and form the monophyletic cluster B on the tree (tMRCA = 36; 95 % HPD, 16–62). The 95 % HPD interval of tMRCA of cluster B is the narrowest regarding the two other TBEV subtypes and does not overlap with them.

Fig. 1. The location of studied TBEV strains in the Zabaikalsky Krai territory of Eastern Siberia (Russia). Isolation territories are marked in green. District notations (see Table 1): a – Duldurginsky, b – Chitinsky, c – Karymsky, d – Krasnochikoysky, e – Kyrinsky, f – Petrovsk-Zabaikalsky, g – Khiloksky, h – Baleysky, i – Shelopuginsky, j – Borzinsky.

3.3. The molecular clock calibration, comparing prior and posterior distributions, and estimating the TBEV evolutionary rate heterogeneity

The temporal structure of heterochronus sequences (i.e. an association between genetic divergence and time) can be assessed by reconstructing the simple linear relationship between the dates of samples isolation and their genetic distance (Rambaut et al., 2016). For this purpose, we used the TempEst software which rooted a phylogenetic tree by minimising the sum of the squared residuals from the regression line. The temporal signal analysis in TempEst showed the absence of such dependence ($R^2 \ll 0.05$). This could be interpreted the following way: theoretically, in the case of strict clock samples with an identical date of isolation should encamp in one point regarding the X-axis on a phylogenetic tree rooted by minimising the sum of the squared residuals from the regression line. In our case, the samples of different TBEV subtypes with the same isolation dates lie at various intervals on the X-axis (Fig. 3). This indicates that the ancestral lineages of TBEV subtypes possibly have different evolution rates. Also, the residual distribution showed a subtype-mediated clustering of tips (Fig. 4). Further analysis employing the relaxed molecular clock with UCLD showed a coefficient of rate variation of about 0.2, which indicates the adequacy of relaxed clock application (Drummond and Bouckaert, 2015). Based on these conclusions, we will only consider the relaxed molecular clock in the following analysis.

For an evolution rate parameter ('ucld.mean' in BEAST), we set weakly informative prior (lognormal distribution with mean = 1.0E-4, standard deviation = 1.0E-4; 95 % quantile width was 1.3E-5–3.6E-4) accepted from a previously published TBEV nucleotide substitution rate (Adelshin et al., 2019; Subbotina and Loktev, 2012; Suzuki, 2007; Uzcátegui et al., 2012).

Using the relaxed molecular clock under the BSP model showed that the substitution rate estimates have a large amount of uncertainty (a median evolution rate is 6.0E-5 nucleotide substitutions per site per year; 95 % HPD, 1.5E-5–1.3E-4 nucleotide substitutions per site per year) (Table 1). This may be indicative of a lack informative sites in a studied genome region since evolutionary rate estimates based on TBEV full-genome sequences (10,245 nt) have narrower 95 % HPD intervals (2.6E-5–5.3E-5 nucleotide substitutions per site per year (Uzcátegui et al., 2012); 1.0E-5–2.2E-5 nucleotide substitutions per site per year (Adelshin et al., 2019)). Comparing prior and posterior distributions of the root node height (as the oldest node of the tree) and TBEV-B tMRCA (as one of the youngest reliable node, p = 1) in BEAST showed that the temporal signal was absent near the root (95 % quantile width of root height prior, 78–3098; 95 % HPD, 460–4500; Supplemental Fig. S1) and appeared towards the present (95 % quantile width of tMRCA of TBEV-B prior, 12–171; 95 % HPD, 14–101; Supplemental Fig. S2). Thus, more ancient divergence events (such as tree root height and TBEV subtypes emergence) will not be regarded as reliable. Parameters for an additional run with an "sample from prior only" option in BEAST are shown in Supplemental Table 1.

As expected, employing the relaxed clock model ($M_{UCLD+SRD06+BSP}$) revealed that the divergence between TBEV subtypes occurred at a different rate (6.6E-5, 6.1E-5, and 5.47E-5 nucleotide substitutions per site per year for ancestor lineages of TBEV-FE, TBEV-Sib, TBEV-B subtypes, respectively) (Fig. 5); the difference between the TBEV-FE and TBEV-B rate values is most significant (21 %). On the whole, the reconstructed phylogeny substitution rates vary between 5.3E-5 and 6.62E-5 nucleotide substitutions per site per year. A median evolutionary rate is 6.0E-5 nucleotide substitutions per site per year (95 % HPD, 1.3E-5–1.3E-4 nucleotide substitutions per site per year). The highest rate corresponds to the ancestor lineage of TBEV-FE cluster (6.62E-5 nucleotide substitutions per site per year).

Table 1
Information about TBEV strains isolated on the territory of Zabaikalsky Krai (Eastern Siberia, Russia).

№ п/п	Strain	GenBank №	Year	Location	Source	Subtype
1	Cht-653	JN003207	1995	Shelopuginsky District	Human TBE case	Siberian subtype
2	11 – 99	KC414090	1999	Baleyskiy District		
3	Cht-22	JN003208	2002	Chitinsky District (Chita)		
4	8 – 98	MN520109	1998	Krasnochikoysky District	*Ixodes persulcatus*	
5	46 – 99	KF956067	1999	Duldurginsky District		
6	1 – 09	KF826914	2009	Chitinsky District (Chita)		
7	17 – 11	KF956072	2011	Duldurginsky District	*Dermacentor silvarum*	
8	506 – 60	MN520110	1960	Khiloksky/ Petrovsk-Zabaykalsky districts	*Microtus maximowiczi*	
9	516 – 60	MN520120				
10	518 – 60	MN520119				
11	569 – 60	MN520121				
12	562 – 60	MN520113			*Ochotona hyperborea*	
13	68B-00	KC422663.2	2000	Duldurginsky District	*Myodes rutilus*	
14	612 – 60	MN520117	1960	Khiloksky/ Petrovsk-Zabaykalsky districts	*Motacilla alba*	
15	206 – 63	MN520111	1963		*Parus montanus*	
16	210 – 63	MN520112			*Dendrocopos major*	
17	241 – 63	MN520116			*Bonasa bonasia*	
18	253 – 63	MN520114			*Pinicola enucleator*	
19	260 – 63	MN520122			*Loxia curvirostra*	
20	262 – 63	MN520118			*Cinclus cinclus*	
21	1 – 98	JX968560	1998	Baleyskiy District	Human TBE case	Far-Eastern Subtype
22	30 – 00	KC422667.2	2000	Chitinsky District (Chita)		
23	123 – 01	MN520105	2001	Borzinsky District		
24	12 – 02	MN520107	2002	Karymsky District		
25	19 – 07	MN520106	2007	Chitinsky District (Chita)		
26	24 – 07	MN520108		Krasnochikoysky District		
27	6 – 09	KF826915	2009	Chitinsky District		
28	78 – 99	MN520115	1999		*M. maximowiczi*	
29	64 – 00	KC422664	2000	Duldurginsky District		
30	50 – 03	KC422666	2003	Kyrinsky District	*Tamias sibiricus*	
31	67 – 99	KC422665	1999	Chitinsky District	*D. silvarum*	
32	104 – 01	KF956070	2001		*I. persulcatus*	
33	43 – 99	KF956068	1999	Duldurginsky District	*I. persulcatus*	Baikalian subtype
34	110	MH481364	2001			
35	3094 – 9	KF956069	2010		*M. rutilus*	
36	3094 – 18	KF956071				
37	3094 – 29	MH481365				
38	3033 – 1	MH481366		Chitinsky District	*M. gregalis*	

3.4. Past population dynamics

The BSP model reveals two drastic declines of TBEV GD in 1900–2011 (Fig. 6a).

In consideration of the absence of TBEV strains sampled from 1963 to 1995, we decide to extend an initial data set by addition TBEV-FE and TBEV-Sib strains from the territory of Siberia (Novosibirsk region, Irkutsk region, and Zabaikalsky Krai) and the Far East, isolated in the specified period (the total amount of nucleotide sequences was 69). The following analysis also indicated two GD declines, wherein positive dynamics was revealed until 1900 (Fig. 6b).

Applying "classic" statistics such as Tajima's D and Fu and Li's F tests with an outgroup (a TBEV European subtype prototype strain "Neudoerfl", NC_001672.1) showed positive criteria values (2.17 and 2.03 ($p < 0.05$) for Tajima's D and Fu and Li's F tests, respectively; see Supplemental Table 2) which suggest a sudden past population contraction as well.

4. Discussion

4.1. Revealed phylogenetic structure

The aim of our study was to reveal phylogenetic structure and the past population dynamics of TBEV strains isolated in Zabaikalsky Krai

Table 2
Selection of model combinations for the Bayesian phylogenetic analysis of TBEV.

Clock model	Substitution model	Population model	Marginal likelihood		Substitution rate* (95% HPD)	Root height* (95% HPD)
			PS	SS		
UCLD	SRD06	BSP	−4231.6	−4231.3	6.0E-5(2.9E-5–1.0E-4)	2052(925–3600)
		CS	−4236.6	−4236.3	6.6E-5(2.0E-5–1.3E-4)	1733(513–3819)
		EG	−4240.1	−4239.7	8.5E-5(3.2E-5–1.6E-4)	1403(460–2928)
	GTR + I+F_{emp}	BSP	−4478.0	−4477.9	5.8E-5(2.8E-5–1.0E-4)	1981(901–3450)
		CS	−4483.4	−4483.1	6.5E-5(2.2E-5–1.3E-4)	1688(512–3489)
		EG	−4487.1	−4486.9	7.3E-5(2.6E-5–1.3E-4)	1550(547–3218)
	HKY + G+F_{emp}	BSP	−4517.0	−4516.7	6.0E-5(2.8E-5–1.0E-4)	1987(857–3518)
		CS	−4520.9	−4520.7	6.8E-5(2.0E-5–1.4E-4)	1710(494–3791)
		EG	−4525.1	−4524.8	9.1E-5(3.0E-5–2.0E-4)	1264(159–2622)

* median values.

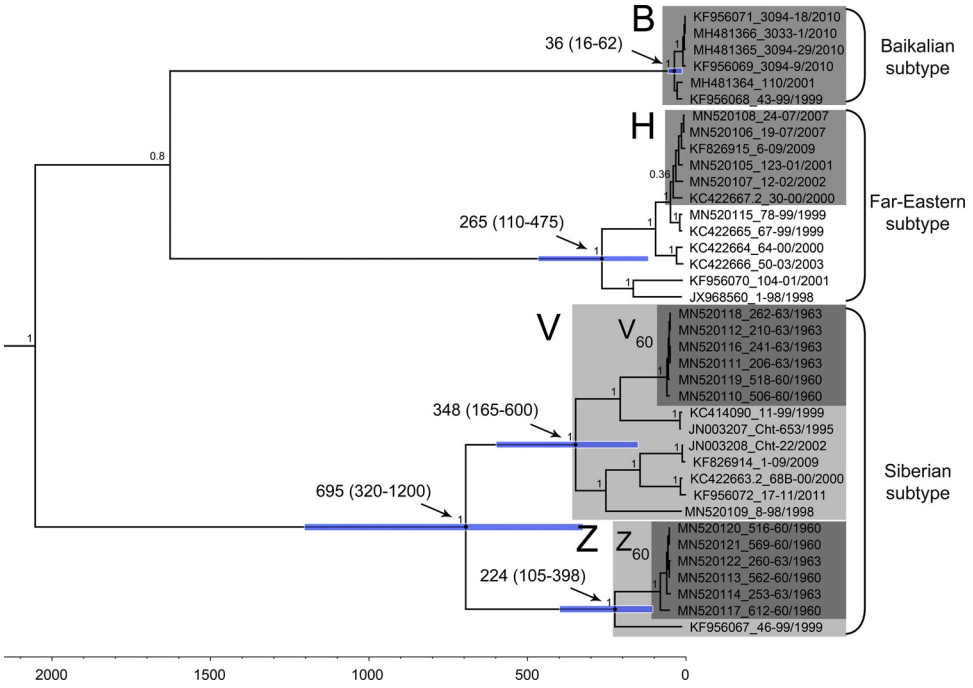

Fig. 2. A phylogenetic tree inferred with partial E gene sequences (1308 nt) of 42 TBEV strains (the names of the strains are reduced to numbers for uniformity). A phylogeny was reconstructed in a BEAST program using a model $M_{UCLD+SRD06+BSP}$ with the highest marginal likelihood value according to the PS/SS. The sampling year of the most recent sample is 2011. Timescale in the X-axis expressed in years before present (2011). Posterior probability values are indicated above the main nodes. TMRCAs and 95 % HPD intervals values are above arrows. Blue horizontal lines illustrate 95 % HPD intervals of the nodes.

(Eastern Siberia, Russia). Towards this aim, we provided the E protein partial gene sequences (1308 nt) of 38 TBEV strains, isolated from 1960–1963 and 1995–2011. Previously, it was shown that three of TBEV subtypes circulated on the territory of Zabaikalsky Krai (TBEV-FE, TBEV-Sib, and the Baikalian subtype (Sidorova et al., 2012)). Phylogenetic analysis revealed that the TBEV-Sib and TBEV-FE strains were isolated from patients with an encephalitic form of the disease, small mammals and taiga ticks; Baikalian subtype strains were isolated from mammals and *I. persulcatus*; all strains isolated from mammals and birds in the 1960s are represented by TBEV-Sib (Table 1).

To perform a phylogenetic analysis, we applied the Bayesian approach implemented in a BEAST program. Intending to increase estimation reliability, we compared 9 model combinations using the PS/SS Bayesian methods.

Revealed phylogenies allowed to consider 'hidden' evolutionary traits of TBEV. Particularly, phylogenetic analysis showed occupancy

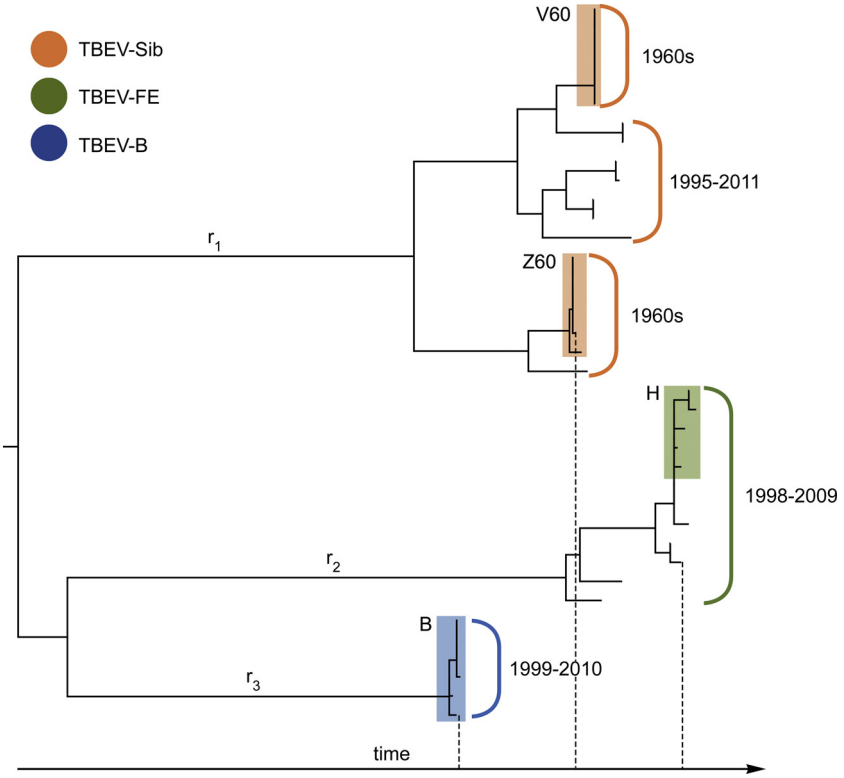

Fig. 3. A maximum likelihood tree rooted in TempEst by calculating the minimum sum of the squared residuals from the regression line. Different subtype positions regarding the X-axis are indicated by dashed lines and reveal variable evolutionary rates of ancestor lineages (r_i).

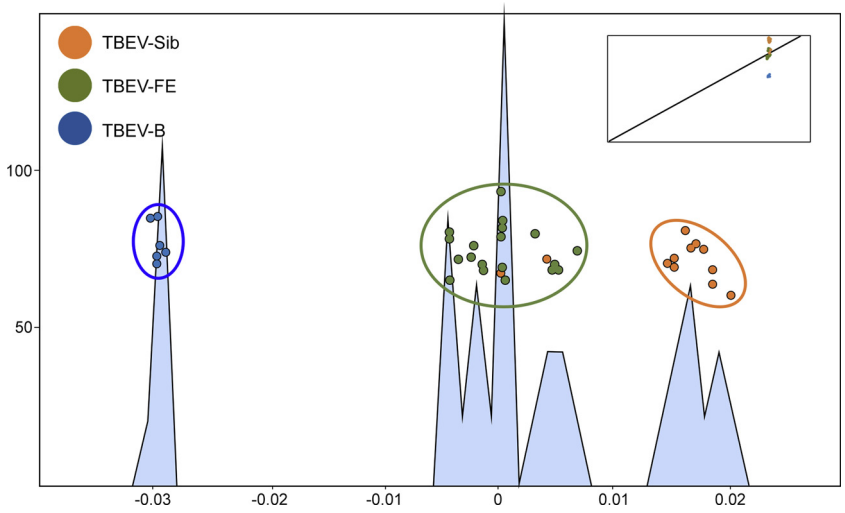

Fig. 4. Histogram and scatterplot of the residuals of the linear regression. The X-axis indicates residual values, and the Y-axis shows sequence proportions. An insert is a plot of the root-to-tip genetic distance against sampling time.

on the tree the cluster H (TBEV-FE), formed only by strains isolated from patients with an encephalitic form of the disease. In the previous study (Belikov et al., 2014), the researchers determined potential amino acid residue substitutions, correlated with the variable pathogenicity of TBEV-FE in humans. According to these data, there is one potential substitution in an E gene (A463 V) related with an encephalitic form of the disease (although there is information on stating that amino acid position 463 probably does not face the surface of the viral particle (Bukin et al., 2017)). However, in our investigation, the cluster H strains carry an Ala residue at position 463, which is typical for a subclinical form of the disease (Belikov et al., 2014). Also, there is another TBEV-FE strain "Zabaikalye 1 – 98" (JX968560.1) outside cluster H isolated from a human brain (encephalitic form) which also carries an Ala residue in the indicated position. It is clear that clinical forms cannot be specified by one amino acid residue substitution in one gene. Thus, sequencing of the complete genome of cluster H strains (except for strains "6 – 09" (KF826915.1) and "30 – 00" (KC422667.2) which are already deposited in GenBank) is of great interest.

4.2. Baikalian subtype

Of particular attention is TBEV-B strains – a relatively new TBEV subtype (Adelshin et al., 2019; Kovalev and Mukhacheva, 2017). The divergence time of TBEV-B was the most accurately estimated (tMRCA = 36; 95 % HPD interval, 16–62). Remarkably, the 95 % HPD interval of the cluster B does not overlap with intervals of other TBEV subtypes. Previously, we assessed the tMRCA of Baikalian subtype strains (including "886 – 84" isolate) using 162 full-genome sequences (10,245 nt) of TBEV (Adelshin et al., 2019). The analysis showed that the evolutionary rates of full genome and E gene were significantly different (1.6E-5 and 6.0E-5 nucleotide substitutions per site per year, respectively). Therefore, the assessed tMRCA values of Baikalian subtype strains were shifted back in time. Comparison of tMRCA values for the full TBEV genome and the E gene is presented in Table 3. Based on

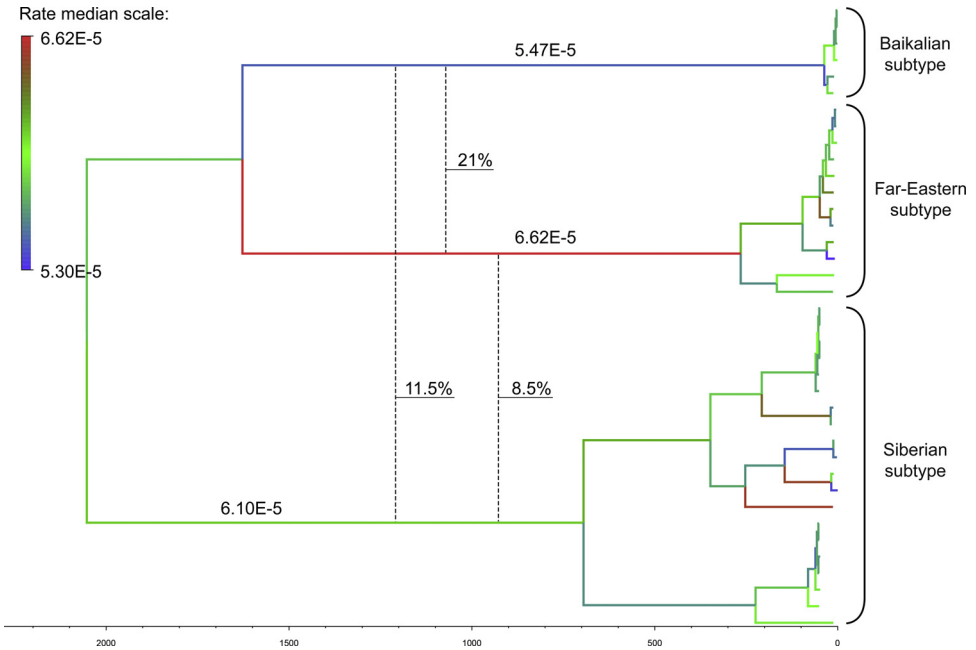

Fig. 5. Substitution rate heterogeneity of TBEV lineages under the relaxed clock ($M_{UCLD + SRD06 + BSP}$). Gradient filling indicates rate variation between branches (blue colour corresponds to a slower rate; red colour corresponds to a higher rate).

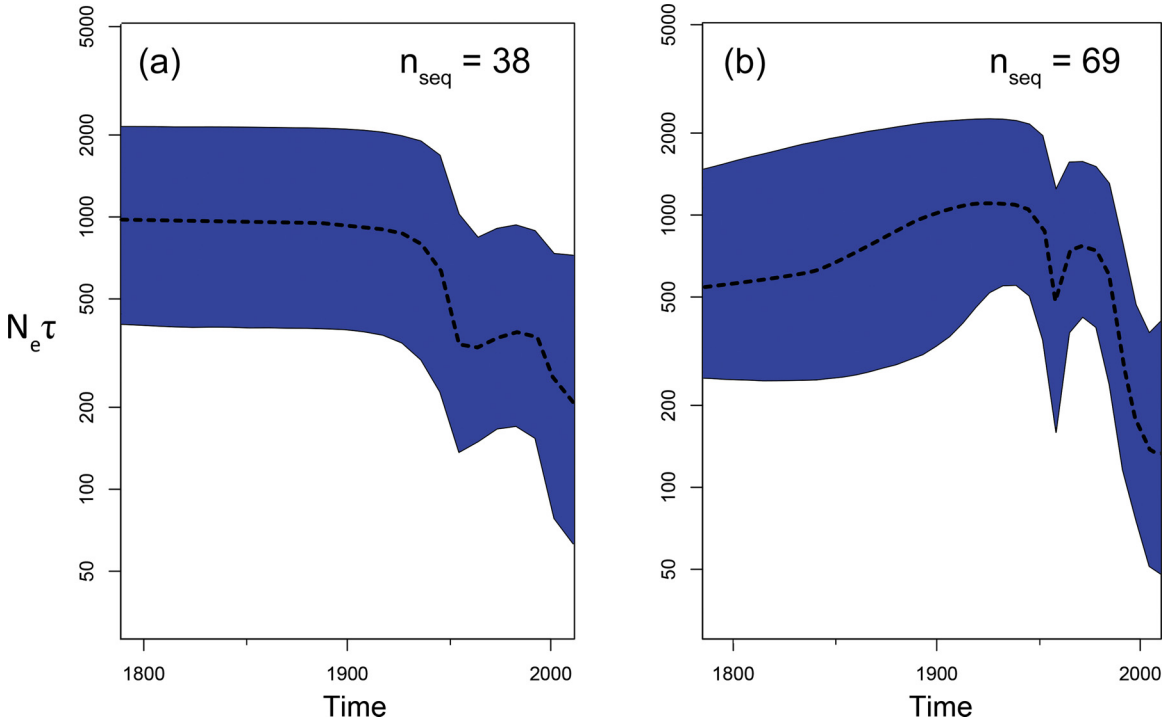

Fig. 6. Past population dynamics of TBEV reconstructed in BEAST for the partial E gene sequences (1308 nt). Y-axis indicates relative GD ($N_e\tau$, where N_e – effective population size; τ – generation time (Drummond et al., 2005)). A dashed line – a median of GD. Blue shaded fields – 95 % HPD interval of GD estimated by Bayesian coalescent inference. Note the logarithmic scale on the Y-axis. **(a)** Past population dynamics inferred with 38 nucleotide sequences of TBEV (Zabaikalsky Krai only). **(b)** Past population dynamics inferred with 69 nucleotide sequences of TBEV (Siberia and the Far East, Russia).

Table 3
Comparison of the tMRCA for the full-genome and E gene sequences.

Main clusters	Full-genome ($M_{UCLD+GTR(G+I)+BSP20}$)		E gene ($M_{UCLD+SRD06(G)+BSP8}$)	
	tMRCA[a]	95 % HPD	tMRCA[a]	95 % HPD
Cluster "B"	104	57–158	36	16–62
Far-Eastern	2232	1467–3216	265	110–475
Siberian	2973	1932–4319	695	320–1200

[a] The age of the most recent sample is 2015 and 2011 for the full-genome and E gene analyses, respectively.

the data obtained, we can assume that the Baikalian subtype circulating on the territory of Zabaikalsky Krai is younger than the Far-Eastern and Siberian TBEV subtypes.

Applying of the relaxed clock model ($M_{UCLD+SRD06+BSP}$) confirmed the specific evolutionary rate of the ancestor lineage of the Baikalian clade. The difference in nucleotide substitution rate for TBEV-FE and TBEV-B was about 21 % (6.62E-5 and 5.47E-5 nucleotide substitutions per site per year, respectively). This possibly indicates specific selection pressure impacting on each TBEV subtype during their evolution.

4.3. Population bottleneck detection

PS/SS methods showed that the BSP model fits our data set better than the simpler CS (a null hypothesis) and EG models.

To the best of our knowledge, the data manifested in this study are one of the first descriptions of past population dynamics of TBEV isolated on the territory of Siberia and the Far East. Reconstruction was performed based on two different data sets consisting of 38 and 69 nucleotide sequences. Using 38 nucleotide sequences showed that GD remained at a constant size of 1.0E3 until the first half of the 20th century, after which GD decreased by about two-fold (Fig. 6a). Subsequently, the second decline was observed. Employing the extended data set (69 sequences) revealed differences in the population dynamics of TBEV. Due to the appearance of additional coalescent events in the sampled trees, the analysis indicated positive dynamics of GD until the beginning of 1900. After the GD increased, two consecutive GD declines were detected (Fig. 6b).

The first GD decline observed at about 1950 in both cases can be explained by a high coalescent rate on the corresponding time interval of the phylogenetic tree (Clusters V60 and Z60, Fig. 2). According to the coalescent theory, the high coalescent rate corresponds to a small effective population size, i.e. these quantities are inversely proportional (Rosenberg and Nordborg, 2002). We supposed that the revealed decline of GD could be caused by extensive dichlorodiphenyltrichloroethane (DDT) and hexachlorocyclohexane (HCH) usage in the second half of the 20th century. In the USSR, studies of DDT and HCH efficiency against ticks (as the main vector of TBEV) on livestock began in 1953. The efficiency of DDT and HCH was shown to be high throughout the Siberian territory. Particularly, Gorchakovskaya stated that DDT and HCH forest floor dusting and spraying led to the extermination of 98–99.9 % of adult ticks. At the same time, DDT significantly reduced numbers of larvae and nymphs (96–98 %). In the Kemerovo region (Western Siberia), the forest was dusted by DDT (30 kg/ha) after which no ticks were indicated seven years later, while numbers of field voles and shrews remained high (Gorchakovskaya, 1962). In the Krasno-Chikoisk district of Chita region (Zabaikalsky Krai now), the efficiency of aerial spraying reached 95–97 % (Zlobin and Gorin, 1996). Our research shows that a significant decrease in the main vectors (*I. persulcatus)* could influence the GD of TBEV. In the Kemerovo region (Western Siberia) the regular DDT forest treatments on a large scale began in 1956 (Gorchakovskaya, 1962). We assumed that the TBEV population on the territory of Siberia went through a genetic bottleneck. Importantly, the estimated date of the beginning of the GD decline does not correspond to the date of the start of DDT forest treatments on the territory of Siberia in 1956. The real GD dynamics after DDT usage would possibly appear as a faint vertical line in 1956

on the Skyline plot. The discrepancy can be explained by the small number of samples.

There are no evident ecological factors (climatic or anthropogenic) that affected the second faint GD decline in the present revealed by the BSP. The BSP model usually loses some accuracy closer to the present due to the hidden population structure that leads to false signals of population decline. In the literature, this is known as “the confounding effect” (Heller et al., 2013). We speculate that the second TBEV GD decline is apparently a consequence of the confounding effect of population structure and cannot be interpreted as a decrease in the real effective population size.

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.

Funding sources

This work was supported by the Federal budget of the Irkutsk Antiplague Research Institute of Siberia and the Far East.

Acknowledgements

We sincerely thank two anonymous reviewers for their helpful comments, suggestions, and corrections. We are also grateful to Andrey V. Lishtva and Nadezhda V. Pomazkova for their help with maps of Zabaikalsky Krai.

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.2020.101496.

References

Adelshin, R.V., Sidorova, E.A., Bondaryuk, A.N., Trukhina, A., Sherbakov, D.Y., White III, R.A., Andaev, E.I., Balakhonov, S.V., Adelshin, R.V., 2019. “886-84-like” tick-borne encephalitis virus strains: intraspecific status elucidated by comparative genomics. Ticks Tick. Dis. 10, 1168–1172. https://doi.org/10.1016/j.ttbdis.2019.06.006.

Baele, G., Lemey, P., Bedford, T., Rambaut, A., Suchard, M.A., Alekseyenko, A.V., 2012. Improving the accuracy of demographic and molecular clock model comparison while accommodating phylogenetic uncertainty. Mol. Biol. Evol. 29, 2157–2167. https://doi.org/10.1093/molbev/mss084.

Baele, G., Li, W.L., Drummond, A.J., Suchard, M.A., Lemey, P., 2013. Accurate model selection of relaxed molecular clocks in bayesian phylogenetics. Mol. Biol. Evol. 30, 239–243. https://doi.org/10.1093/molbev/mss243.

Belikov, S.I., Kondratov, I.G., Potapova, U.V., Leonova, G.N., 2014. The relationship between the structure of the tick-borne encephalitis virus strains and their pathogenic properties. PLoS One 9, e62083. https://doi.org/10.1371/journal.pone.0094946.

Bertrand, Y.J., Johansson, M., Norberg, P., 2016. Revisiting recombination signal in the tick-borne encephalitis virus: a simulation approach. PLoS One 11, e0164435. https://doi.org/10.1371/journal.pone.0164435.

Dellicour, S., Baele, G., Dudas, G., Faria, N.R., Pybus, O.G., Suchard, M.A., Rambaut, A., Lemey, P., 2018. Phylodynamic assessment of intervention strategies for the West African Ebola virus outbreak. Nat. Commun. 9, 2222. https://doi.org/10.1038/s41467-018-03763-2.

Deviatkin, A.A., Kholodilov, I.S., Vakulenko, Y.A., Karganova, G.G., Lukashev, A.N., 2020. Tick-borne encephalitis virus: an emerging ancient zoonosis? Viruses. 12, 247. https://doi.org/10.3390/v12020247.

Drummond, A.J., Bouckaert, R.R., 2015. Bayesian Evolutionary Analysis with BEAST. Cambridge University Press, Cambridge.

Drummond, A.J., Rambaut, A., Shapiro, B., Pybus, O.G., 2005. Bayesian coalescent inference of past population dynamics from molecular sequences. Mol. Biol. Evol. 22, 1185–1192. https://doi.org/10.1093/molbev/msi103.

Gorchakovskaya, N.N., 1962. Direct extermination of ticks in the control of tick-borne encephalitis. Med. Parazitol. 1, 67–72 (in Russian).

Grenfell, B.T., Pybus, O.G., Gog, J.R., Wood, J.L., Daly, J.M., Mumford, J.A., Holmes, E.C., 2004. Unifying the epidemiological and evolutionary dynamics of pathogens. Science 303, 327–332. https://doi.org/10.1126/science.1090727.

Hasegawa, M., Kishino, H., Yano, T., 1985. Dating of the human-ape splitting by a molecular clock of mithohondrial DNA. J. Mol. Evol. 22, 160–174.

Heinze, D.M., Gould, E.A., Forrester, N.L., 2012. Revisiting the clinal concept of evolution and dispersal for the tick-borne flaviviruses by using phylogenetic and biogeographic analyses. J. Virol. 86, 8663–8671. https://doi.org/10.1128/JVI.01013-12.

Heller, R., Chikhi, L., Siegismund, H.R., 2013. The confounding effect of population structure on Bayesian skyline plot inferences of demographic history. PLoS One 8, e62992. https://doi.org/10.1371/journal.pone.0062992.

Henquell, C., Mirand, A., Richter, J., Schuffenecker, I., Böttiger, B., Diedrich, S., Terletskaia-Ladwig, E., Christodoulou, C., Peigue-Lafeuille, H., Bailly, J.L., 2013. Phylogenetic patterns of human coxsackievirus B5 arise from population dynamics between two genogroups and reveal evolutionary factors of molecular adaptation and transmission. J. Virol. 87, 12249–12259. https://doi.org/10.1128/JVI.02075-13.

Ho, S.Y., Shapiro, B., 2011. Skyline-plot methods for estimating demographic history from nucleotide sequences. Mol. Ecol. Resour. 11, 423–434. https://doi.org/10.1111/j.1755-0998.2011.02988.x.

Katoh, K., Rozewicki, J., Yamada, K.D., 2017. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. Brief Bioinform. 20, 1160–1166. https://doi.org/10.1093/bib/bbx108.

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. https://doi.org/10.1016/j.meegid.2017.09.014.

Kuraku, S., Zmasek, C.M., Nishimura, O., Katoh, K., 2013. aLeaves facilitates on-demand exploration of metazoan gene family trees on MAFFT sequence alignment server with enhanced interactivity. Nucleic Acids Res. 41, W22–28. https://doi.org/10.1093/nar/gkt389.

Larsson, A., 2014. Larsson A. AliView: a fast and lightweight alignment viewer and editor for large data sets. Bioinformatics 30, 3276–3278. https://doi.org/10.1093/bioinformatics/btu531.

Nguyen, L.T., Schmidt, H.A., von Haeseler, A., Minh, B.Q., 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol. Biol. Evol. 32, 268–274. https://doi.org/10.1093/molbev/msu300.

Pickett, B.E., Sadat, E.L., Zhang, Y., Noronha, J.M., Squires, R.B., Hunt, V., Liu, M., Kumar, S., Zaremba, S., Gu, Z., Zhou, L., Larson, C.N., Dietrich, J., Klem, E.B., Scheuermann, R.H., 2012. ViPR: an open bioinformatics database and analysis resource for virology research. Nucleic Acids Res. 40, D593–598. https://doi.org/10.1093/nar/gkr859.

Rambaut, A., Lam, T.T., Max Carvalho, L., Pybus, O.G., 2016. Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen). Virus Evol. 2https://doi.org/10.1093/ve/vew007. vew007.

Rambaut, A., Drummond, A.J., Dong, X., Baele, G., Suchard, M.A., 2018. Posterior summarization in bayesian phylogenetics using Tracer 1.7. Syst. Biol. 67, 901–904. https://doi.org/10.1093/sysbio/syy032.

Rosenberg, N.A., Nordborg, M., 2002. Genealogical trees, coalescent theory and the analysis of genetic polymorphisms. Nat. Rev. Genet. 3, 380–390. https://doi.org/10.1038/nrg795.

Rozas, J., Ferrer-Mata, A., Sanchez-DelBarrio, J.C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S.E., Sanchez-Gracia, A., 2017. DnaSP 6: DNA sequence polymorphism analysis of large data sets. Mol. Biol. Evol. 34, 3299–3302. https://doi.org/10.1093/molbev/msx248.

Shapiro, B., Rambaut, A., Drummond, A.J., 2006. Choosing appropriate substitution models for the phylogenetic analysis of protein-coding sequences. Mol. Biol. Evol. 23, 7–9. https://doi.org/10.1093/molbev/msj021.

Shi, J., Hu, Z., Deng, F., Shen, S., 2018. Tick-borne viruses. Virol. Sin. 33, 21–43. https://doi.org/10.1007/s12250-018-0019-0.

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).

Subbotina, E.L., Loktev, V.B., 2012. Molecular evolution of the tick-borne encephalitis and Powassan viruses. Mol. Biol. (Mosk). 46, 82–92 (In Russian).

Suchard, M.A., Lemey, P., Baele, G., Ayres, D.L., Drummond, A.J., Rambaut, A., 2018. Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. Virus Evol. 4https://doi.org/10.1093/ve/vey016. vey016.

Suzuki, Y., 2007. Multiple transmissions of tick-borne encephalitis virus between Japan and Russia. Genes Genet. Syst. 82, 187–195. https://doi.org/10.1266/ggs.82.187.

Tavaré, S., 1986. Some probabilistic and statistical problems in the analysis of DNA sequences. Some mathematical questions in biology. Lectures on Math. Life Sci. 17, 57–86.

Uzcátegui, N.Y., Sironen, T., Golovljova, I., Jääskeläinen, A.E., Välimaa, H., Lundkvist, Å., Plyusnin, A., Vaheri, A., Vapalahti, O., 2012. Rate of evolution and molecular epidemiology of tick-borne encephalitis virus in Europe, including two isolations from the same focus 44 years apart. J. Gen. Virol. 93, 786–796. https://doi.org/10.1099/vir.0.035766-0.

Wei, K., Li, Y., 2017. Global evolutionary history and spatio-temporal dynamics of dengue virus type 2. Sci. Rep. 7, 45505. https://doi.org/10.1038/srep45505.

Zlobin, V.I., Gorin, O.Z., 1996. Tick-borne Encephalitis: Etiology. Epidemiology and Prophylaxis in Siberia. Nauka, Novosibirsk (in Russian).