

Molecular Epidemiology, Evolution and Dispersal of the Genus *Flavivirus*

Paolo M. de A. Zanotto¹ and Ernie A. Gould²

¹LEMB, Departamento de Microbiologia, ICBII - USP, Av. Prof. Lineu Prestes, 1374, CEP
05508-900, São Paulo, Brazil.

²Centre for Ecology and Hydrology, Mansfield Road, Oxford, OX1 3SR, United Kingdom.

1. INTRODUCTION

1.1 *Flavivirus* classification

The genus *Flavivirus* is named after the yellow fever virus (from the Latin, *flavus* = yellow), which is also the prototype of the family Flaviviridae (Strode, 1951; Theiler and Downs, 1973; Markoff and Falgout, 1995). The genus *Flavivirus* contains approximately 70 antigenically related viruses which infect both vertebrate and invertebrate species (Karabatsus, 1985; Calisher *et al.*, 1989). Almost half of the identified flaviviruses do cause disease in humans. Most of the representatives of this genus make use of arthropod hosts, hence their generic name: Arboviruses (*i.e.*, arthropod borne viruses). Moreover, albeit having a similar overall genomic organization and making use of similar replication strategies and virion morphology, they are divergent from other members of the *Flaviviridae* family, such as hepatitis C, hepatitis G and pestiviruses. The level of sequence similarity among the polymerase genes of the members of the *Flaviviridae* does not depart from what would be expected by chance, even at the protein level (Zanotto *et al.*, 1996b).

Detailed information on the place of isolation of flaviviruses and on other important biological data (*e.g.*, antigenic relationships among viruses and with their vertebrate and invertebrate hosts) was systematically collected (Theiler and Downs, 1973; Karabatsos *et al.*, 1985). From the wealth of data amassed so far, it can be said that members of the genus *Flavivirus* are found all around the World. However, some members of the genus have distinctive biogeographic characteristics and distributions. Other than by vicariance, the

distinct geographic distribution of the flaviviruses has to be determined, to a great extent, by the vertebrate hosts and invertebrate vectors that they infect. Hosts and vectors are ultimately distributed according to a multitude of constraints at all ecological levels, from niche to habitat. There are probably far more flaviviruses than are currently recognized, nevertheless, the dispersal strategies employed by the major groups of viruses within the genus *Flavivirus* are remarkably characteristic for each group and reflect very closely the ecological associations of the viruses. Despite the clear evidence that many of the flaviviruses are transmitted between vertebrate hosts by arthropods, it is known that they can also be transmitted orally (Gresikova *et al.*, 1975) and transplacentally (Mathur *et al.*, 1982). The methods by which some of the flaviviruses are transmitted, such as the viruses with unknown vectors, have never been adequately defined, but in addition to those proposed above, other possible routes include urine, feces, aerosols and blood. This wide range of possibilities may help to explain the success with which flaviviruses were able to establish in so many different host species and presumably refine their specific transmission mechanisms to suit the particular environment in which they had become established.

The recognized members of the genus have been classified into twelve distinct groups by the International Committee for the Taxonomy of Viruses (ICTV) (Heinz *et al.*, 2000). This classification was based on their overall biological properties, including the association with vertebrate and invertebrate hosts and disease. In table 1, the flaviviruses are shown according to their vector associations (Gould *et al.*, 2001). Alternatively, the pathogenic flaviviruses can be subdivided according to the clinical syndromes that they cause (Markoff and Falgout, 1995). This last classification does not necessarily consider the evolutionary relationships among viruses, but may be useful from a health care perspective. Notwithstanding, any classificatory proposition based on ancestral relationships (*i.e.*, phylogeny-based) is preferable since it may stand as a sound basis for subsequent comparative studies, which may unveil important biological, ecological and epidemiological correlations. Viruses do not have a sufficient number of distinctive morphologic features that would allow detailed or comprehensive taxonomic propositions. On the other hand, they may be grouped by molecular phylogenies based on their genes and genomic organization. However, most RNA viruses do have high rates of error and undergo drastic fluctuations in their population size, which may cause significant loss of phylogenetic signal in time (Zanotto *et al.*, 1996b). Therefore, viral molecular phylogenies may work only within certain levels of sequence similarity, such as within genus and family. While any current taxonomic proposition for the *Flavivirus* will be possibly revised as more sequence data become available, there has been some agreement between molecular phylogenies (based on viral genes and complete genomes) and serological classificatory schemes (Calisher *et al.*, 1989; Porterfield, 1980).

Tab. 1. Subdivision of the flaviviruses based primarily on sequence data

Virus Vector	Group	Virus/Subtype	Abreviation
1. Tick-borne viruses	Mammalian tick-borne virus group	Louping ill virus	(LI)
		Irish subtype	(ISE)
		British subtype	(BSE)
		Spanish subtype	(SSE)
		Turkish subtype	(TSE)
		Tick-borne encephalitis virus	(TBE)
		European subtype	(WTBE)
		Siberian subtype	(STBE)
		Far Eastern subtype	(FETBE)
		Omsk haemorrhagic fever	(OHF)
		Langat virus	(LGT)
		Kyasanur Forest disease virus	(KFD)
		Alkhurma virus	(ALKV)
		Karshi virus	(KSI)
		Royal Farm virus	(RF)
	Powassan virus	(POW)	
	Gadgets Gully virus	(GGY)	
	Seabird tick-borne virus group	Tyuleny virus	(TYU)
		Meaban virus	(MEA)
		Saumaraez Reef virus	(SRE)
Kadam virus†		(KAD)	
2a. Mosquito-borne viruses (<i>Aedes</i> spp. associated)	Yellow fever virus group	Banzi virus	(BAN)
		Uganda S virus	(UGS)
		Jugra virus	(JUG)
		Potiskum virus	(POT)
		Saboya virus*	(SAB)
		Bouboui virus	(BOU)
		Edge Hill virus	(EH)
		Yellow fever virus	(YF)
		Wesselsbron virus	(WSL)
		Sepik virus	(SEP)
	Kedougou virus group	Kedougou virus†	(KED)
	Dengue virus group	Dengue virus 1	(DEN-1)
		Dengue virus 2	(DEN-2)
		Dengue virus 3	(DEN-3)
		Dengue virus 4	(DEN-4)
Spondweni virus group	Spondweni virus	(SPO)	
	Zika virus	(ZIK)	

* no direct evidence of transmission in mosquitoes

† tentative assignment

Tab.1. (cont.) Subdivision of the flaviviruses based primarily on sequence data

Virus Vector	Group	Virus/Subtype	Abreviation
2b. Mosquito-borne viruses (Culex spp. associated)	Aroa virus group	Aroa virus	(AROA)
		Iguape virus	(IGU)
		Naranjal virus	(NJL)
		Bussuquara virus	(BUS)
	Ntaya virus group	Bagaza virus	(BAG)
		Israel turkey meningoencephalitis virus	(IT)
		Ntaya virus	(NTA)
		Tembusu virus	(TMU)
		Ilheus virus	(ILH)
		Rocio virus	(ROC)
		St Louis encephalitis virus	(SLE)
		Japanese encephalitis virus group	Alfuy virus
	Murray Valley encephalitis virus		(MVE)
	Japanese encephalitis virus		(JE)
	Usutu virus		(USU)
	Koutango virus		(KOU)
	Kunjin virus		(KUN)
	West Nile virus		(WN)
	Yaounde virus		(YAO)
	Cacipacore virus		(CPC)
Kokobera virus group	Kokobera virus	(KOK)	
	Stratford virus	(STR)	
3. No-known vector viruses	Rio Bravo group	Batu Cave virus	(BC)
		Phnom Penh bat virus	(PPB)
		Carey Island virus	(CI)
		Bukalasa bat virus	(BB)
		Dakar bat virus	(DB)
		Rio Bravo virus	(RB)
		Montana myotis leucoencephalitis virus	(MML)
		Modoc virus group	Cowbone Ridge virus
	Modoc virus		(MOD)
	Sal Vieja virus		(SV)
	Jutiapa virus		(JUT)
	San Perlita virus		(SP)
	Apoi virus group†	Apoi virus	(APOI)
	Yokose virus group	Entebbe bat virus	(EB)
		Yokose virus	(YOK)
		Sokoluk virus	(SOK)
	Tentative Species in the genus	Tamana bat virus	(TAB)
		Cell fusing agent	(CFA)

* no direct evidence of transmission in mosquitoes

† tentative assignment

For quite some time it has been understood that the evolution of arboviruses could be under a constraint imposed by their need to alternate replication in both vertebrate and invertebrate hosts. The biology of the genus *Flavivirus* constitutes a nice testing ground for this notion since the genus has members that make obligatory use of arthropod vectors and others, which do not appear to be vectored. Herein, by using molecular phylogenies, we will address the significance of several life history parameters, such as the biology of reservoirs, vectors, mode and tempo of evolution and dispersal of the flavivirus.

1.2 Structure and replication

Members of the *Flaviviridae* have enveloped virions, which are spherical with approximately 60 nm in diameter (Rice, 1996). They contain a single-stranded positive sense RNA genome that is about 11 kb in length and enclosed by a single capsid (C) protein (Rice, 1996; Lindebach and Rice, 2001). After viral entry Ph-mediated fusion takes place in the cytoplasm and replication ensues. The positive-sense RNA serves as the messenger for negative strands which are then used as templates for additional genome-length positive-sense RNA molecules by a semi-conservative mechanism of RNA synthesis. Direct translation of the uncoated genomic RNA yields a single polyprotein, which is immediately processed by cellular and viral proteases to produce the structural (S) and non-structural (NS) viral proteins (Rice, 1996; Lindebach and Rice, 2001). In the members of the genus flavivirus the envelope of mature virions contains two virus-encoded proteins, the membrane (M) protein which is post-translationally cleaved from the pre-membrane protein (immature virions) and the main glycoprotein E, which binds the virions to cell receptors binding. The E glycoprotein is also the main target for antibody neutralization by the vertebrate host (Monath *et al.*, 1996). During the course of infection, seven non-structural proteins; NS1, NS2A, NS2B, NS3, NS4A, NS4B and NS5 are present in the cellular environment. Although the function and precise role of several of the NS gene products have not yet been defined, it is known that the NS1 protein may play a role in virion assembly and/or release (Lee *et al.*, 1989). The NS1 may also play some role in virus replication (Gritsun *et al.*, 1988; 1989; 1990), signal transduction and membrane anchoring (Jacobs *et al.*, 2000). The NS3 protein is the virus-encoded serine protease. The protease becomes active when its N-terminal one-third becomes complexed with the NS2B. The NS3 has the key function of processing of the viral polyprotein. The C-terminal portion has an RNA helicase domain and RNA triphosphatase activity that is probably involved in the formation of the 5'-terminal cap structure of the viral RNA (Rice, 1996). The RNA-dependent-RNA polymerase function is provided by the NS5 protein (Tan *et al.*, 1996). After replication and accumulation of structural proteins,

virus assembly probably takes place in the rough endoplasmic reticulum. Immature virions are then transported through the membrane systems of the host cell to the surface. The last maturation step involves the cleavage of the PrM protein by a furin-like protease. After that, infectious viral particles are released by exocytosis (Rice, 1996; Lindebach and Rice, 2001). Many flaviviruses can replicate in vertebrate and invertebrate cells and in whole animals, which makes them amenable for growth under laboratory conditions.

2. MOLECULAR PHYLOGENIES OF THE *FLAVIVIRUS*

2.1 Phylogenies of the *Flavivirus*: Mosquito- and tick-borne virus trees

Initial phylogenetic comparisons of tick- and mosquito-borne flaviviruses did show some level of congruence between flavivirus relationships based on serological and molecular data (Mandl *et al.*, 1989). Subsequently, trees based on C, M and E genes showed that the tick- and mosquito-borne viruses were phylogenetically distinct lineages (Shiu *et al.*, 1991; Venugopal *et al.*, 1994). It was also noted that the degree of relative similarity among adjacent lineages of TBE complex viruses is higher than that with the mosquito-borne viruses. This suggested that the tick- and mosquito-borne flaviviruses had possibly been subjected to different evolutionary processes during their radiation. Moreover, trees based on the E-NS1 junction of a large number of DEN-2 viruses led to the conclusion that sylvatic West African dengue viruses could be distinguished from epidemic strains and also that some dengue epidemics in the New World had arisen as the result of the introduction of DEN-2 viruses from Asia (Rico Hesse, 1990). Subsequently, phylogenetic trees constructed from each individual gene of 11 different flaviviruses revealed similar topologies (Blok *et al.*, 1992). These data were interpreted as implying that the flaviviruses had diverged along their radiation with minimal or no genetic recombination between the main lineages. The authors also estimated both rates of transitions and transversions of a portion of the genomes of 40 DEN-2 virus isolates (Blok *et al.*, 1992). The results showed that sylvatic DEN-2 strains from West Africa had less sequence diversity compared with the epidemic strains of DEN virus (Blok *et al.*, 1992). Although the authors suggested that this could imply greater constraints for sylvatic viruses, the result could also be interpreted under the light of population genetics. Namely, that the sylvatic viruses had less sequence diversity because they do not undergo the same amount of population growth (and hence increase in sequence diversity) as the cosmopolitan strains that infect millions of individuals. These observations may also be relevant for yellow fever virus, which has distinct sylvatic and urban cycles. Early studies on tick-borne virus E gene sequences helped to elucidate the movement of Louping ill (LI) viruses

from Scotland to Norway, since the viruses appeared to be dispersed along a genetic continuum, (*i.e.*, a progressive genetic variation across eastern Europe) (Gao *et al.*, 1993b).

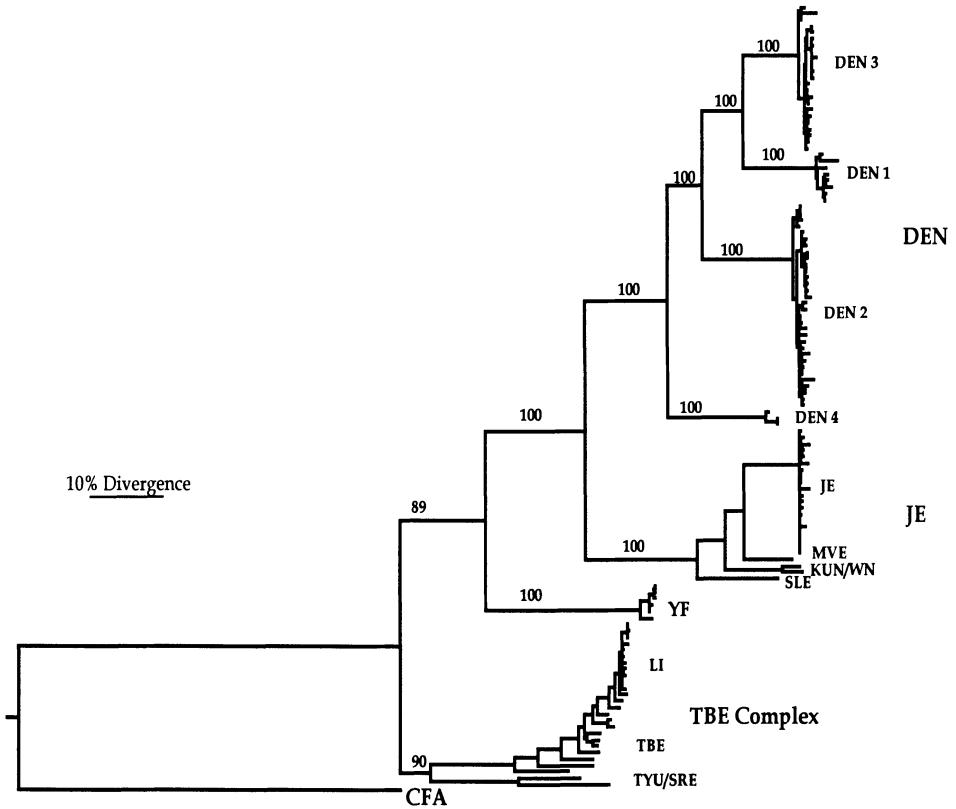


Fig 1. Maximum likelihood phylogenetic tree illustrating the evolutionary relationships of viruses in the genus *Flavivirus* (from Zanotto *et al.*, 1996a). The tree was constructed using the 1st and 2nd codon positions for 123 complete E genes. All horizontal branch lengths are drawn to scale (*i.e.* proportional to the number of accumulated nucleotide substitutions). The root for the genus separating tick- and mosquito-borne groups was determined based on the use of CFA virus as an outgroup.

Phylogenetic analyses of dengue (Rico Hesse, 1990; 1997; Wang *et al.*, 2000), West Nile (Berthet *et al.*, 1997) and yellow fever virus (Lepiniec *et al.*, 1994), showed the extent of genetic variation within the envelope gene of individual virus species and enabled conclusions regarding the origin, persistence and geographic dispersal of these viruses. For example, phylogenetic analysis of DEN-2 virus (Rico Hesse, 1990) showed that lineages clustered according to their geographic origin and that strains of DEN-2 virus responsible for dengue hemorrhagic fever in the Americas originated in Asia. A detailed phylogenetic analysis of the genus *Flavivirus* showed that of the mosquito-borne viruses analysed at that time, YF virus diverged first followed

sequentially by DEN-4, DEN-2, DEN-3 and DEN-1 virus in the *Aedes* spp. group (Fig. 1). Amongst the *Culex* spp. viruses, SLE had the most divergent lineage, followed more recently by WN/KUN and MVE/JE virus (Marin *et al.*, 1995b).

Comparisons of the phylogenetic patterns and on the genetic variation of DEN-2 viruses from Southeast Asia and South America suggested that this serotype dispersed throughout the tropics recently, possibly reflecting commercial activities and the movement of large numbers of military personnel across the major oceans, since the Second World War (Leitmeyer *et al.*, 1999). Phylogenetic trees also grouped YF lineages into East African, Central/West African, West African and South American subtypes (Chang *et al.*, 1995; Liepiniec *et al.*, 1994; Wang *et al.*, 1997). The degree of sequence similarity between South American and African YF viruses also confirmed that YF virus was introduced from Africa to the New World. Moreover, within these YF groups, the amount of genetic variation is lesser than that of the dengue viruses. It is interesting that these observations, based on molecular data have good historical and biological explanations. Since YF virus in recent times has not caused massive urban epidemics, but has been maintained mainly in sylvatic cycles in nature, lesser genetic diversity should be expected when compared to the pandemic DEN viruses.

2.2 Three main groups in the *Flavivirus*

When molecular data on a comprehensive and representative set of members of the genus was obtained, a better picture of the impact of vectors on the evolution of the *Flavivirus* was revealed. Based on partial NS5 sequences, a dendrogram of the *Flavivirus* was done based on a distance matrix from sequences of members of all three main groups of the genus (*i.e.*, the tick-borne, the mosquito-borne and, the no-known vector viruses) (Kuno *et al.* 1998). That important dataset was revisited and a maximum likelihood tree was reconstructed with an increased number of viruses, excluding the third codon position and the hypervariable loop (Fig. 2) (Jenkins *et al.*, 2001). This tree places the KAD virus in the seabird-tick associated virus group, with POW virus as a member of the earliest lineage in the TBE complex viruses group. Gadgets Gully (GGY), Royal Farm (RF) and Karshi (KSI) virus were also shown to be lineages within the TBE complex (Kuno *et al.* 1998). The same conclusions on the overall phylogenetic relationships among members of the genus were derived from phylogenetic trees constructed using sequences representing either the entire flavivirus genome or individual genes (Billoir *et al.*, 2000; Kuno *et al.* 1998). Therefore, the *Flavivirus* are divided into three groups: (i) the tick-borne group, (ii) the mosquito-borne group and, (iii) the no-known vector group, shown in table 1, has a phylogenetic basis (Fig. 2) and

will be used to refer to the viruses in the genus throughout the remaining of this chapter.

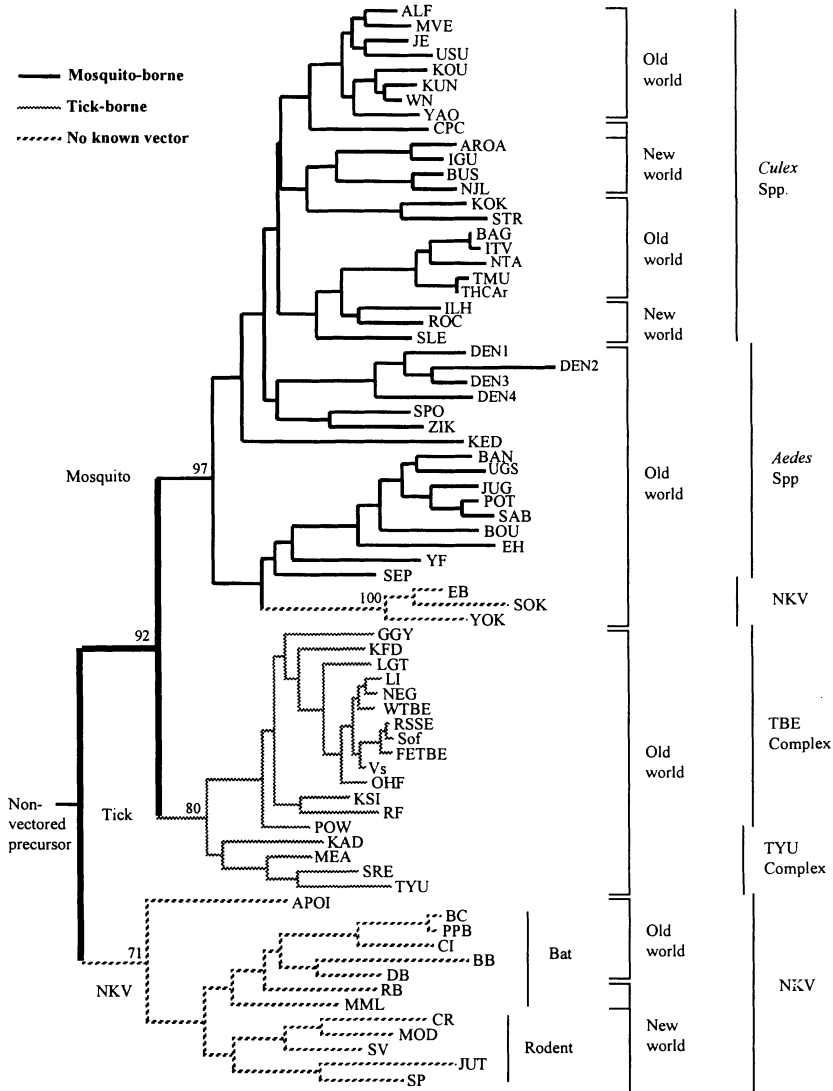


Fig. 2. Maximum likelihood tree of most flaviviruses (from Jenkins *et al.*, 2001 and Gould *et al.*, 2001). This tree is based on partial NS5 sequence data from Kuno *et al.* (1998).

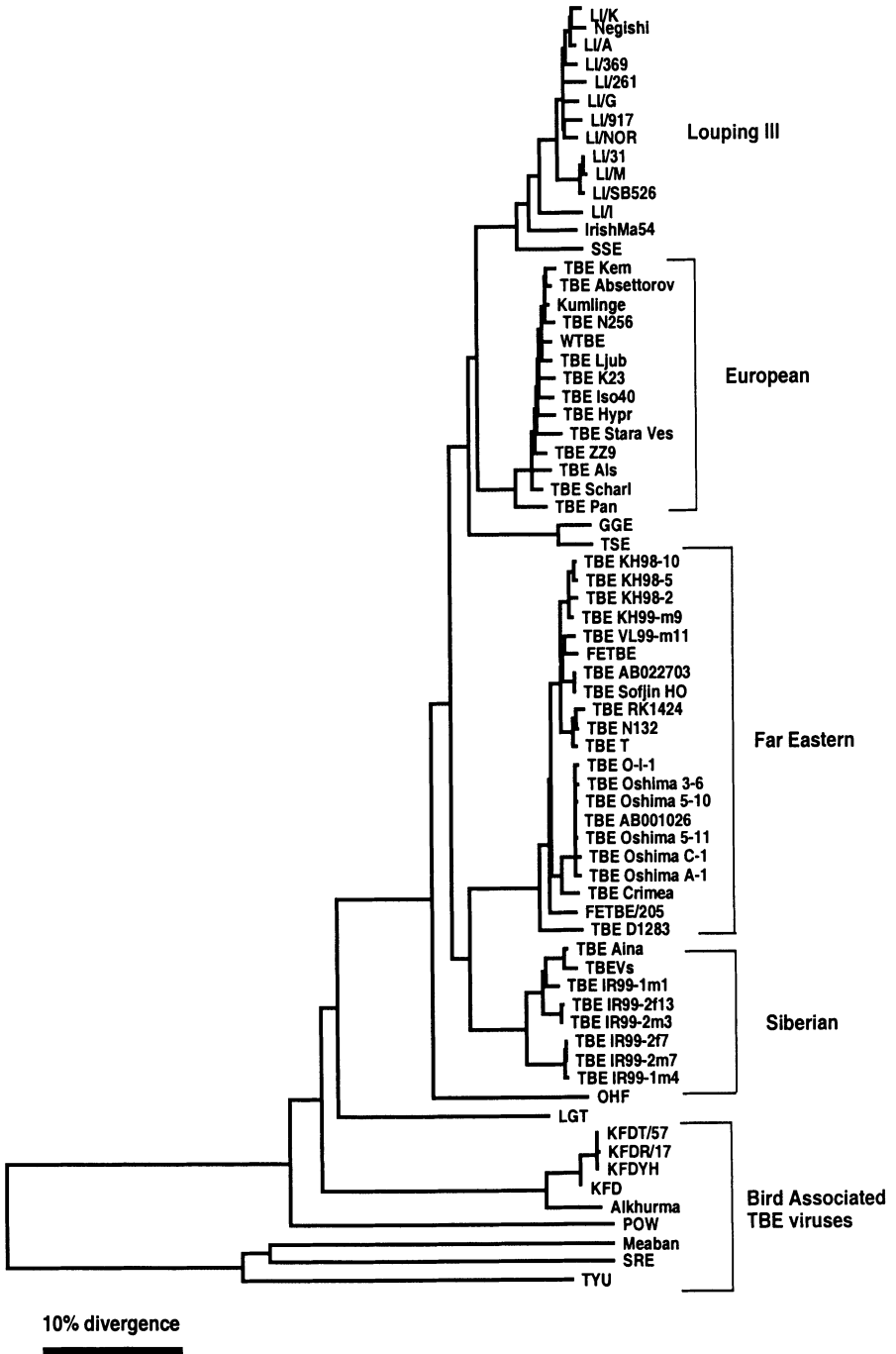


Fig. 3. Maximum likelihood phylogenetic tree for the TBE virus complex including 70 complete E gene sequences. The subtyping of the TBE virus complex was done according with table 1.

3. TICK-BORNE VIRUSES

The tick-borne encephalitis (TBE) viruses cause endemic zoonotic infections of the central nervous system, which have been reported since the 19th century (Gaidamovich, 1995). They can be divided into two main groups: (i) the tick-borne encephalitis (TBE) complex viruses and (ii) the TBE group associated with seabird-birds and their ticks.

The tick-borne encephalitis (TBE) complex viruses are human pathogens and are usually found in infected ticks surviving in vegetation that provides a micro-climate with relative humidity close to saturation throughout the year. The forest undergrowth in many parts of Asia, Europe and North America, as well as the upland sheep-grazing pastures of the UK and parts of southern Europe provide such conditions. Most of the TBE complex viruses infect rodents, other forest animals, humans and are more closely associated with *Ixodes* spp. With the exception of Powassan (POW) virus and a closely related strain designated Three Arch virus, all recognised TBE complex viruses are found only in the Old World. As a general rule, the TBE complex viruses do not share overlapping habitats with other flaviviruses, *i.e.*, they exhibit niche-like characteristics, in the forests of Malaysia, India, Asia, and Europe. This heterogeneity and dispersal pattern of the TBE complex viruses can be explained by a lack of significant overlap of the two main tick species in Eurasia: the *Ixodes ricinus* to the East and the *Ixodes persulcatus* to the West (Gaidamovich, 1995). There are reports that TBE complex viruses were isolated from the brains of ducks that have migratory flight paths between Russia and Slovakia (Ernek, 1960). However the sequences of these bird-associated TBE complex viruses have not been determined so far. However, it will be interesting to see whether the bird-associated viruses are closely related to the Slovakian or Russian strains of the TBE. A sub-group of the TBE complex viruses causes encephalomyelitis in sheep. These viruses are found on the sheep-rearing hillsides in southern Europe (Marin *et al.*, 1995a), Turkey (Gao *et al.*, 1993a), Greece (Marin *et al.*, 1995a) and on the moorlands of the British Isles (McGuire *et al.*, 1998; Reid, 1984) and Norway (Gao *et al.*, 1993b). The British virus, louping ill, is the most comprehensively described of these sheep-associated viruses (McGuire *et al.*, 1998; Gao *et al.*, 1998).

3.1 Phylogenetic patterns: The TBE cline

Viruses in the tick-borne encephalitis (TBE) complex have always been hard to distinguish by any conventional serological methods, since they are highly similar. Using serology almost no relevant ideas about their evolutionary relationships was obtained before sequence data became available. Using the cell fusing agent (CFA) virus (Cammisa Parks *et al.*, 1992) as outgroup, phylogenetic trees on the complete envelope (E) gene sequences

of 22 flaviviruses reinforced the notion of an early split between the mosquito- and tick-borne viruses (Marin *et al.*, 1995b). For the maximum likelihood tree in figure 3, obtained with PAUP (Swofford, 2001), the transition probability matrix was calculated iteratively from the data during a heuristic search with the nearest neighbour perturbation (NNI). This tree for 70 E genes of the tick-borne viruses has basically the same topology as that of Marin *et al.* (1995b) and has kept some of its main features. However, it includes an additional 48 sequences (Ecker *et al.*, 1999, Hayasaka *et al.*, 1999, 2001) plus the partial E gene sequence of the Meaban virus (Gaunt *et al.*, 2001). Its adjacency patterns show that the tick-borne viruses diverged into two sister lineages: the tick-borne encephalitis (TBE) complex viruses and the TYU, Meaban and SRE group. Interestingly, both the adjacent TYU and the Meaban lineages are made of virus found in seabirds and are vectored by the *ornithodoros* spp. ticks. By excluding the 3rd codon position of the gene for the analysis (which was saturated and without reliable phylogenetic signal), the overall branch lengths of the mosquito- and tick-borne groups indicates other interesting aspects of the tick-borne viruses. Firstly, there appears to be a gradual radiation of the TBE complex viruses in time since its split from a lineage that leads to the POW virus, currently found in Maritime Russia and in the North America. The wider geographic distribution of the POW virus compared with other TBE complex viruses may be explained by its zoonotic association with different species of ticks, mosquitoes, small mammals and birds (Luby, 1995). The Louping ill (LI) virus, appears to be a recent lineage in the TBE virus complex, found in the Eastern most part of Europe.

Surprisingly, the gradually increasing genetic distance between the TBE viruses correlates directly with increasing geographic distance of the viruses from each other in the northern hemisphere, measured from a point in Scotland where LI viruses are most frequently found (Zanotto *et al.*, 1995). The TBE complex viruses have an asymmetrical phylogenetic pattern (a grade) that is indicative of a clinal distribution (Hennig, 1966). It was postulated that the TBE-complex viruses form a cline across the northern hemisphere (Fig. 4) by Zanotto *et al.*, (1995). However population level studies are necessary to unveil the fine structure of this cline. Hence, calculations based on estimated substitution rates and dates of virus isolation (MacGuire *et al.*, 1998) suggested that LI virus in the British mainland was probably introduced from Ireland about 400 to 500 years ago and persisted on the northern hillsides of Britain during the past 200 to 300 years. This is further supported by veterinary reports of sheep encephalomyelitis on the Scottish hillsides and on the Devonshire moors following their introduction during the 19th and 20th centuries respectively. The phylogenetic tree in figure 4 was based on sequence data of the E gene, (MacGuire *et al.*, 1998). It includes all recognized sheep encephalomyelitis viruses, *i.e.*, those from Turkey, Greece, Spain, Ireland, Wales, England, Scotland and Norway and illustrates their divergence from the

other TBE complex viruses that are associated with rodents in forests. Specific viral populations in the cline show considerable sequence homogeneity. For example, the LI virus in the United Kingdom vary in only 3% of their amino acids in the E gene, even when comparing strains that were isolated more than 50 years apart (MacGuire *et al.*, 1998). The Spanish equivalent of LI virus, SSE virus, is a distinct lineage 5% divergent from the LI, European TBE and TSE virus and it is found only in the Basque region of Spain. Given the high conservation among the LI viruses it was found that a unique tripeptide sequence in the E gene works as a signature distinguishing each of the antigenically very closely related LI-like viruses found in Turkey, (Gao *et al.*, 1993a), Greece and Spain (Marin *et al.*, 1995a) the United Kingdom and Norway (Gao *et al.*, 1993b), and Ireland (MacGuire *et al.*, 1998). Subsequent population level studies provided evidence for a TBE virus cline. They were also based on additional E genes of isolates from Western and Central Europe (Ecker *et al.*, 1999) and Far East Asia (Hayasaka *et al.*, 1999; 2001) (Figs. 3 and 5).

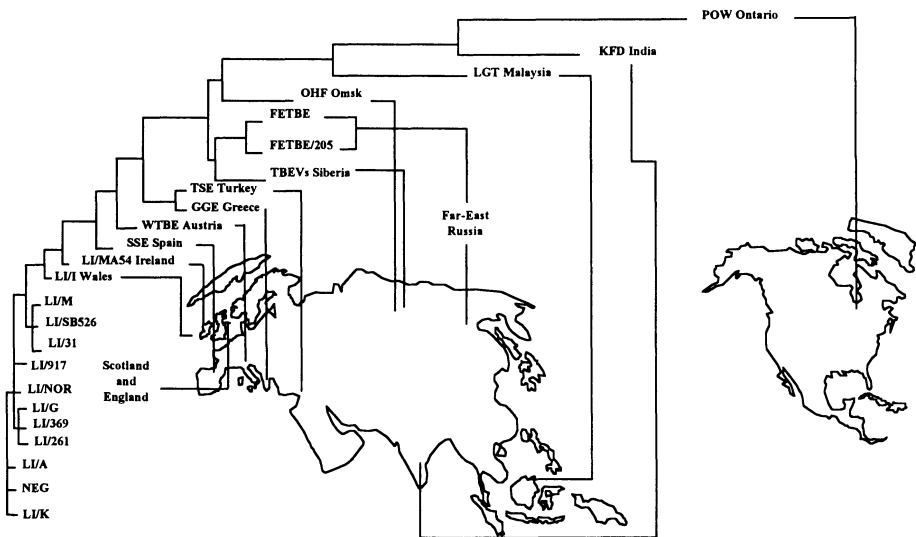


Fig. 4. TBE cline from Zanotto *et al.*, 1995. Tree of the TBE is shown over a map of Eurasia with the tips connected to the place of sampling of each lineage. The actual root of the tree is near the branch leading to the POW virus.

In figure 3, it can be seen that the isolates of TBE tend to group according with the general geographic area where they were isolated from. If the TBE did not form a cline one would expect different adjacency patterns, with a lot of mixing of lineages among viral populations from Eastern European LI virus with those of the TBE of Central Europe (Ecker *et al.*, 1999) and those from Siberia and from the Far-East (Hayasaka *et al.*, 1999; 2001).

Furthermore, the postulated continuous evolution of the TBE complex viruses, north-eastward and then westward across Asia and Europe (Zanotto *et al.*, 1995), was also supported by the sequence data for Karshi (KSI), Royal Farm (RF) and Gadgets Gully (GGY) virus (Kuno *et al.*, 1998).

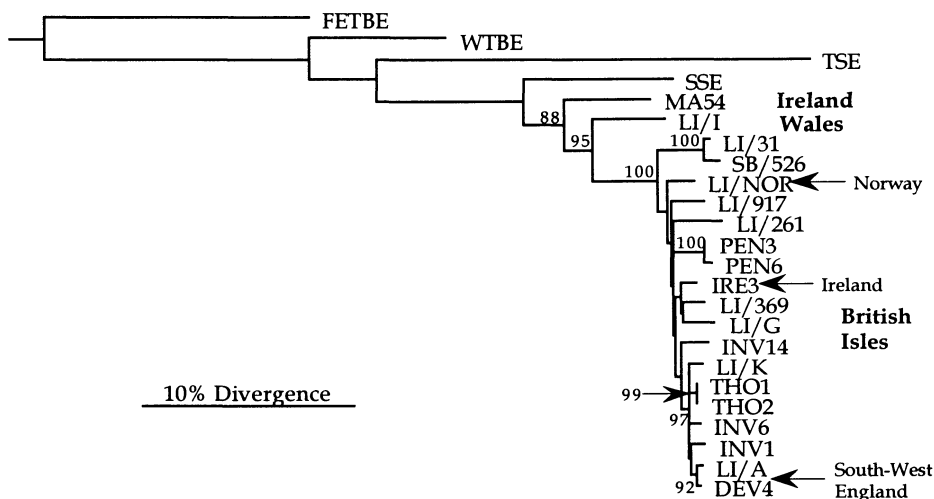


Fig 5. Maximum likelihood phylogenetic tree for the E gene from 24 tick-borne flaviviruses (McGuire *et al.*, 1998). Branch lengths are drawn to scale and all nodes supported by more than 75% bootstrap support are indicated. The tree is rooted with the sequence from FETBE virus, Sofjin strain. The three main populations of virus in the British Isles (Ireland, Wales and Great Britain) are indicated, along with those viruses secondarily introduced into Ireland and Norway, and the viruses found in the south-west of England.

3.2 Wide distribution of seabird-tick associated viruses

A second and smaller group of tick-borne viruses, the TYU and the TBE complex viruses, are dependent for their life cycles on different tick vectors, *i.e.*, *ornithodoros* or *ixodes* spp. that infest the nesting grounds of seabirds and appropriate vertebrate hosts, *i.e.*, seabirds and rodents respectively. These seabird-associated viruses (Tab. 1; Fig. 2) found only in the Old World and the lineages leading to these viruses diverged relatively early in the radiation of the genus. The Tyuleniy (TYU), Saumarez Reef (SRE) and Meaban (MEA) virus are thought not to be human pathogens. They have dispersed into very different regions of the world, *i.e.*, Far Eastern Russia (TYU), the Great Barrier Reef (SRE) and North Western France (MEA). Serological evidence shows that seabirds may become infected by these viruses when they are bitten by infected ticks, but no clinical signs of infection have been noted. This may indicate a long-term association between the birds and the viruses. Given its phylogenetic relationship to the seabird viruses (Fig. 2),

the Kadam (KAD) virus, which was isolated from both tick species, *Rhipicephalus pravus*, feeding on cows in Uganda and *Hyalomma dromedarri*, feeding on camels in Saudi Arabia, was also included in table 1 in this group of viruses. The seabird-tick associated viruses (TYU, SRE and MEA) constitute a monophyletic sister group to the TBE complex viruses. The sequence diversity between each of the seabird-tick associated viruses appear to be greater (approximately 40% amino acid differences) than that between most of the TBE complex viruses (approximately 30% amino acid differences). These viruses can be isolated from ticks that are found in the nests of the migratory seabirds. It is fair to assume that their wide biogeographic distribution, including almost all continents, may be due to their association with migratory seabirds. It is quite possible that some TBE lineages were dispersed by mechanisms similar to those for the TYU serogroup viruses. For example, Macquarie Island in the Southern Ocean, the source of Gadgets Gully (GGY) virus, is situated several hundred miles off Southern Australia, and is visited by very few humans. This virus, also associated with migratory seabirds, can be found under the rocks and debris that are used by the penguins and other seabirds. Because of their antigenic relatedness with the TBE complex viruses in the northern hemisphere, it was suggested that GGY virus could have been introduced to the Island by Storm Petrels (*Oceanites oceanicus*) or the Arctic Tern (*Sterna Paradisaea*) which have reciprocal breeding patterns between the Arctic and Antarctic (St. George *et al.*, 1985). It was suggested that the GGY virus could represent a natural link between subarctic and subantarctic TBE complex viruses and possibly also a more genetically distant link between the seabird associated TYU group and the rodent associated TBE complex viruses (St. George *et al.*, 1985). Possibly during their clinal dispersal, the TBE complex viruses reached the forests and woodlands of Asia (POW, KSI, RF, KFD, LGT, etc) where the opportunity arose for continuous or progressive dispersal along the defined corridors provided by the forests.

4. MOSQUITO-BORNE VIRUSES

In the natural environment the mosquito-borne viruses are known to be primarily, but not exclusively, associated with mosquitoes from either *Culex* spp. or *Aedes* spp. (Tab. 1). This division of the viruses on the basis of their associated mosquito species correlates precisely with phylogenetic inferences (Fig. 2).

Viruses associated with *Culex* spp. are distributed in the Old World, such as the Japanese encephalitis (JE) virus and West Nile virus; or in the New World, such as the Rocio virus, Ilhéus virus, Cacipacore virus and St Louis encephalitis virus. Most of these viruses are possibly ancestrally related, since they form a monophyletic group in phylogenies based on either the E gene

(Zanotto *et al.*, 1996) or in the NS5 gene (Kuno *et al.*, 1998). However, at this time it is not clear where they did come from. The *Culex* spp. viruses, preferentially bite birds and rodents in forest environments and are also attracted to pigs, horses, ducks, etc. that form an essential part of the human food chain in rural Asia. There is debate on if the JE group viruses have moved from the Old to the New World or vice-versa. However, frequently the origin of any specific virus is quite hard to pinpoint. During the last four centuries humans started to move around all continents, with increasing speed and numbers. As a consequence, drastic changes in the natural ecological balance and distributions of many host and vector species took place since then. Today, modern transportation systems diminished global distances to a few hours flight, certainly having a serious impact on the spread of infectious agents and the emergence of viral diseases (Morse, 1993). An example on how a flavivirus may become an unsuspected emergent health threat happened in 1999. During that year the West Nile (WN) virus, previously only detected in Southern Europe and Africa, made its way to North America causing a highly publicized outbreak in the city of New York (Briese *et al.*, 2000; Lanciotti *et al.*, 1999). Several of the *Culex* spp. associated flaviviruses have also been isolated in nature from a wide range of species of bats and rodents. It is interesting that some *Culex* spp. associated flaviviruses such as Murray Valley encephalitis, Alfuy, Kokobera and Stratford are found only in Australia and nearby regions of Southeast Asia.

As shown in figure 2, flaviviruses associated with *Aedes* spp. are found in the Old World. These viruses tend to bite primates in forest or savannah environments and/or humans and herded animals in urban or rural environments. Some of these viruses (yellow fever and dengue virus) are also associated with the development of haemorrhagic disease in humans. In contrast with the *Aedes* spp. viruses, many of the *Culex* spp. viruses (SLE, JE, MVE and WN) are more characteristically associated with encephalitic disease in humans. Whether or not such divisions in disease characteristics have occurred through selection of specific genetic determinants remains to be confirmed. Possibly due to its association with humans in recent historical times, both the yellow fever (YF) and dengue (DEN) virus have also been found in the New World. The role of slave trading on their emergence and establishment in the Americas has always been considered (Strode, 1951; Innis, 1995). There is good historical evidence that the YF virus caused outbreaks in the Caribbean 350 years ago due the introduction of its natural vector, *Aedes aegypti*, which was transported from Africa (Markoff and Falgout, 1995). The YF virus was also the first human disease known to be caused by filterable agent and the first human-infecting arbovirus isolated (Theiler and Downs, 1973). Yellow fever shows a relatively restricted geographic distribution, and can be found in central tropical regions of both the Old World (Africa) and New World (Americas). Yet, the YF virus did not become established in

tropical Asia, even when the virus could make use of susceptible mosquitoes and non-human primates hosts. Several causes to this interesting bias on the pattern of YF virus dispersal have been postulated (Gould *et al.*, 2001). Possibly in Asia, the YF virus may become subjected to competitive exclusion, either by related viruses or by some form of herd immunity. Moreover, the low vector competence of Asian strains of *Aedes aegypti*, and the relatively low frequency and load of virus introduced into Asia, compared with the Americas (there was no significant slave trading between Africa and Asia) may also have reduced its chances to move west. Differences in biological characteristics of viral lineages and on the immunity of vertebrate reservoirs may also be brought to mind. The viral strains dispersed eastward could be significantly different from the strains from central and West Africa. It would be interesting to see if Asians and/or Asian primates may not allow YF viruses to establish competent sylvatic cycles. The role of the sylvatic cycle in the establishment of the YF in the Americas compared to its failure in the West has yet to be fully assessed.

On the other hand, DEN viruses appear to have bypassed the need of a sylvatic cycle. They are now mainly a human disease and have a worldwide distribution along the tropics wherever *Aedes (Stegnomomyia)* becomes established. This is certainly helped by the intense anthrophilic nature of the *Aedes aegypti*, which is its main vector (Innis, 1995). Outbreaks of dengue fever were described 200 years ago and the DEN virus was isolated in 1907 and shown to transmit by arthropods (Markoff and Falgout, 1995). Today, its four known serotypes, which also happen to map onto four distinct groups (Fig. 1) cause 100 million cases of dengue fever (DF) every year, producing at least 250,000 cases of dengue hemorrhagic fever (DHF) with a 5% mortality rate (Monath, 1994). DEN virus frequently produces dual infections in humans (Laille *et al.*, 1991; Lorono Pino *et al.*, 1999), which in some cases result in intratypic genetic recombination (Worobey *et al.*, 1999). The increasingly wide dispersal of DEN viruses throughout the tropics, reflects the severe impact that urbanization, transportation, commercialization and tourism, are having on human disease. Other *Aedes* spp. associated viruses, such as Spondweni, Zika, Wesselsbron, Banzi, Uganda S, etc. appear to have more restricted geographic distributions. They are found in Africa or Asia in forest habitats, or in environments containing wild or farmed animals. As with some of the *Aedes* spp., many of the *Culex* spp., associated viruses also frequently show geographic overlap in Asia, Australasia or the Americas.

4.1 Dispersal of mosquito-borne viruses

Mosquito-transmitted viruses, have dispersal characteristics which are largely understood to be based on the different behavioural patterns of mosquitoes and their vertebrate hosts. Whereas a tick bloodmeal usually takes several days to complete, mosquitoes feed on vertebrate hosts within a few

seconds. If the mosquito becomes infected, the virus then replicates and within 8 to 10 days reaches a high level of infectivity in the brain, body and salivary glands of the mosquito. Within a few weeks, the mosquito may feed again, transferring the virus to a vertebrate host which if susceptible will reproduce the virus to high titres in the target organs, develop a viraemia and transmit this virus to uninfected mosquitoes that feed during the viraemic stage. The mosquito-borne viruses are therefore replicated through many cycles in a relatively short period of time. Mosquitoes are more mobile than ticks and may disperse the virus over significant distances within a few hours. Moreover, the hosts that are infected by the mosquitoes may also be more mobile than the corresponding forest animals that serve as hosts for ticks. Overall, the life cycle and the factors that determine dispersion are much more dynamic in mosquito-transmitted viruses than tick-borne viruses. This is reflected by (i) the wider geographic dispersal of many individual mosquito-transmitted virus species (ii) the fact that mosquito-borne viruses show overlapping distributions and (iii) the mosquito-borne viruses show greater levels of genetic variation and higher estimated evolution rates particularly those that cause human epidemics, the structure of the phylogenetic trees suggests periods of rapid population growth. Despite these generalizations, many of the Old World *Aedes* spp. associated viruses have remained essentially sylvatic and show restricted geographic dispersal presumably because of their adaptation to local vector-vertebrate ecology. Yellow fever virus has a geographically limited distribution. Epidemic outbreaks occur in Central and West Africa, the Caribbean, Central and South America. Cases of YF anywhere else are introduced by individuals infected in a YF region, and then traveling to other parts of the world. Yellow fever introduced in this way, does not lead to subsequent infections. In the natural environment, the virus can be isolated from *Aedes* spp. that bite and infect monkeys living in the tree canopy of the tropical rainforests and in the savannah, which borders the equatorial forests. In Africa, many species of monkey show no clinical symptoms although they replicate and serve as reservoir hosts for the virus that is then transmitted to non-infected mosquitoes feeding on the infected monkeys. A similar virus life cycle, between mosquito and monkey, also occurs in the New World rain forests but the monkeys frequently become sick and die as the result of becoming infected which is taken to indicate that YF virus was introduced into the Americas relatively recently. As mentioned above, the considered wisdom is that the virus was transported from Africa on the slave boats that traveled frequently to the Americas during the past three or four hundred years (Strode, 1951; Innis, 1995; Markoff and Falgout, 1995). The fact that cases of YF were also seen in many seaports in Europe and even in northern ports of the United States of America, where slave boats were frequent visitors, supports this argument. Dispersal of YF out of Africa is therefore very restricted and is almost entirely

attributable to human commercial activities, particularly the slave trade, during the past few centuries.

Human YF infections occur throughout the year in the humid equatorial forests. Humans are bitten by infected sylvatic *Aedes* spp. vectors that usually feed on monkeys. The sylvatic cycle of virus being transmitted between monkeys and mosquitoes is referred to as jungle fever when it involves humans. Yellow fever epidemics in Africa occur towards the end of the rainy season when *Aedes aegypti* densities are at their highest. Humans infected in the savannah or rain forests take the disease to urban dwelling *Aedes aegypti*. These mosquitoes then spread the disease through the human population. This form of YF, *i.e.*, “human-mosquito-human”, is referred to as urban fever (Strode, 1951). In tropical America, the urban form of YF has not been reported for many years although in some of the heavily populated parts of tropical America it seems to be only a matter of time.

The dispersal characteristics of the dengue virus serotypes contrast significantly with YF virus. Whereas YF virus appears to have retained its sylvatic nature, the dengue viruses are better suited to the urban environment and epidemic dengue no longer seems to require a sylvatic reservoir host. All four dengue serotype viruses circulate and cause human epidemics throughout most tropical regions of the World wherever high densities of *Aedes aegypti* are present. On the basis of their phylogenies, the dengue viruses represent more recent evolutionary lineages than YF virus. It is therefore tempting to postulate that an ancestral dengue lineage arose in Africa, where *Aedes* spp. competent for dengue virus transmission were present, and the four recognized serotypes subsequently emerged and diverged as they dispersed across Asia and then into the New World, exploiting modern transportation to aid their dispersal. In common with YF, other viruses in the *Aedes* spp. clades, for example, SEP, EH, UGS show more restricted geographic dispersal presumably because, like YF virus, they are preferentially adapted to the prevailing vectors, vertebrates and local ecology.

Not surprisingly, the dispersal characteristics of the *Culex* spp. associated viruses contrast strongly with those of the *Aedes* spp. viruses mainly due to the fact that most *Culex* spp. viruses have life cycles that involve birds. Strains of many of these viruses, for example SLE, JE, WN, etc. can be isolated over wide geographic areas, because they are carried by migrating birds. Nevertheless, it is guesswork to try defining dispersal patterns for bird-associated viruses. However, in some cases, for example, WN or JE virus, dispersion was very efficient and these viruses can now be isolated over very large geographic regions. In other cases, for example KOK and STR, the viruses were more suitably adapted to local species in Northeast Australia and neighbouring regions of Southeast Asia. Even though there are *Culex* spp. viruses in both the Old and the New World, WN virus is the only one that has become established simultaneously, on both sides of the Atlantic Ocean and

this occurred only very recently. This supports the belief that birds do not routinely disperse *Culex* spp associated flaviviruses across the major Oceans. It is thought that WN virus was inadvertently introduced into the New York area of the United States after being carried by or in (i) an infected human, (ii) an infected mosquito (iii) the luggage of an aeroplane passenger, (iv) an infected bird or animal, imported illegally into North America, and (v) an infected bird blown off course from the Old World. Whilst any of these alternatives is theoretically possible, the fact that WN virus appears to have been successfully introduced to the New World on only one occasion argues against its introduction by an infected bird flying in from the Old World.

5. NO-KNOWN VECTOR VIRUSES

The NKV viruses (Tab. 1) fall into two distinct groups: (i) those associated with bats in the Old World and, (ii) those associated with rodents in the New World. Interestingly, each group is in a distinct phylogenetic lineage. The NKV viruses make a third monophyletic group (Fig. 2), being a sister group to both the mosquito- and tick-borne groups. The individual bat-associated NKV viruses are found either in the New World or in the Old World but none so far has been found in both regions. On the other hand, and with the exception of APOI virus, rodent associated NKV viruses have only been isolated in the New World where they have retained very restricted dispersal patterns. It is known that neither bats nor birds use migratory pathways across the major Oceans, *i.e.*, the most common long distance migratory flights are generally in a northerly and southerly direction. This probably accounts for the apparent lack of mixing between Old and New World NKV viruses even though there have clearly been introductions at some time in the past in one direction or the other, perhaps by rodents on ships. Nevertheless, bats undoubtedly contribute significantly to localized spread of flaviviruses, over major land masses. All the other viruses in the family *Flaviviridae* are non-vector-borne suggesting that this could be the ancestral state for the genus. In this case, we could speculate that flavivirus vector-transmission evolved from non-vector-borne transmission. The fact that some NKV viruses (EB, YOK, SOK) diverged with the mosquito-borne viruses and then separated to form a distinct NKV group, could imply that there has been a secondary loss of vector-borne transmission in these viruses. This is also evident from the fact that several mosquito-borne and even some tick-borne flaviviruses have been isolated from bats in the natural environment. The NKV viruses are an interesting group in the genus, which may hold some interesting clues as to the origin and biology of this genus. Moreover, the tree in figure 2, suggests that the divergence between the rodent and bat NKV viruses may have occurred early in the

evolutionary history of the genus *Flavivirus*, which poses important questions on its origin and ancestral biology.

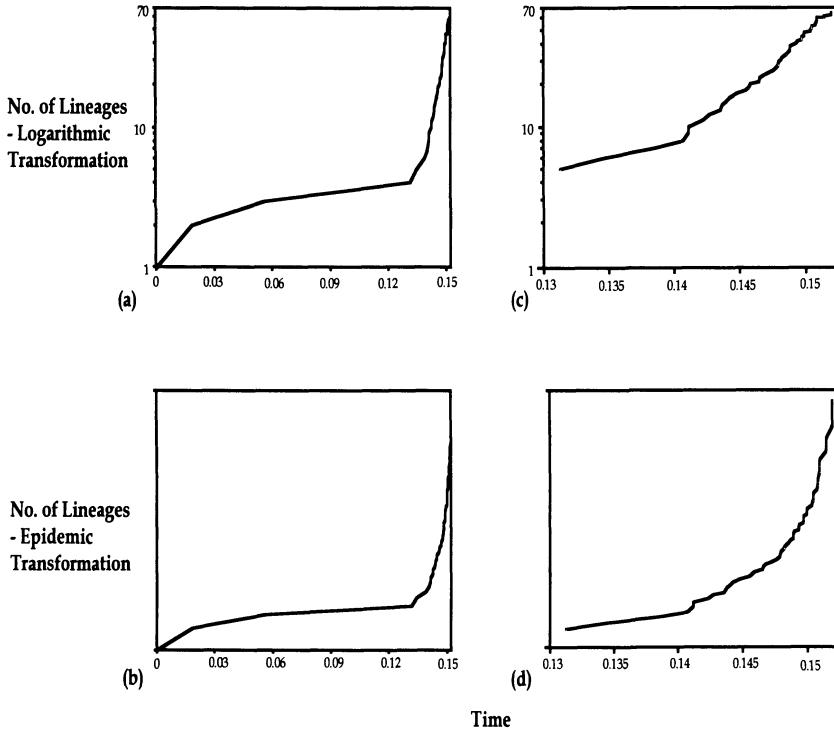


Fig. 6. (a) Lineages-through-time plots of 66 dengue (DEN) viruses (from Zanotto *et al.*, 1996a). The time at which each lineage split (*i.e.*, cladogenetic event) occurs on a constant rate (KITCH) phylogenetic tree is plotted with the time axis scaled as the number of nucleotide substitutions from the root of the tree to the tips. (b) Epidemic transformation of the same data. This transformation of the y -axis determines whether the rate of population growth has been constant through time (straight line), increasing (upward curvature, as here) or decreasing (downward curvature). (c) Lineages-through-time of the most recent 62 nodes of DEN viruses. (d) Epidemic transformation of the same data.

6. METHODOLOGICAL CONSIDERATIONS ABOUT PHYLOGENETIC INFERENCES

6.1 The origin of the Negishi (NEG) virus: A case study

The use of molecular phylogenies for distinguishing flaviviruses allowed solving the puzzle of the Negishi (NEG) virus, which was found to be a LI virus (Venugopal *et al.*, 1992; Marin *et al.*, 1995b; Kuno *et al.*, 1998; Gao *et al.*, 1998). The Negishi virus was apparently isolated from two children during an epidemic of Japanese encephalitis in 1948 in an urban area of Tokyo (Ando

et al., 1952; Okuno *et al.*, 1961). However, historical evidence corroborates the molecular phylogenies. Firstly, the LI virus has been isolated many times in the British Isles and was fully characterized nearly twenty years before the first reports on the Negishi virus. Therefore, the LI virus was already in many different laboratories around the World, being used as a reference strain. Secondly, Negishi virus was never re-isolated in Japan and there is no definitive serological evidence of its presence there. Finally, the concept that LI virus (strain Negishi) exists in Japan is inconsistent with the recognized characteristic of LI virus as the etiological agent of tick-borne sheep encephalomyelitis on the sheep-grazing uplands of the British Isles (Reid, 1984). It remains to be explained how a strain of LI virus apparently traveled thousands of miles from Britain, to cause two fatal infections in children in a region of Japan where there were no sheep and suitable ticks to transmit the virus.

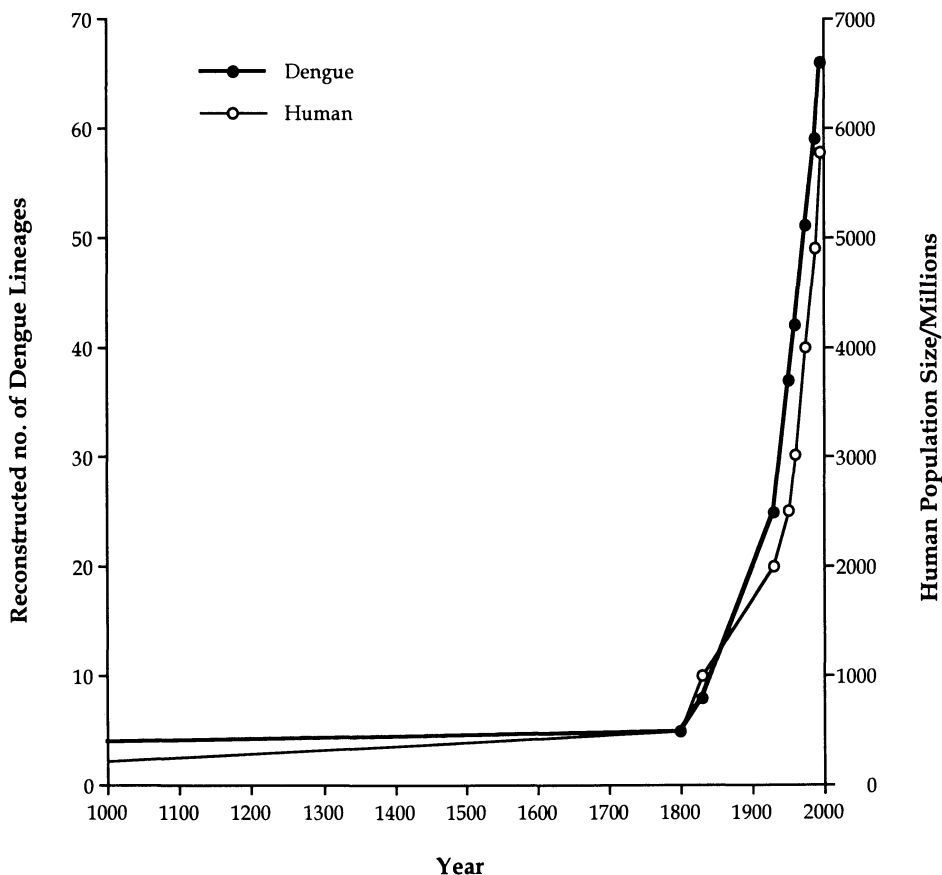


Fig. 7. Plot of the number of reconstructed lineages of dengue (DEN) virus against the time at which they first appear and the corresponding growth of the human population (from Zanotto *et al.*, 1996a).

6.2 Birth and death, rates and dates

The TBE complex viruses appear to have diverged continuously, generating new lineages continually through time. This can be recovered in its asymmetric phylogenetic tree with the characteristic stepwise branching process (Figs. 1, 3, 4 and 5). In contrast, the mosquito-borne viruses, such as the DEN viruses, show a distinct phylogenetic pattern, with long branches followed by intense apical cladogenetic activity (Zanotto *et al.*, 1996a) (Figs. 1 and 6). For the mosquito-borne viruses there are no observable lineages branching deep from within each of the main serotypes until relatively recently, when there is a burst of cladogenesis (Fig. 1). The lack of intermediary lineages in the DEN viruses, which has four distinct serotypes, could be explained by intense replacement of lineages in time, whereas the TBE viruses may not have been subjected to the same stringent birth and death process (Zanotto *et al.*, 1996a). This notion is also supported by the serological data, which shows close antigenic relationships between the TBE complex viruses and greater antigenic and genetic diversity among the mosquito-borne viruses. The simplest explanation for the two-phase branching pattern observed in mosquito-borne viruses, such as DEN viruses, is that they have been enzootic until the last 200 years. Since then, the DEN viruses were able to find an endless transmission web in the exponentially growing human population. This notion was suggested by some rough calculations of the times of divergence of the mosquito-borne viruses, which show that many of the viruses currently causing epidemic outbreaks possibly represent very recent lineages. In the DEN and JE serotypes the period of intense cladogenesis was estimated to have occurred during the past two centuries (Zanotto *et al.*, 1996a; Holmes, 1998), reflecting the recent availability of new and susceptible hosts, *i.e.*, humans, coupled with increased worldwide movement and mixing of vectors, viruses and hosts (Fig. 7). Assuming the CFA as an outgroup for the *Flavivirus* and by knowing the dates of isolation of these viruses, it was possible to speculate on some rough estimates of the rates of evolution of tick-borne and mosquito-borne viruses. These preliminary results indicated that mosquito-borne viruses could be evolving almost twice as fast as the tick-borne viruses (Zanotto *et al.*, 1995). If these rates are different, they could entail the higher turnover of the pandemic mosquito-borne viruses in mosquitoes and humans compared to that of the tick-borne virus, which is mainly zoonotic. Based on these rates, estimations of the times of emergence for the entire TBE complex suggest that the earliest lineage of the TBE complex viruses, *i.e.*, the POW virus, separated from the TYU group between 4000 to 6000 years ago (Zanotto *et al.*, 1996a). However, those estimates can underestimate the amount of change along branches and include a substantial margin of error and uncertainty. Interestingly, this dating is sensible since the most recent glaciation lasted until about 10,000 to 12,000 years ago encompassing extensive areas of the northern

hemisphere. It can be speculated that after the ice age, viruses carried by ticks left the warmer regions of Africa and South East Asia and moved northward into the Far East Asia and westward into Europe forming clines.

6.3 The origin of the *Flavivirus*

It has been postulated that the genus *Flavivirus* originated in the Old World (Gould *et al.*, 2001). However, there is insufficient information to answer unequivocally the question of the origin of the flaviviruses. Nevertheless, most of the tick-borne viruses, all of the *Aedes* spp. associated viruses and many of the NKV viruses, have sylvatic life cycles in the Old World, implying their presence there over a long period of time. Cell fusing agent (CFA), which is a very divergent lineage, has been used extensively to root flavivirus trees (Marin *et al.*, 1995b) on the basis of its similar genome strategy and sequence similarity. This insect virus was isolated from African *Aedes aegypti* and appears to be non-pathogenic for this mosquito species (Cammissa Parks, 1992), suggestive of a possible long term relationships between the virus and the host. The rooting based on the CFA is however problematic for two reasons. Firstly, the level of sequence similarity approaches random and because of this, it may lead to spurious inferences on the temporal relationships among sister groups that are function of its precise placement on the phylogeny. Therefore, as a precautionary measure, the age relationships among the three groups in the genus should be understood as a hard trichotomy, since their association based on the placement of the CFA could be misleading. The proper use of outgroups for the genus is of contrastive nature. For example, in order to study more than the patterns of adjacency of the mosquito-borne viruses, one should use either NKV or tick-borne viral sequences (preferably a comprehensive medley of both) to try making inferences about the relative age of lineages. Certainly the use of consensus sequences as outgroups is highly problematic (Chang *et al.*, 1995). Consensus sequences are based on the frequency and representation of each character state in a dataset, not on the ancestral status of each state for each character. By using proper outgroups, consensus sequences will move outwards in trees, becoming useless when we try to order the cladogenetic events in time. Moreover, phylogenetic studies on the *Flavivirus* genus should take into account the need to correct for superimposed mutations, use of better phylogenetic reconstruction methods, which can estimate among-sites rate variations and optimize the values of nucleotide transition probability matrices.

Another pressing issue is that of the proper estimation of the amount change along lineages (genetic distance) (K). Provided we know the time of isolation of lineages, this can lead to dating estimations (T), since both time (T) and distance (K) are related by the substitution rate (R) and can be approximated by the simple expression $R=K/2T$. It was stated that the fast

mutation of some viruses with RNA genomes has led some virologists to conclude that most RNA viruses we know today probably arose in the last 10,000 years (Gibbs *et al.*, 1995). We know that RNA viruses possibly existed way before then and their presence in all life Kingdoms is a sure proof of this. However, the “Gibbs paradox” is real but has an interesting explanation. The high mutation rate in conjunction with the replacement of lineages through time (*i.e.*, birth-death process) possibly sets a threshold of phylogenetic signal, limiting the extent in which we can look back in time using viral sequences (Zanotto *et al.*, 1996b). Therefore, the viral populations that we sample today may not carry all the phylogenetic information of those viruses from which they arose. The genus *Flavivirus* comprises viruses that are recognizable today and it is the lineages that these viruses represent that we can attempt to date. Many other related flaviviruses may have emerged but subsequently become extinct. Possibly there is a common ancestor for the *Flaviviridae* family, but we may not have access to their ancestral lineages since they are long gone and left no trace in currently available sequences (Zanotto *et al.*, 1996a; 1996b).

However, the datings within closely related members of the genus may be reasonable. Using either logical arguments based on historical facts or by calculating nucleotide substitution rates and then estimating dates of divergence from known times of virus isolation, one can attempt to estimate the divergence times. As a consequence, a few interesting correlations appear. The evidence based on molecular phylogenies for the presence of the LI viruses in the UK for no more than three or four centuries is now compelling (McGuire *et al.*, 1998). The term louping ill is an ancient word and was used in the 18th century to describe a disease of sheep occurring in the Border counties of England and Scotland where sheep were farmed intensively on the hillsides (MacFadzean, 1990). Sheep were then introduced north of the Forth-Clyde valley onto the Scottish moorlands and the disease followed quite rapidly. Phylogenetic trees indicate that Scottish LI viruses appeared more recently than those lineages leading to the Irish and Welsh LI viruses (McGuire *et al.*, 1998). In each of these geographic regions, grouse are highly susceptible to infection by LI virus even though the grouse have clearly been associated with moorland environments for many centuries, as shown by the fact that they have adapted a specialized digestive tract for the digestion of heather (Leslie and Shipley, 1912). The very high virulence of LI virus for grouse could also imply its recent introduction to the moorlands (Chamberlain, 1982). Further support comes from the evidence that sheep could not have been successfully farmed either on moorlands or even on any uplands during most of the 16th century because Britain and Europe experienced a “Little Ice Age”. Moreover, throughout Europe and Asia, *Ixodes ricinus* is essentially a forest and woodland tick and only prospers on moorlands if there are sufficient suitable mammalian hosts, such as sheep or deer. Despite significant investigation, there is no evidence that either LI virus or an antigenically closely related TBE

complex virus is, or ever was, present in the British forests or woodlands. These facts therefore strongly support the phylogenetic explanation for the appearance of LI virus in the British mainland, *i.e.*, an early sheep encephalomyelitic virus was introduced from Europe, probably from Spain, through Ireland during the past few centuries. The increasing amount of data on distinct populations of the TBE complex virus (Ecker *et al.*, 1999; Hayasaka *et al.*, 1999; 2001) could lead to a better understanding on the mode and tempo of the dispersal of these viruses in Eurasia.

In sum, it could be assumed that the viruses in the genus *Flavivirus* are likely to represent lineages that have radiated during the past 5000 to 10,000 years, since the last major Ice Age across the northern hemisphere. Estimates were made for several tick- and mosquito-borne viruses (Zanotto *et al.*, 1996a; McGuire *et al.*, 1998) and they are in general agreement. Whilst it is recognized that the precision with which such estimates can be made, is likely to be low for viruses in the deeper nodes of the tree, it is unlikely that they are incorrect by several orders of magnitude.

7. REFERENCES

- Ando K., Kuratsuka K., Arima S., Hironaka N., Honda Y., and Ishii K. 1952. Studies on the viruses isolated during epidemic of Japanese B encephalitis in 1948 in Tokyo area *Kitasato Arch Exp Med* 24:49-61.
- Berthet F. X., Zeller H.G., Drouet M.T., Rauzier J., Digoutte J.P., and Deubel V. 1997. Extensive nucleotide changes and deletions within the envelope glycoprotein gene of Euro-African West Nile viruses *J Gen Virol* 78:2293-7.
- Billoir F., de Chesse R., Tolou H., de Micco P., Gould E.A., and de Lamballerie X. 2000. Phylogeny of the genus flavivirus using complete coding sequences of arthropod-borne viruses and viruses with no known vector *J Gen Virol* 81:781-90.
- Blok J., McWilliam S.M., Butler H.C., Gibbs A.J., Weiller G., Herring B.L., Hemsley A.C., Aaskov J.G., Yoksan S., and Bhamarapravati N. 1992. Comparison of a dengue-2 virus and its candidate vaccine derivative: Sequence relationships with the flaviviruses and other viruses. *Virology*. 187:573-590.
- Briese T., Glass W.G., and Lipkin W.I. 2000. Detection of West Nile virus sequences in cerebrospinal fluid [letter] *Lancet* 355:1614-5.
- Calisher C.H., Karabatsos N., Dalrymple J.M., Shope R.E., Porterfield J.S., Westaway E.G., and Brandt W.E. 1989. Antigenic relationships between flaviviruses as determined by cross-neutralization tests with polyclonal antisera *J Gen Virol* 70:37-43.
- Camisa Parks H., Cisar L.A., Kane A., and Stollar V. 1992. The complete nucleotide sequence of cell fusing agent (CFA): Homology between the nonstructural proteins encoded by CFA and the nonstructural proteins encoded by arthropod-borne flaviviruses *Virology* 189:511-524.
- Chamberlain R.W. 1982. Arbovirology-then and now. *Am J Trop Med Hyg* 31:430-437.
- Chang J.J., Cropp C.B., Kinney R.M., Trent D.W., and Gubler D.J. 1995. Nucleotide sequence variation of the envelope protein gene identifies two distinct genotypes of yellow fever virus *J Virol* 69:5773-5780.
- Charrel R.N., Zaki A.M., Attoui H., Fakeeh M., Billoir F. Yousef A.I., Chesse R., De Micco P., and Gould E. A. 2001. Complete coding sequence of the alkhurma virus, a tick-borne flavivirus causing severe hemorrhagic fever in humans in Saudi Arabia. *Bioch Biophys Res Commun* 287: 455-461.
- Ecker M., Allison S.L., Meixner T., and Heinz F.X. 1999. Sequence analysis and genetic classification of tick-borne encephalitis viruses from Europe and Asia *J Gen Virol* 80:179-85.
- Ernek E. 1960. Experimental Pathogenicity of the Tick-borne Encephalitis Virus for Domestic Ducks. In *Biology of viruses of the tick-borne encephalitis complex*, p. 387-390. Academic press, New York and London, Smolenice, Slovakia.

- Gaidamovich S.Y. 1995. Tick-borne flavivirus infections. In *Exotic Viral Diseases*, Porterfield J.S., ed. Chapman and Hall Medical, London.
- Gao G.F., Hussain M.H., Reid H.W., and Gould E.A. 1993a. Classification of a new member of the TBE flavivirus subgroup by its immunological, pathogenetic and molecular characteristics: identification of subgroup-specific pentapeptides. *Virus Res* 30:129-44.
- Gao G.F., Jiang W.R., Hussain M.H., Venugopal K., Gritsun T.S., Reid H.W., and Gould E.A. 1993b. Sequencing and antigenic studies of a Norwegian virus isolated from encephalomyelitic sheep confirm the existence of louping ill virus outside Great Britain and Ireland *J Gen Virol* 74:109-14.
- Gao G. F., Zannotto P.M.A., Holmes E.C., Reid H., and Gould E.A. 1998. Molecular variation, evolution and geographical distribution of louping ill virus. *Acta Virologica* 41:259-268
- Gaunt M.W., Sall A.A., de Lamballerie X., Falconar, A.K.I., Dzhanian T.I., and Gould E.A. 2001. Phylogenetic relationships of flaviviruses correlate with their epidemiology, disease association and biogeography. *J Gen Virol* 82:1867-1876.
- Gibbs A., Calisher C.H., and F. 1995. *Molecular Basis of Virus Evolution*. Garcia-Arenal, ed. Cambridge University Press, Cambridge.
- Gould E.A., Lamballerie X., Zannotto P.M.A., and Holmes E.C. 2001. Evolution, epidemiology and dispersal of flaviviruses revealed by molecular phylogenies. *Advances in Virus Research* 57:71-103.
- Gresikova M., Sekeyova M., Stupalova S., and Necas S. 1975. Sheep milk-borne epidemic of tick-borne encephalitis in Slovakia. *Intervirology* 5:57-61.
- Gritsun T.S., Liapustin V.N., Karganova G.G., and Lashkevich V.A. 1988. Nonvirion (soluble) antigen of the tick-borne encephalitis virus. *Vopr Virusol* 33:217-27.
- Gritsun T.S., Liapustin V.N., Shatalov A.G., and Lashkevich V.A. 1990. Multiple forms of the NS1 protein as the main component of the nonvirion ("soluble") antigen of the tick-borne encephalitis virus. *Vopr Virusol* 35:471-4.
- Gritsun T.S., Lisak V.M., Liapustin V.N., Korolev M.B., and Lashkevich V.A. 1989. Slowly-sedimenting hemagglutinin of the tick-borne encephalitis virus. *Vopr Virusol* 34:449-54.
- Gritsun T.S., Venugopal K., Zannotto P.M.A., Mikhailov M.V., Holmes E.C., Polkinghorne I., Sall A.A., Frolova T.V., Pgodina V.V., Lashkevich V.A., and Gould E.A. 1997. Complete sequence of two tick-borne flaviviruses from Siberia and the United Kingdom: analysis and significance of the 5' and 3' UTR's. *Virus Research* 49:27-39.
- Hayasaka D., Suzuki Y., Kariwa H., Ivanov L., Volkov V., Demenev V., Mizutani T., Gojbori T., and Takashima I. 1999. Phylogenetic and virulence analysis of tick-borne encephalitis viruses from Japan and far-Eastern Russia. *J Gen Virol* 80:3127-35.
- Hayasaka D., Ivanov L., Leonova G.N., Goto A., Yoshii K., Mizutani T., Kariwa H., and Takashima I. 2001. Distribution and characterization of tick-borne encephalitis viruses from Siberia and far-eastern Asia. *J Gen Virol* 82:1319-1328.
- Heinz F.X., Collett M.S., Purcell R.H., Gould E.A., Howard C.R., Houghton M., Moorman R.J.M., Rice C.M., and Thiel H.J. 2000. Family *Flaviviridae*. In *Virus Taxonomy*, p. 859-878, Regenmortel M.H.V., Fauquet C.M., Bishop D.H.L., Carstens E., Estes M.K., Lemon S., Maniloff J., Mayo M.A., McGeogch D., Pringle C.R., and Wickner R.B., eds. 7th International committee for the Taxonomy of Viruses, Academic Press, San Diego.
- Hennig W. 1996. *Phylogenetic Systematics*. University of Illinois Press.
- Holmes E.C. 1998. Molecular epidemiology and evolution of emerging infectious diseases. *Brit Med Bull* 54:533-543.
- Innis B.L. 1995. Dengue and dengue Hemorrhagic fever. In *Exotic Viral Diseases*. Porterfield J.S., ed. Chapman and Hall Medical, London.
- Jacobs M.G., Robinson P.J., Bletchly C., Mackenzie J.M., and Young P.R. 2000. Dengue virus nonstructural protein 1 is expressed in a glycosyl-phosphatidylinositol-linked form that is capable of signal transduction. *FASEB J* 14:1603-10.
- Jenkins G.M., Pagel M., Gould E.A., Zannotto P.M.A., and Holmes E.C. 2001. Evolution of base composition and codon usage bias in the genus *Flavivirus*. *J Mol Evol* 52:383-390.
- Karabatsos N. 1985. International Catalogue of Arthropod-borne viruses 3rd edn., San Antonio, Texas: American Society for Tropical Medicine and Hygiene:[Suppl 1]: 137-152.
- Kuno G., Chang G.J., Tsuchiya K.R., Karabatsos N., and Cropp C.B. 1998. Phylogeny of the genus *Flavivirus*. *J Virol* 72:73-83.
- Labuda M., Kozuch O., Zuffova E., Eleskova E., Hails R.S., and Nutall P. 1997. Tick-borne encephalitis virus transmission between ticks cofeeding on specific immune natural hosts. *Virology* 235:138-143.

- Laille M., Deubel V., and Sainte Marie F.F. 1991. Demonstration of concurrent dengue 1 and dengue 3 infection in six patients by the polymerase chain reaction. *J Med Virol* 34:51-4.
- Lanciotti R.S., Roehrig n J.T., Deubel V., Smith J., Parker M., Steele K., Crise B., Volpe K.E., Crabtree M.B., Scherret J.H., Hall R.A., MacKenzie J.S., Cropp C.B., Panigrahy B., Ostlund E., Schmitt B., Malkinson M., Banet C., Weissman J., Komar N., Savage H.M., Stone W., McNamara T., and Gubler D.J. 1999. Origin of the West Nile virus responsible for an outbreak of encephalitis in the northeastern United States. *Science* 286:2333-7.
- Lee J.M., Crooks A.J., and Stephenson J.R. 1989. The synthesis and maturation of a non-structural extracellular antigen from tick-borne encephalitis virus and its relationship to the intracellular NS1 protein. *J Gen Virol* 70:335-43.
- Leitmeyer K.C., Vaughn D.W., Watts D.M., Salas R., Villalobos I., de Chacon, Ramos C., and Rico-Hesse R. 1999. Dengue virus structural differences that correlate with pathogenesis. *J Virol* 73:4738-47.
- Lepiniec L., Dalgarno L., Huong V.T., Monath T.P., Digoutte J.P., and Deubel V. 1994. Geographic distribution and evolution of yellow fever viruses based on direct sequencing of genomic cDNA fragments. *J Gen Virol* 75:417-23.
- Leslie A.S., and Shipley A.E. 1912. *The Grouse in Health and Disease*. Smith, Elder and Co., London.
- Lindebach B.D. and Rice C.M. 2001. Flaviviridae: The viruses and their replication. In *Fields Virology* 4th Edition, Knipe D.M. and Howley D.M., eds. Raven Press, Philadelphia.
- Lorono Pino M.A., Cropp C.B., Farfan J.A., Vorndam A.V., Rodriguez Angulo E.M., Rosado Paredes E.P., Flores Flores L.F., Beaty B.J., and Gubler D.J. 1999. Common occurrence of concurrent infections by multiple dengue virus serotypes. *Am J Trop Med Hyg* 61:725-30.
- Luby J.P. 1995. Powassan Encephalitis. In *Exotic Viral Diseases*. Porterfield J.S., ed. Chapman and Hall Medical, London.
- Mandl C.W., Heinz F.X., Stockl E., and Kunz C. 1989. Genome sequence of tick-borne encephalitis virus (Western subtype) and comparative analysis of nonstructural proteins with other flaviviruses. *Virology* 173:291-301.
- Marin M.S., McKenzie J., Gao G.F., Reid H.W., Antoniadis A., and Gould E.A. 1995a. The virus causing encephalomyelitis in sheep in Spain: a new member of the tick-borne encephalitis group. *Res Vet Sci* 58:11-3.
- Marin M.S., Zanotto P.M., Gritsun T.S., and Gould E.A. 1995b. Phylogeny of TYU, SRE, and CFA virus: different evolutionary rates in the genus *Flavivirus*. *Virology* 206:1133-9.
- Markoff L.J. and Falgout B.N. 1995. The family flaviviridae. In *Exotic Viral Diseases*. Porterfield J.S., ed. Chapman and Hall Medical, London.
- Mathur A., Arora K.L., and Chaturvedi U.C. 1982. Transplacental Japanese encephalitis virus (JEV) infection in mice during consecutive pregnancies. *J Gen Virol* 59:213-7.
- McFadzean J. 1900. The etiology of louping ill. *J Comp Pathol Ther* 13:145-154.
- McGuire K., Holmes E.C., Gao G.F., Reid H.W., and Gould E.A. 1988. Tracing the origins of louping ill virus by molecular phylogenetic analysis. *J Gen Virol* 79:981-8.
- Monath T.P. 1994. Dengue: the risk to developed and developing countries. *Proc Natl Acad Sci USA* 91:2395-2400.
- Monath T.P., and Heinz F.X. 1996. Flaviviruses. In *Fields Virology*, p. 961-1034, Fields B. N., ed. Lippincott-Raven, Philadelphia-New York.
- Morse S.S. 1988. *Emerging viruses*. Oxford University Press, New York.
- Okuno T., Oya A., and Ito T. 1961. The identification of Negishi virus a presumably new member of Russian Spring Summer Encephalitis virus family isolated in Japan. *Jap J Med Sci Biol* 14:51-59.
- Porterfield J.S. 1980. Antigenic characteristics and classification of Togaviridae. In *The Togaviruses*, p. 13-46, Schlesinger R.W., ed. Academic Press, New York.
- Randolph S.E., Miklisova D., Lysy J., Rogers D.J., and Labuda M. 1999. Incidence from coincidence: patterns of tick infestations on rodents facilitate transmission of tick-borne encephalitis virus. *Parasitology* 118:177-86.
- Reid H.W. 1984. Epidemiology of louping ill. In *Vectors in virus biology*, p. 161 -178, Mayo M.A., and Harrap K.A., eds. Academic Press, London.
- Rice C.M. 1996. *Flaviviridae: The Viruses and Their Replication*. In *Fields Virology*, p. 931-960. Fields B.N., Knipe D.M., eds. Lippincott-Raven, Philadelphia-New York.
- Rico Hesse R. 1990. Molecular evolution and distribution of dengue viruses type 1 and 2 in nature. *Virology* 174:479-93.
- Rico Hesse R., Harrison L.M., Salas R.A., Tovar D., Nisalak A., Ramos C., Boshell J., de Mesa M.T., Nogueira R.M., and da Rosa A.T. 1997. Origins of dengue type 2 viruses associated with increased pathogenicity in the Americas. *Virology* 230:244-51.

- Shiu S.Y., Ayres M.D., and Gould E.A. 1991. Genomic sequence of the structural proteins of louping ill virus: comparative analysis with tick-borne encephalitis virus. *Virology* 180:411-5.
- St. George T.D., Doherty R.L., Carley J.G., Filippich C., Brescia A., Casals J., Kemp D.H., and Brothers N. 1985. The isolation of arboviruses including a new flavivirus and a new Bunyavirus from *Ixodes (Ceraticxodes) uriae* (Ixodoidea: Ixodidae) collected at Macquarie Island, Australia, 1975-1979. *Am J Trop Med Hyg* 34:406-12.
- Strode G.K. 1951. *Yellow Fever*. MacGraw-Hill, New York.
- Tan B.F., Fu J., Sugrue R.J., Yap E.H., Chan Y.C., and Tan Y.H. 1996. Recombinant dengue type 1 virus NS5 protein expressed in *Escherichia coli* exhibits RNA-dependent RNA polymerase activity. *Virology* 216: 317-325.
- Swofford D.L. 2001. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.
- Theiler M. and Downs W.G. 1973. *The Arthropod-borne Viruses of Vertebrates: An Account of the Rockefeller Foundation Virus Program (1951-1970)*. Yale University Press, London.
- Venugopal K., Buckley A., Reid H.W., and Gould E.A. 1992. Nucleotide sequence of the envelope glycoprotein of Negishi virus shows very close homology to louping ill virus. *Virology* 190:515-21.
- Venugopal K., Gritsun T., Lashkevich V.A., and Gould E.A. 1994. Analysis of the structural protein gene sequence shows Kyasanur Forest disease virus as a distinct member in the tick-borne encephalitis virus serocomplex. *J Gen Virol* 75:227-32.
- Wang E., Ni H., Xu R., Barrett A.D., Watowich S.J., Gubler D.J., and Weaver S.C. 2000. Evolutionary relationships of endemic/epidemic and sylvatic dengue viruses. *J Virol* 74:3227-34.
- Wang H., Jennings A.D., Ryman K.D., Late C.M., Wang E., Ni H., Minor P.D., and Barrett A.D. 1997. Genetic variation among strains of wild-type yellow fever virus from Senegal. *J Gen Virol* 78:1349-52.
- Worobey M., Rambaut A., and Holmes E.C. 1999. Widespread intra-serotype recombination in natural populations of dengue virus. *Proc Natl Acad Sci USA* 96:7352-7.
- Zanotto P.M., Gao G.F., Gritsun T., Marin M.S., Jiang W.R., Venugopal K., Reid H.W., and Gould E.A. 1995. An arbovirus cline across the northern hemisphere. *Virology* 210:152-9.
- Zanotto, P.M., Gould E.A., Gao G.F., Harvey P.H., and Holmes E.C. 1996a. Population dynamics of flaviviruses revealed by molecular phylogenies. *Proc Natl Acad Sci USA* 93:548-53.
- Zanotto P.M., Gibbs M.J., Gould E.A., and Holmes E.C. 1996b. A Reevaluation of the higher taxonomy of viruses based on RNA polymerases. *J Virol* 70:6083-6096.