Journal of General Virology (1999), **80**, 2545–2557. Printed in Great Britain

Ecology and evolution of rabies virus in Europe

Hervé Bourhy,[1] Bachir Kissi,[1] Laurent Audry,[1] Marcin Smreczak,[3] Malgorzata Sadkowska-Todys,[4] Katariina Kulonen,[5]† Noël Tordo,[2] Jan F. Zmudzinski[3] and Edward C. Holmes[6]

[1,2] Unité de la Rage[1] and Laboratoire des Lyssavirus[2], Institut Pasteur, 25 rue du Docteur Roux, 75724 Paris Cedex 15, France

[3] National Veterinary Research Institute, 24-100 Pulawy, Poland

[4] National Institute of Hygiene, 24 Chocimska Str., 00-791 Warsaw, Poland

[5] National Veterinary and Food Research Institute, PL 368, FIN-00231 Helsinki, Finland

[6] Wellcome Trust Centre for the Epidemiology of Infectious Disease, Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, UK

The evolution of rabies viruses of predominantly European origin was studied by comparing nucleotide sequences of the nucleoprotein and glycoprotein genes, and by typing isolates using RFLP. Phylogenetic analysis of the gene sequence data revealed a number of distinct groups, each associated with a particular geographical area. Such a pattern suggests that rabies virus has spread westwards and southwards across Europe during this century, but that physical barriers such as the Vistula river in Poland have enabled localized evolution. During this dispersal process, two species jumps took place – one into red foxes and another into raccoon dogs, although it is unclear whether virus strains are preferentially adapted to particular animal species or whether ecological forces explain the occurrence of the phylogenetic groups.

Introduction

Lyssaviruses, such as that which causes rabies, are negative-strand RNA viruses that can be divided into seven genotypes (Bourhy *et al.*, 1992, 1993; Gould *et al.*, 1998). Viruses of genotypes 1, 5 and 6 are characterized by their natural and stable association with specific mammalian species which act as vectors for their transmission, so that a number of phylogenetic lineages co-circulate among a range of mammalian species (Kissi *et al.*, 1995; Amengual *et al.*, 1997). Infection of an animal with a lyssavirus that originated within a different reservoir population will generally lead to a fatal self-limiting rabies-like infection (a 'spill-over'), as in the case of humans, and only occasionally to a new stable enzootic infection (Blancou *et al.*, 1983, 1991).

Occasionally lyssaviruses gain access to new populations of susceptible hosts, particularly those which are geographically restricted (Rupprecht & Smith, 1994), or evolve to infect previously less susceptible hosts (Sacramento *et al.*, 1992; Smith *et al.*, 1992, 1995; Tordo *et al.*, 1993; Nadin-Davis *et al.*, 1994). It is evident that such an adaptive process took place in Europe during the first decades of this century when rabies virus became established in the red fox following a decline in incidence among urban dogs and wolves (Zeeti & Rosati, 1966; Petrovic, 1987). Although the virus initially failed to adapt to red foxes, as shown in the records of animal deaths (Barbier, 1929; Jaujou, 1949; Steck & Wandeler, 1980; Blancou *et al.*, 1991), by 1940–1945 rabies-infected foxes were regularly found at the former Russian–Polish border (Zunker, 1954) and in the region of Gdansk in northern Poland (Seroka, 1968). Subsequently, the infection of red foxes spread to the rest of Europe, reaching France by 1968 (Atanasiu *et al.*, 1968).

The study described here was designed to determine the evolutionary history of rabies virus in Europe using nucleoprotein (N) and glycoprotein (G) gene sequences. In particular, we wished to determine the level and structure of standing genetic variation within European rabies viruses and reveal what evolutionary processes might have given rise to this structure. Furthermore, as little is known about how the host range of rabies virus is determined at the molecular level, we also aimed to identify those mutations, if any, which might have allowed the virus to infect new species. The N gene was chosen for this analysis because it encodes an internal protein involved in the regulation of transcription and replication and

Author for correspondence: Hervé Bourhy.
Fax +33 1 40 61 30 20. e-mail hbourhy@pasteur.fr

† This paper is dedicated to the memory of the late Katariina Kulonen.

0001-6325

Table 1. Origin of the rabies virus isolates

N, nucleoprotein gene; G, glycoprotein gene.

Country	No.	Species	Year	Reference/source	GenBank accession no. N	GenBank accession no. G
Europe						
Austria	9447AUT	Red fox	1994	W. Schuller	U42708	
Belgium	9352BEL	Red fox	1992		U42709	
	9354BEL	Red fox	1991	F. Costy	U42710	
	9356BEL	Red fox	1991		U42711	
	9432BEL	Cat	1993	D. Peharpe	U42712	
	9433BEL	Red fox	1994	D. Peharpe	U42713	
Bosnia–Herzegovina	8653YOU	Wolf	1986	M. Petrovic	U42704	AF134341
	86111YOU	Red fox	1986	M. Petrovic	U42706	
Czech Republic	9465TCE	Red fox	1994	O. Matouch	U43010	
Estonia	9142EST	Raccoon dog	1985	Kissi *et al.* (1995)	U22476	AF134339
	9144EST	Raccoon dog	1991	A. King	U42714	
	9338EST	Red fox	1992		U42715	
	9339EST	Raccoon dog	1991	K. Kulonen	U42707	AF134335
	9342EST	Raccoon dog	1991		U43432	AF134336
Fed. Rep. Yugoslavia	86106YOU	Red fox	1972	Kissi *et al.* (1995)	U22839	
	86107YOU	Red fox	1976	M. Petrovic	U42703	AF134345
	8658YOU	Cattle	1981	M. Petrovic	U42705	AF134344
Finland	9348FIN	Raccoon dog	1988	K. Kulonen	U42716	
France	9223FRA	Red fox	1974	J. Barrat	U43433	AF134326
	8661FRA	Hedgehog	1984	Institut Pasteur	U43434	
	8663FRA	Red fox	1984	Institut Pasteur	U42605	AF134325
	9004FRA	Red fox	1986	Bourhy *et al.* (1993)	U43435	
	8903FRA	Red fox	1989	Institut Pasteur	U42606	AF134329
	9147FRA	Red fox	1991	Kissi *et al.* (1995)	U22474	AF047516
	9244FRA	Red fox	1992		U42607	AF134342
	9353FRA	Red fox	1993		U42717	AF134327
	9427FRA	Red fox	1994		U42718	AF134330
	94288FRA	Red fox	1994	Institut Pasteur	U42992	AF134331
	9429FRA	Red fox	1994		U42993	AF134333
	9445FRA	Red fox	1994		U42700	AF134332
	9512FRA	Cow	1995		AF033904	
	9616FRA	Sheep	1996		AF033905	AF134328
Germany	9201ALL	Red fox	1991		U42994	
	9202ALL	Red fox	1991		U42701	AF134338
	9203ALL	Red fox	1991	K. Stöhr	U42995	
	9206ALL	Red fox	1991		U42996	
	9210ALL	Red fox	1991		U42997	
	9212ALL	Red fox	1991	Kissi *et al.* (1995)	U22475	AF134337
	9213ALL	Red fox	1991	K. Stöhr	U42702	
Hungary	9215HON	Human	1991	I. Lontai	U43025	AF134343
	9383HON	Red fox	1993		U42998	AF134340
	9384HON	Red fox	1993	E. Moscàri	U42999	AF134340
	9386HON	Red fox	1993		U43000	
	9387HON	Cat	1993		U43001	
Lithuania	9345LIT	Dog	1992	K. Kulonen	U43002	
Poland	8618POL	Raccoon dog	1985	Kissi *et al.* (1995)	U22840	
	9422POL	Raccoon dog	1993	J. F. Zmudzinski	U43003	
	9424POL	Red fox	1993	J. F. Zmudzinski	U43004	

Table 1 (cont.).

N, nucleoprotein gene; G, glycoprotein gene.

Country	No.	Species	Year	Reference/source	GenBank accession no. N	GenBank accession no. G
	9620POL	Water rat	1994		AF033888	
	9621POL	Badger	1986		AF033870	
	9622POL	Hedgehog	1993		AF033889	
	9623POL	Dog	1991		AF033871	
	9624POL	Cat	1992		AF033872	
	9625POL	Rat	1985		AF033873	
	9626POL	Raccoon dog	1986		AF033874	
	9627POL	Fox	1987		AF033875	
	9634POL	Fox	1987		AF033876	
	9637POL	Fox	1996		AF033890	
	9638POL	Cow	1996	D. Seroka	AF033877	
	9639POL	Roe-deer	1996		AF033878	
	9642POL	Fox	1996		AF033879	
	9645POL	Fox	1993		AF033891	
	9646POL	Fox	1992		AF033894	
	9656POL	Fox	1993		AF033892	
	9671POL	Raccoon dog	1993		AF033895	
	9677POL	Raccoon dog	1994		AF033893	
	9680POL	Marten	1993		AF033896	
	9689POL	Fox	1992		AF033881	
	9691POL	Fox	1993		AF933880	
	96104POL	Fox	1995		AF033897	
	96105POL	Fox	1995		AF033898	
	96106POL	Fox	1995		AF033899	
	96140POL	Raccoon dog	1993		AF033900	
	96152POL	Fox	1995		AF033884	
	96156POL	Raccoon dog	1993		AF033885	
	96159POL	Dog	1995	J. F. Zmudzinski	AF033886	
	96161POL	Cat	1993		AF033901	
	96174POL	Fox	1994		AF033882	
	96176POL	Fox	1994		AF033883	
	96250POL	Fox	1996		AF033887	
	96253POL	Raccoon dog	1996		AF033902	
	96256POL	Fox	1996		AF033903	
Switzerland	9336SUI	Red fox	1992	R. Zanoni	U43006	
Slovak Republic	94250SLK	Cat	1994	S. Svreck	U43007	
Slovenia	9494SLN	Red fox	1994	P. Hostnik	U43008	
	94100SLN	Horse	1994	P. Hostnik	U43009	
Middle East						
Turkey (ME territory)	93100TU	Dog	1993	A. Gürel	U43014	
	93101TU	Dog	1992	A. Gürel	U43015	
Iran	89681IRA	Dog	1986	Kissi *et al.* (1995)	U22482	
	9308IRA	Jackal	1993		U43016	
	9309IRA	Jackal	1993	A. Fayaz	U43017	
	9320IRA	Wolf	1993		U43018	
Israel	9329ISR	Dog	1993	Kissi *et al.* (1995)	U22850	
	9332ISR	Jackal	1993	S. Perl	U43022	
	9333ISR	Dog	1993	S. Perl	U43023	
Arctic						
Canada (Ontario)	9105CAN	Red fox	1990	Kissi *et al.* (1995)	U22655	
	A1	Dog	1993	Nadin-Davis *et al.* (1994)		U03767
	HB	Dog	1992	Nadin-Davis *et al.* (1994)		U03766

Table 1 (cont.).

N, nucleoprotein gene; G, glycoprotein gene.

Country	No.	Species	Year	Reference/source	GenBank accession no. N	GenBank accession no. G
Russia	9141RUS	Arctic fox	1988–90	Kissi *et al.* (1995)	U22656	
Greenland	8684GRO	Arctic fox	1981		U22654	
Africa						
Algeria	9137ALG	Dog	1982		U22643	
	ALGER1	Human		Benmansour *et al.* (1992)		M81060
	ALGER2	Dog				M81058
	ALGER3	Human				M81059
Burkina Fasso	8636HAV	Dog	1986	Kissi *et al.* (1995)	U22486	
Ethiopia	8807ETH	Hyena	1987		U22637	
Guinea	8660GUI	Dog	1986		U22487	
Rep. South Africa	8721AFS	Human	1981		U22633	
Laboratory strains	PV			Tordo *et al.* (1986)		
	AVO1			Poch *et al.* (1988)		

could therefore be an important factor in host adaptation (Kissi *et al.*, 1995). The G gene encodes an external protein important in pathogenicity (Dietzschold *et al.*, 1983) and which reacts with cellular receptors of rabies virus, and so may also be important in determining host range (Tuffereau *et al.*, 1998; Thoulouze *et al.*, 1998). To this end, 245 isolates of rabies virus, stemming from a range of mammalian hosts and a variety of geographical locations, particularly within Europe, were analysed either by sequencing or by RFLP.

Methods

■ **Extraction of RNA, PCR and sequencing.** Ninety-five isolates, the original hosts and geographical sources of which are given in Table 1, were chosen as representatives of the spatial and temporal diversity of rabies virus in Europe and to a lesser extent the Middle East. Brains were obtained from naturally infected animals or after a limited number of passages through suckling mice. RNA extraction and cDNA synthesis were performed as previously described (Sacramento *et al.*, 1991). PCR amplification and sequencing were performed according to Amengual *et al.* (1997) with primer sets N7 (nucleotides 55–73) and N8 (nucleotides 1584–1568), GH3 (nucleotides 3891–3908) and GH4 (nucleotides 4621–4602), and G (nucleotides 4665–4687) and L (nucleotides 5520–5539) for the N, G and G–L genomic regions, respectively (positions are described relative to the PV genome; Tordo *et al.*, 1986). These data have been submitted to GenBank and assigned accession numbers (Table 1). Also included in the analysis were 15 previously determined sequences representing laboratory strains and isolates from the Arctic region and Africa (Table 1).

In addition to the sequence data, a further 135 European isolates were amplified by PCR and then analysed by RFLP. This analysis was based on 400 bp of the N gene using primers N53 (5′ GGATGCCGACAAG-ATTGTAT 3′, nucleotides 73–92 of the PV sequence; Tordo *et al.*, 1986) and N55 (5′ CTAAAGACGCATGTTCAGAG 3′, nucleotides 491–472 of the PV sequence; Tordo *et al.*, 1986). The details concerning those isolates typed by RFLP only are available from the authors upon request.

■ **Analysis of isolates by restriction fragment length polymorphism.** One μl of the amplified products was digested by selected restriction endonucleases and run on a 2% agarose gel with ethidium bromide as described previously (Bourhy *et al.*, 1992). On the basis of the alignment of the nucleotide sequences described in this study and by using the MAPSORT program implemented in the GCG package (Version 8.1-UNIX, program manual for the Wisconsin Package, 1995), four restriction endonucleases (*Bsa*BI, *Hin*dIII, *Mbo*II and *Nla*IV) were selected for their ability to differentiate the European isolates.

■ **Sequence analysis**. After the removal of identical sequences, 33 complete N gene sequences (1350 bp), 29 partial G gene sequences (690 bp) and 85 partial N gene sequences (400 bp) were available for analysis. Multiple sequence alignments of these data were generated with the CLUSTALW program (Thompson *et al.*, 1994). For 19 isolates, both N and G gene sequences were available and so were concatenated into a combined alignment of 2040 bp.

Phylogenetic trees were constructed using the maximum likelihood (ML) method available in the 4.0d65 test version of PAUP* kindly provided by David L. Swofford. The HKY85 model of nucleotide substitution was used in all cases, with the transition/transversion (Ts/Tv) ratio and α shape parameter of a gamma distribution (with eight categories) of rate variation among sites estimated from the empirical data. The values of these parameters for each data set are given in Table 2. To gauge how well each node on the trees was supported, a bootstrap analysis was undertaken (1000 replications), although computational constraints meant that this was performed on neighbour-joining trees reconstructed under the ML substitution model. Monte Carlo simulation (the parametric bootstrap) was then used to determine whether trees estimated on different genes had significantly different topologies, with replicate ML trees generated using the Seq-Gen program (Rambaut & Grassly, 1997).

Table 2. Data parameters of the maximum likelihood phylogenetic analysis of rabies viruses

Gene	Sequence length (bp)	Transition/ transversion ratio	α*	ln L†
N	1350	5·682	0·314	−6946·32351
N	400	4·407	0·374	−2591·94597‡
G	690	4·296	0·466	−2639·25011§
G+N	2040	5·437	0·324	−6551·86950

* Alpha parameter of rate variation among sites.
† Log likelihood.
‡ 704 trees of equal likelihood found.
§ 135 trees of equal likelihood found.

The branching structure of the ML trees, in particular the relative rates of cladogenesis, were analysed using the End-Epi package (Rambaut *et al*., 1997), while the parsimony algorithm within the MacClade program (Maddison & Maddison, 1992) was used to reconstruct the unambiguous amino acid changes along each branch of these trees so that substitutions specific to different groups of viruses could be identified. Finally, the number of nucleotide substitutions per synonymous (d_S) and non-synonymous (d_N) site were estimated using the method of Nei & Gojobori (1986), implemented in the MEGA sequence analysis package (Kumar *et al*., 1993) and plotted for individual codons using the SNAP program (available at http://hiv-web.lanl.gov/SNAP/WEBSNAP/SNAP.html).

Results

Phylogenetic analysis of sequence data

The entire N-coding sequence (1350 bp) was studied from 33 virus isolates, 22 of which were of European origin. The topology of the ML tree of these data reflects both the host and the geographical location of virus isolates: viruses from the same host tend to cluster together, as do viruses from the same region (Fig. 1*a*). A group of viruses isolated from red foxes (*Vulpes vulpes*) in Europe receives good bootstrap support (78%), as does a cluster representing viruses isolated from the Arctic fox (*Alopex lagopus*) (100% bootstrap support) and three viruses obtained from raccoon dogs (*Nyctereutes procyonoides*) (89% bootstrap support), although this latter group falls within the larger red fox cluster. A number of smaller groups can be recognized within the red fox clade, all of which correspond to viruses isolated from particular geographical locations. Specifically, fox viruses from eastern Europe (EE) tend to group together, as do those from western Europe (WE) and central Europe (CE), the latter of which was previously described by Stöhr *et al*. (1992) using monoclonal antibodies. Significantly, the raccoon dog viruses are also found within a particular geographical area, north-eastern Europe (NEE), although their precise phylogenetic relationship to the fox strains is uncertain, as is indicated by low bootstrap support for the critical node (shaded in Fig. 1*a*). Within the WE group, one virus (86111YOU) was in fact collected from Bosnia, suggesting that this clade is dispersed over a wider geographical area, and although strain 9213ALL (from Germany) clusters with the CE group, it is more similar to the WE group at the amino acid level, implying that it represents an early offshoot of the CE viruses.

A number of isolates depict more complex associations between host and phylogeny. In particular, the viruses isolated from dogs and jackals fall into two groups – those collected in Africa, and a single strain isolated from the Middle East (8681IRA), which falls closer to the European fox viruses, suggesting that it might represent an ancestral population. Of equal note are two virus strains isolated from eastern Europe, 86107YOU and 9215HON, isolated from red foxes and humans, respectively, and which seem to occupy an 'intermediate' position between the dog and fox viruses. A third virus, 8658YOU, isolated from cattle in eastern Europe is more divergent still.

Because of the strong patterning by geographical location and host species seen in the complete N gene phylogeny, we decided to examine a much larger number of virus isolates using a 400 bp region from the amino terminus of the N gene, previously identified as one of its most variable regions (Kissi *et al*., 1995). This analysis focused on 85 unique isolates from the Arctic, Africa, the Middle East and Europe (including the 22 analysed previously). The ML tree of these data is shown in Fig. 1(*b*) and is congruent to that obtained from the complete N gene in that the geographically distinct clusters of rabies virus in Europe are evident, although often with weaker bootstrap support. One conspicuous difference is that the CE group appears to be derived from the WE group on a long branch. However, the log likelihood of this tree is only marginally better (−2591·94597 versus −2596·47731) than one in which the position of the WE and CE groups have been rearranged to give the ML topology seen in the analysis of the complete N gene.

Of more importance is that the NEE group (95% bootstrap support) is now found to cover a wider geographical area, including Poland, Estonia, Lithuania and Finland, and includes viruses isolated from both red foxes and raccoon dogs, showing that both species are effective reservoirs for this variant of rabies virus. Furthermore, isolates 86107YOU and 9215HON are both placed close to dog rabies viruses (although with weak bootstrap support), suggesting that they may represent early cross-species transmissions from the dog viruses that circulated in Europe early this century, while the 8658YOU cattle strain remains divergent. Finally, the dog and jackal isolates again form two groups, with those from the Middle East more closely related to the European fox viruses than those from Africa.

To determine whether similar evolutionary patterns are found in other genes of rabies virus, we performed a phylogenetic analysis on a 690 bp region encoding the central

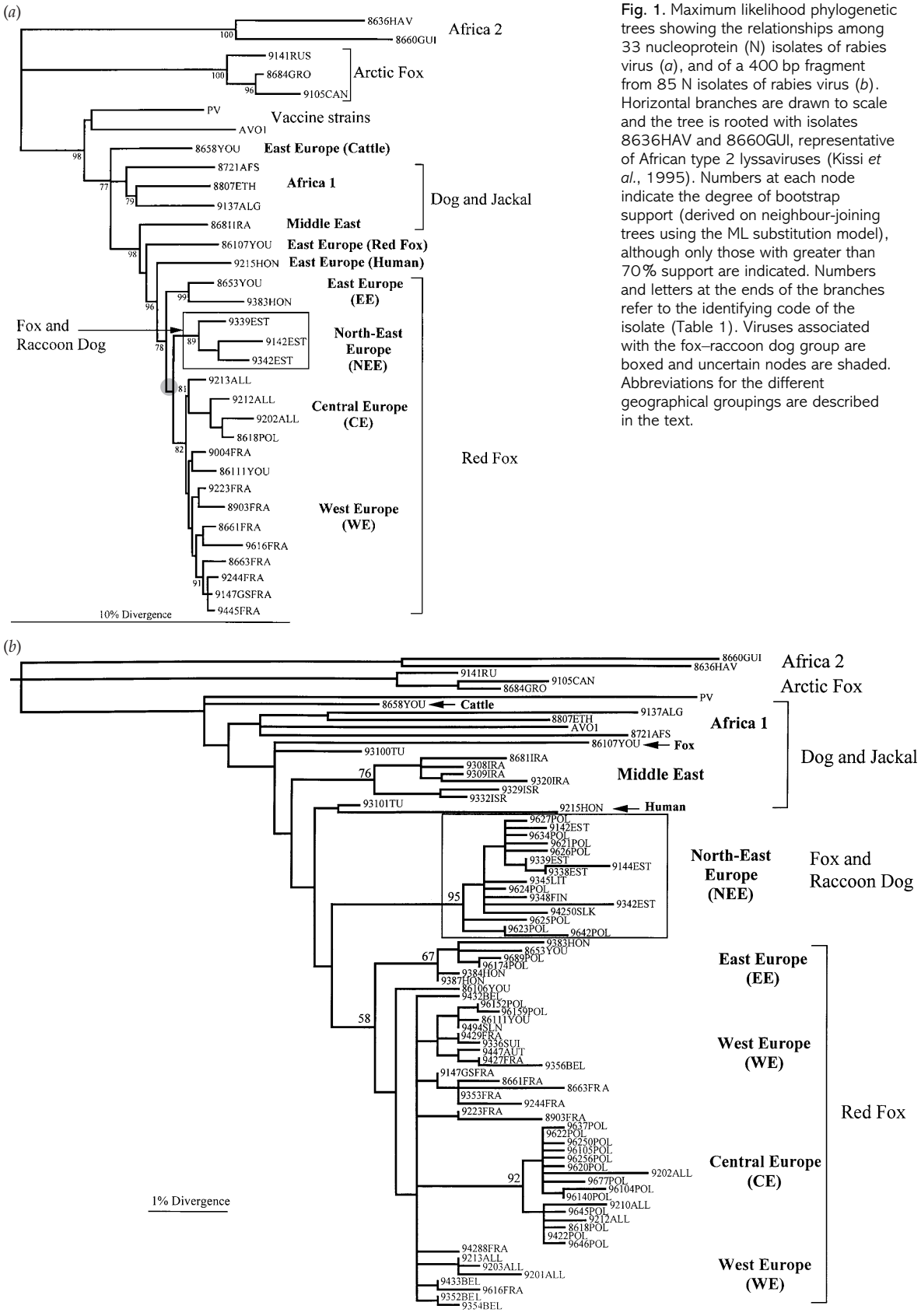

Fig. 1. Maximum likelihood phylogenetic trees showing the relationships among 33 nucleoprotein (N) isolates of rabies virus (*a*), and of a 400 bp fragment from 85 N isolates of rabies virus (*b*). Horizontal branches are drawn to scale and the tree is rooted with isolates 8636HAV and 8660GUI, representative of African type 2 lyssaviruses (Kissi *et al.*, 1995). Numbers at each node indicate the degree of bootstrap support (derived on neighbour-joining trees using the ML substitution model), although only those with greater than 70% support are indicated. Numbers and letters at the ends of the branches refer to the identifying code of the isolate (Table 1). Viruses associated with the fox–raccoon dog group are boxed and uncertain nodes are shaded. Abbreviations for the different geographical groupings are described in the text.

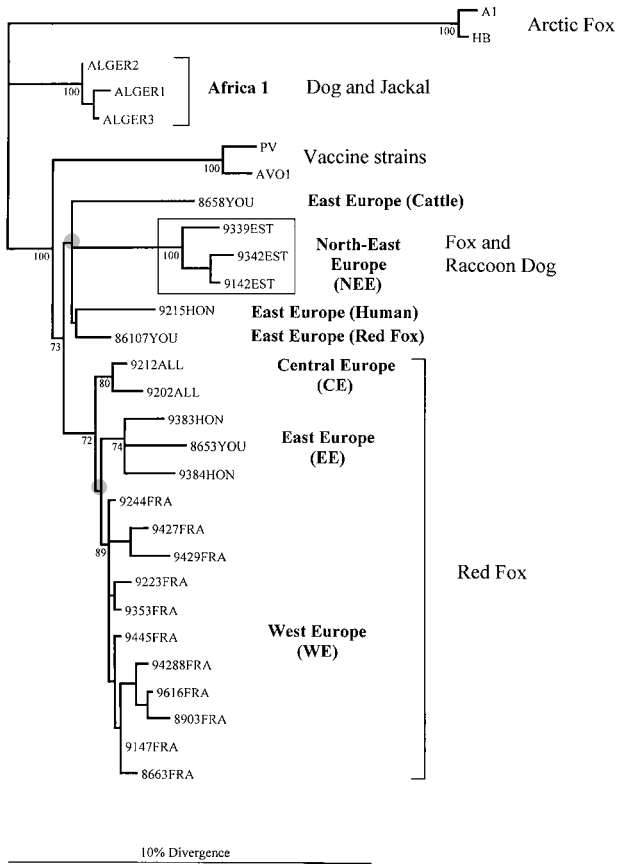

Fig. 2. Maximum likelihood phylogenetic tree of 22 glycoprotein (G) isolates of rabies virus. Horizontal branches are drawn to scale and the tree is rooted with isolates ONT1 and ONT2 obtained from the Arctic fox. Values for the degree of bootstrap support (> 70%) are shown. Viruses associated with the fox–raccoon group are boxed and uncertain nodes are shaded. Abbreviations as in Fig. 1.

part of the ectodomain of the glycoprotein, another region which exhibits a high degree of sequence variation (Tordo *et al.*, 1993). For this purpose, 29 G gene sequences were determined, 22 of which were isolated in Europe with three from Africa (Benmansour *et al.*, 1992). The results of this analysis are presented in Fig. 2. Although many of the phylogenetic relationships depicted are the same as those seen in the N gene tree – that is, there is a general association by host and place of isolation – the EE strains are no longer the most divergent set of fox viruses, instead falling closer to the WE strains, with the fox–raccoon dog NEE strains now more divergent. Whilst some of these nodes are well supported, others are more ambiguous, including those where the G gene phylogeny differs from the N gene phylogeny, such as the divergent position of the NEE isolates (nodes shaded in Fig. 2). The phylogenetic relationships within each group of viruses were also uncertain in places, and were the main reason why 135 trees of equal likelihood were reconstructed on these data.

To assess whether the incongruence between the N and G trees was simply due to a lack of phylogenetic signal or is more fundamental, we undertook a detailed analysis of the 19 virus isolates of European origin (representing all the geographical groups), for which both N and G gene sequences were available. The ML trees for these data closely resembled those constructed previously, with some differences in branching order among the groups of fox viruses (trees not shown, available on request). To determine whether these two trees differed significantly in topology we used Monte Carlo simulation. This first involved comparing the likelihood of the ML tree for the N gene (−4455·25319) to the likelihood of the ML topology for the G gene, but fitted to the N gene sequence data (−4514·82351). Next, 100 replicate data sets of the same length as the real data were evolved along the G gene topology (i.e. the null hypothesis) according to the same evolutionary model (i.e. base composition, Ts/Tv ratio and α value) as in the real data. ML trees were then reconstructed on each replicate to give a null distribution of likelihood scores (range −4308·78899 to −4980·81761). As the likelihood of the ML tree for the N gene falls within this distribution we can say that the N and G tree topologies differ by no more than might be expected by chance. This analysis was then repeated on the G gene data, with the N gene tree topology now assumed to be the null hypothesis (−2054·88972). As before, the likelihood of the G gene ML tree (−2013·37674) fell well within the null distribution produced by the simulated data (range −1800·32235 to −2284·39825).

As we found no evidence that the N and G gene trees differ in topology, we were able to combine the 19 N and G gene sequences in a single phylogenetic analysis. This resulted in an ML tree containing elements of those constructed on the two genes separately (Fig. 3). Specifically, the N + G tree resembles the G gene phylogeny in clearly placing the NEE isolates as more divergent than the red fox groups, but shares a similar topology with the N gene tree in that the CE strains are closer to the WE strains than are the EE strains. Significantly, many of these important nodes now receive strong bootstrap support (Fig. 3). Although the increase in bootstrap support is influenced by the smaller number of taxa compared, it should be noted that all the major groups were represented in analysis, with the reduction in number mainly due to the loss of sequences within groups. We therefore believe that the phylogenetic relationships among the geographical groups of rabies virus in Europe are best represented by the phylogeny of the combined N and G data sets.

As well as considering their topology, we also analysed the branching structures of our phylogenetic trees as these may reveal more about the processes by which rabies viruses have spread across Europe. All the phylogenies appeared to have a highly asymmetrical 'ladder-like' appearance, with the deepest lineages in the north and east of Europe and the most recent towards the west of the continent. To quantify this we determined the extent to which branches differed in their probability of producing daughter lineages, i.e. that they have different rates of cladogenesis. This was done using the 'relative cladogenesis statistic', *Pk*, which measures the

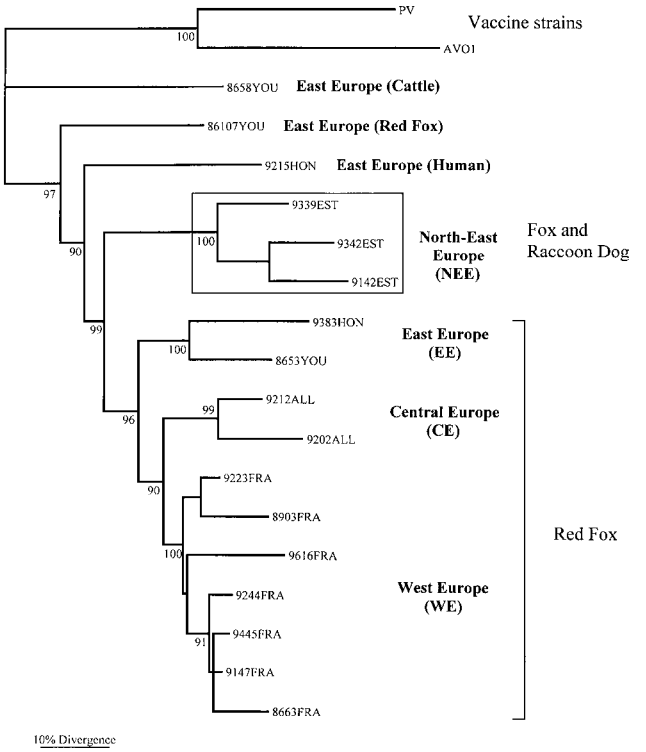

Fig. 3. Maximum likelihood phylogenetic tree of the 19 rabies virus isolates for which both the N and G genes were available. The tree is rooted by the PV and AVO1 vaccine strains and the horizontal branches are drawn to scale. Values for the degree of bootstrap support (> 70%) are shown. Viruses associated with the fox–raccoon group are boxed. Abbreviations as in Fig. 1.

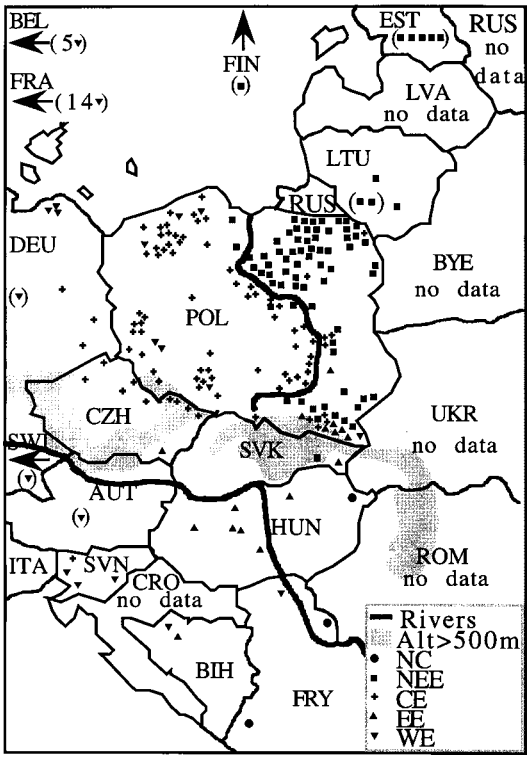

Fig. 4. Geographical distribution of the different phylogenetic groups of rabies virus in central and eastern Europe determined by either gene sequencing or RFLP. The precise sampling location of those isolates in brackets is either not known or not shown on the map. For some specimens it was not possible to indicate the location because of overlap with others. The Vistula and Danube rivers are both denoted by thick lines. The Bohemian and Carpathian mountains (altitude higher than 500 m) are indicated in shadow. Abbreviations of clusters: NC, non-classified, corresponds to isolates 86107YOU, 9215HON and 8658YOU; NEE, north-eastern Europe; CE, central Europe; EE, eastern Europe; WE, western Europe; BEL, Belgium; FRA, France; DEU, Federal Republic of Germany; SWI, Switzerland; ITA, Italy; POL, Poland; CZH, Czech Republic; AUT, Austria; SVN, Slovenia, CRO, Croatia; BIH, Bosnia and Herzegovina; FIN, Finland; RUS, Russia; EST, Estonia, LVA, Latvia; LTU, Lithuania; BYE, Belarus; UKR, Ukraine; SVK, Slovak Republic; HUN, Hungary; ROM, Romania; FRY, Federal Republic of Yugoslavia.

probability that a lineage present at time t in the past, will have k tips (compared with the total number of tips) by time 0 (the present) under a null model of a constant rate of lineage birth and death (Nee *et al.*, 1994; Zanotto *et al.*, 1996).

For the complete N gene tree, 26 tips beginning with the divergence of 8658YOU were found to have produced significantly more daughters than expected under the null model of a uniform rate of cladogenesis ($Pk < 005$). Such a biased branching process was also found in the tree of the partial N gene sequences in which 80 tips, beginning with the divergence of PV and 8658YOU, were linked in an asymmetric fashion ($Pk < 0{\cdot}01$). Similar, although less strong, results were found in the G gene tree, for 22 tips starting with the divergence of the 8658YOU, 9215HON, 86107YOU and NEE clade ($Pk < 0{\cdot}05$), and for the 15 tips in the N + G tree beginning with split of 9215HON ($Pk < 0{\cdot}05$). We therefore conclude that the phylogenetic trees of these isolates are strongly biased in their branching structures.

Geographical distribution of European rabies virus isolates

A further 135 isolates of rabies virus from central-eastern Europe were typed by RFLP according to the geographical (and phylogenetic) groups we describe above. The size of the bands expected in the RFLP profiles of the different phylogenetic clusters were as follows: (i) NEE (*Bsa*BI, 400 bp; *Hin*dIII, 210 and 190 bp; *Nla*IV, 391 or 400 bp); (ii) EE (*Bsa*BI, 400 bp; *Hin*dIII, 400 bp; *Nla*IV, 400 bp; *Mbo*II can be used as a positive control, 271 and 129 bp); (iii) WE (*Bsa*BI, 400 bp; *Hin*dIII, 400 bp; *Nla*IV, 300 and 91 bp); (iv) CE (*Bsa*BI, 283 and 117 bp; *Hin*dIII, 400 bp; *Nla*IV, 300 and 91 bp).

From this analysis (as well as the sequence data), we were able to map the location of the different phylogenetic groups of rabies virus within Europe (Fig. 4). A clear picture of geographical subdivision is revealed with the Vistula (or Wista) river in Poland separating the CE and NEE clusters and, to a lesser extent, the Bohemian and Carpathian mountains reinforced by the Danube river in the Czech Republic, Slovak Republic, Austria and Hungary isolating the EE viruses. To be

Table 3. Animal species involved in the different phylogenetic clusters in Europe

Percentages are shown in parentheses.

Phylogenetic cluster*	Animal species			Total
	Fox	Raccoon dog	Others	
NEE	35 (43·8)	28 (35)	17 (21·2)	80
EE	12 (75)	0	4 (25)	16
WE	31 (75·6)	0	10 (24·4)	41
CE	54 (62·8)	16 (18·6)	16 (18·6)	86
Total	132	44	47	223

* NEE, north-eastern Europe; EE, eastern Europe; WE, western Europe; CE, central Europe.

more specific, the NEE cluster was found in Finland, Estonia, Lithuania, Poland and in the eastern part of the Slovak Republic. In Poland, the NEE group is limited to the eastern side of the Vistula river, with the exception of four isolates found close to the river. In contrast, the CE cluster was isolated mainly in the west and south of Poland (i.e. to the west of the Vistula river), the east of Germany, the Czech Republic and Slovenia. One CE isolate was also found in Poland near the Lithuanian border. The WE cluster was found in a region stretching from France and Belgium to the west and south of Poland. It was also isolated from Switzerland and Austria, and was frequently found in the south of this region, particularly in Slovenia, Bosnia–Herzegovina and the Federal Republic of Yugoslavia. The EE cluster was limited to the south-east of the Czech Republic and Poland, to Bosnia–Herzegovina and to Hungary.

These data also confirmed the host specificity of rabies virus (Table 3). The red fox was the predominant host for the EE, WE and CE groups, with 75%, 75·6% and 62·8% of all specimens analysed stemming from this animal, respectively. The second most important host species was the raccoon dog: 35% of isolates from the NEE group were obtained from raccoon dogs, and if we consider only the 43 samples originating from Estonia, Finland, Lithuania and the administrative subdivisions in northern Poland bordering Russia, Byelorussia and Lithuania (Olsztyn, Bialystok and Suwalki), where the density of raccoon dogs is greatest having been introduced into this region for fur farming between 1927 and 1957 (Nowak & Paradiso, 1983), a predominance of raccoon-dog-associated viruses is apparent (53·5% raccoon dog, 32·5% fox). Since their introduction into NEE, raccoon dogs have gradually dispersed westwards; 18·6% of the CE strains came from this species, although we found no animals infected with viruses from the EE and WE groups (a lack of surveillance of rabies in raccoon dogs could, of course, underestimate the impact of rabies in this species). Similarly, a predominance of dog and jackal isolates ($n = 17$, 70·5%) was found in the samples collected from the Middle East and Africa.

Host adaptation at the molecular level?

Our next task was to determine, for both the N and G genes, the amino acid changes which distinguish each group of viruses in the hope of identifying those which may have facilitated the change in host species and/or geographical location. This was done by reconstructing, using parsimony, the unambiguous amino acid changes along each branch of the ML tree. Surprisingly, perhaps, few amino changes are apparent, suggesting that both proteins are subject to relatively strong selective constraints: 112 amino acid changes were reconstructed on the N gene tree and just 59 on the G gene tree. A single amino acid substitution separated the fox (CE, EE and WE) groups in the N gene, an Asp to Asn (in the case of EE) or to Ala (CE and WE) change at position 101. Another amino acid substitution separated the fox–raccoon dog (NEE) group in the N gene, an Asp to Gly change at position 115 (which has convergently appeared in the Arctic fox group), along with two changes in the G gene, an Ile to Val change at position 357 and a Lys to Arg change at position 361. In contrast, isolates from the Arctic fox are more divergent, being distinguished by six amino acid changes in the N gene and four in the G gene. It is also noteworthy that most amino acid changes are located on the branches leading to each phylogenetic group or to individual isolates, rather than on the internal edges between groups. Such a pattern suggests that the initial spread of the virus through Europe, and to the different species, was achieved with little adjustment to the viral proteins, but that subsequent local evolution (i.e. on the branches leading to each group) has occurred which involved more amino acid changes. This is especially true of isolates 8658YOU, 86107YOU and 9215HON, which have accumulated seven, five and seven amino acid changes, respectively, in the G and N genes.

To determine whether any of these amino acid changes might have been fixed by natural selection, we calculated the mean numbers of synonymous (d_S) and nonsynonymous (d_N) substitutions per site in both the N and G genes (the two African type 2 lyssaviruses were removed from the N gene analysis to make the results more comparable between genes). As expected given the relatively low numbers of amino acid changes, d_S was much greater than d_N in every case, therefore providing no evidence for positive selection (i.e. $d_N > d_S$) at this level, with the N gene (mean $d_N = 0{\cdot}0087 \pm 0{\cdot}0012$, mean $d_S = 0{\cdot}2678 \pm 0{\cdot}0142$; $d_N/d_S = 0{\cdot}032$) apparently under slightly stronger selective constraints than the G gene (mean $d_N = 0{\cdot}0115 \pm 0{\cdot}0020$, mean $d_S = 0{\cdot}2494 \pm 0{\cdot}0204$; $d_N/d_S = 0{\cdot}046$). Although informative, such large-scale pairwise comparisons are unlikely to reveal the affects of natural selection on individual amino acids. To assess this possibility we calculated the mean d_N value for each codon in the N and

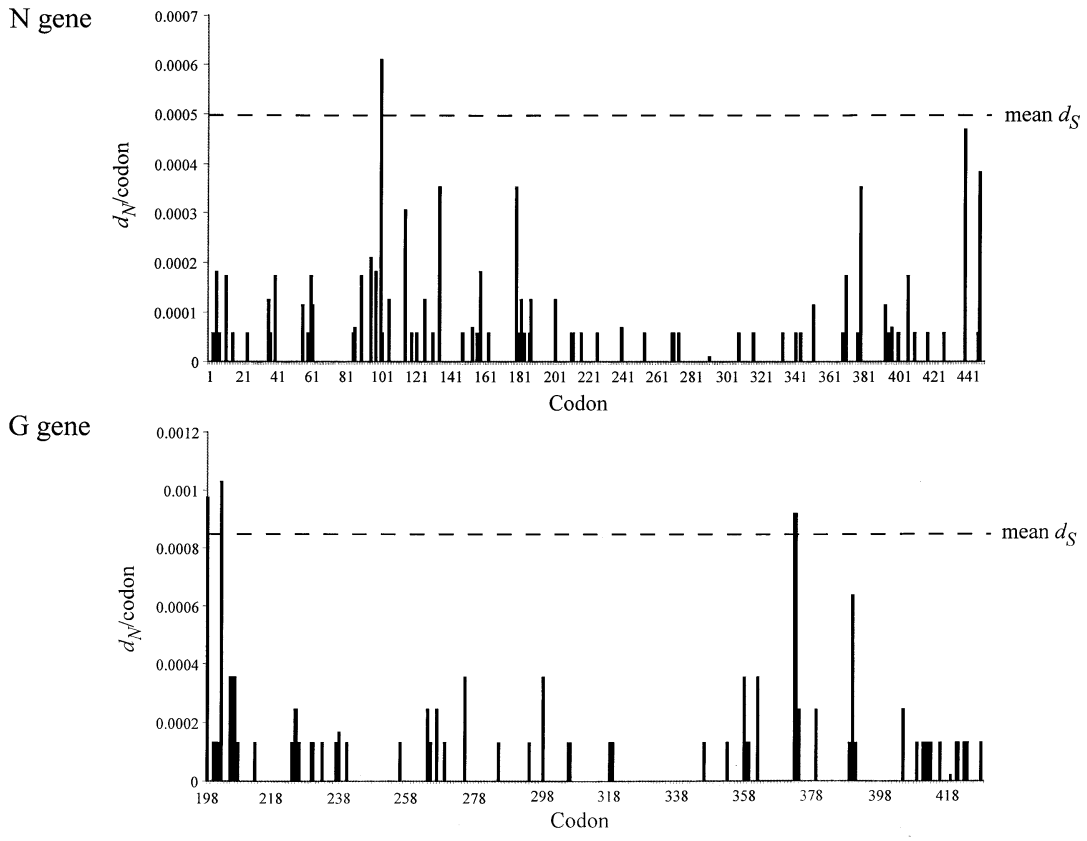

Fig. 5. The mean d_N for each codon (averaged across all pairwise comparisons) in the N and G gene sequences. The mean d_S value across all codons is also shown for both genes (hatched line), with peaks crossing this line indicating those codons with anomalously high rates of nonsynonymous change.

G genes and identified those with higher rates than the mean d_S across all codons, which we assume is a marker of the background (neutral) mutation rate. These results are shown in Fig. 5 and reveal a single codon in the N gene (position 101) and three in the G gene (positions 1, 5 and 175) with elevated rates of nonsynonymous substitution. It is intriguing that N gene codon 101 falls into this category because, as noted above, it contains amino acid substitutions which distinguish the red fox viruses. While is it is possible that selectively advantageous substitutions have been fixed at these sites, the small numbers of changes involved, and hence the likelihood of sampling artefacts, mean that these results should be interpreted with caution.

Discussion

In order to understand the evolution of rabies virus in Europe, particularly the adaptation to new host species, 245 isolates of the virus were either sequenced or typed by RFLP. The European continent represents an ideal opportunity to undertake a study of this kind because several strains of rabies virus co-circulate within this region and infect a range of mammalian species. Our analyses of these data uncovered two distinct, but clearly related, patterns: viruses from the same geographical area tend to group together as do isolates taken from the same host species, although less strongly.

Not only were the phylogenies we obtained strongly ordered by geography, but they also had a strong 'ladder-like' structure, with the deepest branches belonging to viruses collected in the north and east of Europe, and the most recent branches belonging to viruses collected further west and south. From this we conclude that our phylogenetic analysis documents the gradual dispersal of rabies virus from the north-east to the south-west across Europe, as has been previously suggested based on epidemiological data (Blancou *et al.*, 1991). A similar east to west movement was previously documented in tick-borne flaviviruses, a process which perhaps took around 2000 years to unfold (Zanotto *et al.*, 1995). The spread of rabies viruses that we describe clearly occurred much more recently than this, as is evident from the epidemiological records of rabies cases this century (Zunker, 1954; Seroka, 1968; Atanasiu, 1968).

Despite the fluidity of rabies virus transmission in Europe, it is equally clear that its spread can be contained to some extent by natural physical barriers such as the Vistula river in Poland, most likely by restricting the movement of infected hosts. In this respect, our sequencing and RFLP data show that virus isolates in central-eastern Europe have a strong geo-

graphical clustering, suggesting that there is some degree of genetic isolation. In these circumstances it would seem pertinent to continue surveillance of the different populations of fox rabies virus within Europe as we might expect the continued adaptation to local mammalian fauna, as is highlighted by the species jump to raccoon dogs (see below), and to monitor whether the NEE cluster will eventually disperse further westwards as raccoon dogs have themselves done.

During the westwards and southwards movement of rabies virus across Europe two changes of host species took place. The first occurred when the virus initially jumped from dogs to foxes, although it is unclear from our analysis exactly where this took place. However, as the deepest branches of the fox virus tree are found in eastern Europe, we suggest that a species jump in this region seems the most reasonable interpretation of the data. The second change in host took place in north-eastern Europe when rabies viruses colonized raccoon dogs. From our phylogenetic analysis it is not possible to determine precisely whether the source of the virus in raccoon dogs was infected foxes, or whether the virus jumped directly from dogs and was then passed to the local fox population. Nor is it clear what ecological pressure (if any) precipitated this host switch, although it is apparent that raccoon dogs are a common enough wildlife species to be able to sustain such an infection. That the NEE strain is found in the region where the population of raccoon dogs is greatest suggests that the density of susceptible hosts, as well the close proximity of a donor species, are major ecological factors in the establishment of rabies virus in a new host species and also that perhaps the NEE strain is preferentially adapted to this species, although this is clearly an issue that needs to be explored further.

Finally, the status of three virus strains collected from humans, red fox and cattle in eastern Europe and which represent divergent lineages on the trees is unclear, although it seems most likely that they were derived from dog rabies viruses. It is therefore possible that they represent spill-over infections with viruses belonging to lineages which were established early this century when dog viruses were more commonly found in EE and before the red fox was established as the major reservoir of rabies infection. Such a spill-over of *Canidae*-associated viruses into wildlife species is frequently observed (Nel *et al*., 1997).

Given the existence of geographically distinct variants of rabies virus, the next question to address is whether functionally important amino acid changes have accumulated between them, particularly those that might have enabled adaptation to different host species. Strikingly, both the G and N proteins are generally conserved with few amino acid replacements accumulating among the strains studied. In particular, very few amino acid changes were found to accompany the change in transmission from dogs to foxes or raccoon dogs, although it is also possible that key mutations reside in other genes. In a similar vein, an analysis of the relative numbers of synonymous and nonsynonymous substitutions revealed that both the G and N genes are under relatively strong selective constraints, although some codons have experienced much higher rates of nonsynonymous change than others (and higher than the background silent substitution rate), which may signify localized positive selection pressure (Kissi *et al*., 1999). Although the significance of these changes is unclear, they merit careful investigation at the structural–functional level as it is possible that they are of phenotypic importance.

Considering that all strains of rabies virus in Europe are not equally able to infect dogs, foxes and raccoon dogs (Blancou *et al*., 1983; Blancou & Aubert, 1997), our study confirms previous suggestions that the infection of new host species in nature could be caused by a small number of genetic changes in rabies virus, involving just a few amino acid replacements (Tuffereau *et al*., 1989; Kissi *et al*., 1999). However, this does not exclude the possibility that some of the phylogenetic groups we describe reflect instead the ecological separation of individuals belonging to the same or different host species, as has been observed in other viruses (Nickels & Hunt, 1994; Parrish, 1994). Such a conclusion distinguishes the N genes of rabies virus from the capsid genes of some other negative-strand RNA viruses, such as influenza A virus (for a review see Webster *et al*., 1992), and positive-strand RNA viruses like coxsackieviruses and alphaviruses (Villaverde *et al*., 1991; Ishiko *et al*., 1992; Domingo & Holland, 1994; Weaver *et al*., 1994), which often evolve at greatly elevated rates at amino acid changing sites, presumably because of strong positive selection.

To conclude, given that we provide strong evidence that local genetic differentiation is taking place within European rabies viruses, we urge that further studies of virus variation be undertaken so that we may come to a greater understanding of the mechanisms controlling adaptation to new host species, information that is crucial to the greater goal of eliminating terrestrial rabies from Europe.

We are deeply indebted to the many generous people who have provided us with rabies samples. These include J. Barrat, F. Costy, A. Fayaz, A. Gürel, P. Hostnik, A. King, I. Lontai, O. Matouch, E. Moscàri, T. Mustafa, D. Peharpe, S. Perl, M. Petrovic, W. Schuller, K. Stöhr, S. Svreck and R. Zanoni. We also thank Dr S. Messenger for valuable discussions and two anonymous referees for constructive comments.

References

Amengual, B., Whitby, J. E., King, A., Serra Cobo, J. & Bourhy, H. (1997). Evolution of European bat lyssaviruses. *Journal of General Virology* **78**, 2319–2328.

Atanasiu, P., Gamet, A., Gravière, P., Le Guilloux, M., Guillon, J. C. & Vallée, A. (1968). Réapparition de la rage en France. Premier cas chez un renard dans la Moselle. *Bulletin de l'Académie Vétérinaire* **XLI**, 161–163.

Barbier, A. (1929). *Les Sources de la Virulence Rabique. Histoire d'une Epizootie de Rage sur le Renard et le Blaireau dans la Région Dijonnaise*, pp. 253. Dijon: Imprimerie Bernigaud et Privat.

Benmansour, A., Brahimi, M., Tuffereau, C., Coulon, P., Lafay, F. &

Flamand, A. (1992). Rapid sequence evolution of street rabies glycoprotein is related to the highly heterogeneous nature of the viral population. *Virology* **187**, 33–45.

Blancou, J. & Aubert, M. F. A. (1997). Transmission du virus de la rage: importance de la barrière d'espèce. *Bulletin de l'Académie Nationale de Médecine* **181**, 301–312.

Blancou, J., Aubert, M. F. A. & Soulebot, J. P. (1983). Différences dans le pouvoir pathogène de souches de virus rabique adaptées au renard ou au chien. *Annales de l'Institut Pasteur Virology* **134E**, 523–531.

Blancou, J., Aubert, M. F. A. & Artois, M. (1991). Fox rabies. In *The Natural History of Rabies*, pp. 257–290. Edited by G. M. Baer. Boca Raton: CRC Press.

Bourhy, H., Kissi, B., Lafon, M., Sacramento, D. & Tordo, N. (1992). Antigenic and molecular characterization of bat rabies virus in Europe. *Journal of Clinical Microbiology* **30**, 2419–2426.

Bourhy, H., Kissi, B. & Tordo, N. (1993). Molecular diversity of the lyssavirus genus. *Virology* **194**, 70–81.

Dietzschold, B., Wunner, W. H., Wiktor, T. J., Lopes, A. D., Lafon, M., Smith, C. L. & Koprowski, H. (1983). Characterization of an antigenic determinant of the glycoprotein that correlates with pathogenicity of rabies virus. *Proceedings of the National Academy of Sciences, USA* **80**, 70–74.

Domingo, E. & Holland, J. J. (1994). Mutation rates and rapid evolution of RNA viruses. In *The Evolutionary Biology of Viruses*, pp. 161–184. Edited by S. S. Morse. New York: Raven Press.

Gould, A. R., Hyatt, A. D., Lunt, R., Kattenbelt, J. A., Hengstberger, S. & Blacksell, S. D. (1998). Characterization of a novel lyssavirus isolated from *Pteropid* bats in Australia. *Virus Research* **54**, 165–187.

Ishiko, H., Takeda, N., Miyamura, K., Kato, N., Tanimura, M., Lin, K.-H., Yin-Murphy, M., Tam, J. S., Mu, G.-F. & Yamazaki, S. (1992). Phylogenetic analysis of a coxsackievirus A 24 variant: the most recent worldwide pandemic was caused by progenies of a virus prevalent around 1981. *Virology* **187**, 748–759.

Jaujou, M. (1949). L'infection rabique en Corse au cours de l'année 1946. *Académie Nationale de Médecine* **132**, 128–130.

Kissi, B., Tordo, N. & Bourhy, H. (1995). Genetic polymorphism in the rabies virus nucleoprotein gene. *Virology* **209**, 526–537.

Kissi, B., Badrane, H., Audry, L., Lavenu, A., Tordo, N., Brahimi, M. & Bourhy, H. (1999). Dynamics of rabies virus quasispecies during serial passages in heterologous hosts. *Journal of General Virology* **80**, 2041–2050.

Kumar, S., Tamura, K. & Nei, M. (1993). MEGA: molecular evolutionary genetic analysis, version 1.0. The Pennsylvania State University, University Park, PA 16802, USA.

Maddison, W. P. & Maddison, D. R. (1992). MacClade: analysis of phylogeny and character evolution, version 3.0. Sinauer Associates: Sunderland, MA, USA.

Nadin-Davis, S. A., Casey, G. A. & Wandeler, A. I. (1994). A molecular epidemiological study of rabies virus in central Ontario and western Quebec. *Journal of General Virology* **75**, 2575–2583.

Nee, S., May, R. M. & Harvey, P. H. (1994). The reconstructed evolutionary process. *Philosophical Transactions of the Royal Society of London Series B* **344**, 305–311.

Nei, M. & Gojobori, T. (1986). Simple methods for estimating the numbers of synonymous and non synonymous nucleotide substitutions. *Molecular Biology and Evolution* **3**, 418–426.

Nel, L., Jacobs, J., Jaftha, J. & Courteney, M. (1997). Natural spillover of a distinctly Canidae-associated biotype of rabies virus into an expanded wildlife host range in Southern Africa. *Virus Genes* **15**, 79–82.

Nickels, M. S. & Hunt, D. M. (1994). Identification of an amino acid change that affects N protein function in vesicular stomatitis virus. *Journal of General Virology* **75**, 3591–3595.

Nowak, R. M. & Paradiso, J. L. (1983). *Walker's Mammals of the World*, vol. II, 4th edn. Baltimore: The Johns Hopkins University Press.

Parrish, C. R. (1994). The emergence and evolution of canine parvovirus – an example of recent host range mutation. *Virology* **5**, 121–132.

Petrovic, M. (1987). Urban and sylvatic rabies in Yugoslavia. *Rabies Bulletin Europe* **4**, 16–18.

Poch, O., Tordo, N. & Keith, G. (1988). Sequence of the 3386 3′ nucleotides of the genome of the AVO1 strain rabies virus: structural similarities in the protein regions involved in transcription. *Biochimie* **70**, 1018–1029.

Rambaut, A. & Grassly, N. C. (1997). Seq-Gen: an application for the Monte Carlo simulation of DNA sequence evolution along phylogenetic trees. *CABIOS* **13**, 235–238.

Rambaut, A., Harvey, P. H. & Nee, S. (1997). End-Epi: an application for inferring phylogenetic and population dynamical processes from molecular sequences. *CABIOS* **13**, 303–306.

Rupprecht, C. E. & Smith, J. S. (1994). Raccoon rabies: the re-emergence of an epizootic in a densely populated area. *Seminars in Virology* **5**, 155–164.

Sacramento, D., Bourhy, H. & Tordo, N. (1991). PCR technique as an alternative method for diagnosis and molecular epidemiology of rabies virus. *Molecular and Cellular Probes* **6**, 229–240.

Sacramento, D., Badrane, H., Bourhy, H. & Tordo, N. (1992). Molecular epidemiology of rabies in France: comparison with vaccine strains. *Journal of General Virology* **73**, 1149–1158.

Seroka, D. (1968). The distribution of stationary foci of rabies in wild animals in Poland. *Epidemiological Review* (*English Translation of Przeglad Epidemiologiczny*) **22**, 66–75.

Smith, J. S., Orciari, L. A., Yager, P., Seidel, H. D. & Warner, C. K. (1992). Epidemiologic and historical relationships among 87 rabies virus isolates as determined by limited sequence analysis. *Journal of Infectious Diseases* **166**, 296–307.

Smith, J. S., Orciari, L. A. & Yager, P. (1995). Molecular epidemiology of rabies in the United States. *Seminars in Virology* **6**, 387–400.

Steck, F. & Wandeler, A. (1980). The epidemiology of fox rabies in Europe. *Epidemiologic Reviews* **2**, 71–96.

Stöhr, K., Stöhr, P. & Karge, E. P. (1992). Isolierung atypischer Tollwutfeld-Viren in Ostdeutschland. *Tierärztliche Umschau* **47**, 820–824.

Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). CLUSTALW: improving the sensitivity of progressive multiple alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research* **22**, 4673–4680.

Thoulouze, M.-I., Lafage, M., Schachner, M., Hartmann, U., Cremer, H. & Lafon, M. (1998). The neural cell adhesion molecule is a receptor for rabies virus. *Journal of Virology* **72**, 7181–7190.

Tordo, N., Poch, O., Ermine, A., Keith, G. & Rougeon, F. (1986). Walking along the rabies genome: is the large G–L intergenic region a remnant gene? *Proceedings of the National Academy of Sciences, USA* **83**, 3914–3918.

Tordo, N., Badrane, H., Bourhy, H. & Sacramento, D. (1993). Molecular epidemiology of lyssaviruses: focus on the glycoprotein and pseudogenes. *Onderstepoort Journal of Veterinary Research* **60**, 315–323.

Tuffereau, C., Leblois, H., Bénéjean, J., Coulon, P., Lafay, F. & Flamand, A. (1989). Arginine or lysine in position 333 of ERA and CVS glycoprotein is necessary for rabies virulence in adult mice. *Virology* **172**, 206–212.

Tuffereau, C., Bénéjean, J., Roque Alfonso, A. M., Flamand, A. & Fishman, M. C. (1998). Neuronal cell surface molecules mediate specific binding to rabies virus glycoprotein expressed by a recombinant baculovirus on the surfaces of lepidopteran cells. *Journal of Virology* **72**, 1085–1091.

Villaverde, A., Martinez, M. A., Sobrino, F., Dopazo, J., Moya, A. & Domingo, E. (1991). Fixation of mutations at the VP1 gene of foot-and-mouth disease virus. Can quasispecies define a transient molecular clock? *Gene* **103**, 147–153.

Weaver, S. C., Hagenbaugh, A., Bellew, A., Gousset, L., Mallampalli, V., Holland, J. J. & Scott, T. W. (1994). Evolution of alphaviruses in the eastern equine encephalomyelitis complex. *Journal of Virology* **68**, 158–169.

Webster, R. G., Bean, W. J., Gorman, O. T., Chambers, T. M. & Kawaoka, Y. (1992). Evolution and ecology of influenza A viruses. *Microbiological Reviews* **56**, 152–179.

Zanotto, P. M. de A., Gao, G. F., Gritsun, T. S., Marin, M. S., Jiang, W. R., Venugopal, K., Reid, H. W. & Gould, E. A. (1995). An arbovirus cline across the northern hemisphere. *Virology* **210**, 152–159.

Zanotto, P. M. de A., Gould, E. A., Gao, G. F., Harvey, P. H. & Holmes, E. C. (1996). Population dynamics of flaviviruses revealed by molecular phylogenies. *Proceedings of the National Academy of Sciences, USA* **93**, 548–553.

Zeeti, R. & Rosati, T. (1966). Informations relatives à la situation de la rage en Italie et les mesures employées pour la combattre. *Bulletin de l'Office International des Epizooties* **65**, 37–39.

Zunker, M. (1954). L'importance des renards dans la propagation de la rage en Allemagne. *Bulletin de l'Office International des Epizooties* **354**, 1–11.

Received 18 March 1999; Accepted 16 June 1999