Molecular phylogeography of tick-borne encephalitis virus in central Europe

Manfred Weidmann,1 Stefan Frey,2 Caio C. M. Freire,3 Sandra Essbauer,2 Daniel Růžek,4,5 Boris Klempa,6,7 Dana Zubrikova,8 Maria Vögerl,9 Martin Pfeffer,10 Frank T. Hufert,1 Paolo M. Zanotto3 and Gerhard Dobler2

1Department of Virology, University Medical Center, 37075 Göttingen, Germany
2Bundeswehr Institute of Microbiology, 80937 Munich, Germany
3Department of Microbiology, Biomedical Sciences Institute – ICB II University of São Paulo, 05508-000 São Paulo, Brazil
4Department of Virology, Veterinary Research Institute, Hudcova 70, CZ-62100 Brno, Czech Republic
5Institute of Parasitology, Biology Centre of the Czech Academy of Sciences, Branisovska 31, CZ-37005 Ceske Budejovice, Czech Republic
6Institute of Virology, Slovak Academy of Science, Dubravská cesta 9, 84505 Bratislava, Slovakia
7Institute of Virology Charité University Hospital, Berlin, Germany
8Institute of Parasitology, Slovak Academy of Science, Kosice, Slovakia
9Comparative Tropical Medicine and Parasitology, Faculty of Veterinary Medicine, Ludwig-Maximilians-University, Munich, Germany
10Institute of Animal Hygiene and Veterinary Public Health, University of Leipzig, Leipzig, Germany

In order to obtain a better understanding of tick-borne encephalitis virus (TBEV) strain movements in central Europe the E gene sequences of 102 TBEV strains collected from 1953 to 2011 at 38 sites in the Czech Republic, Slovakia, Austria and Germany were determined. Bayesian analysis suggests a 350-year history of evolution and spread in central Europe of two main lineages, A and B. In contrast to the east to west spread at the Eurasian continent level, local central European spreading patterns suggest historic west to east spread followed by more recent east to west spread. The phylogenetic and network analyses indicate TBEV ingressions from the Czech Republic and Slovakia into Germany via landscape features (Danube river system), biogenic factors (birds, red deer) and anthropogenic factors. The identification of endemic foci showing local genetic diversity is of paramount importance to the field as these will be a prerequisite for in-depth analysis of focal TBEV maintenance and long-distance TBEV spread.

INTRODUCTION

In the past decade Europe has experienced an increase in tick-borne encephalitis (TBE) incidence, with more than 3400 human TBE cases annually and more than 3600 human TBE cases in Russia in 2009 (Donoso Mantke et al., 2011). Knowledge of the key factors involved in spread and spatiotemporal diversification of the causative viral agent, tick-borne encephalitis virus (TBEV), however, is limited. Field work on the focal activity of TBEV in ticks in central Europe (Labuda et al., 2002) indicates dependence on microclimate, humidity and the 8 °C isotherm (Labuda et al., 2002; Randolph et al., 2000) and coincidence of tick and host population densities (Cagnacci et al., 2012; Carpi et al., 2008; Randolph et al., 1999; Vor et al., 2010). Several efforts are under way to analyse the tick and animal reservoirs involved in the complex three-host transmission cycle as the tick molts from larva to nymph to adult with a blood meal involved in each step (Achazi et al., 2011; Cagnacci et al., 2012; Klaus et al., 2010; Knap et al., 2012; Labuda & Nuttall, 2004).

As a member of the genus Flavivirus of the family Flaviridae the genome of TBEV consists of a single-stranded 11 kb (+) RNA coding for one large polyprotein yielding three structural proteins (C, prM, E) and seven non-structural proteins (NS1, 2a, 2b, 3, 4a, 4b, 5)
It is commonly accepted that the envelope protein (E) plays a major role for the infection of mammalian and tick cells, and that it is the main determinant of virulence and a main target for neutralizing antibodies (Aberle et al., 1999). A recent study looking at the evolutionary rates of two TBEV isolates from the same focus in Finland sampled 44 years apart concluded that the E gene in particular is subject to purifying selection. In spite of high immune pressure it is nevertheless the most constrained gene in the TBEV genome, thereby being a prime target for nucleotide-based phylogenetic analysis (Uzcategui et al., 2012).

In Germany, TBE infections are mainly reported from the southern federal states of Baden-Württemberg and Bavaria (SurvStat, 2012). Since 2005, detailed data on circulating viruses and corresponding sequences have been under intensive investigation in several foci in Bavaria (Klaus et al., 2010; Kupcˇa et al., 2010; Weidmann et al., 2011). A previously published analysis of 61 central European TBEV-E gene sequences generated from collected ticks unveiled a general east to west spread of TBEV within a small geographical range (Weidmann et al., 2011; Zanotto et al., 1995). On this small scale, a phylogenetic connection between one strain from Nova Rise in the Vysocina region of the Czech Republic (T-730) and one from Germany at Haselmühl (277 km apart) separated by the mountain range of the Bavarian forest (1400 m) implicated a non-continuous distribution pattern and possibly some anthropogenic influence in the TBEV spread. The genetic analysis of the new dataset of 102 TBEV sequences allowed in-depth analysis of possible spread patterns due to landscape features, biogenic and anthropogenic influences of TBEV in central Europe.

RESULTS

Dated tree analysis

In order to obtain a better understanding of TBEV strain spread patterns in central Europe the E gene sequences of an additional 41 TBE viruses collected from 1953 to 2011 from Slovakia (three sites), Austria (two sites) and Germany (14 sites) were determined. This included 26 new TBEV isolates from ticks collected in south-east Germany from 2010 to 2011 (Fig. 1a). In sum, the extended dataset allowed analysis of 102 TBEV-E genes from 38 sites (Table S1, available in JGV Online). The comparison of two alternative schemes for estimating the rate of change (μ) expressed in substitutions per site per year (s⁻¹/y) indicated that a low rate of 4.07 × 10⁻⁵ s⁻¹/y, within the 95% highest posterior probability density (HPD) interval from 4.05 × 10⁻⁵ to 4.08 × 10⁻⁵ s⁻¹/y had a slightly better Bayes factor (BF) of 1.635 compared with a higher rate of 1.69 × 10⁻⁴ s⁻¹/y, with 95% HPD from 1.09 × 10⁻⁴ to 3.37 × 10⁻⁴, suggestive of a TBEV history time frame spanning more than 300 years in central Europe (mean=314 and 95% HPD ranging from 261 to 365 years). The rates (μ) that we obtained are in line with previous estimates showing slow evolution of tick-borne viruses compared with mosquito-borne flavivirus (Twiddy et al., 2003; Zanotto et al., 1996). Moreover, our results indicated that the TBEV lineages are split into two main clades, designated clade A and B (Figs S1, S2, Fig. 1b) and experienced purifying selection in the E gene (Table 1).

TBEV spread patterns

Phylogeographic estimates revealed that discrete transitions explain TBEV spread much better than continuous diffusion, since discrete models had significantly higher likelihood (L) than continuous models. BF values did not indicate significant differences between the L from continuous models (Brownian diffusion and Cauchy distribution) and the L from discrete models (Table 2). Moreover, Mantel’s test for correlation among matrices of patristic and geographical distances indicated that the relationship between them is significant (P=9.9999 × 10⁻⁶). The Mantel’s correlogram lacked homogeneity across all geographical distances (Fig. 2) and significant positive correlation (solid squares above the zero line) was found only for distances less than 100 km. Correlation was significantly negative for distances between 100 and 200 km and positive again close to 300 km.

TBEV diversity

The study identified six sites with multiple (n≥5) E gene sequences. Their local diversity was analysed by bootstrapped maximum-likelihood trees (Fig. 3). TBEV-E nucleotide sequences at Zdar Kaplice (n=7, 1986–1987; Weidmann et al., 2011), Borovany (n=5, 2000–2002), Burglengenfeld (n=6, 2010) and Heselbach (n=5, 2011) had low diversity with sequence identities as high as 99.6–99.9%, while those of sequences at Potěpi (n=7, 1954–1967) and Haselmühl (n=17, 2009–2011) ranged from 94.8 to 99.8%. At Potěpi and Haselmühl both lineage A and B strains were found.

Migratory routes

We analysed local migratory routes over time by estimating discrete state changes among isolation places for all ancestral nodes in time-scaled viral genealogies that were sampled during Markov Chain Monte Carlo (MCMC) convergence. Our results unveiled multiple migratory movements of TBEV from Germany to other places in the East during the past 300 years (Fig. 4, Table 3 and file S4). In the first half of the eighteenth century, multiple lineages of TBEV spread from Germany to Austria, Slovakia, North Bohemia and Central Bohemia. In the second half, the virus went from Germany to North Bohemia and from Central Bohemia to South Bohemia. From there, TBEV was reintroduced to Germany around 1930 and moved to North Moravia around 1970. Moreover, TBEV moved from North Bohemia to South Moravia around 1860.
Fig. 1. (a) Sampling sites: white circles, sites of previous study (Weidmann et al., 2011); blue circles, extended set of sampling sites; light blue circles, new isolates at known sites. (b) Distribution of lineages A and B according to the maximum clade credibility (MCC) tree (Fig. S1). Yellow circles, lineage A; green circles, lineage B; circles with red rims, sites with TBEV genetic diversity (Fig. 3); mixed coloured circles, sites with lineage A and B strains. (c) Hypothesized migration routes of TBEV in the twentieth century. Country codes: D, Germany; A, Austria; CZ, Czech Republic; SK, Slovakia. NB, North Bohemia; CB, Central Bohemia; SB, South Bohemia.
Median joining (MJ) network analysis

Since linear inheritance patterns of viral sequences are amenable to MJ network analysis, we used this approach to test if the TBEV strains might have migrated to south-east Germany by bypassing the Bavarian or Sumava mountains to the south. The resulting MJ network shows a discontinuous pattern of lineage A and B strains spread across central Europe and several introductions of TBEV strains into Germany (Fig. 5) and allowed spreading patterns to be deduced in great detail.

<table>
<thead>
<tr>
<th>Amino acid position</th>
<th>$\omega$</th>
<th>$P$ value</th>
</tr>
</thead>
<tbody>
<tr>
<td>283</td>
<td>−4.80</td>
<td>7.09 x 10^{-3}</td>
</tr>
<tr>
<td>294</td>
<td>−4.08</td>
<td>1.24 x 10^{-2}</td>
</tr>
<tr>
<td>308</td>
<td>−5.00</td>
<td>4.12 x 10^{-3}</td>
</tr>
<tr>
<td>311</td>
<td>−3.00</td>
<td>3.70 x 10^{-2}</td>
</tr>
<tr>
<td>334</td>
<td>−4.00</td>
<td>1.23 x 10^{-2}</td>
</tr>
<tr>
<td>348</td>
<td>−3.00</td>
<td>3.70 x 10^{-2}</td>
</tr>
<tr>
<td>372</td>
<td>−4.71</td>
<td>7.09 x 10^{-3}</td>
</tr>
<tr>
<td>417</td>
<td>−9.13</td>
<td>3.65 x 10^{-2}</td>
</tr>
<tr>
<td>421</td>
<td>−3.00</td>
<td>3.70 x 10^{-2}</td>
</tr>
<tr>
<td>434</td>
<td>−3.53</td>
<td>2.27 x 10^{-2}</td>
</tr>
<tr>
<td>447</td>
<td>−4.00</td>
<td>1.23 x 10^{-2}</td>
</tr>
<tr>
<td>455</td>
<td>−3.00</td>
<td>3.70 x 10^{-2}</td>
</tr>
<tr>
<td>477</td>
<td>−3.00</td>
<td>3.70 x 10^{-2}</td>
</tr>
<tr>
<td>500</td>
<td>−4.71</td>
<td>6.44 x 10^{-3}</td>
</tr>
<tr>
<td>516</td>
<td>−5.88</td>
<td>1.82 x 10^{-3}</td>
</tr>
<tr>
<td>533</td>
<td>−3.53</td>
<td>2.27 x 10^{-2}</td>
</tr>
<tr>
<td>551</td>
<td>−3.00</td>
<td>3.70 x 10^{-2}</td>
</tr>
<tr>
<td>582</td>
<td>−5.00</td>
<td>4.12 x 10^{-3}</td>
</tr>
<tr>
<td>592</td>
<td>−4.71</td>
<td>6.44 x 10^{-3}</td>
</tr>
<tr>
<td>603</td>
<td>−3.53</td>
<td>2.27 x 10^{-2}</td>
</tr>
<tr>
<td>604</td>
<td>−10.59</td>
<td>1.20 x 10^{-5}</td>
</tr>
<tr>
<td>616</td>
<td>−3.33</td>
<td>2.82 x 10^{-2}</td>
</tr>
<tr>
<td>618</td>
<td>−3.53</td>
<td>2.44 x 10^{-2}</td>
</tr>
<tr>
<td>660</td>
<td>−3.53</td>
<td>2.27 x 10^{-2}</td>
</tr>
<tr>
<td>676</td>
<td>−3.00</td>
<td>3.71 x 10^{-2}</td>
</tr>
<tr>
<td>683</td>
<td>−3.53</td>
<td>2.27 x 10^{-2}</td>
</tr>
<tr>
<td>718</td>
<td>−10.70</td>
<td>1.20 x 10^{-5}</td>
</tr>
<tr>
<td>719</td>
<td>−4.04</td>
<td>1.23 x 10^{-2}</td>
</tr>
<tr>
<td>732</td>
<td>−3.03</td>
<td>3.70 x 10^{-2}</td>
</tr>
<tr>
<td>735</td>
<td>−3.82</td>
<td>3.39 x 10^{-2}</td>
</tr>
<tr>
<td>763</td>
<td>−3.03</td>
<td>3.70 x 10^{-2}</td>
</tr>
</tbody>
</table>

*The amino acid position was related to the first codon of polyprotein sequence of isolate Toro-2003 (GenBank ABD62793.3).
†Difference between non-synonymous ($d_N$) and synonymous ($d_S$) rates per site.

Table 2. BF values for comparisons among different models used during phylogeographic inferences

<table>
<thead>
<tr>
<th>Model</th>
<th>lnP (model/data)</th>
<th>sem</th>
<th>BSSVS*</th>
<th>CTMC†</th>
<th>Brownian diffusion</th>
<th>Cauchy distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>BSSVS</td>
<td>−5408.819</td>
<td>0.402</td>
<td>0</td>
<td>−1.062</td>
<td>11.679</td>
<td>11.478</td>
</tr>
<tr>
<td>CTMC</td>
<td>−5407.758</td>
<td>0.348</td>
<td>1.062</td>
<td>0</td>
<td>12.74</td>
<td>12.539</td>
</tr>
<tr>
<td>Brownian diffusion</td>
<td>−5420.498</td>
<td>0.999</td>
<td>−11.679</td>
<td>−12.74</td>
<td>0</td>
<td>−0.201</td>
</tr>
<tr>
<td>Cauchy distribution</td>
<td>−5420.297</td>
<td>0.879</td>
<td>−11.478</td>
<td>−12.539</td>
<td>0.201</td>
<td>0</td>
</tr>
</tbody>
</table>

†CTMC, Continuous time Markov chain.
*BSSVS, Bayesian stochastic search variable selection.

DISCUSSION

Phylogenetic analysis

The analysis of the extended dataset of TBEV-E gene sequences yielded a dated tree spanning more than 300 years and consisting of mainly two subclades designated lineages A and B. About half (48) of the strains analysed are from Germany (mainly Bavaria, Fig. 1a). Lineage A strains branching in a shallow forked subclade, suggest recent introduction at Haselmühl (Fig. 3, Fig. S1). Indeed in both clades, deeply forked subclades intermingle with shallow forked subclades, and could be explained by fairly recent evolutionary events, such as bottlenecks, pruning older lineages from both main clades, a phenomenon also described for mosquito-borne viruses (Zanotto et al., 1996). Moreover, we identified several sites under purifying selection in the E gene (Table 1) and because of their potential bias towards underestimation of branch lengths on time-calibrated phylogenetics, we employed a substitution model with gamma correction (GTR+$\Gamma$) to better estimate hidden variation (Wertheim & Kosakovsky Pond, 2011). Nevertheless, patterns of molecular evolution from populations under strong purifying selection were shown to be very similar under simulated conditions to those of populations under neutral selection (Nicolaisen & Desai, 2012; Walczak et al., 2012).

An additional analysis of patristic and geographical distances suggested that TBEV foci appear to cohere below a radius of 100 km, while continuous diffusion breaks apart for greater distances. This leads to suggest two distinct modes of dispersion of TBEV: (i) on a local scale TBEV disperses continuously with positive correlation between low genetic and geographical distances, and (ii) on a regional scale (distances greater than 100 km) TBEV spreads through migration events, since genetic and geographical distances...
have a negative correlation. This observation is strengthened by the identification of six genetically diverse TBEV foci in total with at least two evolutionary rates and two foci with mixed lineages (Fig. 3).

The migration analysis indicates TBEV spread from Germany to other areas in Europe early in the eighteenth century and it was carried back to it in the twentieth century (Fig. 4). As recently confirmed, the main movement of TBEV appears to follow a clinal distribution oriented primarily from east to west (Heinze et al., 2012). Nevertheless, at a fine grain during short spans of space and time, rather complex spread patterns are expected to emerge. Considering the complex history of this central European region with trade, traffic, economic development and wars going on for centuries, these migrations may only be understandable in greater detail by using larger datasets, possibly comprising more comprehensive viral sequence information and a larger sample.

**Landscape-dependent spread patterns**

The combination of geographical and genetic information in the MJ network allowed for the analysis of detailed spreading patterns of TBEV. The introduction of one set of lineage A strains to Passau and onward to Amberg-Sulzbach directly linked to strains from Slovakia supports the idea of TBEV strains being introduced to Germany by circumventing the Bavarian mountains to the south (Fig. 1c) via the Danube river system.

Another set of lineage A strains appears to have migrated southward from the Heselbach and Burglengenfeld area in the northern part of south-east Germany. All sites are in proximity to tributaries to the Danube, or in the respective river plains or valleys [Heselbach (Naab), Burglengenfeld (Naab), Renholding (Geisfa), Fürstenstein (Ilz), Rosenheim (Inn)].

As humidity and appropriate temperature are important for the life cycle of *Ixodes ricinus* (Randolph & Storey, 1999), it seems very likely that TBEV-infected ticks could easily establish foci along river valleys independently of the means of migration. *I. ricinus* and *Dermacentor reticulatus* have been shown to occur sympatrically (Silaghi et al., 2012), and the latter is currently spreading from east to west along river valleys in Poland and from south to north along the corridor of the Dyje and Morava rivers in the Czech Republic (Široký et al., 2011; Zygner et al., 2009). The Morava flows to the south, directly passing TBEV isolation sites at Zahorska Ves and Malacky in Slovakia to meet the Danube. We therefore hypothesize that the Danube river system supports spread and establishment of new TBEV foci by providing favourable climatic conditions.

**Biogenic spread patterns**

The migration of lineage A strains towards the south of south-east Bavaria extends directly from strains described in investigations following an emergence of TBE cases in the Usti nad Labem (UNL) region (North Bohemia) in the 1990s (Figs 1b, 5; Weidmann et al., 2011).

A recent comparative study on network calculation methods pointed out that parsimony approaches are less error-prone than MJ methods, with genes showing a high substitution rate as is the case for TBEV gene (Woolley et al., 2008). We therefore also performed a parsimony analysis (Fig. S3). It confirmed the ingress patterns into Germany and particularly the ingress from North Bohemia observed in the MJ analysis.

Public health data in the Czech Republic indicate a general spread of TBE northward and westward (Daniel et al., 2011; Kriz et al., 2012). After eruption of TBE at the end of the 1990s in the UNL region, incidence increased first in the Kralov Vary region (2001–2010) to the west and then even further west in south-east Bavaria (2001–2012; SurvStat, 2012). The regions Plzen and South Bohemia with the highest morbidity rates in the Czech Republic (10.9 and 23.2 per 10000) cover the whole remaining expanse of the German–Czech border. Rodent and roe deer (*Capreolus capreolus*) ranges are too small to explain the spread of TBEV-infected ticks over long distances. Red deer (*Cervus elaphus*), however, have much larger resident ranges (<300 ha for females, <700 ha for males; Licoppe & Lievens, 2011) and stags that cannot acquire a herd of females migrate for long distances (<100 km) looking for a herd elsewhere. A population analysis of red deer populations in Sumava Park, on the eastern hillsides of

---

**Fig. 2.** Mantel’s correlogram of patristic and geographical distances. Geographical distances (horizontal axis) plotted against Spearman’s rank correlation coefficient (vertical axis). Zero line, no relationship between geographical and patristic distances. Points above the zero line, positive correlation; points below the zero line, negative correlation. Points with statistical significance (*P*<0.05) are marked with solid (black) squares.
the Bavarian forest in Plzen and South Bohemia, provides some additional evidence for tick dispersal via red deer migration. During the cold war the red deer populations on either side of the Bavarian–Bohemian ecosystem were discreetly separated by border fortifications, and this still shows up in matrilineal mtDNA differentiation into predominantly west-European and Czech female red deer populations. Microsatellite-based population analyses, however, reveal that the populations have intermingled since the fortifications were removed two decades ago (Fickel et al., 2012). The concomitant free migration of stags documented in these genetic signatures may have been a route of an accompanying ingression of TBEV-infested ticks into Germany.

In the MJ network, Central Bohemia itself acts as a sort of hub for introductions of TBEV lineage A strains to south-east Bavaria (Mühlendorf, Asbach) and in south-west Germany (Salem, Karlsruhe; Fig. 5). Prime candidate mechanisms for this type of spread could involve migrations of passerine bird species, which have already been implicated in TBEV spread (Hoogstraal et al., 1963; Waldenström et al., 2007).

**Anthropogenic spread patterns**

Lineage B strains were introduced to Hauzenberg and Haselmühl. Introduction into Haselmühl had stood out at the end of our last phylogeographic analysis (Weidmann et al., 2011) and the MJ network suggests a route from Brno via Nova Rise. This introduction was successful as evident from the many isolates from this site (Fig. 1). The most divergent strain at this site (HM666), however (Fig. 4), is directly connected to two strains from North Bohemia (T-754/II, T-740) indicating the additional arrival of a lineage A strain at this site. Lineage A strains appear also to have been introduced from South Bohemia (e.g. Zdar Kaplice to Müldorf and on to Asbach). Anthropogenic spread is a prime candidate for all of these introductions, but at this stage cannot be demonstrated as clear cut as described.

---

**Fig. 3. TBEV-E gene diversity.** Upper row: Starplot of TBEV-E gene sequences of isolates from Borovany, Burglengenfeld, Haselmühl (one bar for all three sites 9.9×10⁻⁴ substitutions per site). Lower row: Haselmühl and Potepli (individual bars 1.0×10⁻² substitutions per site). The bootstrap percentage values are derived from 1000 bootstrap trials. A and B indicate the two main lineages.
along the Siberian railway thoroughfare in Russia (Kovalev et al., 2009). Prime candidates could be rodents on long-haul lorry transport.

**Focal activity**

The spread of tick-borne flaviviruses across the Eurasian continent has recently been reanalysed and it appears that the TBEV-Sib and TBEV-Fe types branched off the TBEV-Eur type to migrate towards the east about 3000 years ago in central Asia, whereas TBEV-Eu migrated towards the west (Heinze et al., 2012). The novelty of our data lies in the close examination of natural active TBEV foci, identified by searching for ticks in the area of the residency of confirmed TBE cases.

Focal activity of TBEV is subject to many factors and currently subject to investigation and discussion (Durmiši et al., 2011; Fajs et al., 2012; Lommano et al., 2012). A reset of tick phenology due to conditions during diapause and warming conditions in spring allowing synchronous appearance and co-feeding of at least two of the three tick

---

**Table 3. Estimated dates of migratory events of TBEV in central Europe**

<table>
<thead>
<tr>
<th>Origin</th>
<th>Destination</th>
<th>Median (years)</th>
<th>95 % HPD (years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Germany</td>
<td>Central Bohemia</td>
<td>1731</td>
<td>1682–1779</td>
</tr>
<tr>
<td>Germany</td>
<td>Central Bohemia</td>
<td>1744</td>
<td>1669–1872</td>
</tr>
<tr>
<td>Germany</td>
<td>North Bohemia</td>
<td>1744</td>
<td>1669–1872</td>
</tr>
<tr>
<td>Germany</td>
<td>Austria</td>
<td>1748</td>
<td>1695–1803</td>
</tr>
<tr>
<td>Germany</td>
<td>Slovakia</td>
<td>1754</td>
<td>1701–1807</td>
</tr>
<tr>
<td>Germany</td>
<td>North Bohemia</td>
<td>1767</td>
<td>1743–1867</td>
</tr>
<tr>
<td>Germany</td>
<td>South Bohemia</td>
<td>1791</td>
<td>1734–1839</td>
</tr>
<tr>
<td>Germany</td>
<td>Austria</td>
<td>1809</td>
<td>1708–1867</td>
</tr>
<tr>
<td>Germany</td>
<td>Central Bohemia</td>
<td>1815</td>
<td>1755–1873</td>
</tr>
<tr>
<td>North Bohemia</td>
<td>South Moravia</td>
<td>1865</td>
<td>1755–1873</td>
</tr>
<tr>
<td>South Bohemia</td>
<td>Germany</td>
<td>1922</td>
<td>1894–1943</td>
</tr>
</tbody>
</table>
stages (larva, nymph, adult) is essential for saliva-assisted non-viraemic TBEV transmission (Labuda & Randolph, 1999; Randolph et al., 1999). In contrast TBEV emergence at higher altitudes due to climate warming appears to depend mainly on vertical transmission among ticks (Daniel et al., 2004; Danielová et al., 2006).

Long-term studies on TBEV in Russia reveal that reproductive adult male voles with a suppressed immune system due to high testosterone levels are the best amplifiers of TBEV and that the virulence of TBEV increases in tick–host–tick passages (Moshkin et al., 2009). Infection rates in host rodents in turn depend on abundance of tick larvae and nymphs feeding on the adult reproductive voles whose springtime activity coincides with tick activity, thus establishing horizontal transmission of TBEV among the rodents (Moshkin et al., 2009). More recently, new data on the role of rodents in maintenance of TBEV transmission in Slovenia (Knap et al., 2012) and Scandinavia (Tonteri et al., 2011) support the view that rodents also play a significant role in natural transmission cycles in Europe. Additionally, empirical data supported by mathematical modelling implicate roe deer population densities in sustaining tick populations and therefore TBEV transmission in Slovakia and Italy (Bolzoni et al., 2012; Cagnacci et al., 2012).

Our phylogenetic and network analysis of 102 TBEV-E gene sequences allowed us to infer the diverse recent evolutionary history of TBEV-E gene sequences in central

**Fig. 5.** Reduced MJ network. Plot of a 1228 character set of 102 TBEV-E gene sequences. Yellow circles, lineage A according to MCC tree in Fig. S1; green circles, lineage B. German TBEV-E sequences in the network are depicted in light blue. Central-, South- and North Bohemian clades as defined in Weidmann et al. (2011) are highlighted with areas shaded in light red. Foci with diversity (Fig. 3) are highlighted in bright red.
Europe. It identified local (focal) continuous and long-distance discontinuous spreading patterns of two lineages of TBEV in central Europe and TBEV foci with differing evolutionary rates.

Ecological assessment in respect of climate data, tick and host density data for these genetically diverse TBEV foci will be paramount to obtaining a detailed understanding of the maintenance and spread of TBEV.

The question of why lineages A and B differ in their spread pattern should also be analysed by virological means to characterize replication efficiencies in various cell tissues, dendritic cells and macrophages of relevant host species and in ticks. Since most virological data published to date were presumably generated with lineage B strains (e.g., TBEV strain Hypr71), a comparative analysis may yield important insights into these two lineages and ultimately explain if and how their spreading patterns differ.

**METHODS**

**Virus strains.** Viral sequences were obtained from TBEV strains from ticks (*Ixodes ricinus*, *I. hexagonus*), or mammalian hosts (*strain 465: Sciurus vulgaris*; strains V-352, V-361; *Apodemus sylvaticus*; M2 A104: *Apodemus flavicolliis*; strains V-364, V-540, M5 CIG223: *Myodes glareolus* in the Czech Republic (19 sites), Slovakia (two sites), Austria (three sites) and Germany (14 sites) over a period of more than 67 years. In the years 2010 and 2011 questing ticks were collected by flagging the low vegetation at seven sites in south Germany (Hasselmühl, Burglenfeld, Amberg-Sulzbach, Asbach, Neustadt a. d. W., Passau and Hesselbach), in the vicinity of recent TBE cases reported by the local health authorities. Ticks were separated according to their sex and developmental stage, morphologically identified to the species level (Hillyard, 1996) and subsequently homogenized and processed for reverse transcription (RT)-PCR as described previously (Frey et al., 2012). The year, host source and geographical location of the particular virus samples are summarized in Table S1.

**RNA extraction PCR amplification and sequencing.** All sequences generated since 2005 were obtained directly from RNA isolated from tick homogenates, without any isolation or attenuation in cell cultures or mouse brain.

Total nucleic acids were extracted from 200 μL sample using the MagNA Pure LC Total Nucleic Acid (NA) Isolation kit (Roche) and the MagNA Pure LC instrument (Roche). Isolated NA (50 μL) was stored at −80 °C. Screening real-time RT-PCR, subsequent envelope gene (E) gene RT-PCR and sequencing of the screened positive tick samples were performed as described previously (Frey et al., 2012; Weidmann et al., 2011).

**Phylogenetic analysis**

**Sequence alignment:**

Nucleotide sequence alignments were performed using CLUSTAL W and bootstrapping (1000 repetitions) in DNASTAR LASERGENE MEGALIGN. Dendrograms were built using DENDROSCOPE (Huson et al., 2007).

**Bayesian phylodynamics analysis**

For the analysis of the evolutionary history of the TBEV, a 1488 character-long alignment (including gaps of 106 E genes, connected to isolation year and sampling site (latitude and longitude), was used for a maximum-likelihood (ML) optimization in GARLI BUILD 0.951 (Zwickl, 2006) using the General Time Reversible (GTR) model of nucleotide substitution (Krejčí, 1949) and optimizing topology, branch lengths, rate heterogeneity and model parameters. The dated samples, allowed to infer a maximum clade credibility (MCC) tree with dated tips and internal nodes using a MCC Bayesian approach and a GTR model with gamma-distributed rate variation (Γ) and a proportion of invariant sites (I) using a relaxed (uncorrelated lognormal) molecular clock in BEAST v1.5.4 (Drummond & Rambaut, 2007). The GTR+Γ+I substitution model was used, since it was the best model obtained with MODELLTEST (Posada, 2008). A relaxed molecular clock was used since flavivirus evolution generally approximates a molecular clock, but some minor rate differences may occur (Twiddy et al., 2003). Four independent MCMC runs of four chains each were run for 100 million states. The convergence of parameters during MCMC runs was assessed by their effective sample size (ESS) reaching values above 200, as calculated with TRACER v1.5 (http://tree.bio.ed.ac.uk/software/tracer/). Since Bayesian inference is highly sensitive to prior selection and because new sequences were being analysed, we tested two different substitution rates to infer TBEV changes in time. Firstly we used an initial substitution rate (μ) of 8.0 × 10−4 s⁻¹’y⁻¹ in agreement with our previous estimate (Weidmann et al., 2011) and we also used a prior based on a coarse rate estimate obtained with the Path-O-Gen v1.3 program (http://tree.bio.ed.ac.uk/software/pathogen/), which uses the best ML tree and sampling dates. To assess the best-fit rate to our data, we did a BF test (Suchard et al., 2001) which compares the harmonic mean of the marginal likelihoods for each model. To investigate selection regimens on the E gene, we estimated the difference (ω = dN − dS) between the non-synonymous (dN) and synonymous (dS) rates per codon sites, using the single likelihood ancestor counting (SLAC) algorithm with HyPhy v2.11 (Pond et al., 2005), assuming a significance level of 5 % (α = 0.05). Thus, ω greater than zero suggests directional selection, while values below zero are indicative of purifying selection.

**Phylogeography of TBEV**

We tested two major hypotheses of TBEV spread in our area of study: dispersion following a continuous diffusion model (Lemey et al., 2010), and spread by discrete events of migration (Lemey et al., 2009). We also tested if the diffusion rates between localities were homogeneous (Brownian diffusion), or if the diffusion rates were independent following a Cauchy distribution. We fed the continuous models with geographical coordinates from the places of viral isolation. The second scenario was inferred using a continuous time Markov chain (CTMC) model (Lemey et al., 2009) and under the special case of a Bayesian stochastic search variable selection (BSSVS) model. We assigned to each sequence a discrete label in agreement with its place of isolation (Central Bohemia, South Bohemia, North Bohemia, Germany, Slovakia and Austria). The time-scaled phylogenies to reconstruct spread patterns were obtained during the stationarity of four MCMC using BEAST. The MCMC was run for 2 × 10⁸ generations for continuous models and 5 × 10⁷ for discrete models to obtain model convergence. Our sample from the posterior included 10 000 trees that were used to construct the MCC tree. We used the BF test (Suchard et al., 2001) to compare the marginal likelihood obtained from the MCMC for each model, after 1000 bootstrap replicates, to test the best-fit model to the data. The MCC tree from the best model with ancestral states in nodes was plotted in Google Earth software (http://www.google.com/earth/index.html), using SPREAD v1.05 (Bielejec et al., 2011). In addition, to investigate the relationships between geographical and genetic distances we calculated, from the phylogenetic tree with the best likelihood obtained after 100 independent runs of the GARLI program, the patristic distances using PATRISTIC software (Fourment & Gibbs, 2005), assuming a significance level of 5 % (α = 0.05). Thus, ω greater than zero suggests directional selection, while values below zero are indicative of purifying selection.
2006); the geographical law distances between two points were estimated using the spherical law of cosines with a Perl script. The correlation between the matrices of geographical and patristic distances was inspected using the Mantel test implemented in the APE package v3.0.2 (Paradis et al., 2004) from the R-project (http://www.r-project.org/). We investigated relationships between geographical and patristic distances with a Mantel correlogram (Carrel et al., 2010), which was generated with the Bonferroni multiple correction and Spearman’s rank correlation coefficient (p), using VEGAN package v2.05 (Oksanen et al., 2012) from the R-project.

**MJ network analysis**

A network was constructed from a 102 TBEV-E gene 1222 character alignment set stripped of all homogeneous character using SPLITS TREE 4.0 with Epsilon 1 and 2000 spring embedded iterations. A parsimony splits analysis was also used.

**ACKNOWLEDGEMENTS**

Supporting grants: Czech Science Foundation, no. P502/11/2116; Academy of Sciences of the Czech Republic, grant Z60220518; the AdmireVet project, no. CZ.1.05/2.1.00/01.006 (ED006/01/01) of the Ministry of Education, Youth and Sports of the Czech Republic; German Federal Ministry of Education and Research, grant 01KI0710. In addition, we acknowledge CAPES and CNPq-PQ for scholarships received by C.C.M.F. and P.M.A.Z. We are grateful for computational resources provided by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), projects 2010/19341-4 and 2011/17120-3. We thank F. X. Heinz for providing the TBEV Scharl case site.

**REFERENCES**


Klaus, C., Hoffmann, B., Hering, U., Mielke, B., Sachse, K., Beer, M. & Süß, J. (2010). Tick-borne encephalitis (TBE) virus prevalence and virus genome characterization in field-collected ticks (*Ixodes ricinus*) from risk, non-risk and former risk areas of TBE, and in ticks removed from humans in Germany. *Clin Microbiol Infect* 16, 238–244.


