Quantifying antigenic relationships among the Lyssaviruses

Quantifying Lyssavirus antigenic relationships

Authors

Horton, D.L.,1,2,3*; McElhinney, L.M.,1,7; Marston, D.A.,1; Wood, J.L.N.,2; Russell, C.A.,3,6; Lewis, N.,2,3; Kuzmin, I.V.,4; Fouchier, R.A.M.,5; Osterhaus, A.D.M.E.,5; Fooks, A.R.,1,7; Smith, D. J.,3,5,6

1Rabies and Wildlife Zoonoses, Veterinary Laboratories Agency Weybridge, New Haw, Addlestone, Surrey KT15 3NB, UK.

2Cambridge Infectious Diseases Consortium, Department of Veterinary Medicine, Cambridge, CB2 1DQ, UK.

3Department of Zoology, Cambridge University, Downing Street, Cambridge, CB2 3EJ, UK.

4Centers for Disease Control and Prevention, Division of Viral and Rickettsial Diseases, 1600 Clifton Road, MSG-33, Atlanta, GA 30333, USA.

5Department of Virology, Erasmus Medical Center, Dr. Molewaterplein 50, 3015GE Rotterdam, Netherlands

6Fogarty International Center, National Institutes of Health, Bethesda, Maryland 20892, USA

7The National Centre for Zoonosis Research, University of Liverpool, Leahurst, Chester High Road, Neston, CH64 7TE, UK.

*Corresponding author, Address: Rabies and Wildlife Zoonoses, Veterinary Laboratories Agency Weybridge, New Haw, Addlestone, Surrey KT15 3NB, UK.

Email: d.horton@vla.defra.gsi.gov.uk Telephone +44 (0) 1932 341111 Fax +44(0)1932 347046

Abstract

All lyssaviruses cause fatal encephalitis in mammals. There is sufficient antigenic variation within the genus to cause variable vaccine efficacy but this variation is difficult to characterize quantitatively: sequence analysis cannot yet provide detailed antigenic information, and antigenic neutralization data have been refractory to high-resolution robust interpretation. Here, we address these issues by using state-of-the-art antigenic analyses to generate a high-resolution antigenic map of a global panel of 25 lyssaviruses. We compared the calculated antigenic distances with viral glycoprotein ectodomain sequence data. Although 67% of antigenic variation was predictable from glycoprotein amino acid sequence, there are in some cases substantial differences between genetic and antigenic distances thus highlighting the risk of inferring antigenic relationships solely from sequence data at this time. These differences included epidemiologically important antigenic differences between vaccine strains and wild-type rabies viruses. Further, we quantitatively assessed the antigenic relationships measured by rabbit, mouse, and human sera, validating the use of non-
human experimental animals as a model for determining antigenic variation in 
humans. Passive immune globulin use is a crucial component of rabies post-exposure 
prophylaxis, and here we also show that it is possible to predict the reactivity of 
immune globulin against divergent lyssaviruses.

Introduction
Rabies remains a globally important zoonosis, despite being one of the oldest 
recognized infectious diseases (55),(27). The majority of rabies in terrestrial animals 
and humans is caused by classical rabies virus (RABV), a lyssavirus in the family 
Rhabdoviridae. Since the 1950s, many related lyssaviruses have been identified which 
are capable of causing clinical rabies. The majority of those viruses have been isolated 
from bats (Chiroptera), including four divergent viruses, isolated in separate 
geographic locations throughout Eurasia in the past eighteen years (2, 29, 31). The 
Chiroptera, therefore, represent a global reservoir for lyssaviruses, creating the 
potential for ‘spill-over’ infection to terrestrial mammals, including humans. 
Occasionally transmission will occur between members of a new host species with 
potential for a subsequent adaptation in that species (35). Phylogenetic evidence 
suggests that one or more host switching events from bats into terrestrial mammals 
are originally responsible for the on-going global epidemic of terrestrial RABV (6).

Pre- or post-exposure prophylaxis, using vaccination and passive immune globulin 
administration according to World Health Organisation (WHO) guidelines is currently 
the only effective way to prevent rabies after infection with a lyssavirus (1). The 
efficacy of both active and passive immunisation is likely to be affected by antigenic 
differences between viruses. The lyssavirus trimeric glycoprotein is the primary 
surface antigen, the major target for neutralizing antibodies (8), and is involved in cell 
binding and entry (34, 36, 53). Antigenic sites on the glycoprotein have been 
described using monoclonal antibody escape mutants (47, 8, 51,16). These studies 
have elucidated two major sites (sites II and III) and multiple minor sites. Although 
estimates can be made of antigenic differences using information regarding these 
known antigenic sites, protein structure and amino acid properties, predictions of the 
relative importance of those sites, and specific mutations within those sites cannot be 
quantitatively tested without a method to reliably measure antigenic effect.
Use of serological cross neutralization data to measure antigenic difference is limited by the reliability of the serological test, and more importantly by paradoxes, or irregularities in the data. These irregularities include higher heterologous than homologous titers and individual variations between sera raised against the same antigen (22, 52). Hence serological data is considered to have low resolution, and is often only used qualitatively. Despite these difficulties, studies have attempted to further quantify antigenic differences among lyssaviruses. Badrane and Tordo (5) showed a linear correlation between the glycoprotein amino acid identity and four cross neutralization titers. Other studies have demonstrated variable serological cross reactivity between European Bat Lyssaviruses (EBLV) and RABVs (10-11), and suggested that antigenic relationships between EBLV-1 and EBLV-2 may not be fully reflected in the genetic relationships (41). Recent investigations into the efficacy of biologics against the Eurasian lyssaviruses showed an array of relatedness between lyssavirus species with, for example, a murine anti-ARAV serum neutralizing KHUV and ARAV equally, but an anti-KHUV serum less effective at neutralizing ARAV than KHUV (22). Until recently, however, there were no established tools for the quantitative analysis of antigenic data.

Here we resolve the issue of quantitative interpretation of antigenic data using antigenic cartography. Antigenic cartography is a theory and associated computational method that resolves the paradoxes in the interpretation of antigenic data and makes possible high-resolution quantitative analyses, and visualizations, of binding assay data (52, 15, 49, 25, 20, 44).

Integrating antigenic data with direct sequencing data, here we quantify the antigenic and genetic variation among a global panel of lyssaviruses, including representatives from all lyssavirus species. Furthermore we address two key issues in the development of anti-lyssavirus biologics: the appropriateness of animal models, and the development of efficacious alternatives to human rabies immune globulins.

**Materials and Methods**
Virus isolates. A panel of 25 lyssaviruses (Table 1) were grown in baby hamster kidney (BHK-21) or murine neuroblastoma (N2A) cell culture. Viruses were passaged in tissue culture between two and eight times to obtain satisfactory titers. Tissue culture supernatant (TCSN) was harvested, aliquoted and stored at -80°C.

Sequence data. Complete glycoprotein ectodomain sequences were obtained for each virus. Nine of 25 sequences were obtained from GenBank and the remaining 16 samples were sequenced as part of this study as previously described (39). RNA was extracted from the brains of mice infected with the TCSN used in the neutralization tests, or directly from the TCSN, to minimise effect of any amino acid mutations that may have occurred during passage in tissue culture (51), and to ensure consistent comparison between antigenic and genetic data. Primer sequences were designed using published virus sequences (39), and are available on request.

Phylogenetic analyses. The most appropriate model of nucleotide substitution was determined to be the general time-reversible model, with the gamma distribution of rate variation among sites determined by the data (ModelTestv3.7). Phylogenetic trees were then inferred using maximum likelihood (ML) methods (PAUP, v4.0) and tree bisection-reconnection branch swapping was used to search for the optimal tree. The sequences were bootstrap re-sampled 100 times to assess robustness of each tree node.

Production of antisera. Polyclonal antisera were raised in 29 naïve adult male New Zealand White rabbits against 18 viruses (one or two rabbits per virus) using betapropiolactone-inactivated TCSN and either Freund’s incomplete adjuvant, or Montanide ISA50V(Seppic). TCSN combined with adjuvant was injected at multiple sites subcutaneously and intramuscularly on days 0, 21, 28, and 35. Sedation, euthanasia and exsanguination were carried out on day 42. The inactivation of viruses prior to inoculation was confirmed with a mouse inoculation test (MIT)(28). Blood samples were taken from selected mice used in the MIT by cardiac puncture under anaesthesia at the time of euthanasia (n=7). After centrifugation of blood samples, serum was removed from the blood pellet, heat inactivated at 56 °C for 30 mins, and then stored at -20 °C. Human rabies immune globulin (HRIG) (Bio Products Laboratory, Herts., UK) was diluted to 20IU/ml in phosphate buffered saline.
Neutralization data. The neutralizing ability of the sera was assessed using fluorescent antibody virus neutralization (FAVN) tests on BHK-21 cells (13), modified for each virus under test. A constant volume of virus (100 TCID$_{50}$/50µl) was added to serial two-fold dilutions of serum in duplicate. The 50% endpoint dilution, where neutralization of the virus ceased, was calculated with the Spearman-Kärber method (3). The virus dose was checked by back-titration on every test and results were rejected if virus dose was <64 or >256 TCID$_{50}$. Sera from a rabbit inoculated with uninfected, beta-propiolactone treated BHK cell TCSN was used as a negative control.

Antigenic cartography. Lyssavirus antigenic maps were generated from the FAVN data using antigenic techniques described previously (15, 52). Briefly; a target distance from a serum to each virus is derived by calculating the difference between the logarithm (log$_2$) reciprocal neutralization titer for that particular virus and the log$_2$ reciprocal maximum titer achieved by that serum (against any virus). Thus the higher the reciprocal titer is, the shorter the target distance. As the log$_2$ of the reciprocal titers are used, a two-fold change in titer will equate to a fixed change in target distance whatever the magnitude of the actual titers (denoted one antigenic unit, AU).

Antigenic cartography (52) was then used to optimize the positions of the viruses and sera relative to each other on a map, minimizing the sum-squared error between map distance and target distance. Each virus is therefore positioned by multiple sera, and the sera themselves are also positioned only by their distances to the viruses. Hence sera with different neutralization profiles to the virus panel are in separate locations on the map but contribute equally to positioning of the viruses. To increase the likelihood of obtaining the best fit of map distances to the neutralization data, multiple random restart optimisations (25-100) were carried out, creating multiple maps which we ranked in order of total error and quantitatively compared for self consistency. The position of points on the map can have more than three co-ordinates, and therefore mathematically it is possible for the maps occupy more than the conventional three dimensions. Although one cannot easily visualize maps in more than three dimensions, all quantitative measurements, such as the euclidean or (‘straight-line’) distance between points, are as straightforward in more than three dimensions as they are in three dimensions or fewer (see (52) supplementary online material). To
determine the optimum dimension (D) for these data, the resolution of the maps in each dimension was determined by first making maps in 1, 2, 3, 4 and 5D with a random 10% of titers omitted, and then predicting those omitted titers using the distances from the maps.

Comparison of sera. In addition to the rabbit sera used to make the antigenic map, seven mouse sera (from MITs) and HRIG were tested against a subset of viruses. To compare the mouse and rabbit sera, a separate antigenic map was made using only mouse serum titers against a subset of viruses (n=9). Virus to virus distances measured from that mouse map were compared with virus to virus distances on the rabbit serum map, to quantify any difference between antigenic relationships determined by mice and rabbits. There were too few sera to make separate maps from HRIG so we used an indirect method to compare HRIG with the rabbit sera. HRIG was tested for its ability to neutralise a subset of the viruses. An antigenic map was then constructed using the rabbit serum titers and the HRIG titers. In this case, however, the virus positions were determined by the rabbit titers only, and not HRIG titers. When distances are measured from the map, we can compare how effectively the HRIG titers fit with the antigenic positions of the viruses as determined by the rabbit serum titers. If HRIG and rabbit sera have the same breadth of neutralization titers and therefore determine antigenic differences at the same scale, then the rabbit-serum antigenic map should be a good characterization of the HRIG titers. This characterization was assessed quantitatively by plotting target distance as a function of map distance, allowing comparison of the slope and intercept of linear models applied to the data. This indirect method was then also applied to compare mouse and rabbit sera.

‘Genetic’ maps were also made of the viruses based on their amino acid sequence variation. The same multidimensional scaling techniques used for antigenic data were applied to a distance matrix generated from the number of glycoprotein ectodomain amino acid substitutions between viruses (52).

Results
**Genetic variation.** Phylogenetic analysis of the glycoprotein ectodomain of the panel of viruses (Figure 1) shows evolutionary relationships among the lyssaviruses, inferred using maximum likelihood methods. The majority of evolutionary relationships are well supported by bootstrap analysis. The close relationship of the vaccine strains Flury-LEP, Pittman Moore (PM) and challenge virus standard (CVS) is clear, as is the distinction between New-World (RV108, RV51) and Old-World RABV (RV102, RV313, RV410, RV437 and RV1237). The ABLV (RV634) is the closest of the lyssavirus genotypes to RABV, followed by EBLV-2s (RV628, RV1332 and RV1787), then EBLV-1 (RV9, RV1423 and RV20). MOKV (RV4) and LBV (RV1) are approximately equally distinct from the remaining viruses. The tree suggests that IRKV virus is most closely related to EBLV-1 and KHUV is most closely related to EBLV-2.

The lyssavirus amino acid map (Figure 2) shows the relationships of all viruses based on the glycoprotein amino acid homology. All eleven lyssavirus species are distinct from each other. A good correlation between table distance and map distance (r=0.96, p<0.001) indicates two dimensions are adequate for representing the amino acid distances.

**Antigenic variation.**

Antigenic maps were made in multiple dimensions (1D, 2D, 3D, 4D and 5D). To determine the appropriate number of dimensions and the resolution of the map 25 antigenic maps were made, each with a different random 10% of titers omitted, and those titers predicted using the maps. The average prediction error was 1.36 (SE 0.16) antigenic units in 2D, 1.22 (SE 0.17) in 3D, 1.20 (SE 0.16) in 4D, and 1.20 (SE 0.16) in 5D. Resolution therefore increases with increasing dimension, but incremental increase in precision becomes negligible beyond three dimensions for this data set, suggesting that 3D maps are sufficient for visualizing the lyssavirus antigenic data.

Figure 3 is a three dimensional antigenic map, showing the antigenic relationships among the panel of 25 lyssaviruses, color coded according to species. Sera are represented by open boxes. Antigenic distances can be measured from the antigenic map. Table 2 shows the mean and range of antigenic distances between
Despite their wide host range and geographical origin (Table 1), the wild-type RABV (WT-RABV) included in the study are antigenically similar. However, the laboratory adapted strains CVS and PM are antigenically distinct from the WT-RABV (Figure 3), despite being derived from WT-RABVs and genetically closely related to them. Conversely, ABLV is antigenically indistinguishable from the WT-RABVs, but genetically divergent. Of the remaining genotypes, EBLV-1 and EBLV-2 are closest to RABV, followed by DUVV. Previously, EBLV-1 and 2 were suggested as equally antigenically distinct from RABV (41), which is reflected in our findings. However, EBLV-1 viruses appear antigenically further than EBLV-2 from CVS and PM (Table 2). The viruses IRKV and ARAV are antigenically more similar to EBLV-2 than EBLV-1, and KHUV is closer to RABV than either of the EBLVs.

The LBV and MOKV have been placed in a separate phylogroup, based on genetic distance, limited cross neutralization, and differences in pathogenicity (5, 7). In our study, as in others (22), we saw a low but detectable degree of cross reactivity between phylogroup I and II for a small number of sera, hence we were able to position the phylogroup II viruses relative to the others on the antigenic maps. Both LBV and MOKV are twice as far from RABV as the EBLVs on the antigenic map. Neutralization of WCBV by all sera raised with both phylogroup I and II viruses was very weak, indicating that WCBV is distinct from all other viruses, on average 9 AUs from RABV, equivalent to approximately a 500 fold difference in neutralizing antibody titer.

Sera comparisons. Distances from virus to virus on a map made using mouse sera (not shown) correlated closely with distances from the rabbit serum antigenic map (r=0.86 [95%CI 0.74-0.92], p<0.01, residual standard error 0.76). The correlation between the target distances and map distances for the mouse sera, and HRIG, are all similar to that for the rabbit sera. Neither the slopes nor intercepts of linear models applied to the data are significantly different from those applied to the rabbit data (Figure 4), showing that for the data tested, antigenic distances determined by sera from animal models are the same as those determined by human sera.
Predicting antigenic distance from amino acid sequence homology

Pair-wise comparison of all antigenic distances with genetic distances shows a statistically significant correlation (Pearson product-moment correlation coefficient, \( r = 0.82 \) [95% CI 0.78 to 0.85], \( p < 0.001 \), residual standard error 1.75) (Figure 5). In addition, a linear regression model predicts on average, one unit of antigenic change (95% CI 0.93-1.07, \( p < 0.001 \)), equivalent to a two-fold change in antibody titer, for every 21 amino acid substitutions (4.8%) in the glycoprotein ectodomain. The correlation between antigenic and genetic distances when calculated only for viruses in the same phylogroup is lower than the correlation for all viruses, but remains statistically significant (\( r = 0.64 \) [95% CI 0.56-0.71], \( p < 0.001 \), residual standard error 1.63).

Discussion

We have described the antigenic relationships among a panel of 25 lyssaviruses using serological binding assay data and antigenic cartography. This approach has quantified clinically important antigenic differences between lyssaviruses, shown those differences are equivalent for mouse, rabbit, and pooled human sera and allowed integration of quantitative antigenic data with genetic distances.

These data give a precise estimate for the correlation between genetic and antigenic distances (95% CI for \( r = 0.81 \) to 0.88), an improvement in accuracy over the current estimate (95% CI for \( r = 0.39 \) to 1.00, \( p = 0.08 \) calculated from published data (5)). Increased accuracy allows the evaluation glycoprotein amino acid sequence homology as a predictor of antigenic difference. Fitting a linear regression model to our data demonstrates that on average, a 4.8% change in the glycoprotein ectodomain amino
acid sequence will cause one antigenic unit of difference between viruses (equivalent
to a two-fold change in antibody titer (95% CI 0.93-1.07 AU, p <0.001) (Figure 5). A
linear regression model applied to the log2 of previously published data, gives a
similar mean of a two-fold change in antibody titer per 5.5 percent change in
glycoprotein amino acid sequence homology, but with a much larger confidence
interval (-0.8 to 4.4 two fold dilutions, p=0.296). Despite a good correlation between
genetic and antigenic distance among the lyssaviruses, over 30% of the variance in
antigenic distance cannot be predicted by the number of amino acid substitutions
between viral glycoprotein ectodomains, illustrating the difficulty in interpreting
antigenic differences using gene sequence alone. Although these and previous studies
have used the entire glycoprotein sequence, the techniques applied here could be
applied to specific regions of the glycoprotein, for example previously reported
antigenic sites.

Integration of genetic and antigenic data allows identification of viruses where there
are differences between genetic and antigenic relationships. For example, the
antigenic positions of IRKV and ARAV, closer to EBLV-2; and KHUV, closer to
RABV would not be expected from genetic relationships. Phylogenetic analysis using
the glycoprotein ectodomain here, similar to that using the entire glycoprotein and
entire nucleoprotein gene elsewhere (29), suggests that KHUV is more closely related
to EBLV-2, than ARAV or IRKV are to EBLV-2. However, both genetic and
antigenic studies are limited by the existence of only one representative of some
viruses, a problem that also applies to divergent classified lyssavirus species such as
LBV (38).

Comparison with other RNA viruses illustrates the close antigenic relationships
among the lyssaviruses. Studies of human influenza A(H3N2) viruses showed a
similar correlation between amino acid sequence of the main antigenic component of
the virus (the haemagglutinin HA1 domain) and antigenic distance between viruses
(r= 0.81), but a different slope. On average, only 2.9 amino acid substitutions (<1% of
haemagglutinin HA1 domain) caused a 1-unit change in antigenic difference (52).
Antigenic distances measured by different binding assays and on different pathogens
may not be directly comparable, meaning that one antigenic unit (two-fold dilution)
derived for influenza virus will not equal one antigenic unit (two-fold dilution) from lyssavirus studies. Influenza viruses and lyssaviruses both have trimeric surface glycoproteins. However, an average of 13 amino acid substitutions in influenza virus haemagglutinin will cause lack of cross reactivity, compared with lyssaviruses, where viruses with as many as 100 amino acid substitutions still show significant cross reactivity. Such conservation in the phenotype of the key antigenic determinant of lyssaviruses is consistent with a low immune selective pressure upon lyssaviruses (9, 14, 24) in comparison with pathogens such as influenza that are under large selection pressure (42, 52). Although low immune selective pressure could be expected from the natural history of RABV in terrestrial animals, where infection classically leads to death, the dynamics of RABV in terrestrial animals may not be applicable to bats, particularly in light of evidence of high seroprevalence against RABV and other lyssaviruses in apparently healthy conspecifics (23, 30, 54).

A limitation of all in-vitro antigenic studies is the potential effect of adaptation to cell culture on viral genomic sequence. Sequences generated for this study were taken from cell culture supernatant used in neutralisation assays or after single passage of supernatant in mouse brain, to ensure valid comparison of genetic and antigenic data. Although we cannot rule out potential differences between those viruses in cell culture and the original isolate, previous studies using rabies virus have shown no change in glycoprotein consensus sequence despite 20 passages in cell culture (26).

All current rabies vaccine virus strains are based on classical rabies viruses (18). Evidence suggests these vaccines are fully effective against virtually all RABVs tested to date (4, 11, 33, 37, 46, 50) but not against distantly related viruses in phylogroup II or WCBV (5, 43). In addition, the reported variable efficacy of vaccines against the EBLVs and DUVV (17), (32), (11), and recent evidence of variable efficacy against ARAV, KHUV and IRKV (22), suggest that there may be a gradual loss of protection as viruses become antigenically distant from vaccine strains. This gradation in protection is the case for other viruses (42), (52). Although protection can ultimately only be tested in challenge models, reliably quantifying the antigenic differences among divergent lyssaviruses is an important step towards predicting differences in vaccine protection (5).
In addition to active vaccination, passive immunisation against rabies remains a critical part of post exposure prophylaxis (1). The gold standard, human rabies immune globulin (HRIG), is expensive and in short supply, prompting attempts to find cheaper alternatives including cocktails of monoclonal antibodies (21, 40). In some studies dose dependent survival has been demonstrated following immune globulin administration in animal models (21, 48) suggesting the potency of passive immune globulin may be related to the neutralizing titer. If true, antigenic differences as measured by neutralizing antibodies are of direct relevance to protection provided by passive immune globulin treatments. HRIG titers correspond to rabbit titers and are therefore predictable using antigenic maps. Further investigation is necessary to determine whether candidate monoclonal antibody cocktails can be predicted with similar accuracy.

Sera from animal models are widely used to investigate antigenic differences between lyssaviruses (5, 22). However, all require extrapolation to species of clinical interest, which has thus far been largely unsubstantiated. Here we have demonstrated that antigenic differences between viruses determined by a variety of species are equivalent. These findings validate the use of non-human experimental animals as a model for determining antigenic variation relevant to humans.

The Chiroptera are increasingly implicated as reservoirs for many zoonotic viral diseases (12). The lyssaviruses provide one globally widespread example with at least one antigenically distinct lyssavirus isolated from bats on all continents except Antarctica. Future elimination of lyssaviruses from bats by human intervention is at best very optimistic. The threat posed by lyssaviruses in bats is therefore global and likely to be continuous. Recent human deaths due to DUVV (45) and EBLV-2 (19) highlight the significance of spill over events into humans. More alarming is the possibility of a bat lyssavirus adapting to a terrestrial host, as hypothesized to have created the current global epidemic of terrestrial RABV (6) with widespread and prolonged consequences. A recent example of a bat RABV host-switching to skunks (35), illustrates the ease with which this adaptation can occur. With knowledge of the alarming clinical manifestations and extremely high mortality rates caused by
lyssaviruses, along with increasing globalization and the altered interface between humans and wildlife, a better understanding of the antigenic as well as genetic relatedness among lyssaviruses is vital. The methods presented here provide a quantitative method to test predictions regarding the antigenic effect of amino acid substitutions, and those antigenic effects in turn can be used to make predictions regarding efficacy of biologicals.

Acknowledgements

D.L.H., J.L.N.W. and N.L. were supported by the Cambridge Infectious Diseases Consortium, a DEFRA Veterinary Training and Research Initiative. D.A.M., L.M.M. and A.R.F were supported by DEFRA grant numbers SEO420, SEO423, and SEO424. D.J.S., and C.A.R. were supported by an NIH Director’s Pioneer Award, part of the NIH roadmap for medical research, through grant DP1-OD000490-01; 223498 EMPERIE, an FP7 grant from the European Union; and program grant P0050/2008 from the Human Frontier Science Program. JLNW was also supported by the Alborada Trust and the RAPIDD program of the Science & Technology Directorate, Department of Homeland Security. C.A.R. was also supported by a University Research Fellowship from the Royal Society, London and a research fellowship from Clare College, Cambridge. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. For technical support, we are indebted to VLA Reagent production unit, VLA Sequencing department, Colin Black, Hooman Goharriz, Erasmus Medical Centre, and Eugene Skepner. We also thank Charles Rupprecht, Nicholas Johnson, Thomas Mueller, David Hayman and Sharon Brookes, for helpful discussions and contribution of reagents.


### Table 1. Virus panel

<table>
<thead>
<tr>
<th>ID</th>
<th>Species</th>
<th>Host</th>
<th>Country</th>
<th>Year</th>
<th>Original ID</th>
<th>Genbank accession no.</th>
</tr>
</thead>
<tbody>
<tr>
<td>CVS-11</td>
<td>RABV</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>EU352767</td>
</tr>
<tr>
<td>PM</td>
<td>RABV</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>AJ871962</td>
</tr>
<tr>
<td>LEP</td>
<td>RABV</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>GU936866</td>
</tr>
<tr>
<td>RV51</td>
<td>RABV</td>
<td>Fox</td>
<td>USA (Michigan)</td>
<td>1986</td>
<td>86-1393</td>
<td>GU936880</td>
</tr>
<tr>
<td>RV108</td>
<td>RABV</td>
<td>Bat</td>
<td>Chile</td>
<td>1985</td>
<td>85/120</td>
<td>GU936868</td>
</tr>
<tr>
<td>RV313</td>
<td>RABV</td>
<td>Red fox</td>
<td>Germany</td>
<td>1990</td>
<td>G1455</td>
<td>GU936875</td>
</tr>
<tr>
<td>RV410</td>
<td>RABV</td>
<td>Mongoose</td>
<td>South Africa</td>
<td>1990</td>
<td>332/90</td>
<td>GU936878</td>
</tr>
<tr>
<td>RV437</td>
<td>RABV</td>
<td>Raccoon</td>
<td>Dog Estonia</td>
<td>-</td>
<td>269</td>
<td>GU936879</td>
</tr>
<tr>
<td>RV1237</td>
<td>RABV</td>
<td>Deer</td>
<td>Serbia</td>
<td>1999</td>
<td>329/99</td>
<td>GU936869</td>
</tr>
<tr>
<td>RV102</td>
<td>RABV</td>
<td>Fox</td>
<td>Morocco</td>
<td>1988*</td>
<td>RV102</td>
<td>GU936867</td>
</tr>
<tr>
<td>RV334</td>
<td>RABV</td>
<td>Vaccine</td>
<td>China (Beijing)</td>
<td>-</td>
<td>SRFV 289</td>
<td>GU936876</td>
</tr>
<tr>
<td>RV1</td>
<td>LBV</td>
<td>Bat</td>
<td>Nigeria</td>
<td>1956</td>
<td>LBV</td>
<td>EF547431</td>
</tr>
<tr>
<td>RV4</td>
<td>MOKV</td>
<td>Shrew</td>
<td>Nigeria</td>
<td>1968</td>
<td>MOKV</td>
<td>GU936877</td>
</tr>
<tr>
<td>RV131</td>
<td>DUVV</td>
<td>Bat</td>
<td>Zimbabwe</td>
<td>1986</td>
<td>RS16(Z12)</td>
<td>GU936870</td>
</tr>
<tr>
<td>RV1423</td>
<td>EBLV1</td>
<td>Bat</td>
<td>Germany</td>
<td>1998</td>
<td>934</td>
<td>GU936872</td>
</tr>
<tr>
<td>RV20</td>
<td>EBLV1</td>
<td>Bat</td>
<td>Denmark</td>
<td>1986</td>
<td>RA552(86)</td>
<td>GU936874</td>
</tr>
<tr>
<td>RV9</td>
<td>EBLV1</td>
<td>Bat</td>
<td>Germany</td>
<td>1968</td>
<td>RV9</td>
<td>EU352768</td>
</tr>
<tr>
<td>RV1332</td>
<td>EBLV2</td>
<td>Bat</td>
<td>UK</td>
<td>2002</td>
<td>105/02</td>
<td>GU936871</td>
</tr>
<tr>
<td>RV628</td>
<td>EBLV2</td>
<td>Bat</td>
<td>UK</td>
<td>1996</td>
<td>96/18</td>
<td>GU936882</td>
</tr>
<tr>
<td>RV1787</td>
<td>EBLV2</td>
<td>Bat</td>
<td>UK</td>
<td>2004</td>
<td>603-04</td>
<td>EU352769</td>
</tr>
<tr>
<td>RV634</td>
<td>ABLV</td>
<td>Bat</td>
<td>Australia</td>
<td>1996</td>
<td>96-0648</td>
<td>GU936883</td>
</tr>
<tr>
<td>IRKV</td>
<td>ABLV</td>
<td>Bat</td>
<td>Siberia</td>
<td>2002</td>
<td>-</td>
<td>EF614260</td>
</tr>
<tr>
<td>KHUV</td>
<td>ABLV</td>
<td>Bat</td>
<td>Tajikistan</td>
<td>2001</td>
<td>-</td>
<td>EF614261</td>
</tr>
<tr>
<td>ARAV</td>
<td>ABLV</td>
<td>Bat</td>
<td>Kyrgyzstan</td>
<td>1991</td>
<td>-</td>
<td>EF614259</td>
</tr>
<tr>
<td>WCBV</td>
<td>ABLV</td>
<td>Bat</td>
<td>Russia</td>
<td>2002</td>
<td>-</td>
<td>EF614258</td>
</tr>
</tbody>
</table>

Abbreviations: CVS, Challenge virus standard; LEP, Flury low egg passage; PM, Pittman Moore; RABV, Rabies virus; LBV, Lagos Bat virus; MOKV, Mokola virus; DUVV, Duvenhage virus; EBLV, European bat lyssavirus; ABLV, Australian bat lyssavirus; IRKV, Irkut virus; KHUV, Khujand virus; ARAV, Aravan virus; WCBV, West Caucasian Bat Virus; -, data not available; *, year first passaged.
Table 2. Intergenotype antigenic and glycoprotein amino acid distances

<table>
<thead>
<tr>
<th></th>
<th>CVS-11</th>
<th>RABV</th>
<th>LBV</th>
<th>MOKV</th>
<th>DUVV</th>
<th>EBLV1</th>
<th>EBLV2</th>
<th>ABLV</th>
<th>IRKV</th>
<th>KHUV</th>
<th>ARAV</th>
<th>WCBV</th>
</tr>
</thead>
<tbody>
<tr>
<td>CVS-11</td>
<td>5.1</td>
<td>7.4</td>
<td>8.5</td>
<td>7.1</td>
<td>8.9</td>
<td>5</td>
<td>4.5</td>
<td>7.8</td>
<td>4.6</td>
<td>5.4</td>
<td>13.1</td>
<td></td>
</tr>
<tr>
<td>RABV</td>
<td>30</td>
<td>9.3</td>
<td>7.1</td>
<td>8.3</td>
<td>5.9</td>
<td>4.7</td>
<td>1.6</td>
<td>4.5</td>
<td>2.2</td>
<td>5.1</td>
<td>8.3</td>
<td></td>
</tr>
<tr>
<td>LBV</td>
<td>162</td>
<td>159</td>
<td>6.6</td>
<td>12.4</td>
<td>13</td>
<td>9</td>
<td>8.8</td>
<td>10.3</td>
<td>8.2</td>
<td>8.3</td>
<td>14.2</td>
<td></td>
</tr>
<tr>
<td>MOKV</td>
<td>162</td>
<td>160</td>
<td>87</td>
<td>14.1</td>
<td>12.1</td>
<td>9.9</td>
<td>7.5</td>
<td>9.3</td>
<td>7.3</td>
<td>9.6</td>
<td>9.1</td>
<td></td>
</tr>
<tr>
<td>DUVV</td>
<td>101</td>
<td>101</td>
<td>167</td>
<td>168</td>
<td>6.6</td>
<td>4.7</td>
<td>7.3</td>
<td>7.7</td>
<td>7.3</td>
<td>5.4</td>
<td>15.4</td>
<td></td>
</tr>
<tr>
<td>EBLV1</td>
<td>101</td>
<td>99</td>
<td>161</td>
<td>160</td>
<td>72</td>
<td>4.8</td>
<td>5.3</td>
<td>3.3</td>
<td>5.4</td>
<td>5.2</td>
<td>9.8</td>
<td></td>
</tr>
<tr>
<td>EBLV2</td>
<td>86</td>
<td>84</td>
<td>164</td>
<td>166</td>
<td>71</td>
<td>71</td>
<td>3.4</td>
<td>3.9</td>
<td>3.1</td>
<td>1.5</td>
<td>11.5</td>
<td></td>
</tr>
<tr>
<td>ABLV</td>
<td>74</td>
<td>72</td>
<td>155</td>
<td>161</td>
<td>92</td>
<td>84</td>
<td>65</td>
<td>3.8</td>
<td>0.9</td>
<td>3.8</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>IRKV</td>
<td>113</td>
<td>111</td>
<td>158</td>
<td>156</td>
<td>92</td>
<td>71</td>
<td>87</td>
<td>102</td>
<td>3.4</td>
<td>3.7</td>
<td>8.2</td>
<td></td>
</tr>
<tr>
<td>KHUV</td>
<td>79</td>
<td>76</td>
<td>165</td>
<td>168</td>
<td>78</td>
<td>71</td>
<td>39</td>
<td>59</td>
<td>85</td>
<td>3.2</td>
<td>9.1</td>
<td></td>
</tr>
<tr>
<td>ARAV</td>
<td>87</td>
<td>85</td>
<td>159</td>
<td>164</td>
<td>69</td>
<td>52</td>
<td>50</td>
<td>69</td>
<td>79</td>
<td>51</td>
<td>11.4</td>
<td></td>
</tr>
<tr>
<td>WCBV</td>
<td>197</td>
<td>200</td>
<td>182</td>
<td>192</td>
<td>201</td>
<td>196</td>
<td>198</td>
<td>195</td>
<td>197</td>
<td>198</td>
<td>195</td>
<td></td>
</tr>
</tbody>
</table>

Above diagonal: Antigenic distances measured from antigenic map, in antigenic units.
Below diagonal: number of amino acid substitutions in glycoprotein ectodomain.
Where there are multiple representatives for each genotype on the antigenic map, the mean antigenic distance and number of amino acid substitutions is given.
Figure 1. Phylogenetic tree of 25 Lyssavirus glycoprotein ectodomain nucleotide sequences. Sequences were aligned using Clustal X2. Trees were produced using a maximum likelihood algorithm using a GTR+I+Γ4 model; the consensus was generated and is presented with bootstrap values at key nodes. Trees were rooted using WCBV as an outgroup. Isolates are colored by species (Table 1 and Figure 3).

Figure 2. Amino acid sequence variation map. Viruses are positioned relative to each other using multidimensional scaling and target distances derived from the number of amino acid substitutions between viruses (Table 2). The map is in two dimensions, and is oriented and colored according to the antigenic map (Figure 3). The scale bar represents 50 amino acid substitutions.

Figure 3. Three dimensional antigenic map. Viruses (spheres) and sera (open boxes) are positioned such that the distance from each serum to each virus is determined by the neutralization titer. Multidimensional scaling is used to position both sera and viruses relative to each other, so orientation of the map within the axes is free. Scale bar shows one antigenic unit (AU), equivalent to a two-fold dilution in antibody titer. (Viewed using Pymol, DeLano Scientific LLC, San Carlos, California, U.S.A).

Figure 4. Comparison of species’ responses. Target distance derived from titers of rabbit sera (red, A-B), HRIG (blue, A), and mouse sera (blue, B) as a function of antigenic map distance determined only by the rabbit sera. Correlation between rabbit map distance and target distances from other species sera gives an indication of whether antigenic relationships determined by each species are similar. Correlation between the target distances and rabbit-map distances for the mouse sera (r= 0.83 [95%CI 0.72-0.90], p<0.001, residual standard error 1.09), and HRIG (r=0.88 [95%CI 0.55-0.97], p<0.001, residual standard error 1.60), are similar to that for the rabbit sera (r=0.92 [95%CI 0.90-0.93], p<0.001, residual standard error 0.82), and neither the slopes nor intercepts of linear models applied to the data are significantly different from those applied to the rabbit data. ‘i’=intercept, ‘s’=slope.
Figure 5. Plot of pair-wise antigenic and genetic distances between viruses. Antigenic distances are measured from the antigenic map, in antigenic units. Genetic distances between viruses are represented as number of amino acid substitutions in the glycoprotein ectodomain. The line represents a linear regression model ($R^2 = 0.67$, $p < 0.01$), which predicts, on average, an antigenic distance change of one antigenic unit, equivalent to a two-fold change in neutralization titer, per 21 amino acid substitutions (4.8% of glycoprotein ectodomain).