Journal of General Virology (2003), **84**, 485–496 DOI 10.1099/vir.0.18652-0

A molecular epidemiological study of Australian bat lyssavirus

Kimberley J. Guyatt,[1,2] Jimmy Twin,[1] Patricia Davis,[3,4] Edward C. Holmes,[3] Greg A. Smith,[5] Ina L. Smith,[5] John S. Mackenzie[1] and Peter L. Young[2]

[1]Department of Microbiology and Parasitology, University of Queensland, St Lucia, Queensland, Australia

[2]Department of Primary Industries, Queensland Agricultural Biotechnology Centre, Level 4, Gehrmann Laboratories, University of Queensland, St Lucia, Queensland, Australia

[3]Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, UK

[4]Unité de la Rage, Institut Pasteur, Paris Cedex 15, France

[5]Public Health Virology, Queensland Health Scientific Services, Coopers Plains, Queensland, Australia

Correspondence
Kimberley Guyatt (at Queensland Agricultural Biotechnology Centre)
k.guyatt@mailbox.uq.edu.au

The genetic diversity of Australian bat lyssavirus (ABL) was investigated by comparing 24 ABL isolate glycoprotein (G) gene nucleotide sequences with those of 37 lyssaviruses representing *Lyssavirus* genotypes 1–6. Phylogenetic analyses indicated that ABL forms a monophyletic group separate from other lyssaviruses. This group differentiates into two clades: one associated with *Pteropus* (flying fox) species, the other with the insectivorous bat *Saccolaimus flaviventris*. Calculation of percentage nucleotide identities between isolates of the two clades revealed up to 18·7 % nucleotide sequence divergence between the two ABL variants. These observations suggest that ABL is a separate lyssavirus species with a similar epidemiology to chiropteran rabies virus (RV), where two distinct ABL variants co-exist in Australia in bat species with dissimilar ecology. Analyses of selection pressures in ABL G gene sequences provided some evidence of weak positive selection within the endodomain at amino acids 499 and 501, although in general the dominant evolutionary process observed was purifying selection. This intimates that, in nature, isolates of ABL, like those of RV, are subject to relatively strong selective constraints, suggesting a stability of host species, cell tropisms and ecological conditions.

Received 12 June 2002
Accepted 1 October 2002

INTRODUCTION

Prior to May 1996, Australia had been considered free of rabies and rabies-like viruses. This status was challenged, however, when a virus with a high degree of antigenic and genetic similarity to rabies virus (RV), subsequently named Australian bat lyssavirus (ABL), was isolated from a juvenile black flying fox (Fraser *et al.*, 1996). Since that time, all four common species of Australian flying fox (*Pteropus alecto*, *P. poliocephalus*, *P. scapulatus* and *P. conspicullatus*) and an insectivorous bat species, *Saccolaimus flaviventris* (the yellow-bellied sheath-tailed bat), have been found to be reservoirs of ABL in Australia (Field *et al.*, 1999). ABL has been detected in flying foxes and insectivorous bats collected in the states of Queensland, New South Wales, Victoria and the Northern Territory (Hooper *et al.*, 1997), which broadly correlates with the geographical range of these animals in mainland Australia (Tidemann *et al.*, 1997).

A recent study has also reported the presence of neutralizing antibodies against ABL in six bat species in the Philippines, although none of the seropositive animals was found to have an active ABL infection (Arguin *et al.*, 2002).

Two human deaths in Queensland have been attributed to ABL infection, where ABL is thought to have been transmitted to the susceptible hosts in saliva following bites from infected bat species (Allworth *et al.*, 1996; Hanna *et al.*, 2000). Case studies of both flying fox and human ABL infections have indicated that ABL follows a course of clinical disease similar to that of rabies, presenting as a severe non-suppurative encephalitis with symptoms of aggression, anxiety, hypersalivation and agitation (Allworth *et al.*, 1996; Fraser *et al.*, 1996; Hanna *et al.*, 2000; McColl *et al.*, 2000). These ABL spillover events have been a cause for concern as bat-associated RVs and their variants have been responsible for many of the recent human deaths related to rabies in the USA, which in many cases have had no clear exposure to rabid bat species.

The GenBank accession numbers of the sequences reported in this paper are AF426290–AF426311.

ABL belongs to the *Lyssavirus* genus within the family *Rhabdoviridae* (Hooper *et al.*, 1997). Like other lyssaviruses, ABL has a negative-sense, single-stranded RNA genome, which encodes a nucleocapsid protein (N), phosphoprotein (P), matrix protein (M), glycoprotein (G) and RNA polymerase (L), in the order 3′-N-P-M-G-L-5′ (Gould *et al.*, 1998; Warrilow *et al.*, 2002). A further feature of the ABL genome, also common to other lyssaviruses, is the presence of a non-coding region between the G and L genes, known as the G–L intergenic or pseudogene (ψ) region. Several molecular epidemiological studies have been conducted to determine phylogenetic relationships between lyssavirus isolates, their strains, host species and geographical distribution, with most investigations based on nucleotide sequence comparison of the N gene, G gene and ψ regions (Badrane & Tordo, 2001; Badrane *et al.*, 2001; Nadin-Davis *et al.*, 2001).

Comparison of the N gene, G gene and ψ sequences of numerous lyssavirus isolates and strains has delineated the *Lyssavirus* genus into seven genotypes (GTs): RV (GT1), Lagos bat virus (LAG, GT2), Mokola virus (MOK, GT3), Duvenhage virus (DUV, GT4), European bat lyssavirus 1 (EBLV1, GT5), European bat lyssavirus 2 (EBLV2, GT6) and ABL (GT7) (Badrane *et al.*, 2001). However, since its discovery, there has been some contention as to whether ABL is an Australian rabies virus variant or a distinct *Lyssavirus* species. Virus cross-neutralization and serological assays have indicated that ABL and RV are very similar antigenically (Fraser *et al.*, 1996; Gould *et al.*, 1998). Yet, preliminary phylogenetic analyses of the nucleotide sequences of the ABL N, P, M and G genes with other lyssaviruses have suggested that, although ABL is the most closely related lyssavirus to RV, it is sufficiently different at the nucleotide level to be assigned to its own genotype (Badrane *et al.*, 2001; Gould *et al.*, 1998; Hooper *et al.*, 1997). Two distinct ABL strains appear to be circulating within Australia currently, one found in flying foxes (Gould *et al.*, 1998), the other in the insectivorous bat species *S. flaviventris* (Hooper *et al.*, 1997). However, as yet very little is known regarding the interrelationships between flying fox and insectivorous ABL isolates, as well as ABL and other lyssavirus species, as few ABL isolates have been studied or sequenced.

Several phylogenetic analyses conducted on lyssaviruses have focused on the G gene, which presents the advantage of encoding the G glycoprotein containing determinants of lyssavirus pathogenicity and host specificity. The lyssavirus G glycoprotein ranges in size from 524 to 535 amino acids and consists of four domains: a 19-amino-acid signal peptide, which is cleaved from the protein's N terminus to produce the mature glycoprotein, a surface-exposed 439-amino acid highly conserved ectodomain, a 22-amino-acid transmembrane domain and a C-terminal endodomain, which varies in length from 42 to 53 amino acids. The G glycoprotein forms a trimer on the surface of lyssavirus virions, such that during infection it is the first viral protein to come into contact with the host. Apart from containing domains responsible for host-cell receptor recognition (Lentz *et al.*, 1982; Thoulouze *et al.*, 1998; Tuffereau *et al.*, 1998, 2001) and membrane fusion (reviewed by Gaudin *et al.*, 1999), the G glycoprotein is the major target for host neutralizing antibody responses (Wiktor *et al.*, 1973). Selection of RV mutants with neutralizing anti-G monoclonal antibodies has shown that amino acid changes mapping to specific sites in G often result in RV attenuation (Dietzschold *et al.*, 1983; Flamand *et al.*, 1980). It has also been demonstrated by reverse genetics that sequence variation within the G gene provides a basis for RV pathogenesis (Ito *et al.*, 2001a; Morimoto *et al.*, 2000). Accordingly, changes in lyssavirus G gene sequences may influence not only viral pathogenicity and host susceptibility, but may also allow the viruses to adapt to new hosts resulting in the emergence of new lyssavirus species.

The aims of this study were therefore threefold. First, it was of interest to determine whether correlations exist between particular ABL isolates, geographical locations and ABL infection of pteropid and insectivorous bat populations by comparing their G gene sequences. Such information can be used to establish epidemiological links and potential risk factors that may increase the likelihood of human ABL exposure. Secondly, we wished to investigate whether the ABL G gene is subject to positive selective pressure (adaptive evolution) in nature, as genetic variability within this region of the ABL genome may allow evasion of immune responses and/or produce viruses with an extended host range and specificity. Finally, we wished to clarify the position of ABL within the *Lyssavirus* genus. Most previous studies have included only a single pteropid or insectivorous ABL sequence in their phylogenetic analyses, which has limited the possibility of identifying intra- and interspecies relationships between ABL isolates and other lyssaviruses. We decided to pursue this further by comparing the G gene sequences of a number of ABL isolates with the corresponding regions of other lyssavirus species.

METHODS

Source of specimens. Between the years 1996 and 2000, bat species showing aggressive or abnormal behaviour towards humans and domestic animals in the Australian state of Queensland were submitted to the Queensland Department of Primary Industries, Animal Research Institute, for ABL testing. Specimens of brain or salivary gland material taken from 22 flying foxes and insectivorous bats that had tested positive to ABL by fluorescent antibody test were chosen for nucleotide sequence analysis. The species origin, tissue selected for analysis, year of isolation and geographical distribution of the ABL samples used in this study are described in Fig. 1 and Table 1.

Viral RNA extraction and cDNA synthesis. Samples of ABL-infected brain and salivary gland material were homogenized using QIAshredder spin columns (Qiagen) and total RNA was extracted from tissue lysates using a RNeasy mini kit (Qiagen). Extracted RNA was eluted in a final volume of 60 μl nuclease-free water. Reverse transcription (RT) of viral RNA was performed in a 10 μl

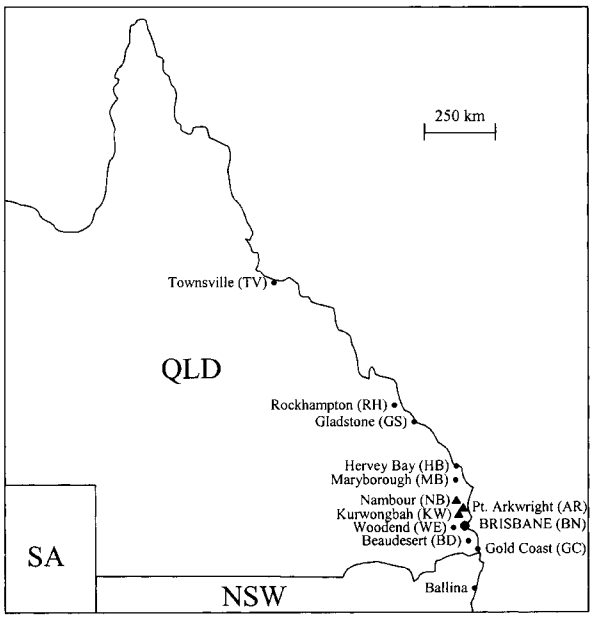


Fig. 1. Locations in Queensland and New South Wales where ABL isolates used in this study were collected. Abbreviations for the areas where samples were obtained are shown in brackets after the location name. Filled triangles and circles represent sites where insectivorous and pteropid ABL isolates were collected, respectively. QLD, Queensland; NSW, New South Wales; SA, South Australia.

reaction volume using 2–4 μl extracted RNA, 100 units Superscript II Reverse Transcriptase (Life Technologies), 30 units RNasin ribonuclease inhibitor (Promega) and 100 pmol of the ABL-specific positive-sense forward primer M-FOR (Table 2) to prime cDNA synthesis from the negative-sense viral genome. RT reactions were incubated at 42 °C for 45–75 min, before heat-inactivating the Superscript II Reverse Transcriptase at 95 °C for 5 min.

PCR and sequencing. PCR amplification was performed in a volume of 50 μl containing 1·5 mM $MgCl_2$, pH 8·0, 250 μM each dATP, dCTP, dGTP and dTTP, 200 pmol forward and reverse primers, 1–2 μl cDNA and 2·5 units *Taq* DNA polymerase (Roche). ABL G gene sequences were amplified as two separate PCR products, G5′HALF and G3′HALF, according to the following PCR programme: 5 min hold at 94 °C; 30 s at 94 °C, 45 s at 50 °C, 90 s at 72 °C for three cycles; 30 s at 94 °C, 45 s at 55 °C, 90 s at 72 °C for 30 cycles; 7 min at 72 °C; hold at 4 °C. The 1012 bp G5′HALF PCR product was amplified using oligonucleotides G5′FOR and G5′REV (Table 2), whereas the 947 bp G3′HALF PCR product was amplified with primer pair G3′FOR and G3′REV (Table 2). All PCR products were analysed on 1·0 % agarose gels in TAE containing ethidium bromide, and amplicons were purified from gel slices using a QIAquick gel extraction kit (Qiagen). Purified PCR products were directly sequenced on both strands using a series of oligonucleotides encompassing the M–G intergenic region to the 5′ extremity of the ψ region (Table 2). Sequencing was performed using an ABI PRISM BigDye Terminator Cycle Sequencing kit (Applied Biosystems) and an ABI 373A automated sequencer (Applied Biosystems).

Phylogenetic analysis. Nucleotide sequences of the 22 ABL isolate G gene regions were edited using the Sequencher 3.1.1 program (Gene Codes Corporation). Glycoprotein amino acid sequences were deduced using ETRANSLATE and compared using the multiple sequence alignment program ECLUSTALW (Thompson *et al.*, 1994), and percentage nucleotide and amino acid identities were calculated by the HOMOLOGIES program (Jack Leunissen CAOS/CAMM Center, University of Nijmegen, The Netherlands). These programs are part of the Genetics Group Wisconsin Inc. (GCG) suite of sequence analysis programs maintained by the Australian National Genomic Information Service (ANGIS, http://www.angis.org.au). Nucleotide and amino acid sequence alignments prepared using ECLUSTALW were edited using the SEAVIEW program and phylogenetic analyses were performed using the Linux-based PHYLO_WIN program (Galtier *et al.*, 1996) or programs of the PHYLIP package (Felsenstein, 1993). Phylogenetic trees were constructed by: (i) maximum-parsimony (MP) using algorithms from the DNAPARS and PROTPARS programs of the PHYLIP package; (ii) neighbour-joining (NJ) using the evolutionary distance correction statistics of Kimura (1980) and Tajima & Nei (1984); and (iii) maximum-likelihood (ML) using the PAUP* phylogenetic program (Swofford, 2001). Bootstrap resampling analysis (Felsenstein, 1985) was undertaken using 1000 data replications to evaluate the robustness of the phylogenetic groupings observed. Bootstrap values of greater than 70 % confidence were regarded as giving strong evidence for a particular phylogenetic grouping (Hillis & Bull, 1993). All ABL nucleotide sequences obtained in this study have been submitted to GenBank and their accession numbers are listed in Table 1. All other lyssavirus nucleotide sequences used for phylogenetic analysis and sequence comparison were obtained from GenBank; their accession numbers and appropriate references are listed in Table 1.

Selection analysis. An ML approach was used to analyse selection pressures in the ABL G gene sequences. In this method, the numbers of synonomous (silent, d_S) and non-synonomous (amino acid-changing, d_N) substitutions per site were determined using various models of codon substitution that differ in how the ratio d_N/d_S varies among codons (Yang *et al.*, 2000). Evidence for positive selection is provided when a model of codon evolution that shows $d_N/d_S>1$ at a number of codons is significantly favoured (in a likelihood ratio test) over a competing model in which d_N/d_S ratios are constrained to be <1 at all codons. If positive selection is found, those sites that are positively selected (i.e. with $d_N/d_S>1$) can be individually identified using Bayesian methods. This analysis was undertaken using the CODEML program of the PAML package (Yang, 1997). More details of this approach applied to RNA viruses are given in Holmes *et al.* (2002).

RESULTS

Nucleotide sequence comparison of ABL isolates and non-ABL lyssaviruses

To determine whether ABL demonstrated a high degree of genetic variability across its bat hosts and geographical range, 22 ABL isolates obtained from Queensland yellow-bellied sheath-tailed bats and three flying fox species (Table 1) were sequenced across a 1960 nucleotide region encompassing the G gene. Alignment of the ABL isolate G gene regions with those of two other ABL sequences available from the GenBank database (one pteropid and one insectivorous bat ABL isolate sequence) showed that nucleotide similarity between pteropid ABL isolate G genes ranged from 96·3 to 100·0 % (Table 3). Likewise, insectivorous ABL isolates also displayed a high degree of nucleotide identity within their G gene sequences (96·8–99·9 %).

Table 1. Origin of lyssavirus isolates used in this study

Virus code	Genotype or variant host*	Animal (tissue) isolated from	Year†	Location of isolate‡	Reference/source of isolate and GenBank accession no.
ABLBallina	Pteropid, GT7	*Pteropus alecto* (kidney)	1996	Ballina, NSW	Gould *et al.* (1998); AF006497
ABLPA22BN	Pteropid, GT7	*P. alecto* (brain)	1997	Brisbane, QLD	This study; AF426304
ABLPA05GC	Pteropid, GT7	*P. alecto* (brain)	1997	Gold Coast, QLD	This study; AF426293
ABLPP02BD	Pteropid, GT7	*P. poliocephalus* (brain)	1997	Beaudesert, QLD	This study; AF426291
ABLPP07BN	Pteropid, GT7	*P. poliocephalus* (brain)	1997	Brisbane, QLD	This study; AF426295
ABLPP08GC	Pteropid, GT7	*P. poliocephalus* (brain)	1997	Gold Coast, QLD	This study; AF426296
ABLPS19WE	Pteropid, GT7	*P. scapulatus* (brain)	1997	Woodend, QLD	This study; AF426303
ABLPA06BN	Pteropid, GT7	*P. alecto* (salivary gland)	1998	Brisbane, QLD	This study; AF426294
ABLPA26GC	Pteropid, GT7	*P. alecto* (brain)	1998	Gold Coast, QLD	This study; AF426306
ABLPA35HB	Pteropid, GT7	*P. alecto* (brain)	1998	Hervey Bay, QLD	This study; AF426309
ABLPA01MB	Pteropid, GT7	*P. alecto* (salivary gland)	1998	Maryborough, QLD	This study; AF426290
ABLPA14RH	Pteropid, GT7	*P. alecto* (brain)	1998	Rockhampton, QLD	This study; AF426300
ABLPA17TV	Pteropid, GT7	*P. alecto* (brain)	1998	Townsville, QLD	This study; AF426302
ABLPA32TV	Pteropid, GT7	*P. alecto* (brain)	1998	Townsville, QLD	This study; AF426307
ABLPP13BN	Pteropid, GT7	*P. poliocephalus* (brain)	1998	Brisbane, QLD	This study; AF426299
ABLPS03TV	Pteropid, GT7	*P. scapulatus* (salivary gland)	1998	Townsville, QLD	This study; AF426292
ABLPS24WE	Pteropid, GT7	*P. scapulatus* (brain)	1998	Woodend, QLD	This study; AF426305
ABLPA34GS	Pteropid, GT7	*P. alecto* (brain)	1999	Gladstone, QLD	This study; AF426308
ABLPA36BN	Pteropid, GT7	*P. alecto* (brain)	2000	Brisbane, QLD	This study; AF426310
ABLPA37GS	Pteropid, GT7	*P. alecto* (brain)	2000	Gladstone, QLD	This study; AF426311
ABLInsect	Insectivorous bat, GT7	*S. flaviventris* (brain)	?	?	AF081020; NC_003243
ABLSF11KW	Insectivorous bat, GT7	*S. flaviventris* (salivary gland)	1997	Kurwongbah, QLD	This study; AF426297
ABLSF12NB	Insectivorous bat, GT7	*S. flaviventris* (brain)	1997	Nambour, QLD	This study; AF426298
ABLSF15AR	Insectivorous bat, GT7	*S. flaviventris* (brain)	1998	Pt Arkwright, QLD	This study; AF426301
RabDrGUY	Bat (*Desmodus rotundus*)	Cow	1985	French Guyana	Badrane & Tordo (2001); AF325490
RabDrBRA	Bat (*D. rotundus*)	Cow	1986	Brazil	Badrane & Tordo (2001); AF325491
RabDrMEX	Bat (*D. rotundus*)	*D. rotundus*	1987	Mexico	Badrane & Tordo (2001); AF325492
RabBtARG	Bat	Bat	1991	Argentina	Badrane & Tordo (2001); AF325493
RabLsUSA	Bat (*Lasiuris*)	*Myotis* sp.	1979	Montana, USA	Badrane *et al.* (2001); AF298141
RabMyUSA1	Bat (*Myotis* sp.)	*Myotis* sp.	1981	Montana, USA	Badrane & Tordo (2001); AF325494
RabMyUSA2	Bat (*Myotis* sp.)	*Myotis* sp.	1982	Montana, USA	Badrane & Tordo (2001); AF325495
RabLnUSA	Bat (*Lasionycteris noctivagans*)	Human (bat bite)	1994	California, USA	Morimoto *et al.* (1996); U52946
RabCoUSA	Coyote	Coyote	1993	Texas, USA	Morimoto *et al.* (1996); U52947
RabDgCHI	Dog	Buck	1986	China	Badrane & Tordo (2001); AF325471
RabDgMAD	Dog	Dog	1985	Madagascar	Badrane & Tordo (2001); AF325478
RabDgMAU	Dog	Camel	1986	Mauritania	Badrane & Tordo (2001); AF325483
RabDgTHA	Dog	Human	1983	Thailand	Badrane & Tordo (2001); AF325488
RabFxFRA	Fox	Fox	1991	France	Badrane & Tordo (2001); AF325461
RabMgSAF	Mongoose	Mongoose	1987	South Africa	Badrane and Tordo (2001); AF325485
RabPfARC	Polar fox	Polar fox	<1981	Arctic Circle	Badrane & Tordo (2001); AF325486
RabRcUSA	Raccoon	Raccoon	?	New York, USA	Nadin-Davis *et al.* (1996); U27214
RabSkUSA1	Skunk	Skunk	1981	Montana, USA	Badrane & Tordo (2001); AF325473
RabSkUSA2	Skunk	Sheep	1981	Montana, USA	Badrane & Tordo (2001); AF325475
RabWfYUG	Wolf	Bovine	1984	Yugoslavia	Badrane & Tordo (2001); AF325463
RabCVS	Laboratory RV	PAS derivative§	1882	USA branch PAS	Sacramento *et al.* (1992)
RabPV	Vaccine RV	PAS derivative§	1882	USA branch PAS	Tordo *et al.* (1986, 1988); A14671
RabNishigahara	Laboratory RV	PAS derivative§	1915	Japanese branch PAS	Ito *et al.* (2001c); AB044824
RabRCHL	Vaccine RV	Nishigahara derivative	1918	Attenuated Nishigahara	Ito *et al.* (2001c); AB009663

Table 1. Origin of lyssavirus isolates used in this study *(cont.)*

Virus code	Genotype or variant host*	Animal (tissue) isolated from	Year†	Location of isolate‡	Reference/source of isolate and GenBank accession no.
RabSADB19	Vaccine RV	Dog	1935	Alabama, USA	Conzelmann *et al.* (1990)
RabERA	Laboratory RV	SAD derivative	1935	Attenuated SAD	Sacramento *et al.* (1992)
RabHEPFlury	Laboratory RV	Human	1939	Georgia, USA	Morimoto *et al.* (1989); M32751
DuvBtSAF1	GT4	Human (bat bite)	1970	South Africa	Badrane *et al.* (2001); AF298146
DuvMsSAF2	GT4	Bat (*Miniopterus schreibersii*)	1981	South Africa	Badrane *et al.* (2001); AF298147
EBL1EsFRA	GT5	Bat (*Eptesicus serotinus*)	1985	France	Badrane *et al.* (2001); AF298143
EBL1EsPOL	GT5	Bat (*E. serotinus*)	1989	Poland	Badrane *et al.* (2001); AF298142
EBL2BtFIN	GT6	Human (bat bite)	1986	Finland	Badrane *et al.* (2001); AF298144
EBL2MdHOL	GT6	Bat (*Myotis dasycneme*)	1986	Holland	Badrane *et al.* (2001); AF298145
LagMpCAR	GT2	Bat (*Micropterus pusillus*)	1974	Central African Republic	Badrane *et al.* (2001); AF298149
LagEhNGA	GT2	Bat (*Eidolon helvum*)	1956	Nigeria	Badrane *et al.* (2001); AF298148
MokCaETH	GT3	Cat	1990	Ethiopia	Mebatsion *et al.* (1995); U17064
MokCaZIM	GT3	Cat	1982	Zimbabwe	Le Mercier *et al.* (1997); Y09762

*All viruses belong to genotype (GT) 1 unless stated otherwise.
†The year in which the virus was isolated. A question mark denotes that the year of isolation was unknown.
‡The location at which the virus was isolated. A question mark denotes that the location of isolation was unknown.
§PAS derivatives are tissue culture adapted variants of the original rabies virus Pasteur strain isolated from a rabid cow in 1882 (Sacramento *et al.*, 1992).

Table 2. Oligonucleotide primers used for PCR and sequencing of ABL G gene regions

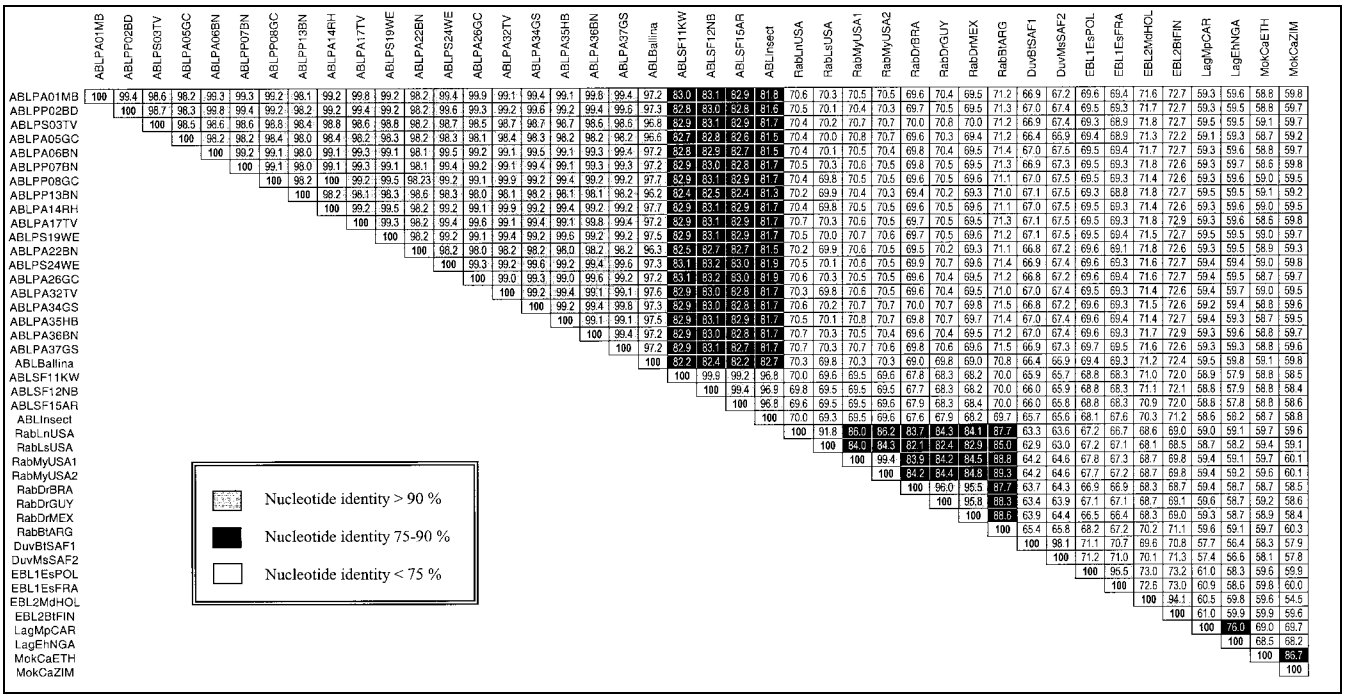
Name	Sequence (5′→3′)*	Nucleotide position†
M-FOR	GCTATGGTCTGACATGTCTCT	2964–2984
G5′FOR	CAAGATACCTGATTACATTTAC	3164–3185
G-REV	TCGGGTATTGTGTAGAGAGGG	rc 3308–3328
G5′SEQ1F	AGGTATGAGGAGTCTTTGCAC	3621–3641
G5′SEQ1R	GTGCAAAGACTCCTCATACC	rc 3622–3641
G3′FOR	AGTTGTGCGGAATCTCTGGTCTC	3979–4001
G5′REV	CATTATTGACTCCAGTGCATC	rc 4155–4175
G3′SEQ1F	GTCTTAATTCCGGAGATGCAG	4413–4433
G3′SEQ1R	GGATGGATAAGAGGTATTACTG	rc 4474–4495
G-FOR	GTGTCAGTAACATCTCAAAGTGGG	4743–4766
G3′REV	ATCTCAGTATTGCATATGGCTC	rc 4905–4926

*Oligonucleotide primers were designed from positive-sense ABL Ballina and insectivorous isolate nucleotide sequences obtained from GenBank (as described in Table 1).
†Nucleotide positions are those according to the numbering of the ABL Ballina (Gould *et al.*, 1998) and ABL insectivorous (GenBank accession nos AF081020, NC_003243) isolate nucleotide sequences.
rc, Reverse complement.

Comparison of ABL pteropid isolate G gene regions with their insectivorous counterparts, however, demonstrated a much greater sequence divergence, with nucleotide sequences differing by up to 18·7 % (Table 3).

Further comparison of the ABL G gene sequences with those of 18 other chiropteran lyssaviruses and Mokola virus isolates indicated that the ABL isolates shared the greatest nucleotide identity with European bat lyssavirus 2 and chiropteran RV (67·9–72·9 %) and the lowest level of nucleotide sequence conservation with Lagos bat and Mokola viruses (58·2–59·8 %) (Table 3). While it was observed that isolates within each of the lyssavirus genotypes shared at least 76·0 % nucleotide sequence conservation between their respective G genes, isolates obtained from the same bat species or previously shown to belong to a specific ABL, chiropteran RV or European bat lyssavirus 1 variant had G gene sequences that were more than

Table 3. Percentage nucleotide identity of lyssavirus isolate G gene regions



95·0 % identical at the nucleotide level (Table 3). Likewise, Duvenhage and European bat lyssavirus 2 isolates shared 98·1 % and 94·1 % nucleotide sequence similarity, respectively, although it was unknown whether the isolates within each of these genotypes were of the same variant or collected from the same bat species.

For the chiropteran RVs, it was additionally observed that the level of nucleotide sequence identity among different RV variants (82·1–91·8 %) was intermediate between that observed for isolates belonging to the same RV variant and viruses belonging to different lyssavirus species. This observation is interesting in light of the fact that pteropid ABL isolates share only 81·3–83·2 % nucleotide sequence conservation with ABL isolated from *S. flaviventris* and that the two Lagos bat isolates originating from different bat species also display a similarly reduced level of sequence similarity (76·0 %).

Evolutionary relationships of ABL isolates inferred from G gene sequences

To investigate the intragenotypic variability of the ABL isolates further, phylogenetic analyses were performed using the ABL G gene nucleotide sequence alignment data. Phylogenetic trees were generated from the ABL G gene alignments using both MP and NJ with two Mokola isolates as an outgroup (Fig. 2). In both trees, ABL segregated into two clusters with 100 % bootstrap support, indicating the presence of two distinct ABL variants derived from a single progenitor. One variant lineage was associated with flying foxes (pteropid ABL) and the other associated with *S. flaviventris* (insectivorous ABL). The short branches observed in the NJ tree for the pteropid ABL cluster indicate that there is a very low level of genetic divergence between the isolates. Likewise, there appears to be little sequence variation among insectivorous ABL isolates.

To clarify the position of ABL within the *Lyssavirus* genus, ABL G gene sequences were aligned with those of 37 other lyssaviruses (representing GTs 1–6), including seven RV laboratory strains and 20 terrestrial and chiropteran RV 'street' isolates. Phylogenetic trees constructed from the G gene alignment using both MP and NJ displayed similar overall topologies, as did trees constructed from G ectodomain nucleotide sequence alignments. Fig. 3 shows the tree generated from the G gene sequences by NJ analysis, which was representative of those obtained for each nucleotide sequence alignment using either NJ or MP. Each phylogenetic method consistently segregated the ABL isolates into a monophyletic group, once again with distinct pteropid and insectivorous ABL lineages (with 100 % bootstrap support), separate from RV and rabies-like lyssavirus genotypes. Furthermore, both the MP and NJ analyses placed ABL as the closest relative to classical RV, a relationship with 82 % bootstrap support (Fig. 3). Similar clustering of *Lyssavirus* species was found when trees were constructed from amino acid sequences and G gene data with the third positions of codons removed in order to reduce the possibility of multiple substitution events (trees not shown, available from the authors on request). In the former, the MP method implemented in PROTPARS was used and the bootstrap support for the ABL–RV grouping

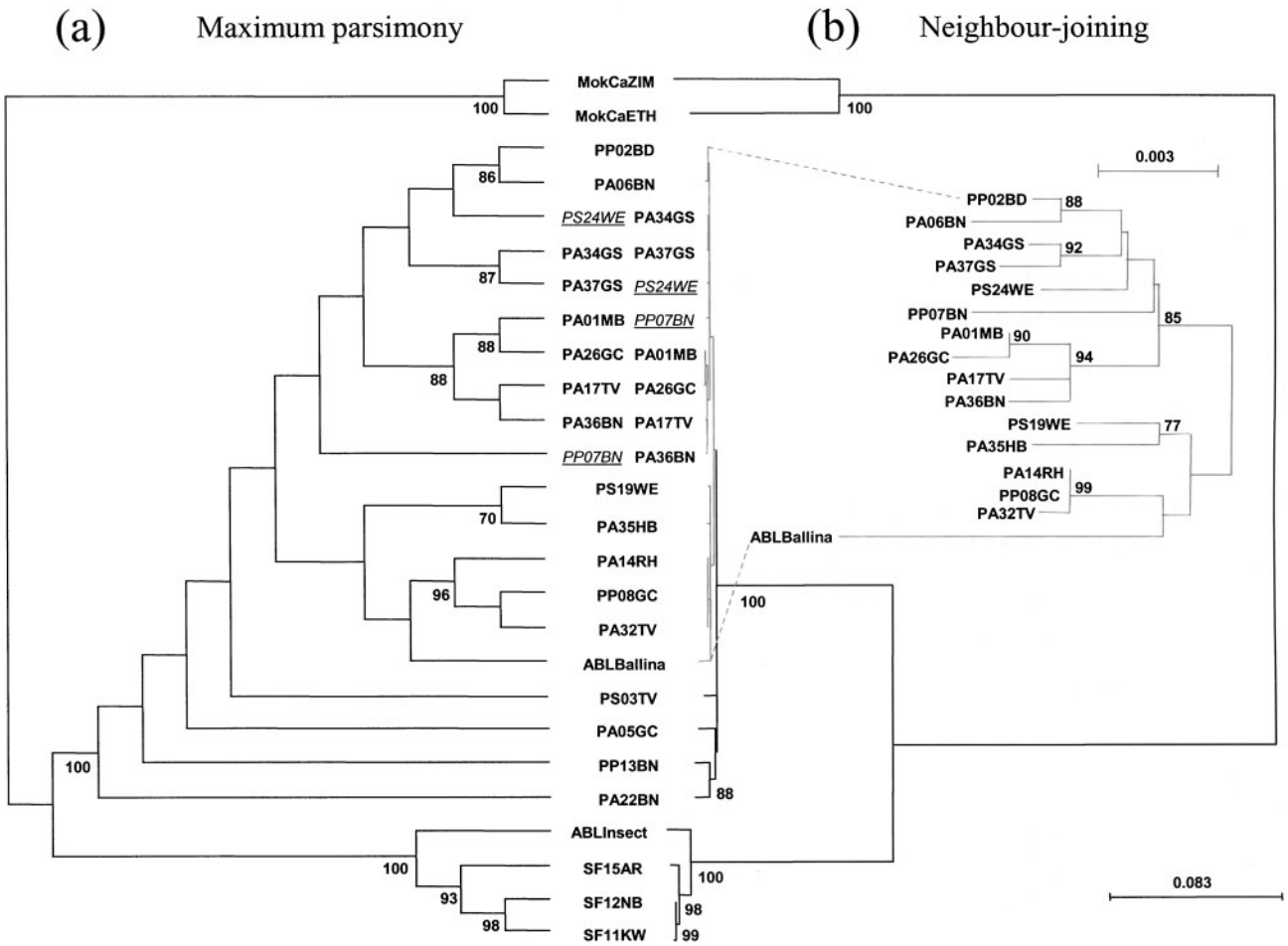


Fig. 2. Phylogeny of Australian bat lyssavirus isolates. Phylogenetic trees were determined using (a) maximum-parsimony (MP) and (b) neighbour-joining (NJ). Numbers at each node represent the percentage bootstrap support (1000 replicates). Only bootstrap values $\geqslant$70 % are indicated. For the NJ phylogenetic tree, branch length is proportional to evolutionary distance between the different ABL isolates (nucleotide substitutions per site), as shown by the distance bar (bottom right). Likewise, the region encompassing the ABL pteropid isolates indicated by dashed lines has been magnified to show the relationships between these isolates more clearly. The distance bar for the magnified portion of the tree is shown at the top right. Identifying codes for each isolate (according to Table 1) are at the ends of each branch. Isolate names that are italicised and underlined represent isolates that have been grouped differently by the MP and NJ analyses. Mokola virus isolates MokCaETH and MokCaZIM were used as an outgroup.

was 24 %. In the latter, the ML method was used incorporating the GTR+I+G model of nucleotide substitution (program PAUP*; Swofford, 2001). The bootstrap support for the ABL–RV grouping in this analysis (1000 neighbour-joining bootstrap replications) was 55 %. Although neither MP nor ML analyses of the lyssavirus G amino acid sequences gave robust evidence for the separate clustering of ABL and RV, the trees generated had similar topologies to those predicted for the G gene nucleotide sequences such that ABL and RV appear to be phylogenetically distinct lyssaviruses. Overall, this analysis supports previous work (Badrane & Tordo, 2001; Badrane *et al.*, 2001; McColl *et al.*, 2000), where ABL was assigned its own species designation (although only on the basis of one ABL sequence).

One further common feature of the *Lyssavirus* genus reflected by the phylogenetic trees presented here and previously noted by McColl *et al.* (2000) is that bat lyssaviruses from each genotype tend to be associated with particular continental regions, i.e. bat RV (GT1), LAG/DUV (GT2/GT4), EBL1/EBL2 (GT5/GT6) and ABL (GT7) have only been found in the Americas, Africa, Europe and Australasia, respectively (Fig. 3). Additionally, each lyssavirus genotype (apart from GT3) can be seen to have characteristic variants allied with specific bat hosts.

Selection pressures in the ABL G gene

Although phylogenetic analyses indicated that the ABL isolates were closely related, it was still of interest to

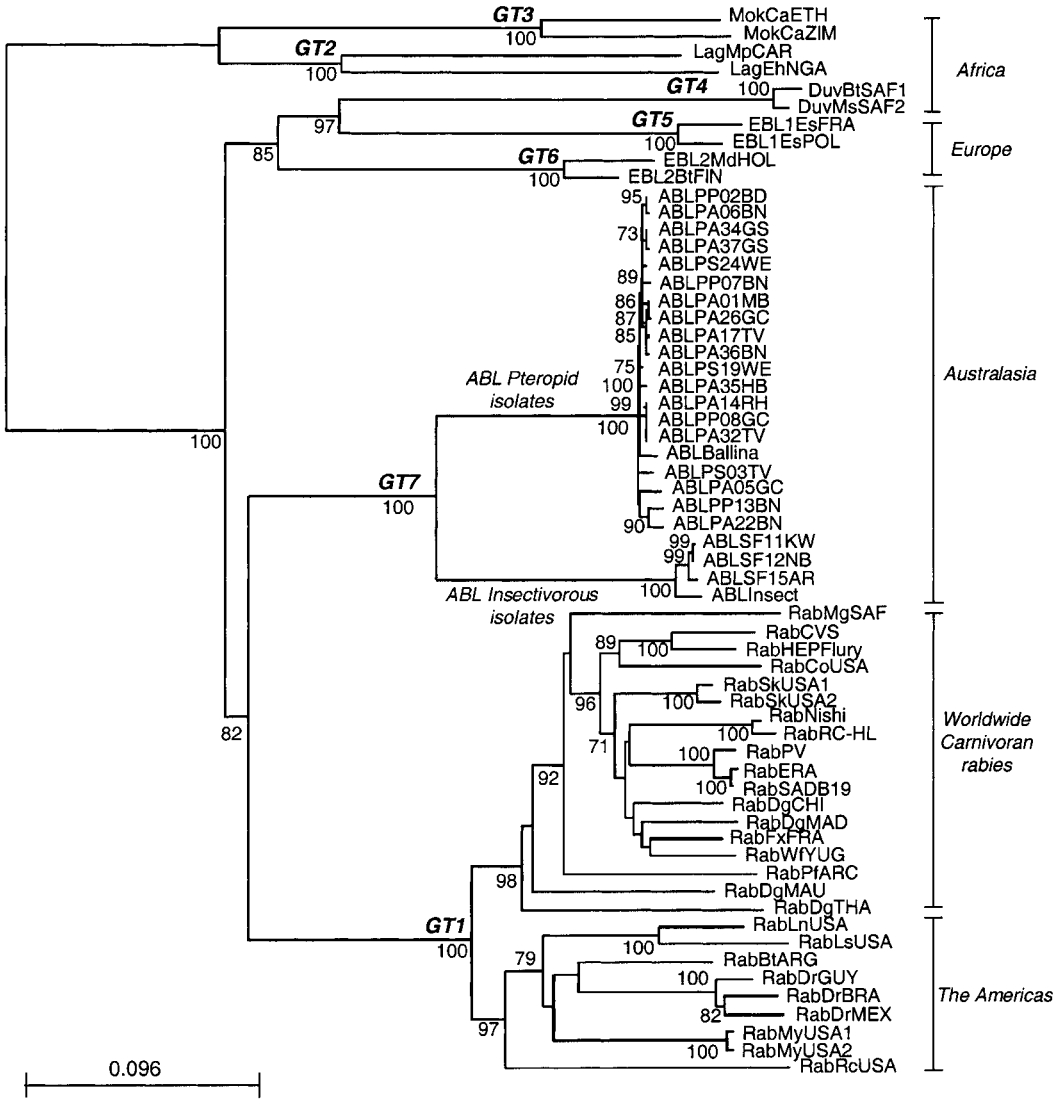


Fig. 3. Neighbour-joining phylogenetic tree demonstrating the genetic relationship of the *Lyssavirus* genus based on the G gene regions of 61 lyssavirus isolates. Numbers at each node represent the percentage bootstrap support (1000 replicates). Only bootstrap values $\geqslant$70 % are indicated. Branch length is proportional to evolutionary distance between the different lyssavirus isolates. Identifying codes for each isolate (according to Table 1) are at the ends of each branch. The main phylogenetic groupings and their geographical locations are shown to the right of the figure.

determine whether these viruses had been subject to positive genetic selection. The G gene of natural ABL isolates was targeted for this analysis, as the G glycoprotein is the major pathogenic determinant responsible for lyssavirus virulence and host-cell tropism, making it perhaps the most likely candidate for the presence of sites experiencing adaptive evolution. ML analyses of selection pressures acting on the ABL G gene sequences provided some evidence, albeit weak, for positive selection. Specifically, a d_N/d_S value of 4·46 was found at amino acids 499 and 501 under the 'M8' model of codon evolution, and this model was significantly favoured over a competing neutral counterpart ('M7'), in which d_N/d_S was constrained to be <1 at all codons (full results available from the authors on request). Codons 499 and 501 were also found to be selected under the 'M3' model, although in this case a competing neutral model ('M2') could not be significantly rejected. Despite this evidence for adaptive evolution, the strongest selection pressure acting on these sequences was clearly purifying (negative) selection, as the numbers of amino acid changes among the sequences were low, so that the vast majority of codons had very low d_N/d_S ratios; for example, 93 % of codons had a mean d_N/d_S of 0·03 under the 'M3' model.

Codon positions 499 and 501 fall within the endodomain of the viral glycoprotein, and alignment of this region revealed some amino acid variation between ABL isolates

at these positions (Fig. 4). For the ABL Ballina and ABL insectivorous isolates previously sequenced by Gould *et al.* (1998), the sequences were quite different, with three nucleotide deletions resulting in a glycoprotein endomain one amino acid shorter at the C terminus than the ABL isolates sequenced in this study. This sequence variation is of interest (assuming the differences observed for the ABL Ballina and insectivorous isolates are not sequencing artifacts), as selectively favoured mutations may have occurred in the G ectodomain region of these viruses when they were passaged in mice and in *in vitro* cell cultures (mouse neuroblastoma 2A and baby hamster kidney cell lines) prior to sequencing.

DISCUSSION

Phylogenetic analyses of lyssavirus nucleotide sequences suggest that ABL is a separate *Lyssavirus* species consisting of two distinct variants

Since the advent of molecular biology, which has allowed rapid acquisition of sequence data for whole viral genomes, the definition of what constitutes a virus species has often been brought into question or left open to interpretation. Delineation of virus species within the *Lyssavirus* genus has been based on classical virology, such that virus species designation has depended on monoclonal antibody reactivity profiles or virus performance in cross-neutralization studies (van Regenmortel *et al.*, 2000). In this respect, ABL has been assigned to the same lyssavirus serotype as RV (serotype 1), as it is neutralized by antibodies raised against RV proteins, and commercial RV vaccines are cross-protective for ABL in mouse challenge experiments (Hooper *et al.*, 1997). Recent studies by Badrane *et al.* (2001), however, have indicated that designation of lyssavirus species should also take into account the phylogenetic relationships of viral gene sequences. As such, the phylogenetic analyses conducted in this investigation support the view that Australian bat lyssavirus is a separate *Lyssavirus* species, with two distinct variants maintained in *S. flaviventris* and pteropid hosts, respectively.

In the Americas, the bat species *Desmodus rotundus* (vampire bat), *Tadarida brasiliensis* (Brazilian free-tailed bat), *Eptesicus fuscus* (big brown bat), *Lasiurus* species (*L. borealiz*, *L. cinereus*), *Lasionycteris noctivagans* (silver-haired bat), *Pipistrellus subflavus* (eastern pipistrelle) and *Myotis* species (*M. lucifugus*; *M. yumanensis*, *M. californicus*; *M. evotis*) have all been identified as rabies virus reservoirs that harbour distinct RV variants (Morimoto *et al.*, 1996; Nadin-Davis *et al.*, 2001; Smith, 1996). These RV variants generally separate into phylogenetic divisions that represent the lifestyle of their chiropteran hosts, i.e. migratory versus non-migratory, colonial versus solitary, insectivorous versus haematophagus (Nadin-Davis *et al.*, 2001). While most of these variants co-segregate only with their specific host reservoirs making spillover events to terrestrial animals uncommon, some have been associated with infection of non-chiropteran species, especially humans and domestic animals (Crawford-Miksza *et al.*, 1999; de Mattos *et al.*, 2000; Favi *et al.*, 2002; Ito *et al.*, 2001b; Morimoto *et al.*, 1996; Nadin-Davis *et al.*, 2001; Noah *et al.*, 1998; Smith *et al.*, 1995).

From this study, it appears that ABL follows a similar pattern of epidemiology to bat RV, in that separate pteropid and insectivorous ABL variants co-circulate in Australia within select bat hosts that display different behavioural patterns. Likewise, although spillover of ABL to domestic animals has not been observed, both ABL variants have been responsible for human deaths (Allworth *et al.*, 1996; Hanna *et al.*, 2000). Phylogenetic analyses indicate that very little sequence divergence exists between the G gene sequences obtained for ABL isolates from the three flying fox species *P. alecto*, *P. poliocephalus* and *P. scapulatus*. Additionally, although the topology of the pteropid ABL

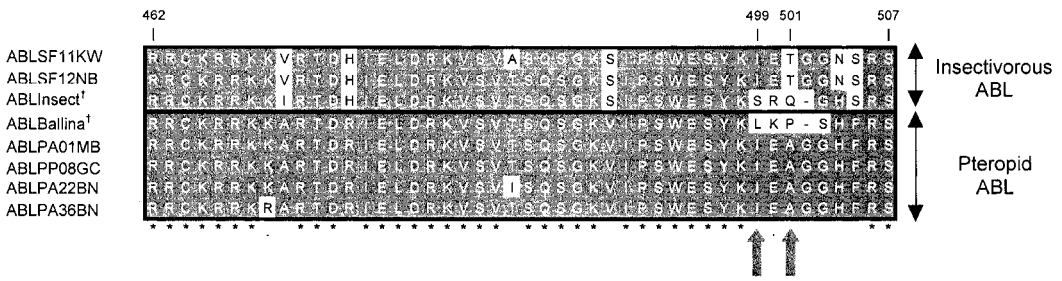


Fig. 4. CLUSTALW alignment of ABL glycoprotein endodomain amino acid sequences. Shaded residues indicate amino acids strictly conserved between pteropid and insectivorous ABL isolates. Asterisks and full stops represent the positions of amino acids that are strictly conserved or similar across ABL, respectively. The two ABL G glycoprotein sequences obtained from the GenBank database and used in the alignment are indicated by the symbol †. Identifying codes for each isolate (according to Table 1) are shown to the left of the figure. Numbers above the alignment represent the amino acid positions within the mature glycoprotein, where the N-terminal amino acid of the ectodomain is assigned position 1. Note that the isolates shown in the alignment are representative of the different ABL sequences obtained in this region, as there is some amino acid sequence redundancy between isolates within the glycoprotein domain.

cluster was found to vary slightly between the MP and NJ trees (Fig. 2), it appears that delineation of the pteropid ABL isolates does not correlate with flying fox species or geographical distribution as isolates from three separate *Pteropus* species collected over a distance of greater than 1400 km group together. Indeed ABL isolates ABLPP08GC and ABLPA14RH obtained from *P. poliocephalus* and *P. alecto*, respectively, at locations over 700 km apart were found to be 100 % identical (Table 3). Such minimal sequence variation of pteropid ABL isolates may be explained in part by the migratory behaviour and gregarious nature of these flying fox species.

During the day, flying foxes roost in trees in colonies known as camps. Depending on the species, flying fox camps may consist of many thousands of individuals (Hall & Richards, 2000). In the border region of Queensland and New South Wales, the distribution range of *P. alecto*, *P. poliocephalus* and *P. scapulatus* overlaps, and it is not uncommon for all three species to share the same camp, especially during periods of migration (Hall & Richards, 2000). This provides both an opportunity for the different flying fox species to interact, as well as conditions that are ideal for a viral pathogen to become adapted to more than one pteropid species simultaneously, the outcome being that ABL isolated from each species of flying fox is nearly identical. A similar situation occurs in the USA, where isolates from RV variants of the migratory bat species *L. noctivagans* and *T. brasiliensis* are found to have a very high degree of nucleotide similarity, even if they have been isolated from locations separated by thousands of kilometres (Smith, 1996).

At this stage it is unclear whether geographical segregation coincides with genetic relatedness of the insectivorous ABL cluster isolates, as the *S. flaviventris* ABL isolates used in this study were collected at sites separated by a distance of only 100 km. However, as it appears that the lifestyle of the bat host affects transmission and spread of each lyssavirus species, it would be interesting to investigate ABL infection of *S. flaviventris* further as these bats live in much smaller colonies than pteropid species (30 individuals at most), eat insects rather than fruit and nectar, and migrate over much shorter distances compared with flying foxes. Therefore, greater sequence variation may be found between insectivorous ABL isolates separated by large distances than that demonstrated by pteropid ABL such that specific subvariants may be identified, as is the case for RV in the sedentary bat host *E. fuscus* (Nadin-Davis *et al.*, 2001; Smith, 1996).

Although no pteropid ABL variants were isolated from ABL-infected *S. flaviventris* and vice versa, it is unclear whether flying foxes and insectivorous bats are susceptible to each other's ABL variants due to the small numbers of ABL-infected bats used in this study. Likewise, it is also not known whether other ABL variants exist in transmission cycles with Australian chiropteran species other than *S. flaviventris* or the common flying foxes. By pursuing these aspects of ABL epidemiology further, ecological factors controlling emergence of ABL within Australia, such as increased interactions between bat species and non-chiropteran wildlife, can be identified and targeted to prevent successful cross-species transmission events and initiation of ABL epidemics in new host species.

Purifying selection of ABL in nature

Although some evidence was found for weak positive selection at two codons within the G glycoprotein of natural ABL isolates, in general it appears that levels of non-synonomous diversity within the G gene of ABL are low. This also appears to be true of RV in nature, where only a few sites (codons 175, 183 and 370) within the conserved glycoprotein ectodomain have been found to be subject to modest positive selection pressure (Badrane & Tordo, 2001; Bourhy *et al.*, 1999; Holmes *et al.*, 2002). Such strong purifying selection is perhaps surprising considering the plasticity of RNA virus genomes (Domingo & Holland, 1997), the abundance of positive selection pressure in the glycoproteins of other negative-strand RNA viruses such as respiratory syncytial virus (Woelk & Holmes, 2001), measles virus (Woelk *et al.*, 2001, 2002), Marburg virus (Sanchez *et al.*, 1998) and influenza A virus (Suzuki & Nei, 2002), and the fact that RV mutants can be readily generated in the laboratory by passage through different host species (Kissi *et al.*, 1999). Consequently, the lack of positive selection in the G glycoprotein suggests either that the virus is not subject to strong immune selection or that it is exposed to relatively constant host, cellular and ecological environments so that most amino acid changes are not adaptively useful. As a case in point, although RV replication is almost exclusively limited to neurons, it has been shown that muscle tissue fibrocytes, acinar cells of the salivary gland and epidermal cells are also capable of supporting RV replication (Morimoto *et al.*, 1996; Charlton *et al.*, 1997). Holmes *et al.* (2002) have suggested that the limited sequence variation within RV N and G genes may be a consequence of simultaneous adaptation to a wide variety of cell types. It is therefore possible that ABL, and indeed all other naturally isolated lyssavirus species, are also viral 'generalists' that are adapted to a variety of cell types, although this hypothesis clearly requires further investigation. Indeed, such work may also give an indication of which viruses have the greatest propensity to produce adaptively useful genetic variation, and hence are most likely to adapt to new host species.

Furthermore, it is also uncertain why amino acids 499 and 501 within the glycoprotein endodomain may be subject to genetic variation. As these amino acids are not surface-exposed, it is unlikely that they play a direct role in either host-cell binding or fusion, and it is therefore doubtful that selection pressure would be of an immunological nature. It is possible, however, that changes within the endodomain amino acids may alter the folding of G such that association of the glycoprotein with other proteins or host membranes during virion synthesis may also be

modified to suit new environmental conditions. If this is indeed the case, the development of an ABL reverse genetics system may provide a means to mutagenize the ABL glycoprotein and investigate which of its domains contain determinants for host specificity and adaptation.

ACKNOWLEDGEMENTS

We are indebted to Craig Smith, Janine Barrett and Hume Field for donating both their time and ABL-infected flying fox and insectivorous bat brain and salivary gland tissue specimens submitted to the Queensland Department of Primary Industries, ARI, for ABL diagnosis. We also wish to thank Hassan Badrane for providing us with lyssavirus nucleotide sequences before their release on GenBank and Robert Rice for his invaluable assistance with the phylogenetic analyses and critical appraisal of the manuscript. This work was supported by grant 102528 from the National Health and Medical Research Council of Australia.

REFERENCES

Allworth, A., Murray, K. & Morgan, J. (1996). A human case of encephalitis due to a lyssavirus recently identified in fruit bats. *Commun Dis Intell* **20**, 504.

Arguin, P. M., Murray-Lillibridge, K., Miranda, M. E. G., Smith, J. S., Calaor, A. B. & Rupprecht, C. E. (2002). Serologic evidence of *Lyssavirus* infections among bats, the Philippines. *Emerg Infect Dis* **8**, 258–262.

Badrane, H. & Tordo, N. (2001). Host switching in *Lyssavirus* history from the Chiroptera to the Carnivora orders. *J Virol* **75**, 8096–8104.

Badrane, H., Bahloul, C., Perrin, P. & Tordo, N. (2001). Evidence of two *Lyssavirus* phylogroups with distinct pathogenicity and immunogenicity. *J Virol* **75**, 3268–3276.

Bourhy, H., Kissi, B., Audry, L., Smreczak, M., Sadkowska-Todys, M., Kulonen, K., Tordo, N., Zmudzinski, J. F. & Holmes, E. C. (1999). Ecology and evolution of rabies virus in Europe. *J Gen Virol* **80**, 2545–2557.

Charlton, K. M., Nadin-Davis, S., Casey, G. A. & Wandeler, A. I. (1997). The long incubation period in rabies: delayed progression of infection in muscle at the site of exposure. *Acta Neuropathol* **94**, 73–77.

Conzelmann, K. K., Cox, J. H., Schneider, L. G. & Thiel, H. J. (1990). Molecular cloning and complete nucleotide sequence of the attenuated rabies virus SAD B19. *Virology* **175**, 485–499.

Crawford-Miksza, L. K., Wadford, D. A. & Schnurr, D. P. (1999). Molecular epidemiology of enzootic rabies in California. *J Clin Virol* **14**, 207–219.

de Mattos, C. A., Favi, M., Yung, V., Pavletic, C. & de Mattos, C. C. (2000). Bat rabies in urban centers in Chile. *J Wildl Dis* **36**, 231–240.

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. *Proc Natl Acad Sci U S A* **80**, 70–74.

Domingo, E. & Holland, J. J. (1997). RNA virus mutations and fitness for survival. *Annu Rev Microbiol* **51**, 151–178.

Favi, M., de Mattos, C. A., Yung, V., Chala, E., Lopez, L. R. & de Mattos, C. C. (2002). First case of human rabies in Chile caused by an insectivorous bat virus variant. *Emerg Infect Dis* **8**, 79–81.

Felsenstein, J. (1985). Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39**, 783–791.

Felsenstein, J. (1993). PHYLIP: Phylogeny Inference Package, version 3.52c. University of Washington, Seattle, WA, USA.

Field, H., McCall, B. & Barrett, J. (1999). Australian bat lyssavirus infection in a captive juvenile black flying fox. *Emerg Infect Dis* **5**, 438–40.

Flamand, A., Wiktor, T. J. & Koprowski, H. (1980). Use of hybridoma monoclonal antibodies in the detection of antigenic differences between rabies and rabies-related virus proteins. II. The glycoprotein. *J Gen Virol* **48**, 105–109.

Fraser, G. C., Hooper, P. T., Lunt, R. A., Gould, A. R., Gleeson, L. J., Hyatt, A. D., Russell, G. M. & Kattenbelt, J. A. (1996). Encephalitis caused by a Lyssavirus in fruit bats in Australia. *Emerg Infect Dis* **2**, 327–331.

Galtier, N., Gouy, M. & Gautier, C. (1996). SEAVIEW and PHYLO_WIN: two graphic tools for sequence alignment and molecular phylogeny. *Comput Appl Biosci* **12**, 543–548.

Gaudin, Y., Tuffereau, C., Durrer, P., Brunner, J., Flamand, A. & Ruigrok, R. (1999). Rabies virus-induced membrane fusion. *Mol Membr Biol* **16**, 21–31.

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

Hall, L. & Richards, G. (2000). *Flying Foxes: Fruit and Blossom Bats of Australia*. Edited by T. J. Dawson. Australian Natural History Series. Sydney: UNSW Press.

Hanna, J. N., Carney, I. K., Smith, G. A. & 7 other authors (2000). Australian bat lyssavirus infection: a second human case, with a long incubation period. *Med J Aust* **172**, 597–599.

Hillis, D. M. & Bull, J. J. (1993). An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. *Syst Biol* **42**, 182–192.

Holmes, E. C., Woelk, C. H., Kassis, R. & Bourhy, H. (2002). Genetic constraints and the adaptive evolution of rabies virus in nature. *Virology* **292**, 247–257.

Hooper, P. T., Lunt, R. A., Gould, A. R. & 7 other authors (1997). A new lyssavirus – the first endemic rabies-related virus recognized in Australia. *Bull Inst Pasteur* **95**, 209–218.

Ito, N., Takayama, M., Yamada, K., Sugiyama, M. & Minamoto, N. (2001a). Rescue of rabies virus from cloned cDNA and identification of the pathogenicity-related gene: glycoprotein gene is associated with virulence for adult mice. *J Virol* **75**, 9121–9128.

Ito, M., Arai, Y. T., Itou, T., Sakai, T., Ito, F. H., Takasaki, T. & Kurane, I. (2001b). Genetic characterization and geographic distribution of rabies virus isolates in Brazil: identification of two reservoirs, dogs and vampire bats. *Virology* **284**, 214–222.

Ito, N., Kakemizu, M., Ito, K. A., Yamamoto, A., Yoshida, Y., Sugiyama, M. & Minamoto, N. (2001c). A comparison of complete genome sequences of the attenuated RC-HL strain of rabies virus used for production of animal vaccine in Japan, and the parental Nishigahara strain. *Microbiol Immunol* **45**, 51–58.

Kimura, M. (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J Mol Evol* **16**, 111–120.

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. *J Gen Virol* **80**, 2041–2050.

Le Mercier, P., Jacob, Y. & Tordo, N. (1997). The complete Mokola virus genome sequence: structure of the RNA-dependent RNA polymerase. *J Gen Virol* **78**, 1571–1576.

Lentz, T. L., Burrage, T. G., Smith, A. L., Crick, J. & Tignor, G. H. (1982). Is the acetylcholine receptor a rabies virus receptor? *Science* **215**, 182–184.

McColl, K. A., Tordo, N. & Aguilar Setien, A. A. (2000). Bat lyssavirus infections. *Rev Sci Tech* **19**, 177–196.

Mebatsion, T., Schnell, M. J. & Conzelmann, K. K. (1995). Mokola virus glycoprotein and chimeric proteins can replace rabies virus glycoprotein in the rescue of infectious defective rabies virus particles. *J Virol* **69**, 1444–1451.

Morimoto, K., Ohkubo, A. & Kawai, A. (1989). Structure and transcription of the glycoprotein gene of attenuated HEP-Flury strain of rabies virus. *Virology* **173**, 465–477.

Morimoto, K., Patel, M., Corisdeo, S., Hooper, D. C., Fu, Z. F., Rupprecht, C. E., Koprowski, H. & Dietzschold, B. (1996). Characterization of a unique variant of bat rabies virus responsible for newly emerging human cases in North America. *Proc Natl Acad Sci U S A* **93**, 5653–5658.

Morimoto, K., Foley, H. D., McGettigan, J. P., Schnell, M. J. & Dietzschold, B. (2000). Reinvestigation of the role of the rabies virus glycoprotein in viral pathogenesis using a reverse genetics approach. *J Neurovirol* **6**, 373–381.

Nadin-Davis, S. A., Huang, W. & Wandeler, A. I. (1996). The design of strain-specific polymerase chain reactions for discrimination of the racoon rabies virus strain from indigenous rabies viruses of Ontario. *J Virol Methods* **57**, 1–14.

Nadin-Davis, S. A., Huang, W., Armstrong, J., Casey, G. A., Bahloul, C., Tordo, N. & Wandeler, A. I. (2001). Antigenic and genetic divergence of rabies viruses from bat species indigenous to Canada. *Virus Res* **74**, 139–156.

Noah, D. L., Drenzek, C. L., Smith, J. S. & 9 other authors (1998). Epidemiology of human rabies in the United States, 1980 to 1996. *Ann Intern Med* **128**, 922–930.

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

Sanchez, A., Trappier, S. G., Stroher, U., Nichol, S. T., Bowen, M. D. & Feldmann, H. (1998). Variation in the glycoprotein and VP35 genes of Marburg virus strains. *Virology* **240**, 138–146.

Smith, J. S. (1996). New aspects of rabies with emphasis on epidemiology, diagnosis, and prevention of the disease in the United States. *Clin Microbiol Rev* **9**, 166–176.

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

Suzuki, Y. & Nei, M. (2002). Origin and evolution of influenza virus hemagglutinin genes. *Mol Biol Evol* **19**, 501–509.

Swofford, D. L. (2001). PAUP*. Phylogenetic analysis using parsimony (*and other methods), version 4. Sunderland, MA: Sinauer Associates.

Tajima, F. & Nei, M. (1984). Estimation of evolutionary distance between nucleotide sequences. *Mol Biol Evol* **1**, 269–285.

Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **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. *J Virol* **72**, 7181–7190.

Tidemann, C. R., Vardon, M. J., Nelson, J. E., Speare, R. & Gleeson, L. J. (1997). Health and conservation implications of Australian bat lyssavirus. *Aust Zool* **30**, 369–376.

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? *Proc Natl Acad Sci U S A* **83**, 3914–3918.

Tordo, N., Poch, O., Ermine, A., Keith, G. & Rougeon, F. (1988). Completion of the rabies virus genome sequence determination: highly conserved domains among the L (polymerase) proteins of unsegmented negative-strand RNA viruses. *Virology* **165**, 565–576.

Tuffereau, C., Benejean, J., Blondel, D., Kieffer, B. & Flamand, A. (1998). Low-affinity nerve-growth factor receptor (P75NTR) can serve as a receptor for rabies virus. *EMBO J* **17**, 7250–7259.

Tuffereau, C., Desmezieres, E., Benejean, J., Jallet, C., Flamand, A., Tordo, N. & Perrin, P. (2001). Interaction of lyssaviruses with the low-affinity nerve-growth factor receptor $p75^{NTR}$. *J Gen Virol* **82**, 2861–2867.

van Regenmortel, M. H. V., Fauquet, C. M., Bishop, D. H. L., Carstens, E. B., Estes, M. K., Lemon, S. M., Maniloff, J., Mayo, M. A., McGeoch, D. J., Pringle, C. R. & Wickner, R. B. (2000). *Virus Taxonomy. Seventh Report of the International Committee on Taxonomy of Viruses*. San Diego: Academic Press.

Warrilow, D., Serafin, I., Harrower, B. & Smith, G. A. (2002). Sequence analysis of an isolate from a fatal human infection with Australian bat lyssavirus. *Virology* **297**, 109–119.

Wiktor, T. J., Gyorgy, E., Schlumberger, D., Sokol, F. & Koprowski, H. (1973). Antigenic properties of rabies virus components. *J Immunol* **110**, 269–276.

Woelk, C. H. & Holmes, E. C. (2001). Variable immune-driven natural selection in the attachment (G) glycoprotein of respiratory syncytial virus (RSV). *J Mol Evol* **52**, 182–192.

Woelk, C. H., Jin, L., Holmes, E. C. & Brown, D. W. (2001). Immune and artificial selection in the haemagglutinin (H) glycoprotein of measles virus. *J Gen Virol* **82**, 2463–2474.

Woelk, C. H., Pybus, O. G., Jin, L., Brown, D. W. & Holmes, E. C. (2002). Increased positive selection pressure in persistent (SSPE) versus acute measles virus infections. *J Gen Virol* **83**, 1419–1430.

Yang, Z. (1997). PAML: a program package for phylogenetic analysis by maximum likelihood. *Comput Appl Biosci* **13**, 555–556.

Yang, Z., Nielsen, R., Goldman, N. & Pedersen, A. M. K. (2000). Codon-substitution models for heterogeneous selection pressure at amino acid sites. *Genetics* **153**, 431–449.