Onderstepoort Journal of Veterinary Research, 60:277–282 (1993)

Taxonomy and evolutionary studies on Lyssaviruses with special reference to Africa

H. BOURHY[1], B. KISSI[1] and N. TORDO[2]

ABSTRACT

BOURHY, H., KISSI, B. & TORDO, N. 1993. Taxonomy and evolutionary studies on Lyssaviruses with special reference to Africa. *Onderstepoort Journal of Veterinary Research*, 60:277–282

Currently, the *Lyssavirus* genus is composed of four serotypes and the ungrouped European bat lyssaviruses (EBL). Using previously described PCR amplification and direct sequencing methods, the N genes of 70 representative lyssaviruses collected from 40 different countries were examined. From the results, a phylogenetic tree with six branches (genotypes) was constructed: genotypes 1–4 correlated with the classification of serotypes 1–4 and in addition EBL 1 and EBL 2 were identified as two further genotypes (5 and 6). Analysis of the 51 rabies (genotype 1) viruses in the study identified nine distinct groups which could be correlated with their geographical origins.

A comparison of the six genotypes with four principal *Vesiculovirus* serotypes showed that their relationship was closer than that of even the two most closely related *Vesiculovirus* serotypes.

Antigenic site sequences of isolates from vaccine/treatment failures were similar to vaccine and field strains, suggesting that failures were not due to genetic variation at the N protein level.

INTRODUCTION

Three *Lyssavirus*, *Vesiculovirus* and plant rhabdovirus genera/groups make up the family Rhabdoviridae. On the basis of serum-neutralization and monoclonal antibody (Mab) studies, the lyssaviruses were subdivided into four serotypes: classical rabies virus strains (serotype 1), Lagos bat virus (serotype 2), Mokola virus (serotype 3) and Duvenhage virus (serotype 4) (WHO 1990); European bat lyssaviruses (EBL) were not classified. Rabies virus (serotype 1) has an almost worldwide prevalence, except for several protected islands; rabies-related viruses (serotypes 2, 3, 4 and EBL) have a large geographical distribution in Africa and in Europe. With the exception of Lagos bat virus which has not been isolated from man, all of the rabies and rabies-related viruses are pathogenic for mammals including man and may lead to a rabies-like encephalitis (King & Crick 1988).

In the last few years, antigenic analyses have played a major role in viral epidemiology by discriminating between variants (Smith 1989; Rupprecht, Dietzschold, Wunner & Koprowski 1991); the advent of molecular biology has introduced tools which achieve this purpose more rapidly and precisely. PCR applied to lyssaviruses provides a progressive technique, starting from diagnosis and ending in a precise genetic characterization by the sequencing of limited areas of the viral genome, thus advantageously replacing serum-neutralization, antinucleocapsid monoclonal antibody (Mab-N) studies and classical cloning in plasmids (Sacramento, Bourhy & Tordo 1991). Comparison between the sequences of Mokola virus and the PV rabies strain allowed the delineation of conserved regions within the lyssavirus genomes that

[1] Unite de la Rage, Institut Pasteur, 25 Rue du Dr Roux, 75724 Paris Cedex 15 France

[2] Laboratoire des Lyssavirus, Institut Pasteur, 25 Rue du Dr Roux, 75724 Paris Cedex 15, France

can be used as sequencing primers (Bourhy, Tordo, Lafon & Surreau 1989). More recently, primers allowing the amplification of the whole nucleoprotein (N) (Bourhy, Kissi, Lafon, Sacramento & Tordo 1992), glycoprotein (G) or the pseudo-gene (Sacramento *et al.* 1991; Tordo, Bourhy & Sacramento 1992) have been defined. Each of these three target regions possesses its own usefulness; the N gene can be used for diagnosis, taxonomy, typing, epidemiology and immunological studies; the glycoprotein gene is useful for epidemiology and for immunological studies and the pseudo-gene can also be used for typing and epidemiology.

In this paper the N gene was chosen for molecular epidemiological studies for several reasons:

- For reasons of comparison, since most of the rabies-related viruses were identified according to their reactivity with Mab-Ns.
- Because of the important role of N-protein in inducing immunity, in particular against infection with heterologous lyssaviruses (Dietzschold, Wang, Rupprecht, Celis, Tollis, Ertl, Heber-Katz, Koprowski 1987; 1990).
- The N gene appears to be a good target for comparison of isolates over relatively long evolutionary periods.
- A study of the N gene demonstrates the adaptability of the PCR technique for both simple diagnosis in which amplification of the N gene has already been achieved to very precise typing by the determination of its nucleotide sequence.

MATERIALS AND METHODS

The viral RNA extraction, cDNA synthesis, PCR amplification, purification of the amplified fragment on 0,7 % NuSieve GTG agarose and its direct sequencing were performed as previously described (Sacramento *et al.* 1991). Where possible, we worked with the original infected brain. Alternatively, suckling mouse brains infected with the original virus at the lowest passage number available were used. Alignment of the deduced amino acid sequence of the N-proteins was performed by the ClustalV package of multiple alignment programmes (Higgins & Sharp 1989). Phylogenetic trees were also generated by the ClustalV package of multiple alignment programmes according to the neighbour joining method (Saitou & Nei 1987).

RESULTS AND DISCUSSION

Taxonomic studies on Lyssaviruses

The N genes of 58 rabies and rabies-related viruses, representative of the diversity of the Lyssavirus genus, were analyzed at the amino acid level and a phylogenetic tree was constructed. Six tight genetic clusters, or genotypes, were distinguished:

- Rabies viruses
- Lagos bat virus
- Mokola virus
- Duvenhage virus
- European bat lyssavirus biotype 1 (EBL1) and
- European bat lyssavirus biotype 2 (EBL2) (Fig. 1).

Genotypes 1–4 corroborated the previous classification of serotypes. However, the genetic grouping appeared more sensitive than the serum-neutralization and Mab-N studies in establishing that EBL1 and EBL2 are considered as independent genotypes, namely genotypes 5 and 6. Genetic grouping also demonstrated more clearly the intergenotype relationships: genotypes 4 and 5 are related and genotypes 2 and 3 are the most phylogenetically distant from the vaccinal and classical rabies viruses of genotype 1. The threshold of similarity below which a new genotype should be defined is in the interval of 97–93 % amino acid similarity (the lowest percentage of amino acid similarity found within one genotype and the highest percentage found between two genotypes) (Table I).

To further define the relationships within the Rhabdoviridae, we performed an alignment of nine lyssavirus N-proteins representative of the six genotypes with those of four principal serotypes of the *Vesiculovirus* genus (VSV Indiana and New Jersey, Chandipura virus and Piry virus) (Banerjee, Rhodes & Gill 1984; Crysler, Lee, Reinder & Prevec 1990; Gallione,

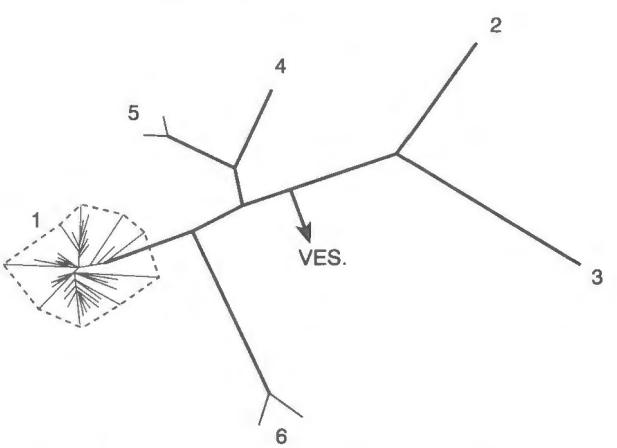

FIG. 1 Radial phylogenetic tree showing the relationships between the different genotypes of lyssaviruses. The alignment of the amino acid sequences of the N protein were performed by the ClustalV package of multiple alignment programmes (Higgins & Sharp 1989). The tree was generated according to the neighbour joining method (Saitou & Nei 1987). The dotted line surrounding the branches corresponding to the strains of genotype 1. The lengths of the branches indicate the phylogenetic distance between the different viruses

TABLE 1 Similarity matrix of the N coding region of representative lyssavirus isolates of genotypes 1, 2, 3, 4, 5 and 6. The percentages of amino acid (top) and nucleotide (bottom) similarity are given

Virus	No.	Mokola	L.b. 8619 NGA	Duv. DUV1 AS	EBL1 8615 POL	EBL1 8918 FRA	EBL2 9007 FIN	EBL2 9018 HOL	Fox 9004 FRA	PV
Mokola			87,1	83,1	81,3	81,6	78,0	77,8	79,6	79,8
L.b.	8619NGA	76,2		85,8	83,1	83,8	79,1	79,1	82,0	81,6
Duv.	DUV1AS	72,7	73,3		92,7	93,3	86,2	85,8	87,1	87,6
EBL1	8615POL	71,4	74,4	69,3		98,7	87,3	86,4	87,6	88,0
EBL1	8918FRA	71,0	74,4	79,8	95,7		88,0	87,6	87,8	88,2
EBL2	9007FIN	71,1	72,4	76,0	77,2	78,1		97,8	86,9	86,9
EBL2	9018HOL	69,9	73,5	75,8	76,1	76,9	96,0		86,4	86,4
Fox	9004FRA	71,0	73,6	74,9	75,6	75,5	74,4	74,5		97,1
PV		72,1	73,2	74,5	75,4	75,3	74,4	74,6	92,8	

L.b. = Lagos bat
NGA = Nigeria
FRA = France

DUV/Duv. = Duvenhage
AS = South Africa
FIN = Finland

EBL = European bat lyssavirus
POL = Poland
HOL = Holland

Greene, Iverson & Rose 1981; Masters & Banerjee 1987) (Fig. 2). In the resulting phylogenetic tree, the branching clearly divides the two genera. The Lyssavirus genus appears more compact than the *Vesiculovirus* genus. Indeed, the whole *Lyssavirus* genus is even less divergent than the two closest serotypes (Indiana and New Jersey) of VSV. Therefore, at the present level of understanding, at the N gene level, the variability is higher among vesiculoviruses than in lyssaviruses.

Determination of geographic lineages

An extensive analysis of genotype 1 rabies viruses collected from all over the world was conducted in order to further evaluate the diversity within the genotype. Nucleotide sequences (1443b) of the whole N-protein gene and the intergenic nucleoprotein-phosphoprotein regions of 52 strains originating from 40 countries were determined. The lineages distinguished by the evolutionary analysis correlated

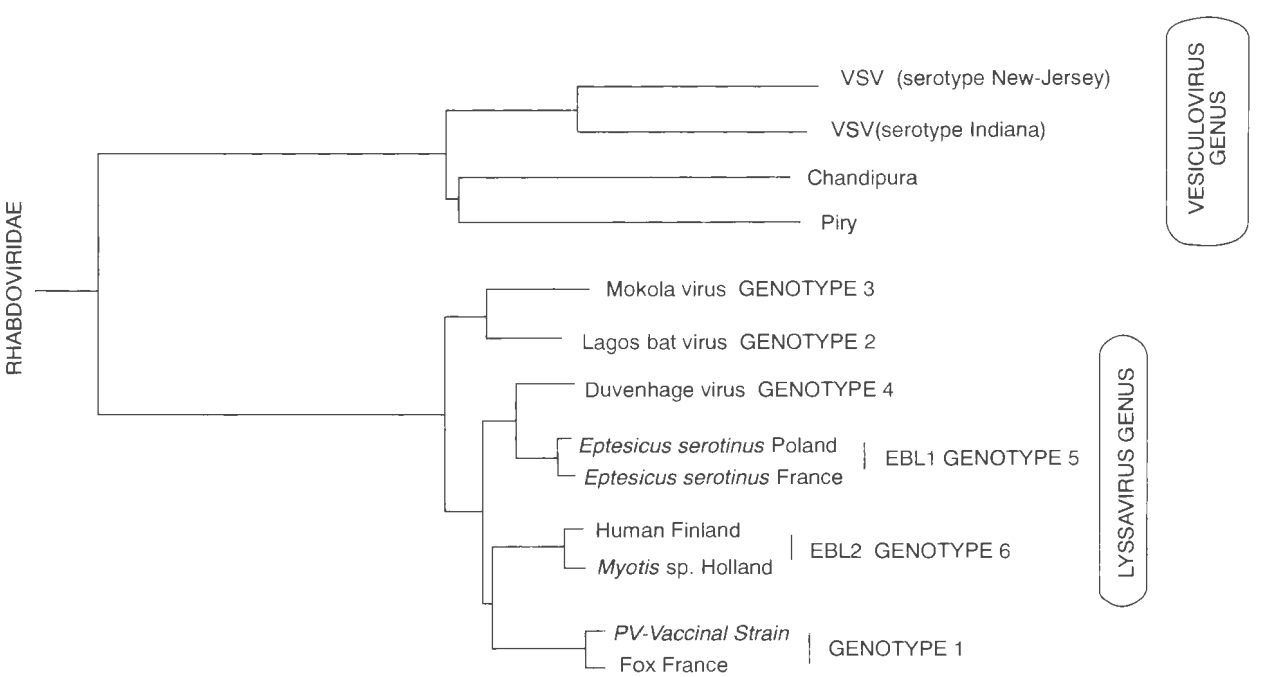

FIG. 2 Phylogenetic tree of lyssaviruses and vesiculoviruses based on the alignment of the amino acid sequence of the N protein. The tree was generated according to the neighbour joining method (Saitou & Nei 1987). The horizontal branch lengths are indicative of evolutionary distances. The vertical lines are for clarity

with the geographic origins (Fig. 3). This analysis allowed us to distinguish nine groups of phylogenetically distinct viruses of genotype 1: Africa 1, Asia (Thailand), Arctic rabies, Europe-Middle East, Latin America I (9110 MEX.nc, 9126 MEX.nc), Latin America 2 (vampire bats) and two groups of vaccinal strains. Viruses of group Europe-Middle East and Africa 1 are closely related. The group Africa 1 can be divided into Africa 1a, circulating in North and East Africa and Africa 1b, circulating in southern Africa. One could postulate the origin of the common ancestor of viruses of groups Africa 1 and Europe-Middle East, but any hypothesis would be speculative. Nevertheless, colonization and the introduction of European dogs in Africa certainly would have favoured the dissemination of European rabies virus to Africa during the last century. Conversely, viruses of the group Africa 2 are not related to any other groups and seem to be indigenous to Africa—they are distributed in the central part of this continent. Nucleotide sequence evidence suggests a common ancestor for fox, raccoon-dog, dog and wolf isolates in Europe. This confirms the hypothesis that the transmission of rabies to foxes and its subsequent adaptation to this species was the source of the rabies outbreak that started in foxes in Europe around 1940.

To complete our study, 12 additional isolates were sequenced only in the region of the 400 nucleotides coding for the amino terminus end of the N-protein. This region is the most variable part of the N-protein gene at both the nucleotide and at the amino acid levels (Fig. 4) and for this reason has an advantage for limited sequence analysis. This allowed us to be more precise in the circulation zone of each virus group. The geographic distribution of 35 viruses groups Africa I and 2 is shown in Fig. 5. One isolate from Egypt could not be related to any group. Sequencing of more strains from this part of Africa should help to characterize viruses related to that of Egypt and perhaps define a third group of viruses circulating in this continent.

Comparison with vaccinal strains

Genetic variability may influence the efficacy of a vaccine intended to prevent rabies. If this proves to be the case, genetic heterogeneity, rapid evolution and consequently antigenic diversity may offer a rabies virus ample opportunity to evade the host immunity induced by rabies vaccines. In fact, homogeneity at the amino acid level between the N-protein of vaccinal strains and wild isolates is relatively high (Fig. 6). This suggests a close evolutionary relationship between these strains and consequently a good level of cross-protection may be expected. Nevertheless, particular attention was given to some strains isolated from reported cases of so-called vaccination or treatment failures in animals or in humans. In these cases the sequences of the antigenic sites were also very similar to those of the vaccinal strains and not different from field strains collected in the same area. Therefore, these failures do not appear to be due to a particular genetic variation, at least at the N-protein level; sequencing of the G-protein antigenic sites may confirm these findings. This is in accordance with previous laboratory immunological investigations which did not reproduce vaccination failures (Bourhy, Lafon, Berthonneau, Renner, Rollin & Sureau 1988). Nevertheless, laboratory studies have shown that current vaccines offer an imperfect protection against genotypes 2 (Tignor & Smith 1972), 4 (Tignor, Murphy, Clark, Shope, Madore, Bauer, Buckley & Meredith 1977) and 5 (Fekadu, Shaddock, Sanderlin & Smith 1988; Lafon, Bourhy & Sureau 1988) and are ineffective against genotype 3 (Koprowski, Wiktor & Abelseth 1985). These distant antigenic relationships can be related to the variability of the T-cell and B-cell antigenic sites on the N-protein and the G-protein (Bourhy, Kissi & Tordo 1993; Tordo, Bourhy, Sather & Ollo 1993). It is presently speculative to draw a direct correlation between conservation of amino acid sequences, particularly of the antigenic sites, and the level of cross-protection between field and vaccinal strains. Nevertheless in the future such comparisons will certainly help to define the molecular basis of cross-protection between lyssaviruses.

CONCLUSION

PCR applied to rabies provides a technique which can be applied progressively to rabies investigation starting from diagnosis and leading to precise genetic characterization by sequencing of limited areas of the viral genome, thus advantageously replacing serum-neutralization and Mab-N studies. For this purpose the study of the N-protein gene appears to be an efficient method for discriminating between genotypes. Genotypic classification is consistent with that of serology but is more rapid, sensitive and precise. Six genotypes may be identified: genotype 2 (Lagos bat virus) and genotype 3 (Mokola virus) are the most phylogenetically distant from the vaccinal and classical rabies viruses of genotype I; genotype 4 (Duvenhage virus) and genotype 5 (EBLI) are closely related to each other; genotype 6 (EBL2) is independent of genotype 5. The threshold of similarity below which a new genotype should be defined is in the interval 97–93 % amino acid similarity. The lineages distinguished by the evolutionary analyses correlate with geographic origin and define phylogenetic relatedness between wild isolates. This technique also permits rapid analysis of mutations in the antigenic sites of strains isolated from reported cases of vaccination or treatment failures in animals or in humans.

ACKNOWLEDGEMENTS

We are grateful to all our correspondents who provided us with many specimens and particularly to Drr

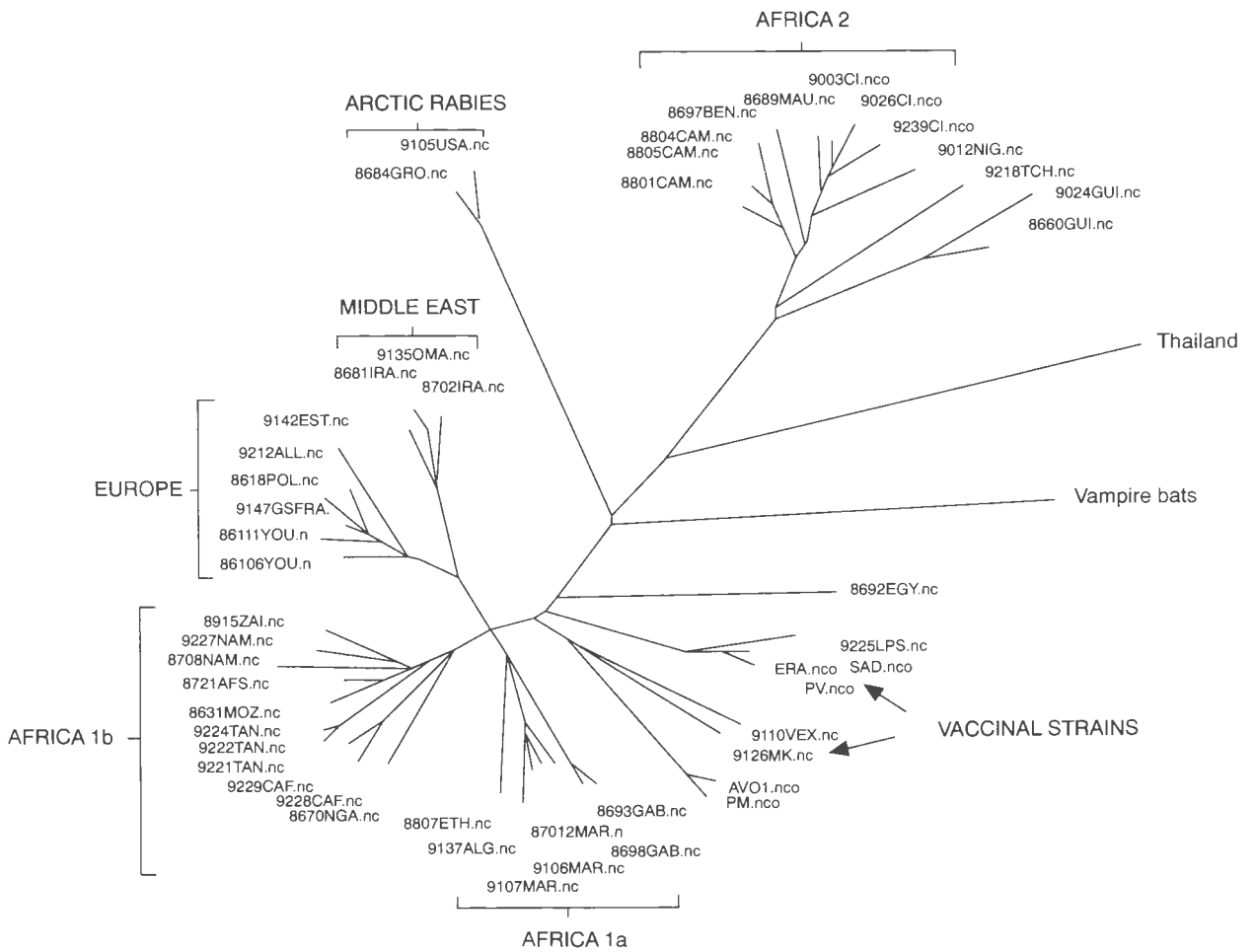

FIG. 3 Radial phylogenetic tree showing the relationships between the different geographic lineages of lyssaviruses. Alignments and tree generation were performed as described in Fig. 1

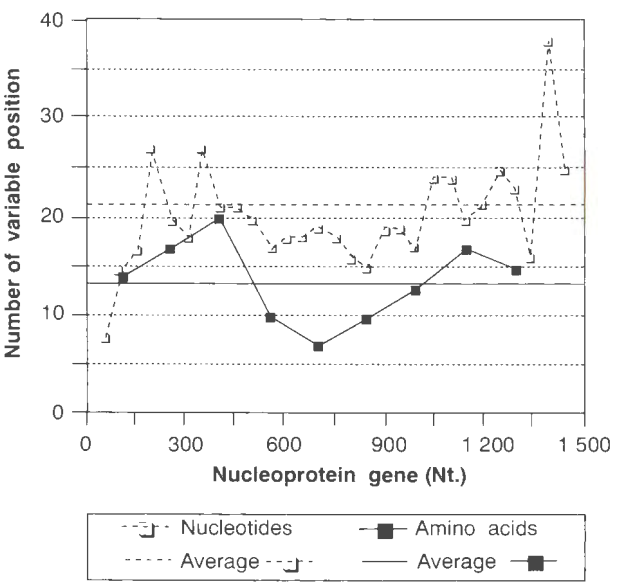

FIG. 4 Schematic representation of the distribution of the variability among the N genes of 59 rabies and rabies-related viruses

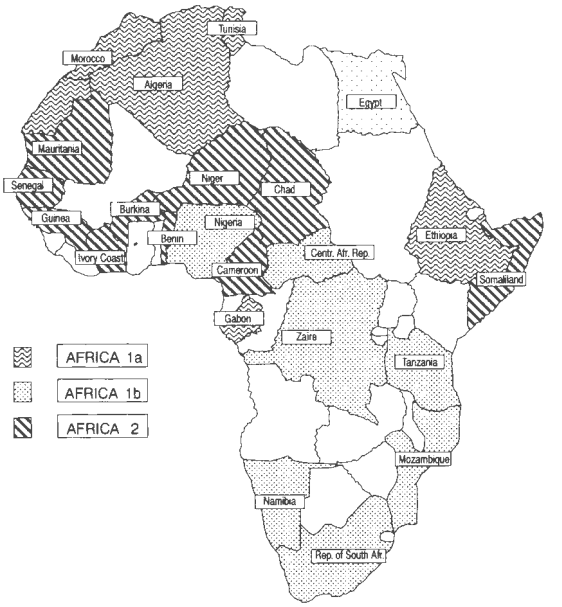

FIG. 5 Different groups of lyssaviruses of genotype 1 and their geographic distribution in Africa

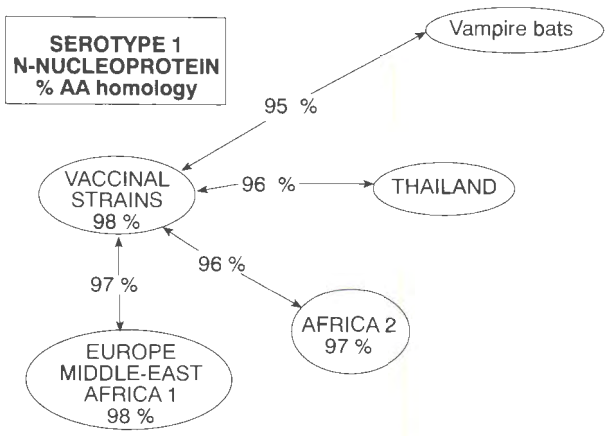

FIG. 6 Schematic representation of the percentage of amino acid similarity between the different virus groups identified in Fig. 3

J. Barrat, R. Bercion, M.C. Berthonneau, N. Bloch, L. Delestre, A. Fayaz, P. Formenty, K. Depner, J. Domenech, S. Gascoyne, P. Giraud, J. Godenir, L. Guerre, J. Morvan, M. Ravinet, Y. Renner, J.J. Salaun, P. Sarrazin, R. Swanepoel, J. Vincent and H. Zeller. We thank L. Audry for expert technical assistance. We are also grateful to A. King for critical reading of the manuscript.

REFERENCES

BANERJEE, A.K., RHODES, D.P. & GILL, D.S. 1984. Complete nucleotide sequence of the mRNA coding for the N protein of vesicular stomatitis virus (New Jersey serotype). *Virology*, 137: 432–438.

BOURHY, H., LAFON, M., BERTHONNEAU, M.C., RENNER, Y., ROLLIN, P.E. & SUREAU, P. 1988. Rabies in vaccinated dogs in Gabon. *Veterinary Record*, 122:361–362.

BOURHY, H., TORDO, N., LAFON, M. & SURREAU, P. 1989. Complete cloning and molecular organization of rabies-related virus: Mokola virus. *Journal of General Virology*, 70:2063–2074.

BOURHY, H., KISSI, B., LAFON, M., SACRAMENTO, D. & TORDO, N. 1992. Antigenic and molecular characterisation of bat rabies virus in Europe. *Journal of Clinical Microbiology*, 30: 2419–2426.

BOURHY, H., KISSI, B. & TORDO, N. 1993. Molecular diversity of the Lyssavirus genes. *Virology*, 194:70–81.

CRYSLER, J.G., LEE, P., REINDER, M. & PREVEC, L. 1990. The complete sequence of the nucleocapsid protein (N) of Piry virus: possible domains in the N protein of vesiculoviruses. *Journal of General Virology*, 71:2191–2194.

DIETZSCHOLD, B., WANG, H., RUPPRECHT, C.E., CELIS, E., TOLLIS, M., ERTL, H., HEBER-KATZ, E. & KOPROWSKI, H. 1987. Introduction of protective immunity against rabies by immunization with rabies ribonucleoprotein. *Proceedings of the National Academy of Science, USA*, 48:9165–9169.

DIETZSCHOLD, B., GORE, M., ERTL, H., CELIS, E., OTVOS, L & KOPROWSKI, H. 1989. Analysis of protective immune mechanisms induced by rabies nucleoprotein, in *Genetics and pathogenicity of negative strand viruses*, edited by B.W.J. Mahy & D. Kolakofsky. New York: Elsevier Science: 295–309.

FEKADU, M., SHADDOCK, J.H., SANDERLIN, D.W. & SMITH, J.S. 1988. Efficacy of rabies vaccines against Duvenhage virus isolated from European house bats (*Eptesicus serotinus*), classical rabies and rabies-related viruses. *Vaccine*, 6:533–539.

GALLIONE, C.J., GREENE, J.R., IVERSON, L.E. & ROSE, J.K. 1981. Nucleotide sequences of the mRNAs encoding the vesicular stomatitis virus N and NS proteins. *Journal of Virology*, 39:529–535.

HIGGINS, D.G. & SHARP, P.M. 1989. Fast and sensitive multiple sequence alignments on a microcomputer. *Cabios*, 5:151–152.

KING, A. & CRICK, J. 1988. Rabies-related viruses, in *Rabies*, edited by J.B. Campbell & K.M. Charlton. Boston: Kluwer Academic Publishers: 177–200.

KOPROWSKI, H., WIKTOR, T.J. & ABELSETH, M.K. 1985. Cross-reactivity and cross-protection: rabies and rabies-related viruses, in *Rabies in the Tropics*, edited by E. Kuwert, C. Merieux, H. Koprowski & K. Bögel. New York: Springer-Verlag: 30–39.

LAFON, M., BOURHY, H. & SUREAU, P. 1988. Immunity against the European bat rabies (Duvenhage) virus induced by rabies vaccines: an experimental study in mice. *Vaccine*, 6:362–368.

MASTERS, P.S. & BANNERJEE, A.K. 1987. Sequences of Chandipura virus N and NS genes: Evidence for high mutability of the NS gene within Vesiculoviruses. *Virology*, 157:298–306.

RUPPRECHT, C.E. DIETZSCHOLD, B., WUNNER, W.H. & KOPROWSKI, H. 1991. Antigenic relationships of Lyssaviruses, in *The natural history of rabies*, edited by G.M. Baer. Boca Raton: CRC Press: 69–100.

SACRAMENTO, D., BOURHY, H. & TORDO, N. 1991. PCR as an alternative method for diagnosis and molecular epidemiology of rabies virus. *Molecular and Cellular Probes*, 6:229–240.

SAITOU, N. & NEI, M. 1987. The neighbour joining methods: a new method for reconstructing phylogenetic trees. *Molecular and Biological Evolution*, 4:406–425.

SMITH, J.S. 1989. Rabies virus epitopic variation: use in ecologic studies. *Advanced Virus Research*, 36:215–253.

TIGNOR, G.H. & SMITH, A.L. 1972. Vaccination and challenge of mice with viruses of the rabies group. *Journal of Infectious Diseases*, 125:322–324.

TIGNOR, G.H., MURPHY, F.A., CLARK, H.F., SHOPE, R.E., MADORE, P., BAUER, S.P., BUCKLEY, S.A. & MEREDITH, C.D. 1977. Duvenhage virus: morphological, biochemical, histopathological and antigenic relationships to the rabies serogroup. *Journal of General Virology*, 37:595–611.

TORDO, N., BOURHY, H. & SACRAMENTO, D. 1992. Polymerase chain reaction technology for rabies virus, in *Diagnosis of human viruses by polymerase chain reaction technology*, edited by Y. Becker & G. Darai. Berlin: Springer-Verlag: 389–405.

TORDO, N., BOURHY, H., SATHER, S. & OLLO, R. 1993. Structure and expression in baculovirus of the Mokola virus glycoprotein: an efficient recombinant vaccine. *Virology*, 194:59–69.

WHO 1990. *Report of the sixth WHO consultation on monoclonal antibodies in rabies diagnosis and research*. Geneva: World Health Organization (WHO/Rab.Res./90.34).