CHAPTER V

MOLECULAR EPIDEMIOLOGY OF LAGOS BAT VIRUS

5.1 Introduction

With the inclusion of this study there have been 28 reports of LBV throughout Africa but only 16 virus isolations were reported (Section 2.8.6). Four LBV cases from *E. wahlbergii* frugivorous bats have been reported from South Africa between 2003 to 2006 (This study, Chapter III) from the Durban area of the KwaZulu Natal Province (LagSA2003; LagSA2004; LagSA2005 and LBVSA2006). Virus isolation from the LagSA2005 case was unsuccessful. LBV was also identified in a vaccinated dog in 2003 (LagSA2003canine) and from a mongoose (*Atilax paludinossus*) in 2004 (Mongoose 2004). Virus was only successfully isolated from the Mongoose2004 case and not from the vaccinated dog. All these new cases have been positively linked to LBV using either antigenic or/and genetic characterization.

Several molecular epidemiological studies of gt 1 lyssaviruses have been performed and only a few studies of ABLV (Guyatt *et al*., 2003) and EBLV (Davis *et al*., 2005) are available. Molecular studies focusing on the African lyssaviruses (MOKV, LBV and DUVV) have not been performed. Like all lyssavirus genotypes, LBV has a negative sense, single stranded RNA genome that encodes for a nucleoprotein (N), phosphoprotein (P), matrixprotein (M), glycoprotein (G) and RNA polymerase (L). Comparison of the N and G genes of lyssaviruses allowed for the grouping of the *Lyssavirus* genus into seven genotypes and four putative genotypes (species). Gt 1 (RABV), gt 2 (LBV), gt 3 (MOKV), gt 4 (DUVV), gt 5 (EBLV-1), gt 6 (EBLV-2) and gt 7 (ABLV) constitute the seven *Lyssavirus* genotypes and Irkut virus, Aravan virus, Khujand virus and West Caucasian bat virus are currently added as putative genotypes (species). Currently the criteria suggested for classification of a new *Lyssavirus* genotype are >80% nucleotide difference and >93% amino acid differences of the nucleoprotein. Very little is known regarding gt 2 molecular epidemiology and the interrelationships between LBV isolates as well as between LBV and other lyssavirus genotypes are unknown since very few LBV isolates have been studied. The aim of investigating the molecular epidemiology of LBV was; 1) to determine the relationship between LBV isolates, geographic location and host species by comparing full length N, P, M and G gene sequences and 2) to re-examine the position of LBV isolates within the lyssavirus genus.

5.2 Materials and Methods

5.2.1 Source of LBV isolates

Gt 2 isolates used in this study were obtained from different sources as indicated in Table 5.1.

Table 5.1: Information of gt 2 isolates analysed in this molecular epidemiological study

VIRUS NAME	YEAR OF ISOLATION	GEOGRAPHICAL LOCATION	ORIGIN	SOURCE	PASSAGE HISTORY
LBVNIG1956	1956	Nigeria	Frugivourous bat (E*idolon helvum*)	Dr. C.E. Rupprecht (CDC, Atanta, USA)	Unknown
LBVCAR1974	1974	Central African Republic	Frugivourous bat (*Micropterus pusillus*)	Dr. C.E. Rupprecht (CDC, Atanta, USA)	Passage 2 in mouse brain
LBVSA1980(640)	1980	South Africa	Frugivourous bat (*Epomophorus wahlbergi*)	Dr. C.T. Sabeta (OVI, Rabies Unit, South Africa)	Freeze dried mouse brain-Passaged multiple times in mouse brain
LBVSA1980(1248)	1980	South Africa	Frugivourous bat (*Epomophorus wahlbergi*)	Dr. C.T. Sabeta (OVI, Rabies Unit, South Africa)	Freeze dried mouse brain-Passaged multiple times in mouse brain
LBVSA1980(679)	1980	South Africa	Frugivourous bat (*Epomophorus wahlbergi*)	Dr. C.T. Sabeta (OVI, Rabies Unit, South Africa)	Freeze dried mouse brain-Passaged multiple times in mouse brain
LBVSA1982	1982	South Africa	Frugivourous bat (*Epomophorus wahlbergi*)	Dr. C.T. Sabeta (OVI, Rabies Unit, South Africa)	Freeze dried mouse brain-Passaged multiple times in mouse brain
LBVSEN1985	1985	Senegal	Frugivourous bat (E*idolon helvum*)	Dr. C.E. Rupprecht (CDC, Atanta, USA)	Passage 4 in mouse brain
LBVZIM1986	1986	Zimbabwe	Feline	Dr. C.E. Rupprecht (CDC, Atanta, USA)	Passage 2 in MNA cells
LBVAFR1999	1999	Africa (Probably Egypt/Togo)	Frugivourous bat (*Rousettus aegyptiacus*)	Dr. F. Cliquet (AFSSA, France)	Passage 2 in mouse brain
LagSA2003	2003	South Africa	Frugivourous bat (*Epomophorus wahlbergi*)	This study	Original bat brain
LagSA2004	2004	South Africa	Frugivourous bat (*Epomophorus wahlbergi*)	This study	Original bat brain
Mongoose2004	2004	South Africa		This study	Original bat brain
LBVSA2006	2006	South Africa	Frugivourous bat (*Epomophorus wahlbergi*)	This study	Original bat brain

Before attempting any further characterization, the isolates were first passaged once in suckling mice after which brains were harvested and used in subsequent analysis. Briefly, pregnant female mice were obtained from Harlan Sprague Daly (USA). Animals were housed and handled according to protocols approved by the Institutional Animal Care and Use Committee (IACUC) of the Centers for Disease Control and Prevention (USA). For inoculation, brain suspensions were prepared in Minimal essential medium (MEM-10; GIBCO) supplemented with 10% fetal bovine serum (GIBCO). The mixture was clarified by centrifugation at 3 200 g for 15 minutes. 30 µl of the supernatant was inoculated intracranially into two day old suckling mice with a 0.5 cc tuberculin syringe and 8 mm, 31 gauge needle (Becton Dickinson and Company). Mice were either humanely euthanized or collected upon death. Brain smears of brain removed aseptically from dead mice were analysed with FAT (Section 3.2.2.1) to indicate the presence of lyssavirus antigen. Brain material that tested positive with the FAT was then used in subsequent downstream applications. A 20% brain suspension in MEM-10 (GIBCO) was prepared and centrifuged at 3 200 g for 10 minutes. The supernatant was removed and stored at -70ºC for future inoculation and the brain sediment was used for subsequent RNA extraction.

5.2.2 Total RNA extraction

Total RNA was extracted from brain material collected from suckling mice (Table 5.1) using the TRIzol™ method (Invitrogen) as described by the manufacturers. Briefly, 50-100 mg of brain tissue was added to 1 ml TRIzol™ and homogenized using vigorous pipetting. The mixture was transferred to a microcentrifuge tube and incubated for 5 minutes at room temperature after which 0.2 ml of chloroform was added. The mixture was vortexed for 15 seconds and incubated for a further 3 minutes and then centrifuged at 12 000 g for 30 minutes at 4ºC, The aqueous phase containing the RNA was transferred to a new microcentrifuge tube. Isopropyl alcohol (500 µl) was added to precipitate the RNA for 10 minutes at room temperature and RNA was recovered by centrifugation at 12 000 g for 30 minutes at 4ºC. The RNA precipitate was washed with 1 ml of 75% ethanol, allowed to dry, and dissolved in 50 µl nuclease free water (GIBCO). The RNA preparations were then stored at –70ºC until further use.

5.2.3 Primer design

To determine the nucleotide sequence of the complete N, P, M and G genes several oligonucleotide primers were designed (Table 5.2) with the use of DNA sequence information available from GenBank.

Table 5.2: Oligonucleotide primers used for RT-PCR and DNA sequencing of gt 2 isolates

PRIMER NAME	SEQUENCE (5'-3')	NUCLEOTIDE POSITION †
001LYSF ‡	ACGCTTAACGAMAAA	3' non-coding region (-70 to -57)
550B ‡	GTRCTCCARTTAGCRCACAT	577-596
550F ‡	ATGTGYGCTAAYTGGAGYAC	577-596 rc
LagNF §	GGGCAGATATGACGCGAGA	374-392
DB1F *	ACCAGGACAAAYGTGGARGG	334-353
DB717F *	GGTATCTTTCACTGGATTCATT	717-738
304B §	TTGACAAAGATCTTGCTCAT	1445-1464 rc
SA1240F §	GGCAGGTTGAAGAAATCTC	1240-1258
DB1218F *	AGCAGGATTATGATGAACAATGG	1219-1241
ZIM1240F §	GCAGGTTGAAGAAATCTC	1241-1259
CAR1240F §	GGCAGATTGAAGAAATCTC	1240-1259
LBVPOB §	CAGTTCTTTACTATCTTCC	2460-2478 rc
EGYPT1680F *	GCAGCTCAGAAGATGAGTTC	1678-1697
LM330B *	ACGGAGTTTATAMACCCAATTC	2777-2798 rc
LM504B *	AGAGTTCATATTGATACACCAC	2951-2972 rc
PF2 ‡	ACTGACAARGTGGCCAGRCTCATG	2295-2318
979B ‡	TAAGCYTTCCCATAYCCTGGCAC	4238-4260 rc
571F ‡	AATCATGATTACGCYTTATGG	3830-3850
LBVLB ‡	GGRTCTTCTATCAAAGGAGAGTT	5456-5479 rc
P490F §	GTCATATTCAAGGGAGA	2931-2947
EGYPT2500F *	GTATGGGGCAGCATCAGCTCC	2501-2521
979F ‡	GTGCCAGGRTATGGGAARGCTTA	4238-4260
571B ‡	CCATAARGTGTAATCATGATT	3830-3850 rc
DBM470F *	GAATTGGGTGTATAAACTCCG	2777-2798
DBM670B *	GCCATGTCAGCCCAAAGCTGAC	2980-3001 rc
DB4350B *	GGAAGGCAGTATGTCGACCC	4321-4340 rc
DB3800F *	GAAGACCCTAACTTGAGTTTGAT	3858-3880

‡ Primers used for the characterization of all gt 2 isolates (Table 5.1)

§ Primers used for the characterization of LBVNIG1956, LBVCAR1974, LBVSA1980(640), LBVSA1980(1248), LBVSA1980(679), LBVSA1982, LBVZIM1986, LagSA2003, LagSA2004, Mongoose2004 and LBVSA2006

*** Primers used for the characterization of LBVAFR1999 and LBVSEN1985 isolates**

 # rc, reverse complement

† Position on LBV genome of LBVSA1982 generated in this study (GenBank accession numbers: EF547410,EF547425,EF547439,EF547455)

5.2.4 RT-PCR

Different combinations of primers indicated in Table 5.2 were used to amplify the complete first 5 600 bp of the LBV genome using the same methodology as described in Section 4.2.5.

5.2.5 PCR product purification

PCR amplicons generated as described in Section 5.2.4 were purified using the Wizard® PCR Preps DNA Purification System (Promega) as described in Section 4.2.7.

5.2.6 DNA sequencing

Purified PCR products were sequenced with the BigDye™ Termination Cycle Sequencing Ready Reaction Kit 1.1 (Applied Biosystems) as described in Section 4.2.8

5.2.7 Phylogenetic analyses

Nucleotide sequences obtained as described in Section 5.2.6 were assembled and edited using Vector NTI 9.1.0 (Invitrogen). Different data sets were created for the N, P, M and G genes. Amino acid sequences were deduced using the translate function of this program. Multiple sequence alignments were generated using ClustalX and exported in FASTA format. Phylogenetic and evolutionary analyses were conducted using Mega 3.1 (Kumar *et al.*, 2004). Neighbor-joining (NJ) and maximum parsimony (MP) phylogenetic trees for a variety of data sets were determined: i) Complete N gene nt and aa sequences including all lyssavirus genotypes and all gt 2 isolates; ii) Complete G gene nt and aa sequences including all genotypes and all gt 2 isolates; iii) Complete P gene nt and aa sequences including all lyssavirus genotypes and all gt 2 isolates; iv) Complete M gene nt and aa sequences including all lyssavirus genotypes and all gt 2 isolates. Neighbor-joining trees were constructed using evolutionary distance correction statistics of Kimura, (1980) and Tajima and Nei, (1984). Bootstrap analysis were performed using 1000 data replications. Bootstrap values greater than 70% confidence was regarded as strong evidence for a particular phylogenetic grouping. Maximum parsimony trees were constructed using Mega 3.1 and bootstrap values were determined using a 100 replicates. P-distances between sequences were calculated. The p-distance method of the Mega 3.1 package was used to compute the synonymous and non-synonymous nucleotide substitutions between sequences. Distance matrixes were generated for the full length P, N, M and G genes of all lyssavirus sequences as indicated in Appendix 1.

5.3 Results

5.3.1 Phylogenetic tree construction

In this study, phylogenetic trees were constructed using the N, P, M and G gene sequences and deduced amino acids. The neighbor-joining (NJ) and maximum parsimony (MP) methods were employed.

i) N gene

For all thirteen isolates sequenced in this study, the N gene (1353 nt) coded for a deduced 450 aa nucleoprotein. A set of 67 complete N gene sequences, consisting of 21 gt 1 representatives, 9 gt 5 representatives, 9 gt 7 representatives, all available full-length N gene sequences for gt 2, 3, 4 and 6 and the putative genotypes as well as N gene sequences generated for the thirteen isolates analysed in this study (Table 5.1) were assembled for phylogenetic analysis. The isolates were selected to cover the maximum intrinsic variability of each genotype. The NJ method indicated high bootstrap support (73% and higher (nt)) for all the major clusters representing all the current lyssavirus genotypes and putative genotypes – with the exception of gt 4 (56% bootstrap support) (Figure 5.1 (A)). The same tree topology was obtained for the deduced N amino acids with slightly higher bootstrap values (Figure 5.1 (B)). These tree topologies are in agreement with the current classification of lyssaviruses as described in previous studies (Kuzmin *et al*., 2003; Kuzmin *et al*., 2005). The thirteen isolates analysed in this study all grouped into the same cluster supported by a 93% bootstrap value (Figure 5.1). In this cluster three lineages can be observed, lineage A-C. Lineage A constitutes the LBVAFR1999 and LBVSEN1985 isolates (88% bootstrap support (nt)), lineage B the LBVNIG1956 isolate (98% bootstrap support (nt)) and lineage C the remaining isolates (Figure 5.1). MP phylogenetic analysis also indicated the major clusters representing the different lyssavirus genotypes as well as the distinct lineages (A-C) (Figure 5.2). However, it is noteworthy that lineage A represented a separate cluster in the maximum parsimony N gene analysis and did not form part of the gt 2 cluster (Figure 5.2 (A)).

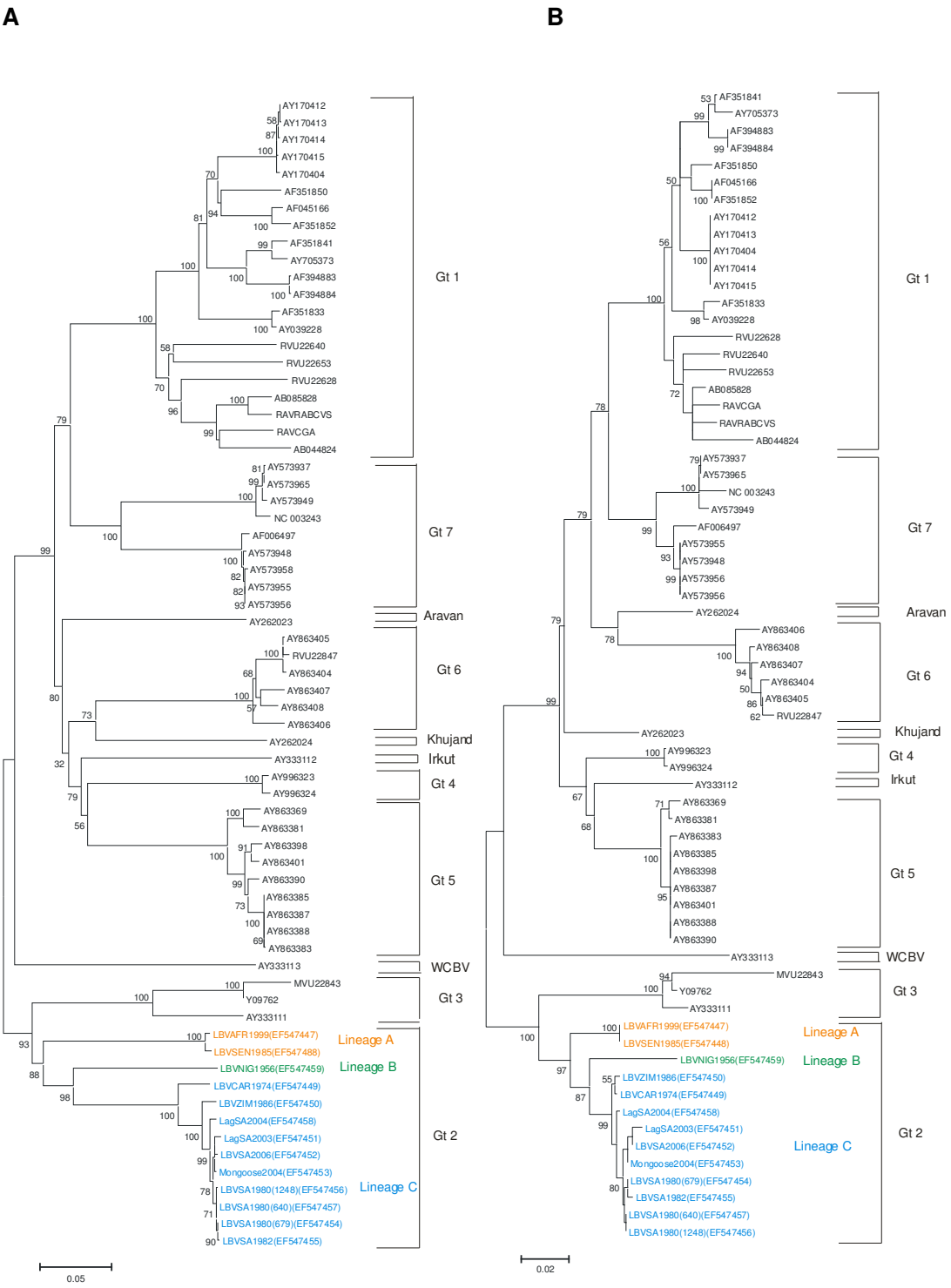


Figure 5.1: NJ phylogenetic tree based upon the 1353 nucleotides of the entire N gene (A) and 450 amino acids of the entire N protein (B) of representatives of the lyssavirus genus, obtained by the neighbor-joining method. GenBank accession numbers are indicated for each isolate. Bootstrap values higher than 50% are indicated at the nodes and branch lengths are drawn to scale.

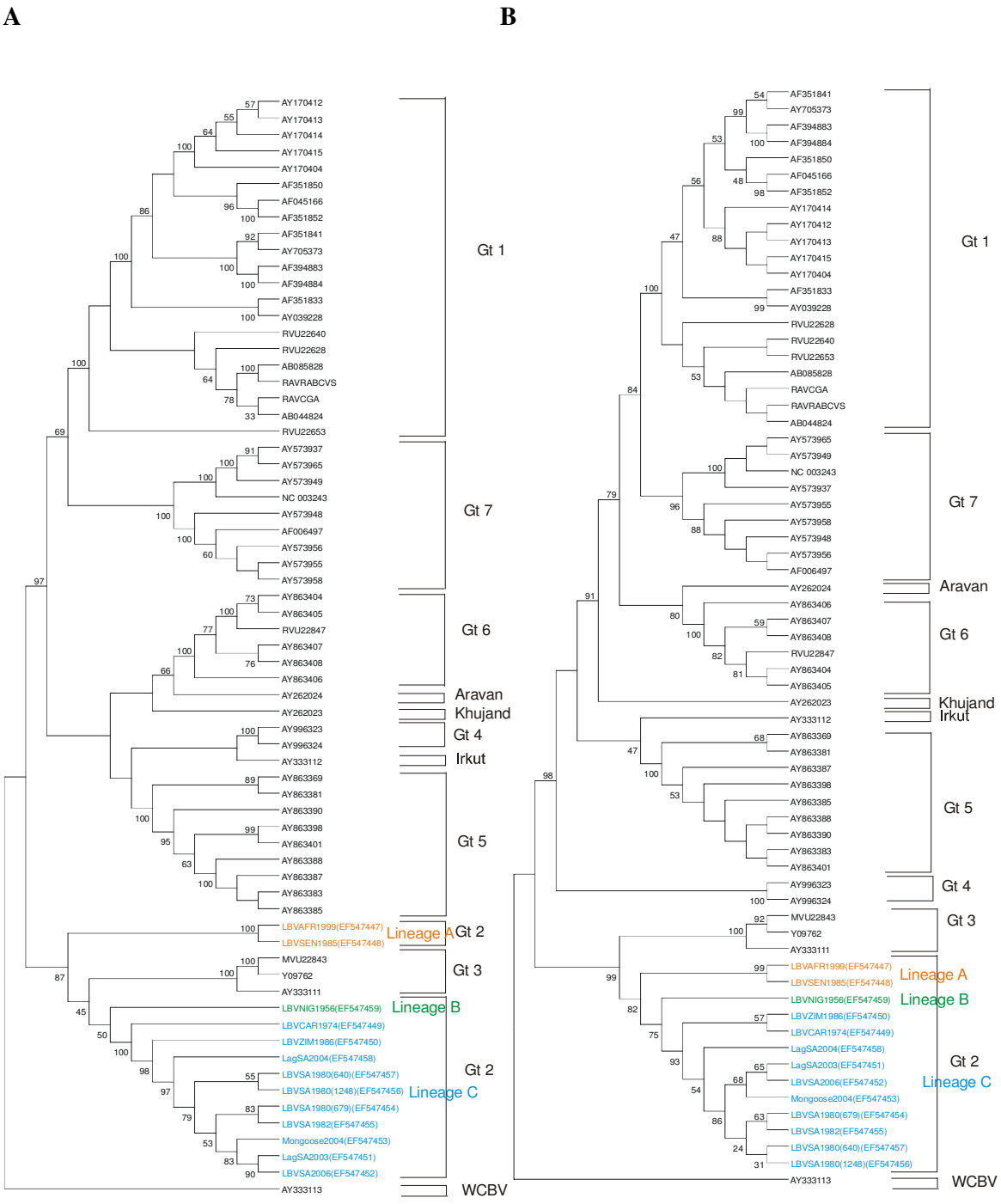


Figure 5.2: MP phylogenetic tree based upon the 1353 nucleotides of the entire N gene (A) and 450 amino acids of the entire N protein (B) of representatives of the lyssavirus genus, obtained by the maximum parsimony method. GenBank accession numbers are indicated for each isolate. Bootstrap values are indicated at the nodes and branch lengths are drawn to scale.

ii) P gene

The P gene of the thirteen isolates analysed in this study (Table 5.1) consists of 918 bp and 305 deduced aa. A set of 66 P gene sequences (Appendix 1) was generated for phylogenetic analysis consisting of 27 gt 1 representatives and all available full-length P gene sequences for gt 2, 3, 4, 5, 6, 7 and the putative genotypes. All isolates were selected to cover the maximum intrinsic variability of each genotype. Neighbor-joining phylogenetic tree construction of the P genes, generated a tree topology with all the major clusters corresponding to the different lyssavirus genotypes (Figure 5.3) as obtained in the N gene analysis. Bootstrap support for the gt 4 cluster (59%) and Aravan and Khujand virus (63%) was low in the nucleotide analysis. These tree topologies are in agreement with previous studies on the classification of lyssaviruses using the P gene (Kuzmin *et al*., 2003; Kuzmin *et al*., 2005; Nadin-Davis *et al*., 2002). In contrast to N gene analysis, Aravan and Khujand virus were joined to the gt 6 cluster and Irkut virus to the gt 4 and gt 5 cluster and did not form distinct groups. P gene and protein NJ analyses of the thirteen isolates analysed in this study, indicated that these isolates cluster together as one group with high bootstrap support (100% nt and 99% aa). In this cluster the same three distinct lineages as observed in N gene analysis were present and supported by high bootstrap values; Lineage A: LBVAFR1999 and LBVSEN1985 (89% bootstrap support (nt)); Lineage B: LBVNIG1956 (94% bootstrap support (nt)) and lineage C (Figure 5.3). MP phylogenetic analysis supported the same clusters representing the different genotypes and also the three lineages (A-C) (Figure 5.4).

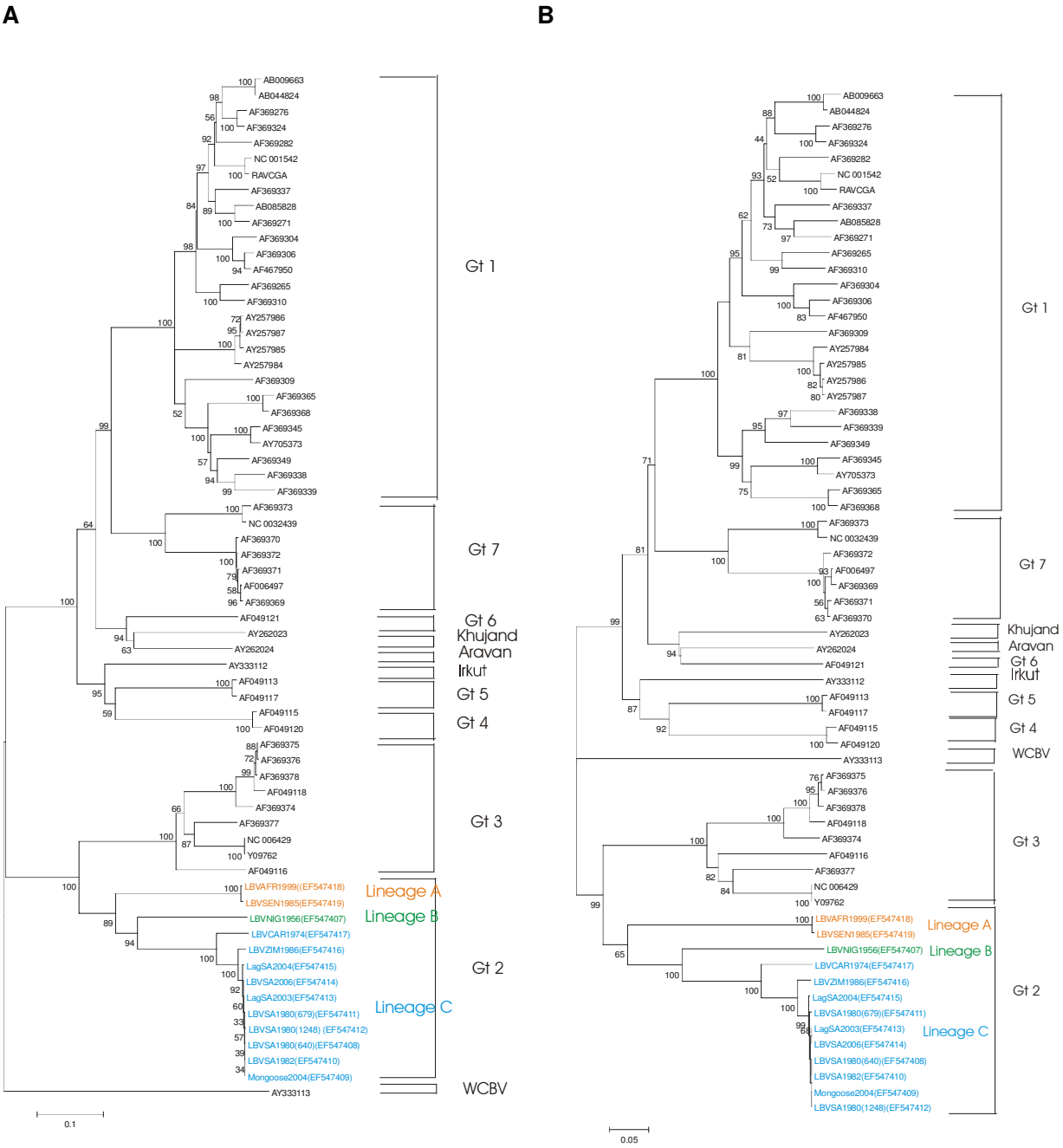


Figure 5.3: NJ phylogenetic tree based upon the 918 nucleotides of the entire P gene (A) and 305 amino acids of the entire P protein (B) of representatives of the lyssavirus genus, obtained by the neighbor-joining method. GenBank accession numbers are indicated for each isolate. Bootstrap values higher than 50% are indicated at the nodes and branch lengths are drawn to scale.

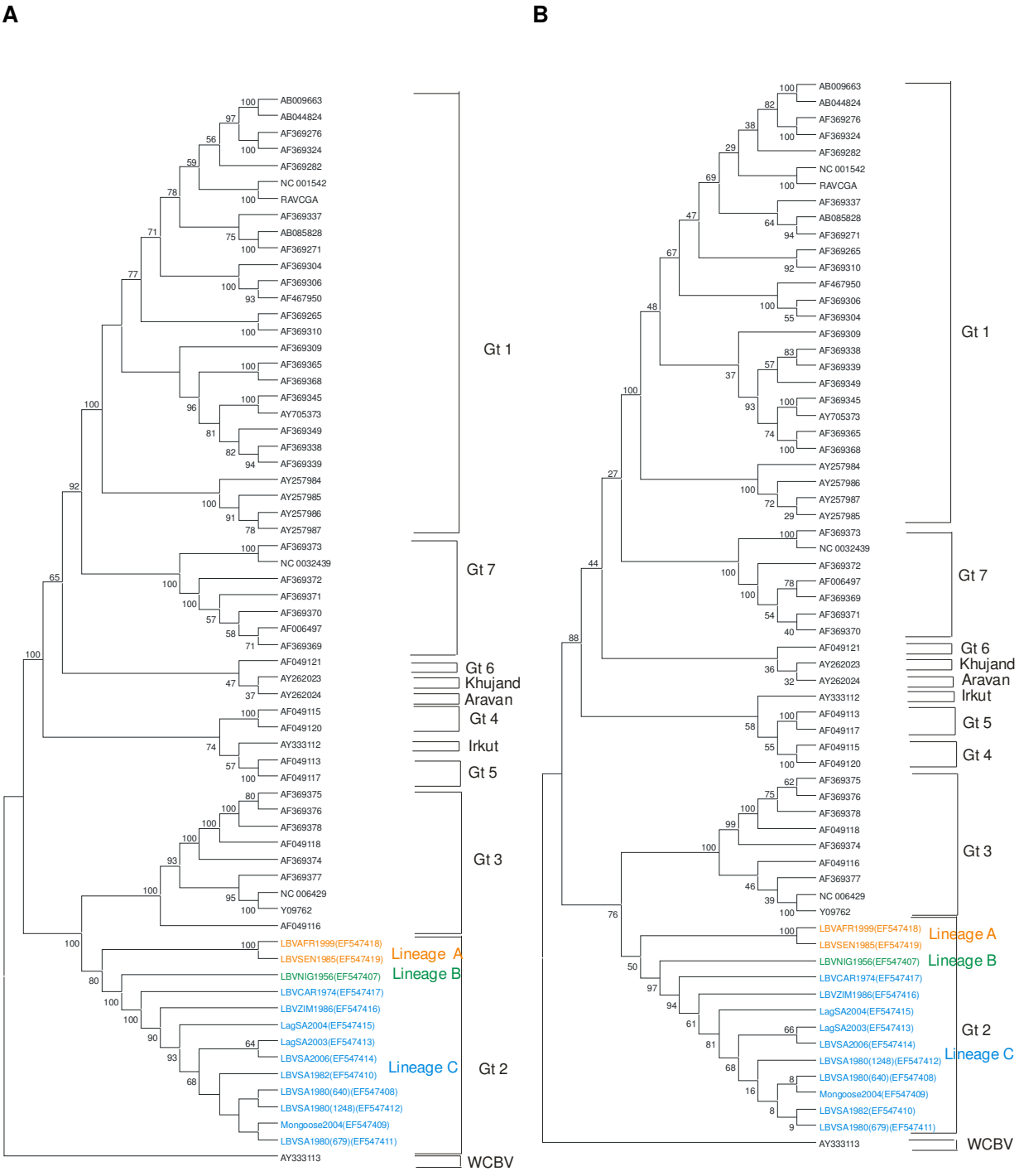


Figure 5.4: MP phylogenetic tree based upon the 918 nucleotides of the entire P gene (A) and 305 amino acids of the entire P protein (B) of representatives of the lyssavirus genus, obtained by the maximum parsimony method. GenBank accession numbers are indicated for each isolate. Bootstrap values are indicated at the nodes and branch lengths are drawn to scale.

iii) M gene

The coding region of the M protein of isolates analysed in this study consisted of 609 bp and coded for a deduced 202 aa protein. A set of 44 M genes was included in the analyses (Appendix 1) representing all lyssavirus genotypes, putative genotypes and isolates analysed in this study. Previous studies did not include the analysis of the M gene when classifying lyssaviruses into genotypes. In this study, the major clusters representing the different lyssavirus genotypes were observed although gt 4, 5, 6, Irkut, Aravan and Khujand virus clustered together and no clear distinction between these groups were observed. Clusters were not supported by high bootstrap values, indicating that the M gene is probably not a good candidate to use for lyssavirus classification. In contrast with N and P gene analyses lineage A and B did not group with the rest of the gt 2 isolates but instead formed part of a cluster consisting of gt 3 isolates (Figure 5.5 (A) and Figure 5.6). However, only lineage A grouped with gt 3 isolates in the NJ deduced amino acid analysis (Figure 5.5 (B)). This phylogenetic analysis again supports lineage A as a distinct cluster. The same tree topology as observed in the NJ nucleotide analysis was supported by MP analysis of nt and deduced aa M gene sequences (Figure 5.6).

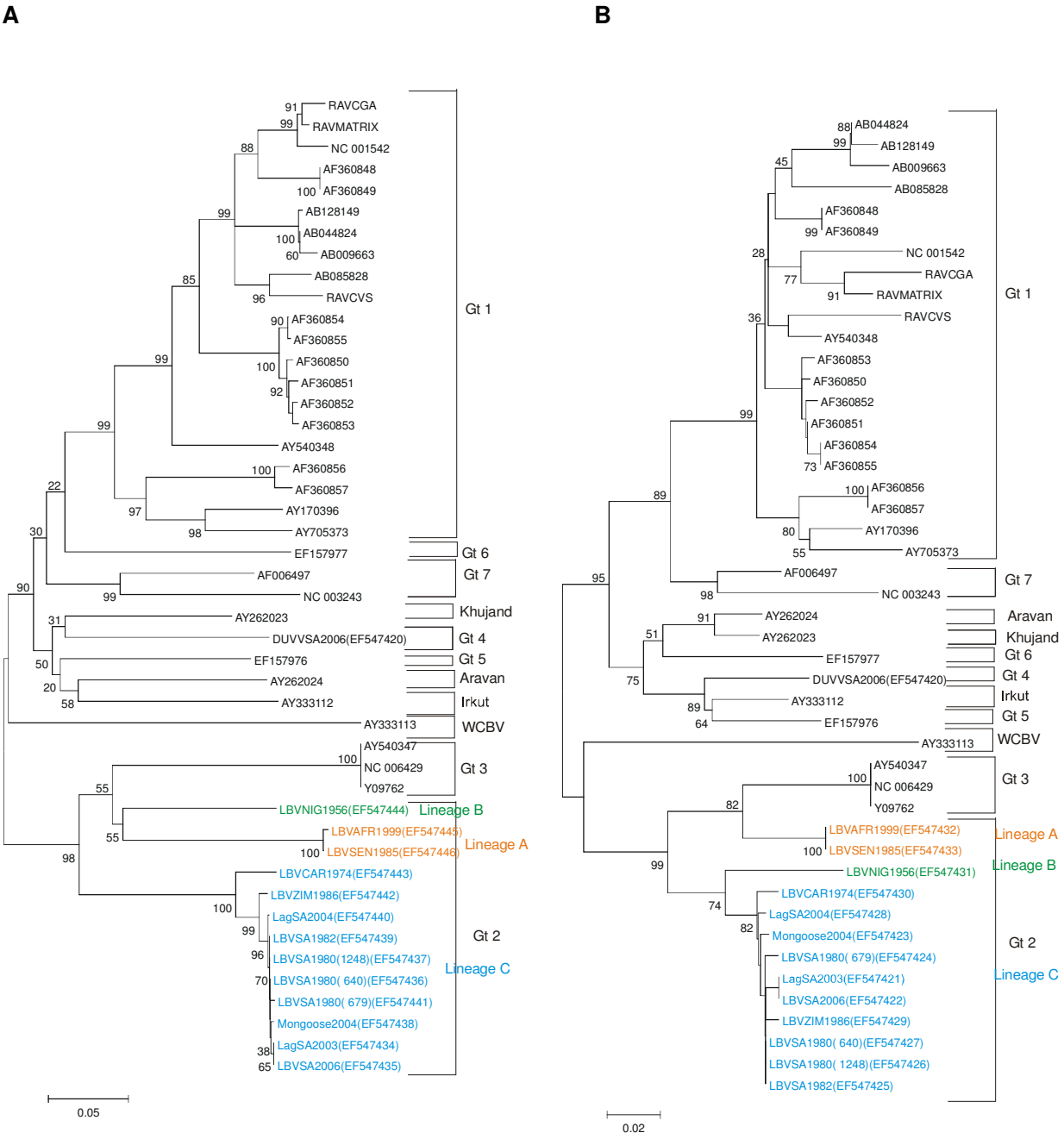


Figure 5.5: NJ phylogenetic tree based upon the 609 nucleotides of the entire M gene (A) and 202 deduced amino acids of the entire M protein (B) of representatives of the lyssavirus genus, obtained by the neighbor-joining method. GenBank accession numbers are indicated for each isolate. Bootstrap values are indicated at the nodes and branch lengths are drawn to scale.

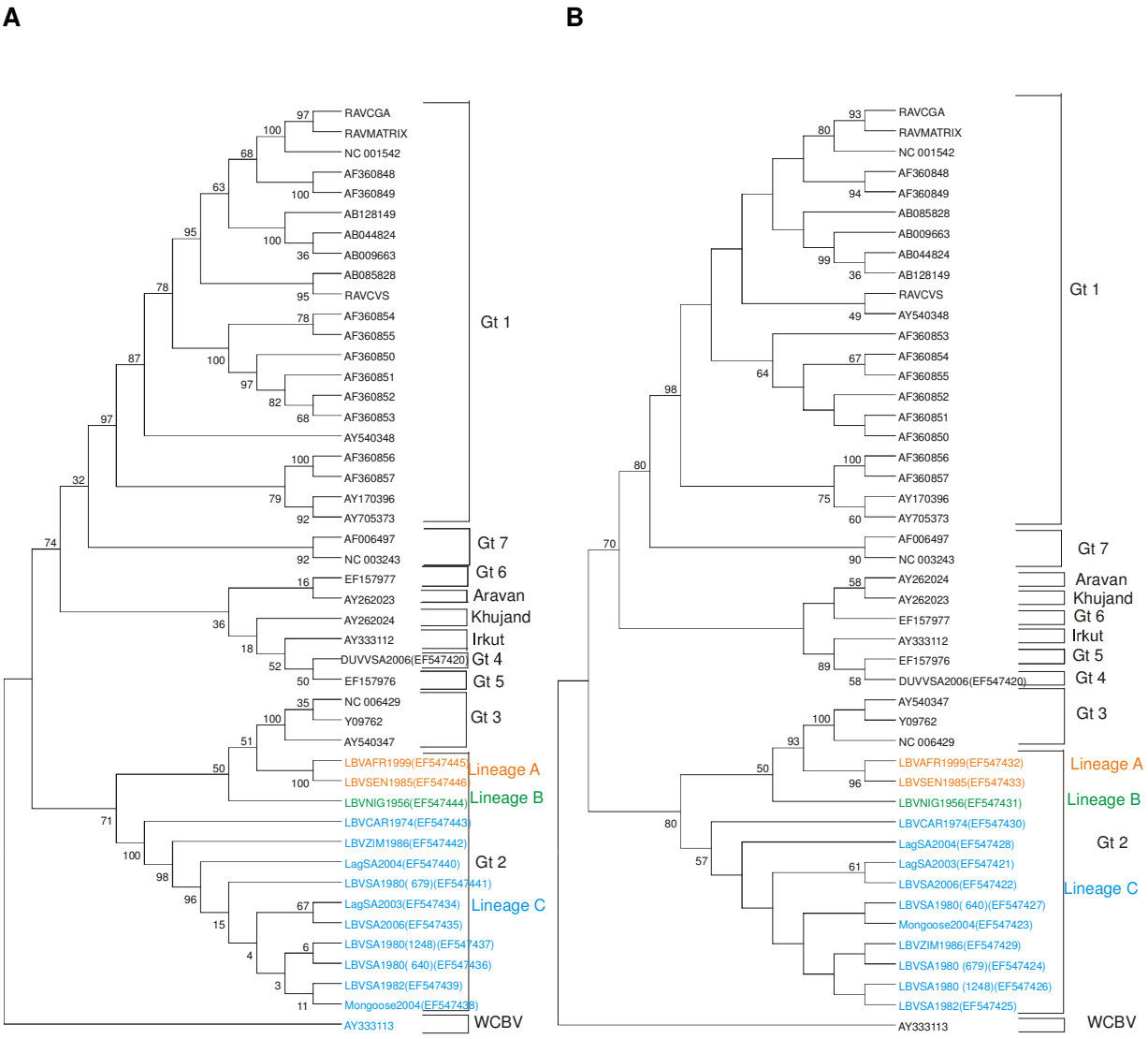


Figure 5.6: MP phylogenetic tree based upon the 609 nucleotides of the entire M gene (A) and 202 amino acids of the entire M protein (B) of representatives of the lyssavirus genus, obtained by the maximum parsimony method. GenBank accession numbers are indicated for each isolate. Bootstrap values are indicated at the nodes and branch lengths are drawn to scale.

iv) G gene

The G gene of isolates analysed in this study consists of 1569 bp and codes for a 502 aa protein (excluding the signal peptide (19 aa)). Phylogenetic analysis of the G gene, including 65 representatives of the lyssavirus genus (Appendix 1), was performed. Isolates were selected to cover the maximum intrinsic genotypic variation. Neigbor-joining (Figure 5.7) and maximum parsimony analysis (Figure 5.8) indicated the same clusters, representing the lyssavirus genotypes and putative genotypes as were indicated by N and P gene analysis. Lineage A, B and C could again be clearly distinguished with high bootstrap support.

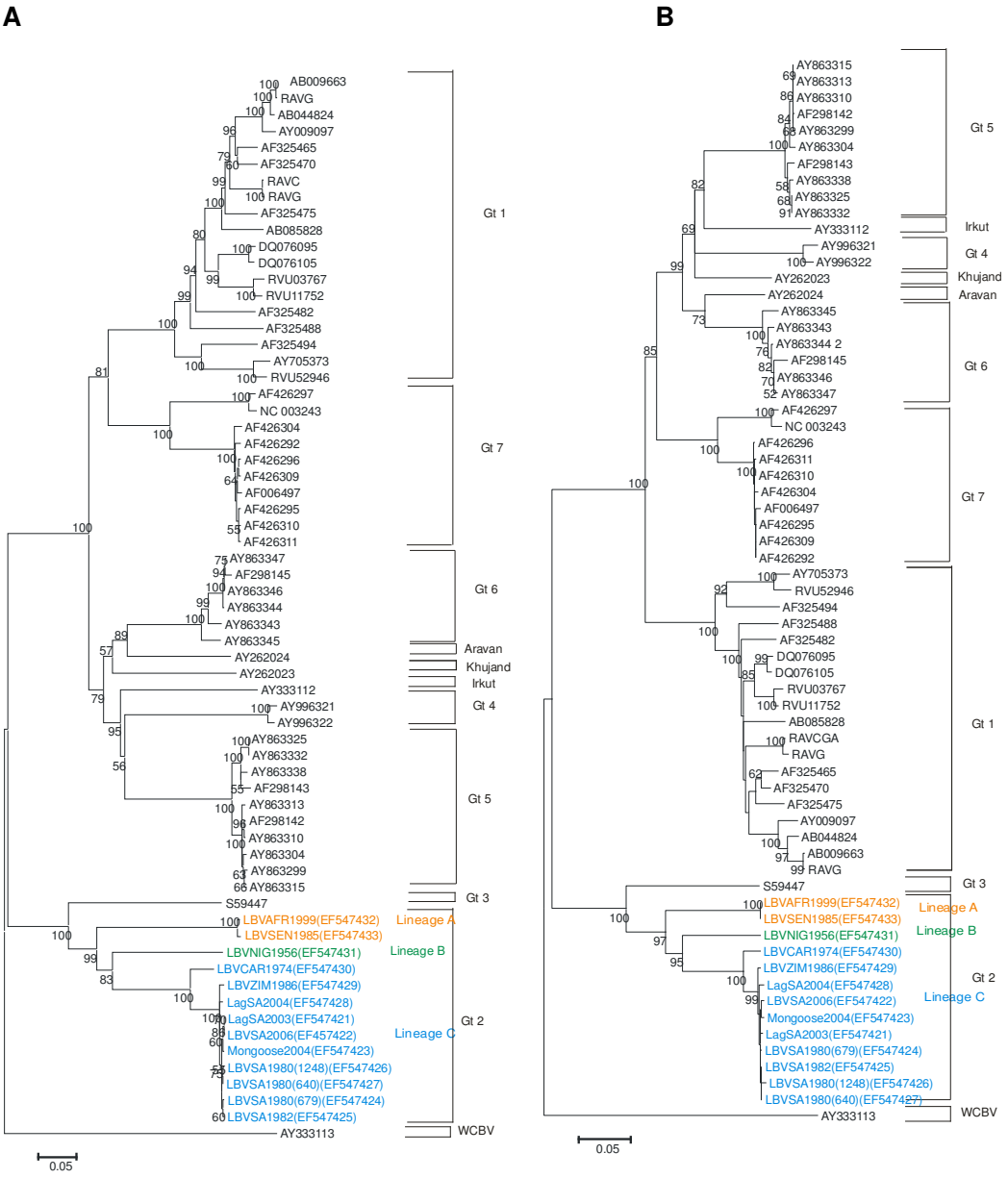


Figure 5.7: NJ phylogenetic tree based upon the 1569 nucleotides of the entire G gene (A) and 522 amino acids of the entire G protein (B) of representatives of the lyssavirus genus, obtained by the neighbor-joining method. GenBank accession numbers are indicated for each isolate. Bootstrap values are indicated at the nodes and branch lengths are drawn to scale.

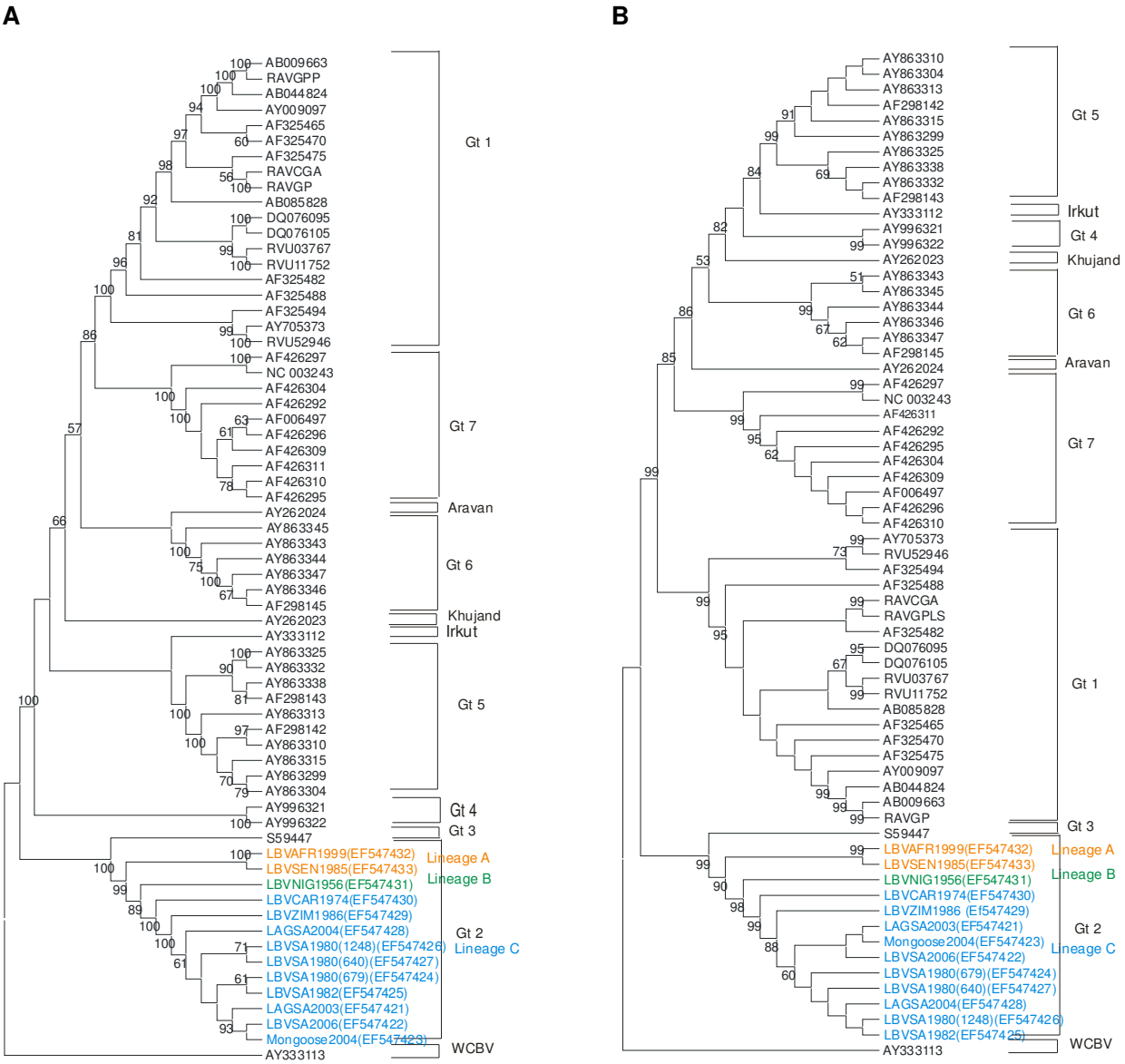


Figure 5.8: MP phylogenetic tree based upon the 1569 nucleotides of the entire G gene (A) and 522 amino acids of the entire G protein (B) of representatives of the lyssavirus genus, obtained by the maximum parsimony method. GenBank accession numbers are indicated for each isolate. Bootstrap values are indicated at the nodes and branch lengths are drawn to scale.

5.3.2 P-distances

Results obtained in Section 5.3.1 already indicated that two of the previously reported gt 2 isolates, LBVAFR1999 and LBVSEN1985 (Lineage A), are distantly related to other gt 2 isolates and this will be further investigated in this section. P-distance matrixes of the nucleoprotein gene and aa sequences, comparing all thirteen previously reported gt 2 isolates, were calculated (Table 5.3). The intrinsic variation between isolates in lineage C was low; 94.8-99.9% nt and 98.4-100% aa identity. The lineage B isolate (LBVNIG1956) had a 82.7-83.3 % nt and 94.2-95.1% aa identity to lineage C. Lineage A isolates (LBVSEN1985 and LBVAFR1999) had a 79.1-80.7% nt and 93.3-96% aa identity to lineage B and C isolates but a high identity (99.3% nt and 100% aa) to each other. The identity of lineage A compared to other gt 2 isolates is <80% when analysing the full-length nucleoprotein gene, indicating that these isolates could be considered as a new lyssavirus genotype. The current proposed criteria for lyssaviruses to be classified in the same genotype is >80% nt identity and >93% aa identity when analysing the nucleoprotein gene (Kissi *et al*., 1999).

Phylogenetic analysis and p-distances of gt 2 isolates (lineage B and C) showed that geographical origin influenced the phylogenetic patterns (Figure 5.1-Figure 5.8). Gt 2 isolates from South Africa had a 98.9-99.9% nt and 99.1-100% aa identity even though these isolates were isolated over several years (1982-2006) (Table 5.3). With the exception of an isolate from a mongoose (Mongoose2004), these viruses were isolated from the same host species (*E. wahlbergi*). Within lineage C, the isolates from South Africa could be distinguished from those retrieved from species in the Central African Republic and Zimbabwe (Table 5.3). The isolate from Nigeria in West Africa (lineage B) had a high sequence diversity compared to other gt 2 isolates (Table 5.3). A molecular evolution pattern associated with host species is unclear due to the limited amount of gt 2 samples available.

Table 5.3: Percentage identity of the N gene (A) and N protein (B) sequences of previously reported gt 2 isolates analysed in this study.

A

	Lineage C										Lineage B	Lineage A	
	LBVSA1982	LBVSA1980(640)	LBVSA1980(1248)	LBVSA1980(679)	LagSA2003	LagSA2004	Mongoose 2004	LBVSA2006	LBVZIM1986	LBVCAR1974	LBVNIG1956	LBVSEN1985	LBVAFR1999
LBVSA1982		99.8	99.6	99.9	99.1	99.2	99.6	99.5	98.2	95.3	82.8	80.1	80
LBVSA1980(640)			99.9	99.9	99.2	99.4	99.7	99.6	98.4	95.4	83	79.3	79.2
LBVSA1980(1248)				99.7	99	99.3	99.6	99.4	98.2	95.3	82.8	80.1	80
LBVSA1980(679)					99.2	99.2	99.7	99.6	98.2	95.4	82.9	79.2	79.1
LagSA2003						98.9	99.3	99.5	97.6	94.8	82.9	79.7	79.7
LagSA2004							99.1	99	98	95.1	83.3	80	80
Mongoose 2004								99.7	98.1	95.3	82.7	80.2	80.1
LBVSA2006									97.9	95.1	82.9	80	79.9
LBVZIM1986										95.1	83.3	80.7	80.6
LBVCAR1974											82.7	80.6	80.5
LBVNIG1956												80	79.8
LBVSEN1985													99.7
LBVAFR1999													

B

	Lineage C										Lineage B	Lineage A	
	LBVSA1982	LBVSA1980(640)	LBVSA1980(1248)	LBVSA1980(679)	LagSA2003	LagSA2004	Mongoose 2004	LBVSA2006	LBVZIM1986	LBVCAR1974	LBVNIG1956	LBVSEN1985	LBVAFR1999
LBVSA1982		99.8	99.8	99.8	98.7	99.3	99.3	99.1	98.7	98.9	94.2	95.1	95.1
LBVSA1980(640)			100	99.8	99.1	99.8	99.7	99.6	99.1	99.3	94.7	95.6	95.6
LBVSA1980(1248)				99.8	99.1	99.6	99.8	99.6	99.1	99.3	94.7	95.6	95.6
LBVSA1980(679)					98.9	99.3	99.6	99.3	98.9	99.1	94.4	95.3	95.3
LagSa2003						98.9	99.3	99.6	98.2	98.4	94.2	94.7	94.7
LagSA2004							99.3	99.1	99.6	99.2	94.9	96	96
Mongoose 2004								99.8	98.9	99.1	94.4	95.3	95.3
LBVSA2006									98.7	98.9	94.7	95.1	95.1
LBVZIM1986										99.8	94.9	95.8	95.8
LBVCAR1974											95.1	96	96
LBVNIG1956												93.3	93.3
LBVSEN1985													100
LBVAFR1999													

The current proposed lyssavirus classification criteria applied by previous studies indicated that intragenotypic identities were greater than intergenotypic identities for gt 1-7, independently of the gene used in the analyses (N, P and G gene) (Kuzmin *et al*., 2003; Kuzmin *et al*., 2005). However, this criteria became problematic when genes other than the N gene was used in an attempt to classify the putative lyssavirus genotypes and overlaps between intragenotypic and intergenotypic identities occurred. In this study the phylogenetic position of lineage A isolates were further analysed with the aim of investigating if these isolates should be considered a new lyssavirus genotype. The shortcomings associated with the current proposed lyssavirus classification criteria were also demonstrated.

Nucleotide and amino acid identities for the complete N, P, M and G gene of lyssaviruses were determined (Appendix 2). Previous studies only included lyssavirus isolates for which sequencing information was available at that specific time. This is the first comprehensive study where all available full length N, P, M and G gene and deduced amino acid sequences for all lyssavirus genotypes and putative genotypes were included in the analysis (Appendix 1). Nucleotide and amino acid identity cannot be less between isolates considered as part of the same lyssavirus genotype (intragenotypic identity) than between isolates considered to belong to separate lyssavirus genotypes (intergenotypic identity). Therefore, the minimum intragenotypic identity must always be higher than the maximum intergenotypic identity (Minimum intragenotypic identity/Maximum intergenotypic identity > 1). This ratio has been analysed for gt 1-7 and all thirteen isolates previously reported to be part of gt 2 (Table 5.4). When lineage A isolates are considered part of gt 2, overlaps occurred between intragenotypic and intergenotypic identities (ratio<1) when analysing the N, P, M and G genes but if considered as a separate lyssavirus genotype these overlaps did not occur (ratio>1) (Table 5.4 and Figure 5.9). Therefore, based on the N, P, M and G gene and amino acid sequence identities, lineage A isolates should be considered a new lyssavirus genotype and lineage B and C considered as part of gt 2.

Table 5.4: Overlaps between intragenotypic and intergenotypic identity between lyssavirus genotypes including Lineage A-C isolates analysed in this study. The ratio of the minimum intragenotypic identity/maximum intergenotypic identity is indicated. A ratio of < 1 indicates an overlap. Where no value is indicated there was only one sequence available and intragenotypic variation could not be determined.

GENOTYPE	N GENE*	N PROTEIN†	P GENE‡	P PROTEIN˙	M GENE#	M PROTEIN¤	G GENE§	G PROTEIN^
1	82.5/80 =1.031	92.4/93.6 =0.987	77.7/71.4 =1.088	67.3/59.7 =1.127	77.7/80.1 =0.970	88.1/92.6 =0.951	79.3/73.2 =1.083	83.4/81.5 =1.023
2 (Lineage A)	99.3/80 =1.241	100/93.6 =1.068	99.6/71.4 =1.395	99.5/59.7 =1.667	99.7/80.1 =1.245	100/92.6 =1.08	99.2/73.2 =1.355	100/81.5 =1.227
2 (Lineage A-C)	79.1/80 =0.989	93.3/93.6 =0.997	67/71.4 =0.938	48.1/59.7 =0.806	77.7/80.1 =0.970	91.1/92.6 =0.984	72.1/73.2 =0.985	80.3/81.5 =0.985
2 (Lineage B and C)	82.7/80 =1.033	94.2/93.6 =1.006	73.4/71.4 =1.028	65.1/59.7 =1.090	82.1/80.1 =1.025	93.6/92.6 =1.011	76.4/73.2 =1.044	84/81.5 =1.031
3	88.5/80 =1.106	94/93.6 =1.004	79.9/71.4 =1.119	68.7/59.7 =1.151				
4	99/80 =1.238	99.2/93.6 =1.06	98/71.4 =1.373	97.3/59.7 =1.63			98.1/73.2 =1.340	97/81.5 =1.190
5	95.1/80 =1.189	98.4/93.6 =1.051	98.7/71.4 =1.382	98.9/59.7 =1.667			94.9/73.2 =1.296	97/81.5 =1.190
6	95.6/80 =1.195	97.3/93.6 =1.04	96/71.4 =1.345	98.3/59.7 =1.647			94/73.2 =1.284	96.5/81.5 =1.184
7	83.7/80 =1.046	96.7/93.6 =1.033	79.9/71.4 =1.119	73.4/59.7 =1.23	83.4/80.1= 1.041	91.6/92.6 =0.989	82.4/73.2 =1.039	89.6/81.5 =1.099

* Maximum intergenotypic identity (80%) observed between gt 4 and gt 5.
† Maximum intergenotypic identity (93.6%) observed between gt 1 (AF045166 and AF351852) and gt 7 (AY573955-58).
‡ Maximum intergenotypic identity (71.4%) observed between gt 1 (AF369365) and gt 7 (AF006497).
˙ Maximum intergenotypic identity (59.7%) observed between gt 6 (AF049121) and gt 7 (AF369371).
Maximum intergenotypic identity (80.1%) observed between gt 4 and gt 5.
¤ Maximum intergenotypic identity (92.6%) observed between gt 4 and gt 5.
§ Maximum intergenotypic identity (73.2%) observed between gt 6 (AY863345) and gt 7 (AF426309/AF426295).
^ Maximum intergenotypic identity (81.5%) observed between gt 5 (AY863338) and gt 6 (AY863344 and AY863345).

This analysis also indicated that variation in gt 1 is high when analysing the nucleoprotein gene and aa sequences (Table 5.4). In gt 1 the most variation were observed in the N gene between a 1980 isolate from an insectivorous bat, *Lasionycteris noctivagans,* from Canada (AF351841) and a 1987 yellow mongoose isolate from South Africa (RVU22628) and for the aa sequence between this yellow mongoose isolate and gt 1 isolates from insectivorous bats (1995 isolate from

Eptesicus fuscus, Canada (AY705373); a SHBV-18 isolate from *Lasionycteris noctivagans*, USA (AF394884) and a 1982 *Lasiurus cinereus* isolate, USA (AF394883). Intragenotypic variation observed in the N protein (aa sequence) of gt 1 isolates was high and indicated an overlap between intergenotypic and intragenotypic identity (ratio<1) (Table 5.4, Table 5.5 and Figure 5.9). Analysis of the M gene and amino acid sequence identities indicated that intergenotypic and intragenotypic identity overlaps for gt 1 and 7 and due to limited sequences available for gt 3-6 this value is unknown for these genotypes. (Table 5.4 and Figure 6.9). This study has shown that N, P and G genes (nucleotide sequence) could be successfully used to classify lyssavirus genotypes but the M gene may be problematic with overlaps occurring (Table 5.5).

Table 5.5: A summary indicating when overlaps between intragenotypic and intergenotypic identity occur, using the N, P, M and G genes to classify lyssavirus genotypes.

	N GENE	P GENE	M GENE	G GENE
Nucleotide	No	No	Yes	No
Amino acid	Yes	No	Yes	No
Overall	Yes	No	Yes	No

*** Yes (Overlap occur); No (No overlap occur)**

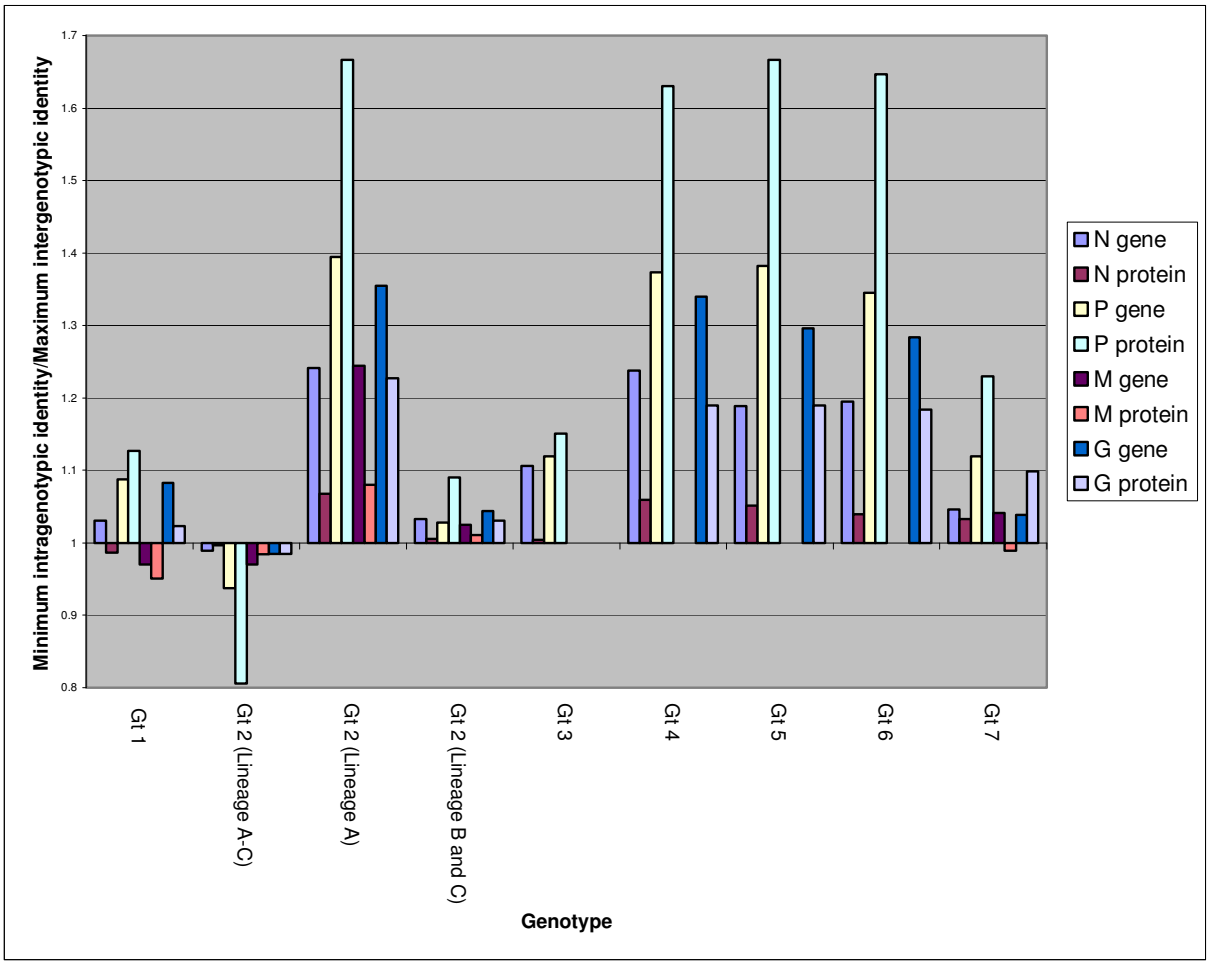


Figure 5.9: Overlaps between minimum intragenotypic and maximum intergenotypic identity observed between lyssavirus genotypes when analysing the nucleotide and amino acid sequence identity of the N, P, M and G genes. The ratio of the minimum intragenotypic identity/maximum intergenotypic identity is indicated. A value of < 1 indicates an overlap. Where no value is indicated there was only one sequence available and intragenotypic variation could not be determined.

The overlap between intragenotypic and intergenotypic identities was analysed when the putative lyssaviruses are considered part of existing lyssaviruses genotypes and not a new lyssavirus genotype (Table 5.6 and Figure 5.10). When analysing only the N gene nucleotide and amino acid identities; Irkut, Aravan, Khujand and WCBV should be considered as new lyssavirus genotypes. However, based on the P and G gene nucleotide and amino acid identities, Aravan and Khujand should be considered as part of gt 6 and Irkut and WCBV could be considered as new lyssavirus genotypes (Figure 5.10).

Table 5.6: Overlaps between intragenotypic and intergenotypic identity for the putative genotypes if considered part of existing lyssavirus genotypes. The ratio of the minimum intragenotypic identity/maximum intergenotypic identity is indicated. A value < 1 indicates an overlap.

	N GENE*	N PROTEIN†	P GENE‡	P PROTEIN•	M GENE#	M PROTEIN¤	G GENE§	G PROTEIN^
Irkut (As part of gt 5)	78.4/80 =0.98	92.7/93.6 =0.990	63.6/71.4 =0.892	56.1/59.7 =0.94	79.6/80.1 =0.994	93.1/92.6 =1.005	70.5/73.2 =0.963	81.3/81.5 =0.998
Aravan (As part of gt 6)	77.5/80 =0.969	88.9/93.6 =0.95	74.2/71.4 =1.039	62/59.7 =1.039	78.8/80.1 =0.984	90/92.6 =0.972	78.7/73.2 =1.075	87.7/81.5 =1.076
Khujand (As part of gt 6)	79.9/80 =0.999	91.1/93.6 =0.973	72.6/71.4 =1.017	65/59.7 =1.089	80.6/80.1 =1.006	90.5/92.6 =0.977	75.8/73.2 =1.036	85.2/81.5 =1.035
WCBV (As part of gt 2)	74.6/80 =0.933	83.8/93.6 =0.895	53/71.4 =0.74	38.2/59.7 =0.64	69.1/80.1 =0.863	80.7/92.6 =0.871	57.5/73.2 =0.786	52.2/81.5 =0.640
WCBV (As part of gt 3)	73/80 =0.913	82.4/93.6 =0.880	53.2/71.4 =0.745	34.7/59.7 =0.581	69.8/80.1 =0.871	76.7/92.6 =0.828	57.9/73.2 =0.791	53/81.5 =0.650

* Maximum intergenotypic identity (80%) observed between gt 4 and gt 5.
† Maximum intergenotypic identity (93.6%) observed between gt 1 (AF045166 and AF351852) and gt 7 (AY573955-58).
‡ Maximum intergenotypic identity (71.4%) observed between gt 1 (AF369365) and gt 7 (AF006497).
• Maximum intergenotypic identity (59.7%) observed between gt 6 (AF049121) and gt 7 (AF369371).
Maximum intergenotypic identity (80.1%) observed between gt 4 and gt 5.
¤ Maximum intergenotypic identity (92.6%) observed between gt 4 and gt 5.
§ Maximum intergenotypic identity (73.2%) observed between gt 6 (AY863345) and gt 7 (AF426309/AF426295).
^ Maximum intergenotypic identity (81.5%) observed between gt 5 (AY863338) and gt 6 (AY863344 and AY863345).

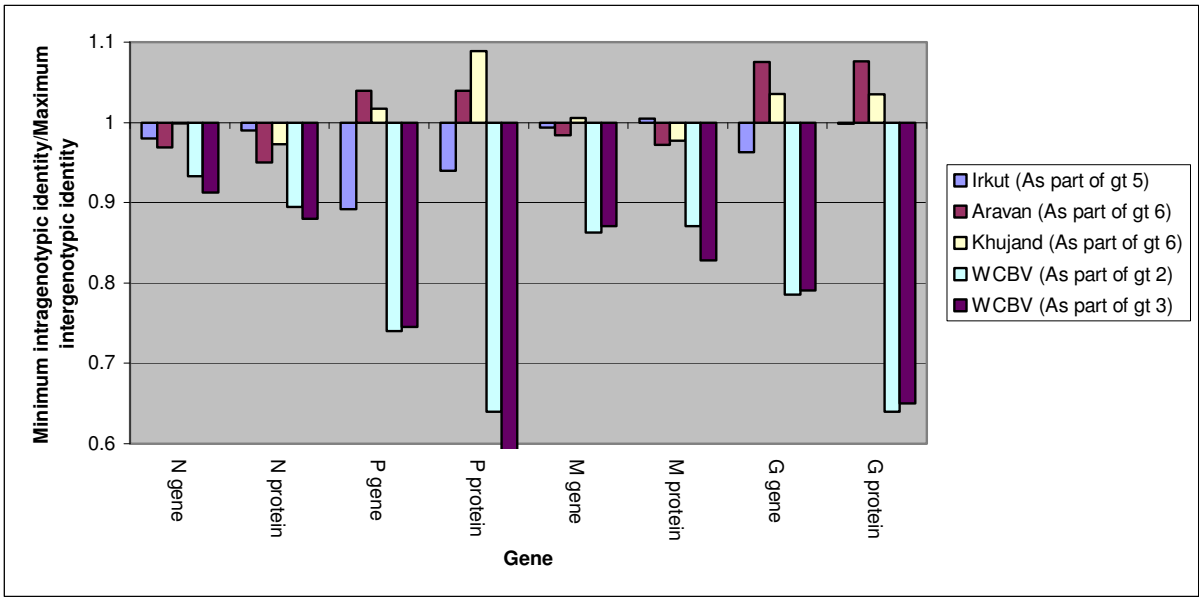


Figure 5.10: Overlaps between intragenotypic and intergenotypic identity for the putative lyssavirus genotypes when analysing the nucleotide and amino acid sequences of the N, P, M and G genes. The ratio of the Minimum intragenotypic identity/Maximum intergenotypic identity is indicated. A value < 1 indicates an overlap.

5.3.3 Antigenic sites

Antigenic sites on the nucleo and glycoproteins of lyssaviruses were previously identified and we have investigated the conservation of these epitopes on the N and G protein of isolates analysed in this study (Figure 5.11 and Figure 5.12).

i) N protein (Figure 5.11)

i) Antigenic site I (aa 355-367) was found to be conserved in gt 2, the new genotype (lineage A) and gt 3. This site was also conserved between gt 1, 4, 5, 7, Irkut and Aravan virus but not in gt 6, Khujand virus and WCBV.

ii) Antigenic site II (aa 313-337), was conserved in all isolates analysed in this study (lineage A and C) except for the LBVNIG1956 isolate/LBU22842 (lineage B) where aa_{322} and aa_{323} were different. This antigenic site was conserved between gt 4 and 5 and between Aravan and Khujand virus but different for other lyssavirus representatives.

iii) Antigenic site III (aa 374-383) was very variable between the different lyssavirus genotypes and between gt 2 isolates and lineage A.

iv) Antigenic site IV (aa 410-413) was conserved between all gt 2, gt 3 and the new genotype (lineage A) isolates and between gt 1, 4, 5, 7, Irkut, Aravan and Khujand virus.

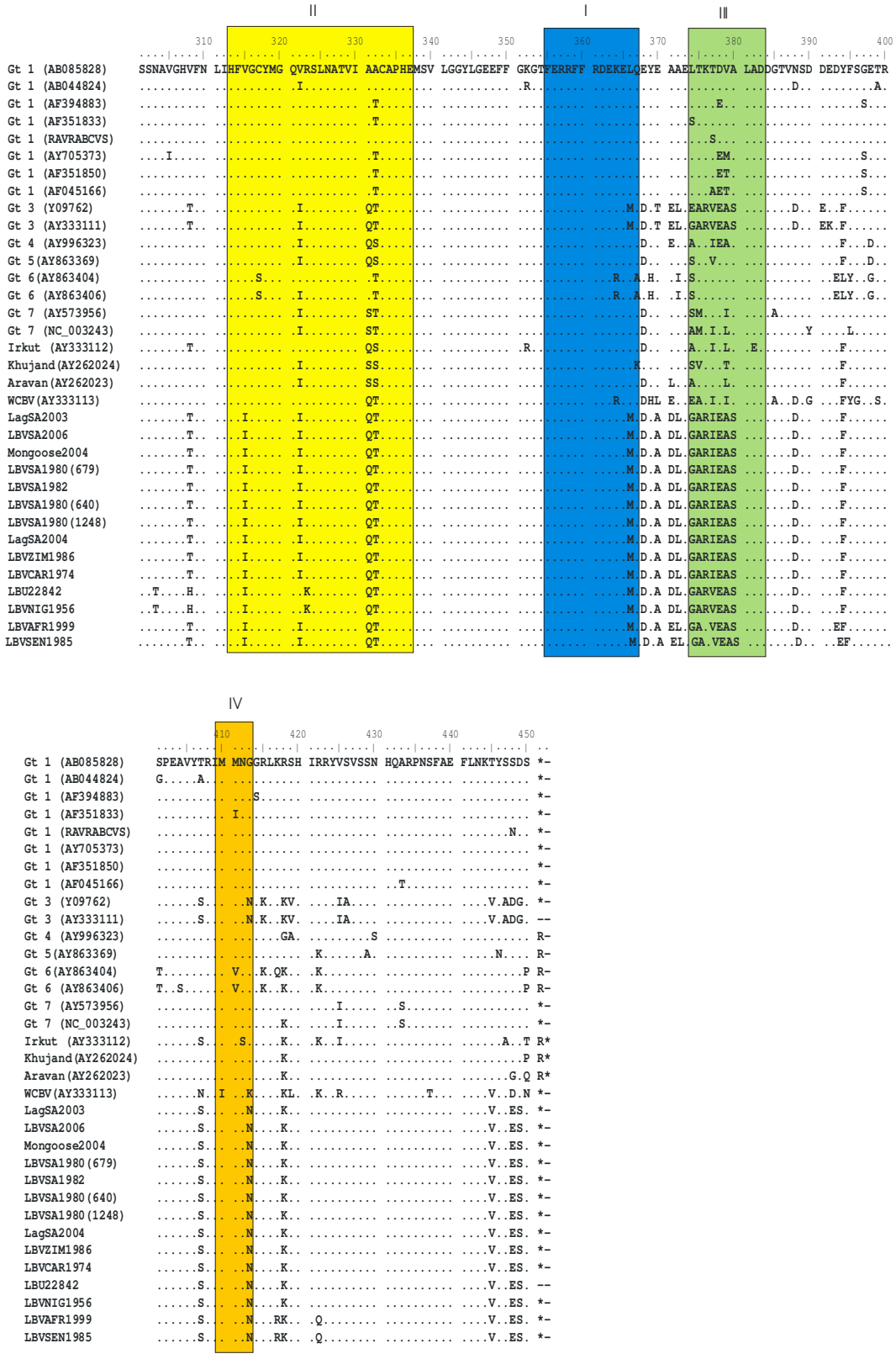


Figure 5.11: Conservation of lyssavirus antigenic sites present on the N protein

ii) G protein

Figure 5.12 indicates an alignment of representatives of all the lyssavirus genotypes and putative genotypes, indicating the antigenic sites present on the G protein.

i) Antigenic site (aa 14-19) was conserved between gt 1, 4, 5, 6, 7, Irkut, Aravan and Khujand virus and conserved between gt 2 and WCBV and different in gt 3.

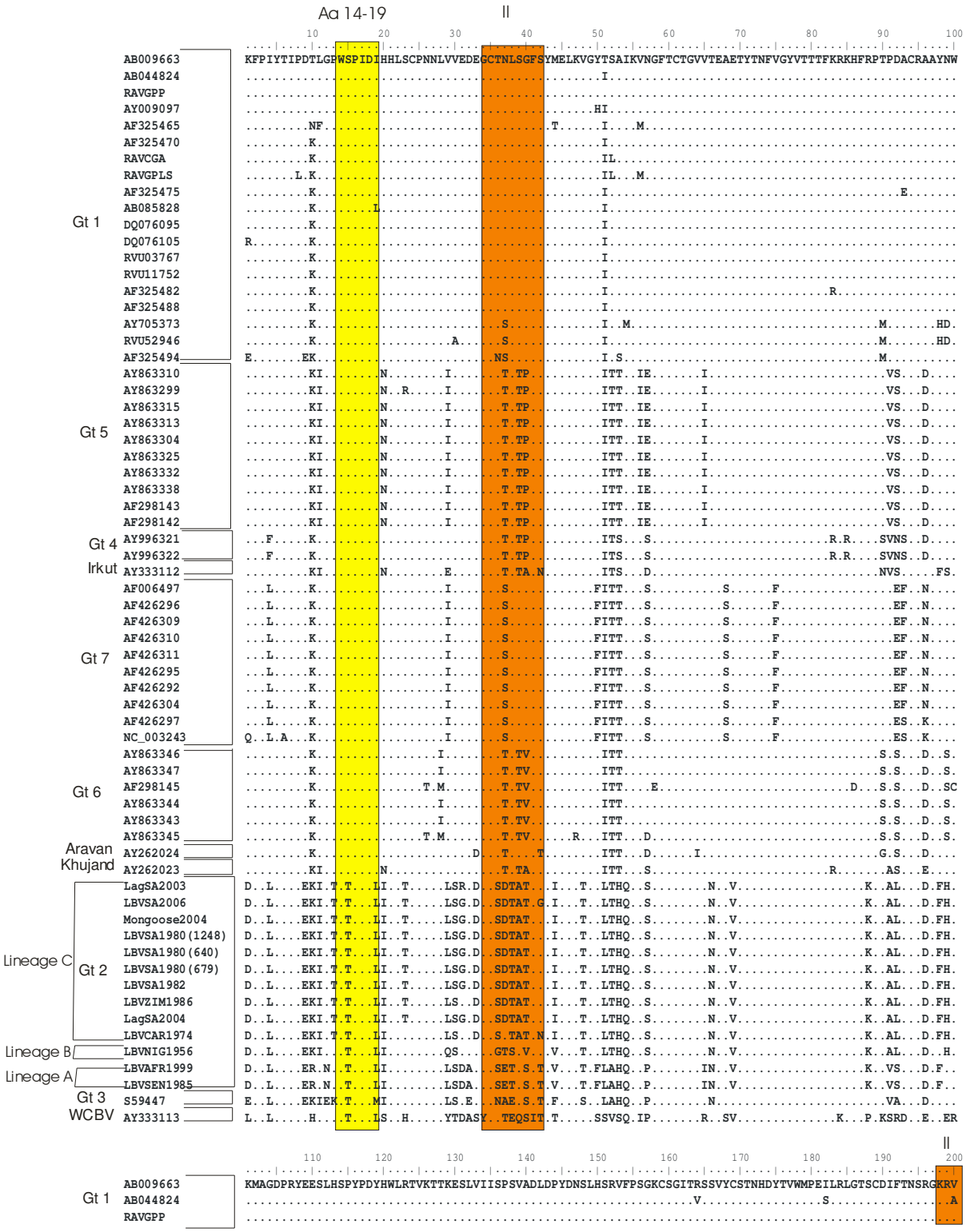
ii) Antigenic site I (aa 231) was conserved in gt 2, 3 and 5 but different in gt 1, 4, 6, 7, Irkut, Aravan, Khujand virus and WCBV.

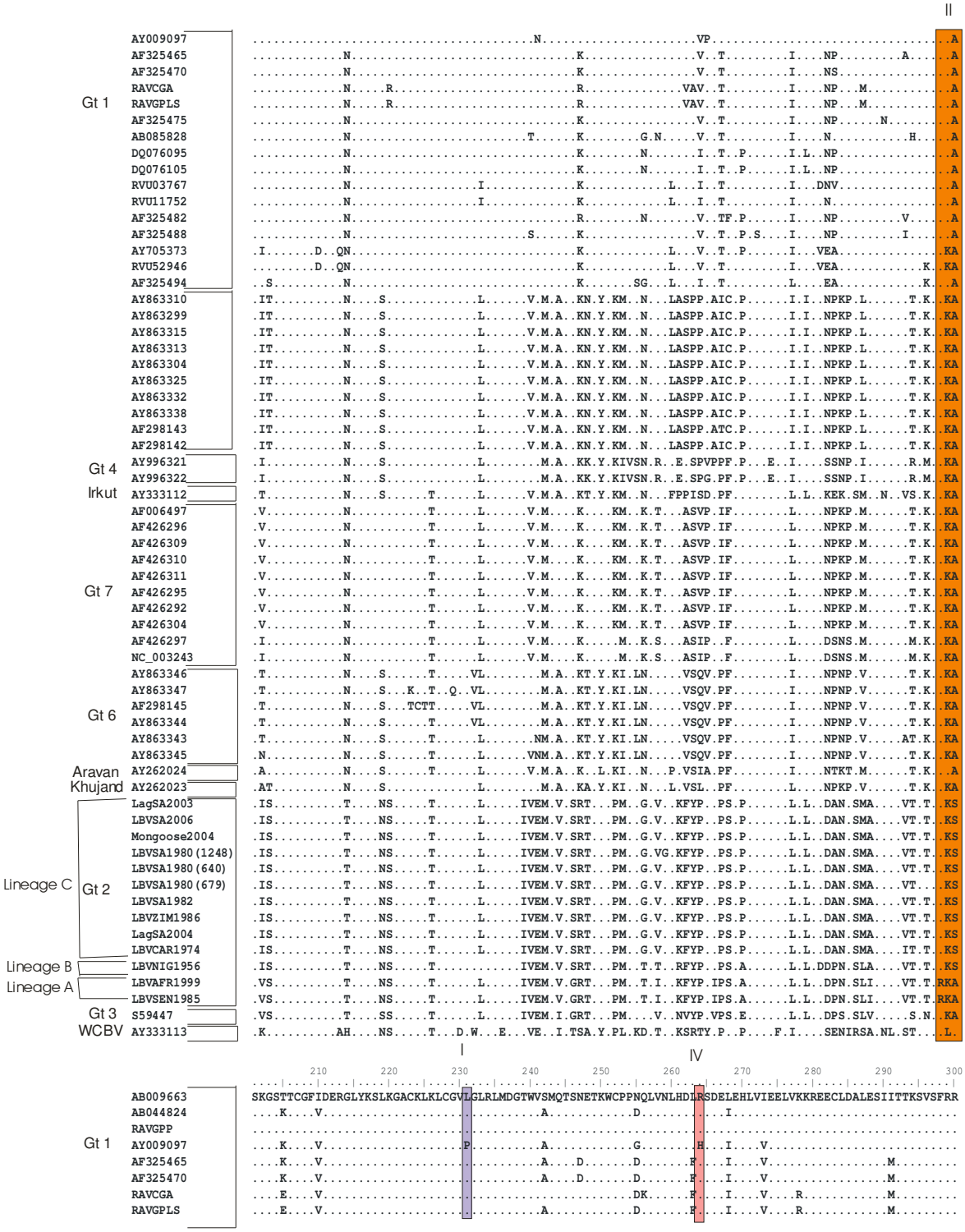
iii) Antigenic site II (aa 34-42 and aa 198-200) was not very well conserved in the lyssavirus genus and was relatively conserved in lineage B and C but different for lineage A isolates (LBVSEN1986 and LBVAFR1999).

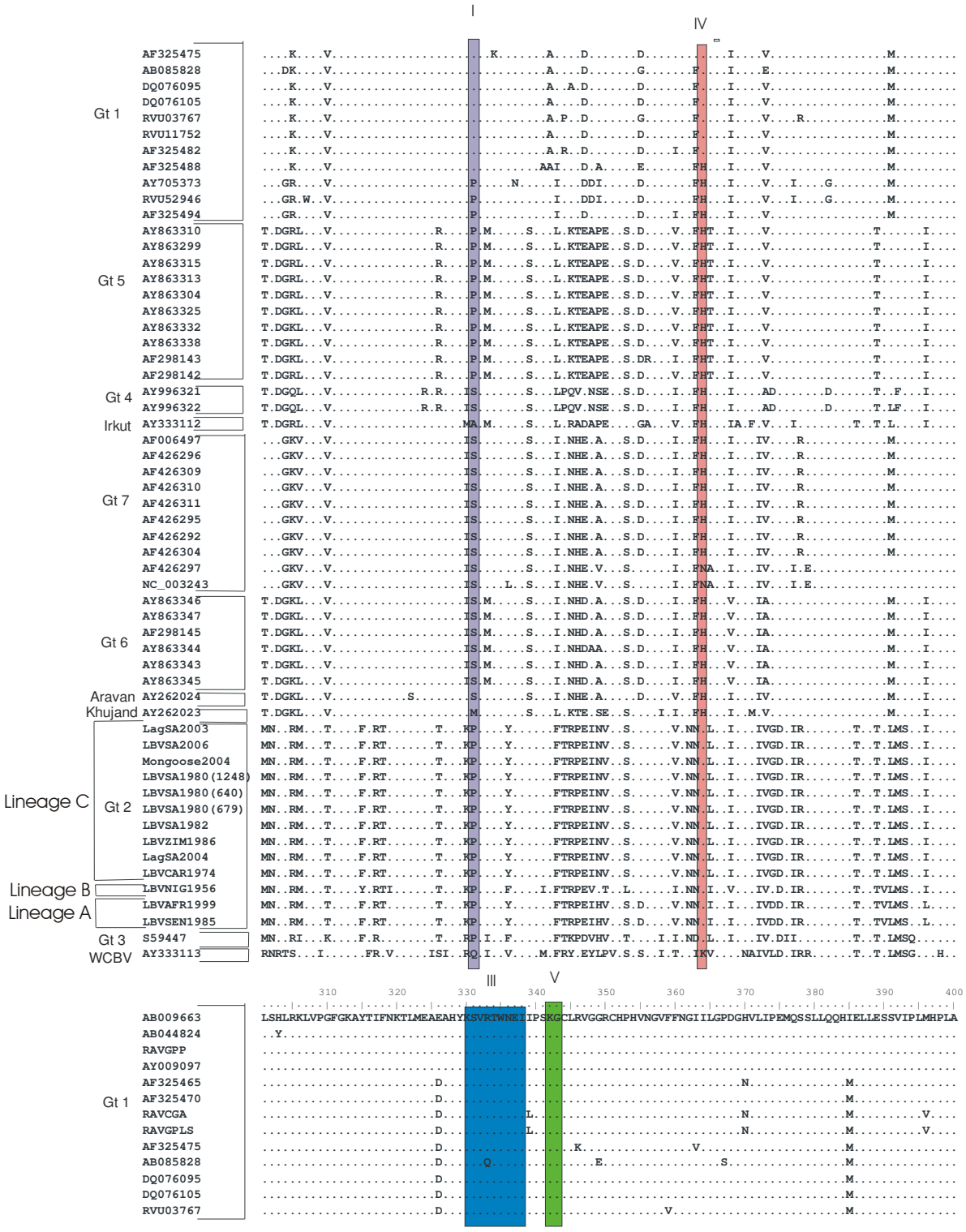
iv) Antigenic site III (aa 330-338) was very different in genotype 2, 3 and WCBV compared to other lyssavirus representatives and also different between lineage C and lineage A and B.

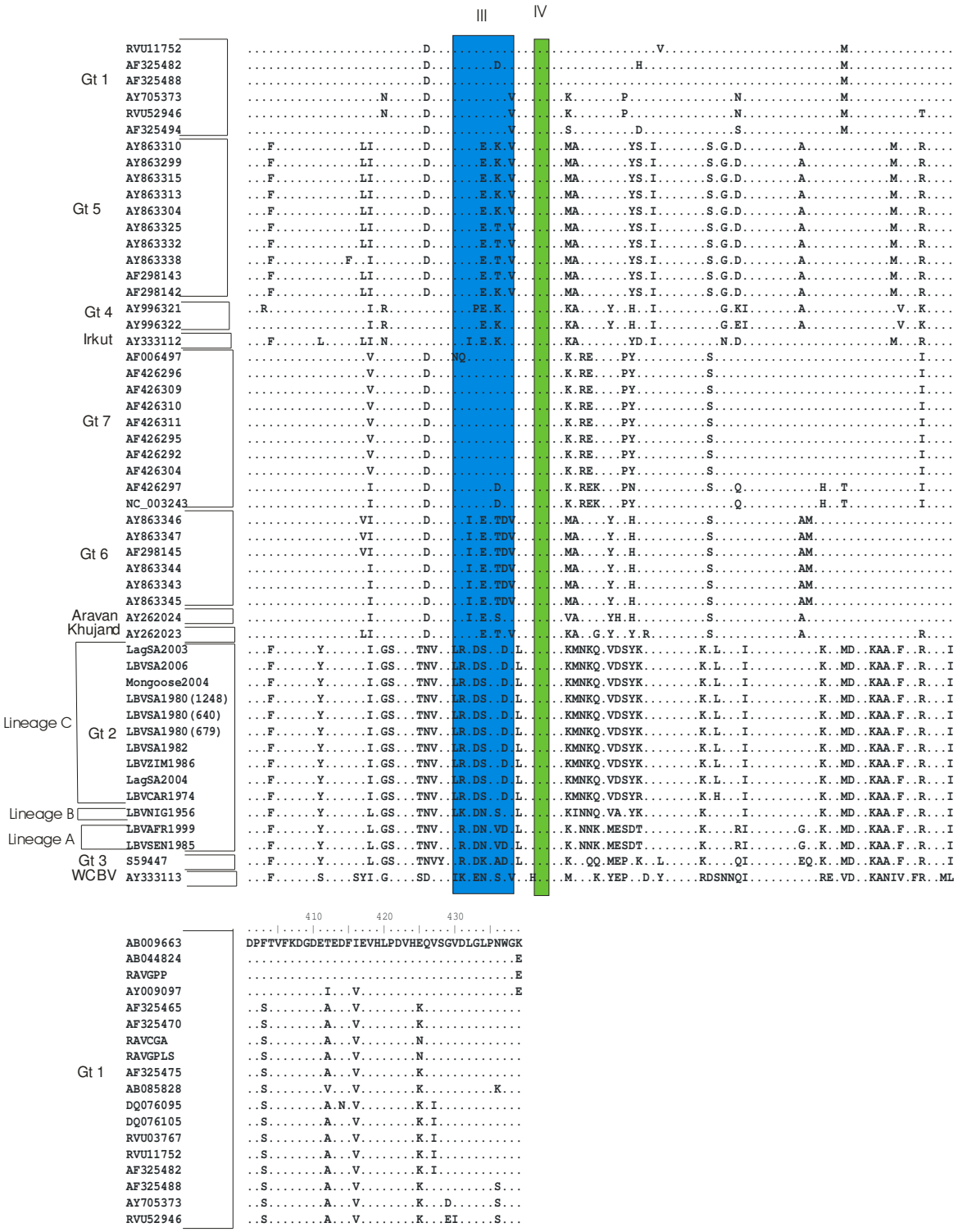
v) Antigenic site IV (aa 264) was relatively conserved in the lyssavirus genus and between gt 2 representatives. Diffferences are observed between gt 1 and gt 7 isolates and this site is also different for WCBV.

vi) Antigenic site V (aa 342-343) was conserved throughout the lyssavirus genus.


Aa 14-19
II
10 20 30 40 50 60 70 80 90 100
AB009663 KFPIYTIPDTLGPWSPIDIHHLSCPNNLVVEDEGCTNLSGFSYMELKVGYTSAIKVNGFTCTGVVTEAETYTNFVGYVTTTFKRKHFRPTPDACRAAYNW
AB044824
RAVGPP
AY009097
AF325465
AF325470
RAVCGA
RAVGPLS
AF325475
AB085828
DQ076095
DQ076105
RVU03767
RVU11752
AF325482
AF325488
AY705373
RVU52946
AF325494
AY863310
AY863299
AY863315
AY863313
AY863304
AY863325
AY863332
AY863338
AF298143
AF298142
AY996321
AY996322
AY333112
AF006497
AF426296
AF426309
AF426310
AF426311
AF426295
AF426292
AF426304
AF426297
NC_003243
AY863346
AY863347
AF298145
AY863344
AY863343
AY863345
AY262024
AY262023
LagSA2003
LBVSA2006
Mongoose2004
LBVSA1980(1248)
LBVSA1980(640)
LBVSA1980(679)
LBVSA1982
LBVZIM1986
LagSA2004
LBVCAR1974
LBVNIG1956
LBVAFR1999
LBVSEN1985
S59447
AY333113
Gt 1
Gt 5
Gt 4
Irkut
Gt 7
Gt 6
Aravan
Khujand
Lineage C
Gt 2
Lineage B
Lineage A
Gt 3
WCBV
II
110 120 130 140 150 160 170 180 190 200
AB009663 KMAGDPRYEESLHSPYPDYHWLRTVKTTKESLVIISPSVADLDPYDNSLHSRVFPSGKCSGITRSSVYCSTNHDYTVWMPEILRLGTSCDIFTNSRGKRV
AB044824
RAVGPP
Gt 1


II
Gt 1
Gt 4
Irkut
Gt 7
Gt 6
Aravan
Khujand
Lineage C
Gt 2
Lineage B
Lineage A
Gt 3
WCBV
I
IV
Gt 1


I
IV
III
V
Gt 1
Gt 5
Gt 4
Irkut
Gt 7
Gt 6
Aravan
Khujand
Lineage C
Gt 2
Lineage B
Lineage A
Gt 3
WCBV
AF325475
AB085828
DQ076095
DQ076105
RVU03767
RVU11752
AF325482
AF325488
AY705373
RVU52946
AF325494
AY863310
AY863299
AY863315
AY863313
AY863304
AY863325
AY863332
AY863338
AF298143
AF298142
AY996321
AY996322
AY333112
AF006497
AF426296
AF426309
AF426310
AF426311
AF426295
AF426292
AF426304
AF426297
NC_003243
AY863346
AY863347
AF298145
AY863344
AY863343
AY863345
AY262024
AY262023
LagSA2003
LBVSA2006
Mongoose2004
LBVSA1980(1248)
LBVSA1980(640)
LBVSA1980(679)
LBVSA1982
LBVZIM1986
LagSA2004
LBVCAR1974
LBVNIG1956
LBVAFR1999
LBVSEN1985
S59447
AY333113
310
320
330
340
350
360
370
380
390
400
AB009663
AB044824
RAVGPP
AY009097
AF325465
AF325470
RAVCGA
RAVGPLS
AF325475
AB085828
DQ076095
DQ076105
RVU03767
LSHLRKLVPGFGKAYTIFNKTLMEAEAHYKSVRTWNEIIPSKGCLRVGGRCHPHVNGVFFNGIILGPDGHVLIPEMQSSLLQQHIELLESSVIPLMHPLA


III
IV
Gt 1
Gt 5
Gt 4
Irkut
Gt 7
Gt 6
Aravan
Khujand
Lineage C
Gt 2
Lineage B
Lineage A
Gt 3
WCBV
410
420
430
Gt 1

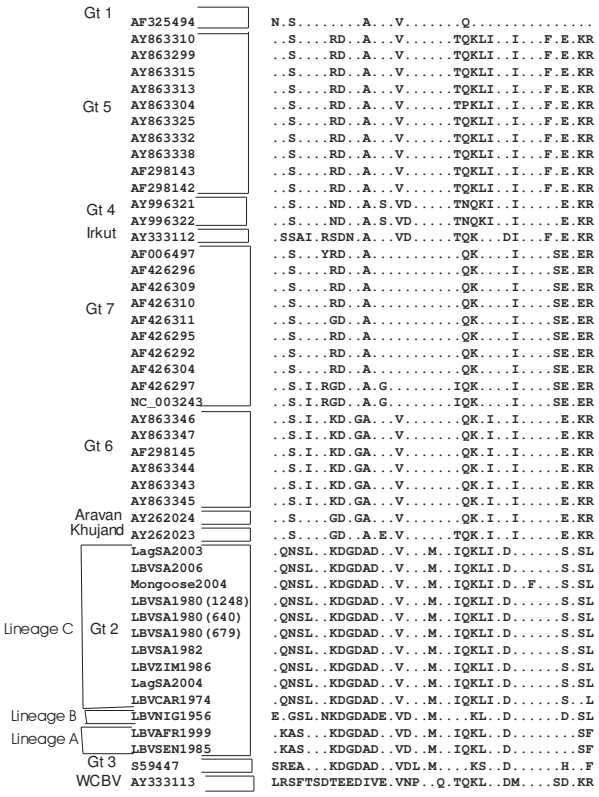


Figure 5.12: Alignment of the deduced amino acids of the ectodomain of the glycoprotein of representatives of the lyssavirus genotypes and putative genotypes. The antigenic sites are indicated.

5.3.4 Nucleoprotein binding motif

The phosphoprotein binds to the nucleoprotein and a conserved lyssavirus nucleoprotein binding motif (aa 209-215) has been identified in previous studies. When analysing the conservation of this motif, a substitution in aa 212 (K→R) was observed between isolates belonging to lineage C of gt 2 compared to isolates belonging to lineages A and B (Figure 5.13). A substitution at aa 209 was observed for WCBV and for gt 7 at aa 215.

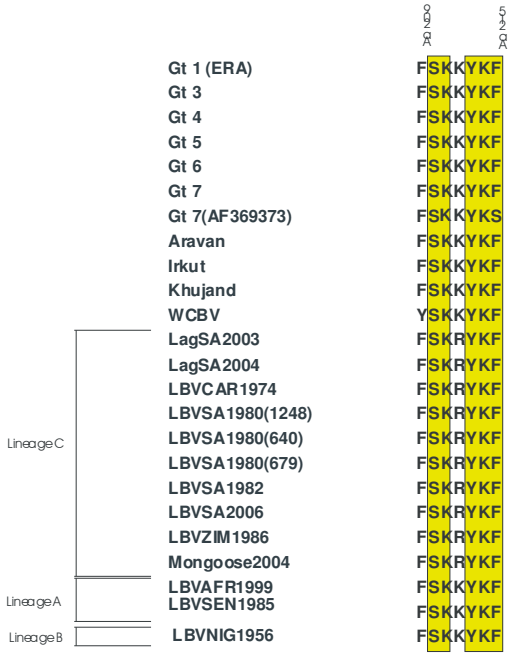


Figure 5.13: Comparison of the C terminal nucleoprotein binding domain (aa 209-215) on the P protein of lyssavirus representatives.

5.3.5 Binding site for the cytoplasmic light chain of dynein, LC8

This binding site was conserved between gt 1, 4, 5, 6, Aravan and Khujand virus but not in gt 3, 7, Irkut virus and WCBV (Figure 5.14). In the isolates studied in this study the site was conserved in lineage C but different in lineage A and B. This domain may play a role in lyssavirus pathogenesis and will be further analysed in Chapter VI.

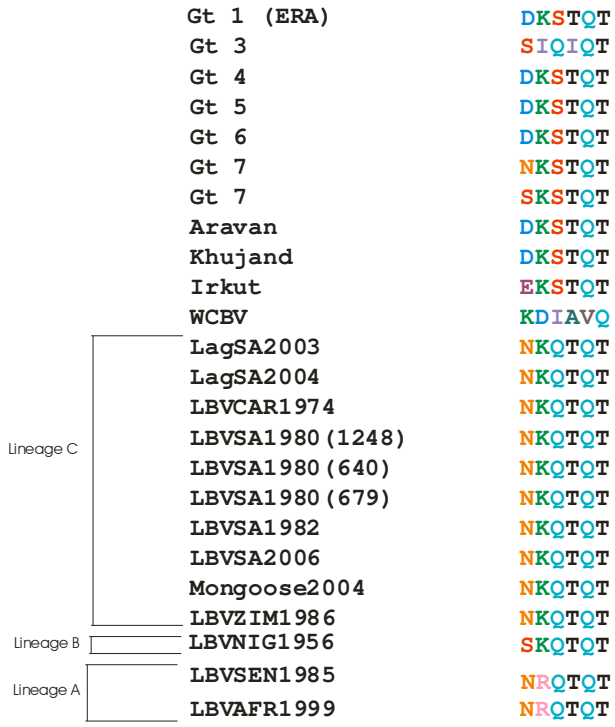


Figure 5.14: Comparison of the binding site for the cytoplasmic light chain of dynein, LC8 (aa 143-148) of the P protein of lyssavirus representatives.

5.4 Discussion

In this study, we provided the first molecular analysis of lyssavirus isolates previously classified as gt 2 using complete N, P, M and G gene and protein sequences. The genetic diversity of this group is certainly underestimated due to only a few isolates available. Nevertheless, molecular phylogeny divided these isolates into three lineages independent of the method or gene used in the analysis. Lineage A consisted of two isolates, one from Senegal and the other from unknown origin but probably also from North-West Africa. Lineage B consisted of one isolate from Nigeria. Within lineage C three distinct separate groups were observed; a) eight isolates from South Africa, b) an isolate from Zimbabwe and c) an isolate from the Central African Republic. Very little sequence variation was observed between the isolates from South Africa even though these isolations were made over several years (1980-2006). Surveillance of LBV in South Africa has focused on one geographical area (Durban, KwaZulu Natal) and isolates from this region were mainly from the fruit bat species, *E. wahlbergi,* implicating that LBV circulates in this fruit bat species in South Africa. Improved surveillance in other areas in South Africa in this and other species will indicate whether genetic relatedness coincides with

geographical location and host species. A very low intrinsic heterogeneity was observed in all the lineages that may suggest an adaptation to host or geographical domain.

It has been shown for other rabies related viruses, EBLV 1 and 2, that viruses from the same host tend to cluster together in phylogenetic analysis (Davis *et al*., 2005). Phylogenetic analysis of RABV of terrestrial mammals indicated that these isolates cluster more by geographical origin than by host species (Kuzmin *et al*., 2004; Nadin-Davis *et al*., 1999; Kissi *et al*., 1995; Bourhy *et al*., 1999; Holmes *et al*., 2002) whereas bat isolates belonging to gt 1 form distinct lineages associated with different bat species (Kobayashi *et al*., 2005) or bat species with similar behavioural traits even when separated by long distances (Davis *et al*., 2006). This has also been indicated for ABLV where separate pteropid and insectivorous ABLV variants co-circulate in specific bat hosts (Gyuatt *et al*., 2003). These observations may also be true for gt 2 isolates.

In this study analysis of the complete N, P, M and G gene nucleotide and amino acid sequences supported the classification of LBVSEN1985 and LBVAFR1999 (lineage A) as a new lyssavirus genotype when using current criteria for lyssavirus genotype classification. The isolate from Senegal (LBVSEN1985) was previously only reported as gt 2 based on complement fixation tests indicating its relatedness to a gt 2 isolate from the Central African Republic (LBVCAR1974) (Institute Pasteur, 1985). The LBVAFR1999 isolate was reported as gt 2 based on very limited nucleoprotein sequencing data that indicated that it was closely related to the LBVSEN1985 isolate (Aubert, 1999). Neither of these isolates was ever further characterized. In phylogenetic analysis performed in this study, lineage A formed a separate cluster with a high bootstrap support independent of the gene used. Comparison of lineage A isolate's N, P, and G nt sequences indicated that they share the highest identity with gt 2 and the lowest level of nucleotide identity with gt 1 and gt 6. P-distance analysis of N, P, M and G gene and amino acid sequences indicated that if lineage A is considered part of genotype 2, overlaps between intragenotypic and intergenotypic identities of lyssavirus genotypes will occur. However, if these isolates are considered a separate lyssavirus genotype, no overlaps occur. Lineage A also indicated a <80% nucleoprotein nucleotide identity to other lyssavirus genotypes and putative genotypes and should therefore be considered a new lyssavirus genotype.

Previous guidelines suggesting that the threshold for defining a new lyssavirus genotype was <80% nucleotide and <93% amino acid similarity of the nucleoprotein, were suggested more than a decade ago (Kissi et al., 1995; Bourhy *et al*., 1993) when limited sequencing information about the lyssavirus genus was available. With the discovery of the four putative lyssavirus genotypes; Irkut, Aravan, Khujand and WCBV, this criteria became problematic since overlaps between intergenotypic and intragenotypic identities occured and it became apparent that with new information available this criteria need to be reviewed. In the data presented here, it was indicated that analysis of N, P and G gene intragenotypic and intergenotypic nucleotide identities supported the classification of gt 1-7, Irkut and WCBV as separate genotypes, however, only N gene analysis supported the classification of Aravan and Khujand virus as separate lyssavirus genotypes. A high intragenotypic variation in amino acid identities were observed between gt 1 isolates and an overlap between intragenotypic and intergenotypic identities occurred when analysing gt 1 nucleoprotein amino acid sequences.

The criteria for classification of lyssaviruses should be revised to accommodate more diversity in the existing lyssavirus genotype classification scheme rather than defining new genotypes (Kuzmin *et al*., 2003; Kuzmin *et al*., 2005). Host species, geographical origin and pathogenic characteristics may contribute towards understanding lyssavirus diversity, but such characteristics should be examined with caution. Among other factors the lack of samples often leads to such properties being defined on the basis of one or at most a few individual isolates. As indicated in this study not all genes provided the same classification for the putative lyssavirus genotypes and overlaps even occurred when analysing the N gene amino acid identities of gt 1 isolates. The M gene is also not a good candidate for lyssavirus classification. To formulate a new criteria for lyssavirus classification it is necessary to obtain more full length lyssavirus gene sequences for representatives of all lyssavirus genotypes to determine the intragenotypic variation in each genotype and formulate a new classification criteria. A specific lyssavirus gene could also be selected to be used for classification purposes. As indicated in this study, only the nucleoprotein gene nucleotide identity (not aa sequences) provided a clear distinction between gt 1-7 and the putative genotypes. When this gene is used for lyssavirus classification, the current criteria suggesting, <80% nucleotide identity of the nucleoprotein gene constitutes a new lyssavirus genotype, still applies. This criteria will, however, lead to more and more lyssavirus genotypes being described as more isolates are discovered and molecularly characterized.

Analysis of antigenic epitopes present on the N and G proteins of lineage A isolates also indicated differences between lineage A and Lineage B and C for aa 410-413 (N-protein) and antigenic site II and antigenic site III of the glycoprotein. This may represent serological differences that may even classify these isolates into different serological groups. For the isolates studied in this study the binding site for the cytoplasmic light chain of dynein, LC8 (aa 143-148) of the P protein was found to be conserved in lineage C but different in lineage A and B. This domain may play a role in lyssavirus pathogenesis and these sequencing differences may be an indication of different pathogenic properties of these isolates. In the conserved lyssavirus nucleoprotein binding motif (aa 209-215), a substitution in aa 212 (K→R) was observed between isolates belonging to lineage A and B and isolates belonging to lineage C, indicating that these lineages may have different nucleoprotein binding properties. These additional sequencing differences in antigenic and pathogenic domains provide further support for lineage A to be considered as a new lyssavirus genotype but need further study to fully understand the biological significance thereof.

CHAPTER VI

PATHOGENESIS OF LAGOS BAT VIRUS IN A MOUSE MODEL

6.1 Introduction

It has been suggested that the *Lyssavirus* genus could be divided in two different phylogroups, phylogroup I and II (Badrane *et al*., 2001b). Phylogroup I was proposed to consist of RABV (gt 1), DUVV (gt 4), EBLV 1 (gt 5), EBLV 2 (gt 6) and ABLV (gt 7) and phylogroup II of MOKV (gt 3) and LBV (gt 2). Phylogroups were described based on the antigenic, genetic and pathogenic differences among the lyssaviruses. It was also proposed that WCBV may represent a new phylogroup III (Kuzmin *et al*., 2005). Within a phylogroup the aa sequence of the G protein ectodomain was found to be 74% identical and between phylogroups only 64.5%. After performing pathogenesis studies in mice with only one representative of each genotype, viruses from phylogroup I were found to be pathogenic when introduced intracerebrally (i.c.) or intramuscularly (i.m.) whereas phylogroup II and III viruses were only pathogenic when introduced into mice via the i.c. route (Badrane *et al*., 2001b). The R333/K333 aa (positively charged aa) in the G protein ectodomain was found to be essential for the virulence of gt 1 viruses, ERA and CVS (Tuffereau *et al*., 1989). Genetic analysis of the G protein ectodomain of a few representatives of the different phylogroups indicated that the R_{333} residue essential for virulence was replaced by a D_{333} in phylogroup II viruses (Badrane *et al*., 2001b). Phylogroup II viruses were suggested to be less pathogenic and therefore considered less of a danger to public health due to the reduced pathogenicity in mice. Specific domains that play an important role in lyssavirus pathogenesis have been identified on the G and P proteins (Table 6.1).

Only a few pathogenesis studies which included the phylogroup II viruses have been carried out. The pathogenicity of MOKV in shrews was investigated and studies indicated that shrews can be infected experimentally via the i.m., oral and subcutaneous (s.c.) route (Kemp *et al*., 1973). Another study investigated the pathogenicity of MOKV and LBV in dogs and monkeys using high concentration virus inoculum introduced via the i.c. and i.m. routes (Percy *et al*., 1973; Tignor *et al*., 1973). Both MOKV and LBV seemed to be less pathogenic when inoculated i.m. and serological responses were observed in animals inoculated i.m. but were insignificant in animals inoculated via the i.c. route.

Table 6.1: Domains on the lyssavirus genome implicated to be important in pathogenesis

REGION ON THE GENOME	FUNCTION	REFERENCE
aa 143-148 (P protein)	Essential for the interaction of the LC8 dynein light chain with the P protein. Dynein is involved in minus end-directed movement of organelles along microtubules and may therefore be involved in retrograde transport of virions through the CNS	Poisson *et al*., 2001
aa 333 (Ectodomain of G protein)	Changes in this region could lead to a less pathogenic or avirulent virus in immune competent mice. Virulence was strongly associated with the presence of a charged aa in this position.	Tuffereau *et al*., 1989; Coulon *et al*., 1989; Seif *et al*., 1985; Dietzschold *et al*., 1983; Takayama-Ito *et al*., 2006a
aa 330 (Ectodomain of the G protein)	Double mutation together with aa 333 led to a further reduction in pathogenicity of the virus compared to only a single aa 333 mutation	Coulon *et al*., 1998
aa 164, 182, 200, 205, 210, 242, 255, 268 and 303 (Ectodomain of the G protein)	These amino acids have been implicated as essential for pathogenicity	Takayama-Ito *et al*., 2004; Takayama-Ito *et al*., 2006b

The present study was designed to investigate the comparative pathogenesis of gt 2, gt 3 and gt 1 in a mouse model. Among other, the dose of inoculum, route of inoculation and serological responses were compared. Although previous studies indicated that phylogroup II viruses may be less pathogenic to mice than phylogroup I viruses (Badrane *et al*., 2001b) this was further investigated with the inclusion of more isolates of these genotypes. The phylogroup II representatives studied so far indicated a high sequence diversity in this group compared to genotypes in phylogroup I (Badrane *et al*., 2001b) and these genotypes may be even more diverse. Amino acid differences on the lyssavirus genome that may play a role in pathogenesis were also analysed.

6.2 Materials and methods

6.2.1 Animals

Four-week-old inbred ICR mice obtained from Harlan Sprague Daly (USA) were used in experimental infections. Each mouse was tagged using an ear tag with a unique number for identification purposes (National band and Tag Co, USA). All animal care and experimental procedures were performed in compliance with the Centers for

Disease Control and Prevention Institutional Animal Care and USE Guidelines (USA).

6.2.2 Viruses

Twelve lyssaviruses were included in this study (Table 6.2) and were first amplified in suckling mouse brain using i.c. inoculation. Brains were collected after mice were either euthanized or succumbed and the presence of lyssavirus antigens was confirmed using the FAT test (Section 3.2.2.1). 10% mouse brain suspensions were prepared in Minimum Essential Medium (MEM-10, GIBCO) supplemented with 10% fetal calf serum. The mixtures were centrifuged at 3 200 g for 15 minutes and the supernatant stored at -70˚C for further use. The titre of the viral inoculum was determined by inoculating 4-week-old ICR mice intracerebrally and the 50% lethal dose (LD_{50}) was calculated using the Spearman-Karber method (Aubert, 1996). Brain material removed from all mice that succumbed to disease during determination of titres was confirmed to be positive for lyssavirus antigen with the FAT (Section 3.2.2.1).

Table 6.2: Information about lyssaviruses used in experimental infections of mice

VIRUS NAME	GENOTYPE	YEAR OF ISOLATION	GEOGRAPHICAL LOCATION	SOURCE	PASSAGE HISTORY
WAmyotis spp	1	2004	Washington, USA	Dr. C.E. Rupprecht (CDC, Atanta, USA)	Passage 3 in MNA cells
MOKVSA(252/97)	3	1997	South Africa	Dr. C.T. Sabeta (OVI, Rabies Unit, South Africa	Freeze dried mouse brain material. Passage multiple times in mouse brain
LBVNIG1956	2	1956	Nigeria	Dr. C.E. Rupprecht (CDC, Atanta, USA)	Unknown
LBVCAR1974	2	1974	Central African Republic	Dr. C.E. Rupprecht (CDC, Atanta, USA)	Passage 2 in mouse brain
LBVSA1982	2	1982	South Africa	Dr. C.T. Sabeta (OVI, Rabies Unit, South Africa)	Freeze dried mouse brain material. Passage multiple times in mouse brain
LBVSEN1985	2	1985	Senegal	Dr. C.E. Rupprecht (CDC, Atanta, USA)	Passage 4 in mouse brain
LBVZIM1986	2	1986	Zimbabwe	Dr. C.E. Rupprecht (CDC, Atanta, USA)	Passage 2 in MNA cells
LBVAFR1999	2	1999	Egypt/Togo	Dr. F. Cliquet (AFSSA, France)	Passage 3 in mouse brain
LagSA2003	2	2003	South Africa	This study	Original bat brain material
LagSA2004	2	2004	South Africa	This study	Original bat brain material
Mongoose2004	2	2004	South Africa	This study	Original bat brain material
LBVSA2006	2	2006	South Africa	This study	Original bat brain material

6.2.3 Experimental infections

Four-week-old female ICR mice were inoculated with different lyssavirus isolates (Table 6.2) using different routes of inoculation and a different dose of inoculum as indicated in Table 6.3. I.c. and i.m. inoculations were performed with an ultrafine II, short, ½ cc, 8 mm, 31 gauge needle (Becton Dickinson, USA). Oral inoculation was performed with a manual pipette and inoculum was introduced into the mouth cavity at 30 µl volumes at a time. Care was taken not to cause lesions in the mouth and that all inoculum was introduced. Groups constituted five mice each. Mice were observed for 56 days and clinical signs and mortality were recorded daily. Blood was collected on day 0, 5, 8, 14, 21, 28, 35 and 56 or till mortality using retro-orbital bleeding with a 250 µl heparinised Natelson blood collecting tube (Chase Instruments, USA). Animals that died within 24 hours after inoculation were excluded from the study.

Table 6.3: Information about the different experiments performed to investigate the pathogenesis of gt 2 isolates using different routes of inoculation and different doses of viral inoculum

GROUP	ROUTE OF INOCULATION	DOSE OF INOCULUM
Group A	i.c.	1000 LD_{50}
Group B	i.c.	LD_{50}
Group C	i.c.	10^{-2} LD_{50}
Group D	i.m. (left hind limb)	100 000 LD_{50}
Group E	i.m. (left hind limb)	1 000 000 LD_{50}/Maximum dose
Group F	Oral	1 000 000 LD_{50}/Maximum dose

6.2.4 Fluorescent antibody test (FAT)

The FAT was performed on mouse brain collected from mice that succumbed to disease or were euthanized at the end of the experiment on day 56. The standard protocol as explained in Section 3.2.2.1 was used.

6.2.5 RT-PCR and DNA sequencing

RNA was extracted from brain material removed from mice that tested positive with the FAT using the methods as described in Section 5.2.2. cDNA synthesis, PCR and DNA sequencing were performed as described in Section 5.2.4-5.2.6. The primers 001lys and 550B as described in Section 5.2.3 were used. Nucleotide sequences

were compared to sequence generated in Chapter V to confirm the identity of the virus isolate in the mouse brain.

6.2.6 Analysis of amino acid sequences

Complete amino acid sequences of P, M and G proteins of isolates used in pathogenesis studies (Table 6.1) were aligned using ClustalW and differences in amino acid sequences that may be involved in pathogenesis were identified.

6.2.7 Determination of the presence of neutralizing antibodies

Blood collected from mice was tested for the presence of neutralizing antibodies using the Rapid Fluorescent Focus Inhibition Test (RFFIT) as described by Smith *et al*, (1996) with modifications adjusting this method to analyse very low amounts of test serum. Blood was collected via the retro-orbital route and sera were separated in Microtainer® serum separation tubes with SST™ (Becton Dickinson and Company, USA) as suggested by the manufacturer, aliquoted and stored at –20°C until analysis.

6.2.7.1 Preparation of challenge viruses

Challenge virus to detect LBV, MOKV and RABV neutralizing antibodies were prepared as described in Section 4.2.11.1. A LBV isolate, isolated in Senegal in 1985*, was used as a representative of gt 2. A MOKV isolate isolated from South Africa in 1997 (252/97) was used as a representative of gt 3. CVS was used to detect gt 1 neutralizing antibodies. Stock suspension of each virus was diluted to contain $50FFD_{50}$ for use in the RFFIT.

6.2.7.2 The Rapid Fluorescent Focus Inhibition Test (RFFIT)

The test was prepared in a 4 well (6 mm in diameter) Teflon coated microtiter plate (Cel-Line/Erie scientific company, USA). The collected serum was heat inactivated at 56°C for 30 minutes. Briefly, 13.7 µl MEM-10 (GIBCO) supplemented with 10% fetal bovine serum (GIBCO) was transferred into the first well and thereafter 12 µl into the remaining 3 wells. 1.3 µl of the test serum were added to the first well and mixed by pipetting after which 3 µl was serially carried over from the first well to the last well on the slide creating a 5 fold dilution of the serum. 3 µl of challenge virus preparation ($50FFD_{50}$) was then added to each well of the test. A control slide was prepared with 12 µl MEM-10 (GIBCO) in each well and setting up a back titration of

* The lineage A strain (LBVSEN1985) was chosen as a challenge virus before sequencing data was available indicating the high sequence diversity compared to other gt 2 isolates. A representative number of serum samples were retested for the presence of antibodies using the LagSA2004 isolate as a challenge virus in the RFFIT and similar results were obtained.

the challenge virus in a 10 fold serial dilution. One well was left uninfected to serve as a cell culture control in the test. The dilutions were incubated at 37℃ and 0.5% CO_2 for 90 minutes in a humidity chamber. After incubation 24 µl (about 5.0 X 10^5 MNA cells per ml) was added to the reactions and incubated again at 37℃ and 0.5% CO_2 for 24 hours. After the incubation the cell culture supernatants were decanted and the slides dip-rinsed in phosphate buffered saline (PBS) (13.7 mM NaCl, 0.27 mM KCl, 0.43 mM $Na_2HPO_4.2H_2O$, 0.14 mM KH_2PO_4, pH 7.3) and transferred to ice-cold acetone for 30 minutes at -20˚C. Slides were then air dried and stained with 12 µl of rabies conjugate/well (Fujirebio Diagnostics Inc) and incubated at 37℃ for 45 minutes. Following the incubation the conjugate was washed from the wells with PBS. The slides were dip-rinsed in distilled water and air dried before reading. The slides were read at 160 to 200 times magnification under a fluorescent microscope and 20 microscope fields per well were observed. The RFFIT results were expressed as endpoint titres.

6.3 Results

6.3.1 Titration of viruses

Viruses were titrated by intracerebral inoculation of four-week-old ICR mice and determined using the Spearman-Karber method. Titres are indicated in Table 6.4. FAT performed on brain material collected from dead mice in titration experiments all tested positive for the presence of lyssavirus antigens

Table 6.4: Titres of viruses used in the pathogenesis experiment as determined by i.c. inoculation of four-week-old ICR mice.

VIRUS NAME	TITRE
WAmyotis spp (gt 1)	4.4 log LD_{50}/0.03 ml
MOKVSA(252/97) (gt 3)	5.4 log LD_{50}/0.03 ml
LBVNIG1956	2.5 log LD_{50}/0.03 ml
LBVCAR1974	5.3 log LD_{50}/0.03 ml
LBVSA1982	5.5 log LD_{50}/0.03 ml
LBVSEN1985	7.1 log LD_{50}/0.03 ml
LBVZIM1986	4.9 log LD_{50}/0.03 ml
LBVAFR1999	8.1 log LD_{50}/0.03 ml
LagSA2003	5.1 log LD_{50}/0.03 ml
LagSA2004	4.3 log LD_{50}/0.03 ml
Mongoose2004	5.1 log LD_{50}/0.03 ml
LBVSA2006	4.1 log LD_{50}/0.03 ml

6.3.2 Susceptibility

Mortality and serological responses were recorded for each animal in the experiment (Appendix 3). To examine the susceptibility of mice to different lyssaviruses isolates, the mortalities after i.c., i.m. and oral inoculation with different concentrations of viral inoculum were compared (Figure 6.1-6.3). Brain tissue collected from all mice that succumbed during the pathogenesis experiments or that survived were tested using FAT. All survivors tested negative and all mice that succumbed to disease tested positive for lyssavirus antigens.

i) Intracerebral inoculation

Intracerebral inoculation with gt 1, 2 and 3 lyssaviruses produced similar results leading to a 100% mortality when inoculated at a 100 LD_{50} with all isolates tested and subsequently to a lower percentage of deaths when lower amounts of virus was introduced. The LBVNIG1956 isolate's results were not included in this experiment due to its low titre after amplification in mice.

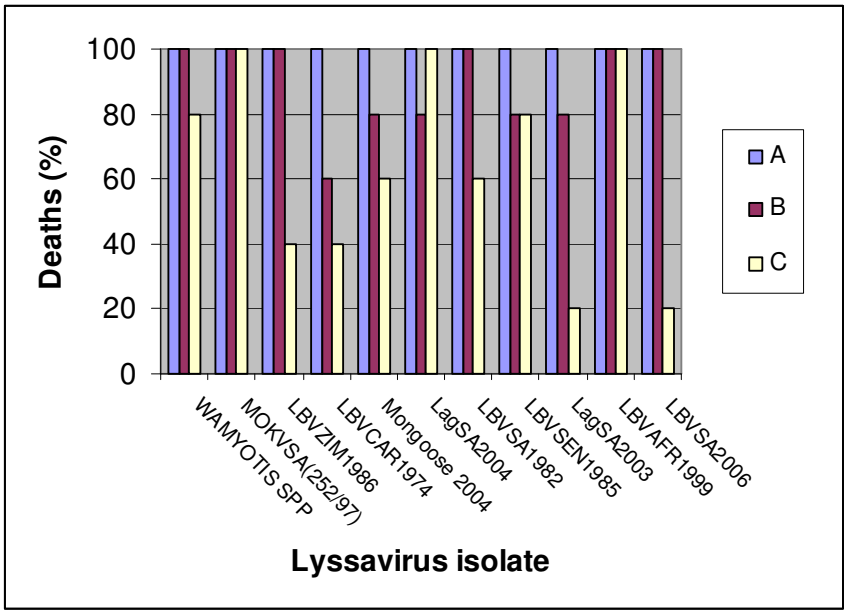


Figure 6.1: Pathogenicity of genotype 1 (WAMYOTIS SPP), gt 2 and gt 3 (MOKVSA(252/97)) lyssaviruses in 4-week-old ICR mice after intracerebral (i.c.) inoculation. Results are expressed as a percentage of dead animals after observation for 56 days. Different viral doses were introduced; A: 1 x 10^2 LD_{50}; B: LD_{50} and C: 10^{-2} LD_{50}.

ii) Intramuscular inoculation

Intramuscular inoculation with gt 1, 2 and 3 lyssaviruses produced different results (Figure 6.2) depending on the isolate used. When virus was introduced i.m. at a high dose (1 x 10^6 LD_{50}), all isolates were able to induce rabies and subsequent death in mice although all were not equally virulent. Inoculation with a gt 1 virus (WAMYOTISspp) led to a 100% mortality in the four-week-old-mice and the same result was obtained for two of the previously reported gt 2 isolates (LBVSEN1985 and LBVAFR1999). When a representative of gt 3 (MOKVSA(252/97)) was inoculated, only 20% of the mice succumbed whereas when gt 2 isolates were introduced 20-60% of mice succumbed. When virus isolates were introduced at a lower concentration (1 x 10^3 LD_{50} dose) the percentage of mortality of the mice decreased and for some isolates there was no mortality observed at this virus dose.

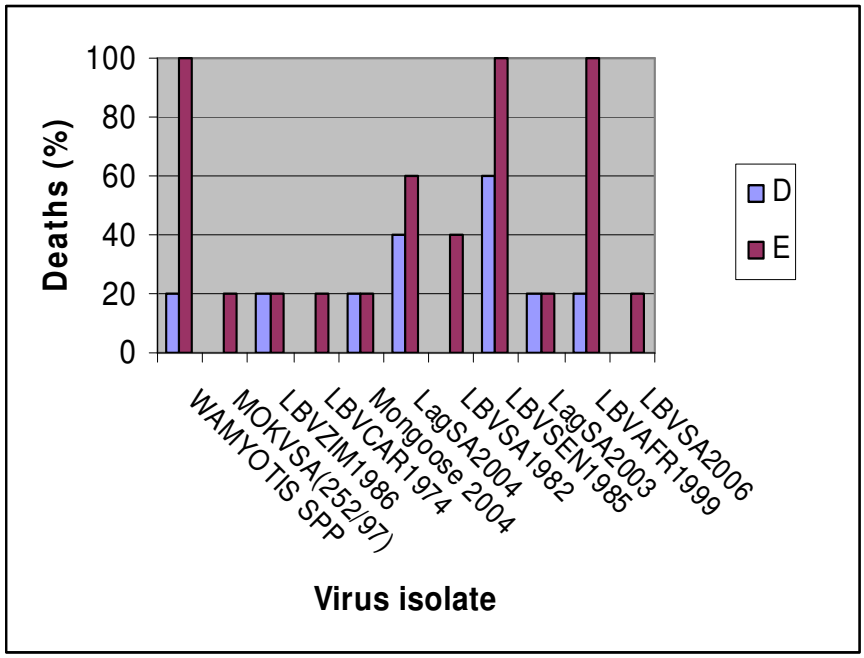


Figure 6.2: Pathogenicity of genotype 1 (WAMYOTIS SPP), gt 2 and gt 3 (MOKVSA(252/97) lyssaviruses in 4-week-old-ICR mice after intramuscular (i.m.) inoculation. Results are expressed as a percentage of dead animals after observation for 56 days. Different viral doses were introduced; D: 1 X 10^3 LD_{50} and E: 1 X $10^6$$LD_{50}$.

iii) Oral inoculation

Mortality was only observed in some mice after oral inoculations of high amounts of virus (Figure 6.3). Only four of the eleven isolates used in this study led to mortality in mice when introduced *via* the oral route.

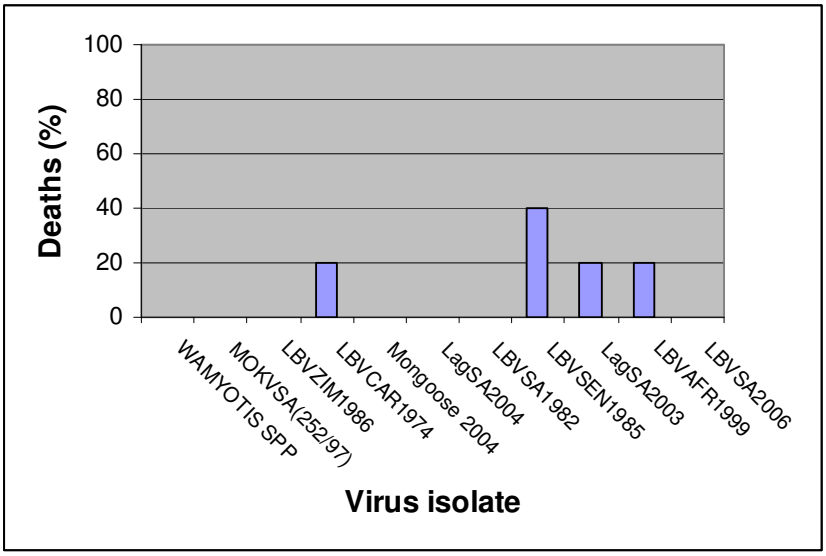


Figure 6.3: Pathogenicity of genotype 1 (WAMYOTIS SPP), gt 2 and gt 3 (MOKVSA(252/97) lyssaviruses in 4-week-old-ICR mice after oral inoculation with 1 X 10^6 LD_{50}. Results are expressed as a percentage of dead animals after observation for 56 days.

The susceptibility of mice to lyssaviruses was found to be related to the route of inoculation, the dose of viral inoculum and the properties of the virus isolate. Percentage of mice that succumbed to the disease when inoculated via the i.c. route was higher compared to mice inoculated via the i.m. or oral route. Lower amounts of viral inoculum led only to mortality in mice inoculated via the i.c. route and not via the i.m. and oral routes.

The mean incubation period before mice succumbed to the disease was different depending on the virus isolate, the route of inoculation and viral dose (Table 6.5 and Figure 6.4). The mean incubation period were found to be proportional to the inoculation dose. Generally, a longer incubation period was demonstrated when lower amounts of virus were introduced. Virus introduced i.c. at a viral dose of 1 x 10^2 LD_{50} produced the shortest mean incubation period in all virus isolates tested. A representative of gt 1 (WAMYOTISspp) did not indicate a significant mean incubation

period difference between different doses of virus inoculum, inoculated via the i.c. route and at high amounts via the i.m. (1 x 10^6 LD_{50}) route but when introduced at a dose of 1 x 10^3 LD_{50} via the i.m. route the mean incubation period increased to 17 days. For a representative of gt 3 (MOKVSA(252/97), the mean incubation period increased when low amounts of virus were introduced i.c. (1 X 10^{-2} LD_{50}) and i.m. at 1 X 10^6 LD_{50}. In general gt 2 representatives demonstrated increased incubation times associated with lower doses independent of introduction via the i.c. or i.m. routes. The shortest mean incubation periods and highest susceptibility were observed for LBVSEN1985 and LBVAFR1999 (Lineage A) for all the routes of inoculation and over the enire dosage spectrum.

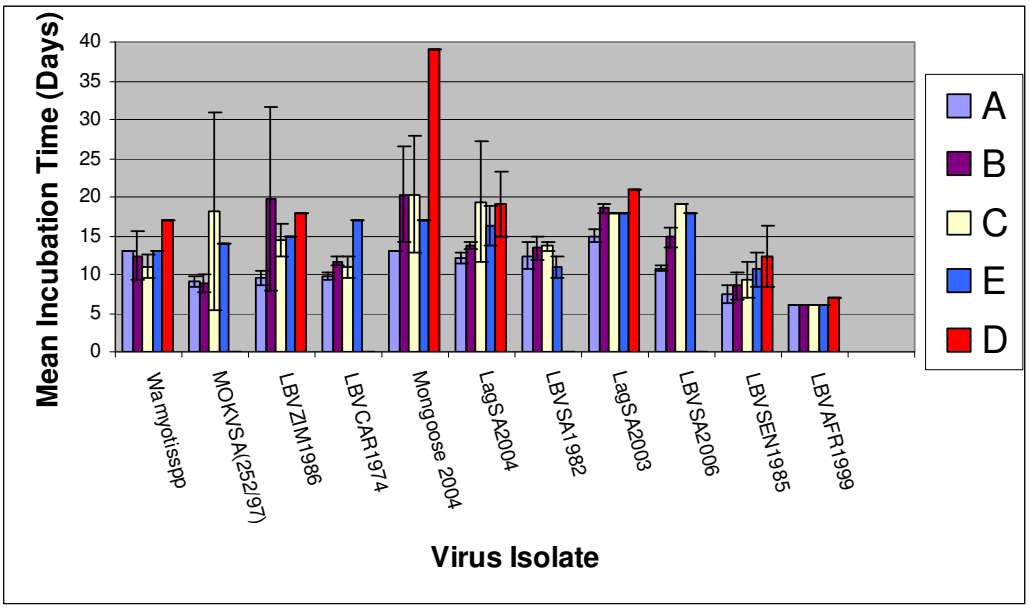


Figure 6.4: Mean Incubation Time (Days) of lyssavirus isolates after different routes of inoculation (i.c. and i.m.) and different viral doses were introduced into 4-week-old-ICR mice. The standard deviation (SD) is indicated. Some SD values were 0. A: i.c. (1 x 10^2 LD_{50}); B: i.c. (LD_{50}); C: i.c. (1 x 10^{-2} LD_{50}); D: i.m. (1 X 10^3 LD_{50}) and E: i.m. (1 X $10^6$$LD_{50}$).

Table 6.5: Effect of route of inoculation (i.c., i.m. and orally) and dose of inoculum on the mean interval between inoculation and death following inoculation of 4-week-old ICR mice with different lyssavirus isolates. The mean incubation period indicated in days ± the standard deviation (SD) followed by the amount of animals that died per group of five mice followed by the range of incubation period are indicated in the table. In groups where no mice died, no mean incubation period or range of incubation period are shown .

ROUTE OF INOCULATION	INOCULUM DOSE (LD50)	INCUBATION PERIOD (DAYS) MEAN ± SD ANIMALS THAT DIED/GROUP OF 5 MICE RANGE OF INCUBATION PERIOD										
		WAMYOTIS SPP	MOKVSA(252/97)	LBVCAR1974	LBVSA1982	LBVSEN1985	LBVZIM1986	LBVAFR1999	LagSA2003	LagSA2004	Mongoose2004	LBVSA2006
i.c.	10^{2} LD_{50}	13 ± 0 (5/5) (13)	9 ± 0.7 (5/5)((8 – 10)	9.8 ± 0.4 (5/5) (9-10)	12.4 ± 1.8 (5/5) (10-14)	7.4 ± 1.1 (5/5) (6-9)	9.6 ± 0.9 (5/5) (8-10)	6 ± 0 (5/5) (6)	15 ± 0.7 (5/5) (14-16)	12 ± 0.7 (5/5) (11-13)	13 ± 0 (5/5) (13)	10.8 ± 0.4 (5/5) (10-11)
i.c.	LD_{50}	12.4 ± 3.1 (5/5) (9-17)	8..8 ± 1.1 (5/5) (10)	11.7 ± 0.6 (3/5) (11-12)	13.4 ± 1.5 (5/5) (11-15)	8.5 ± 1.7 (4/5) (7-11)	19.8 ± 11.9 (5/5) (14-41)	6 ± 0 (5/5) (6)	18.5 ± 0.6 (4/5) (19-19)	13.8 ± 0.5 (4/5) (13-14)	20.3 ± 6.1 (4/5) (13-28)	14.8 ± 1.3 (5/5) (14-17)
i.c.	10^{-2} LD_{50}	11 ± 1.5 (4/5) (13-16)	18.2 ± 12.8 (5/5) (10- 41)	11 ± 1.4 (2/5) (10-12)	13.7 ± 0.6 (3/5) (6-14)	9.3 ± 2.2 (4/5) (7-12)	14.5 ± 2.1 (2/5) (13-16)	6 ± 0 (5/5) (6)	18 (1/5) (18)	19.4 ± 7.8 (5/5) (14-33)	20.3 ± 7.5 (3/5) (13-28)	19 (1/5) (19)
i.m.	10^{3} LD_{50}	13 (1/5) (13)	(0/5)	(0/5)	(0/5)	12.3 ± 4 (3/5) (8-16)	18(1/5) (18)	7 (1/5) (7)	21 (1/5) (21)	19 ± 4.2 (2/5) (16-22)	39 (1/5) (39)	0/5
i.m.	10^{6} LD_{50}	17 ± 0 (5/5) (17)	14 (1/5) (14)	17 (1/5) (17)	11 ± 1.4 (2/5) (10-12)	10.6 ± 2.3 (5/5) (8-13)	15 (1/5) (15)	6 ± 0 (5/5) (6)	18 (1/5) (18)	16.3 ± 2.5 (3/5) (14-19)	17 (1/5) (17)	18 (1/5) (18)
oral	10^{6} LD_{50}	(0/5)	(0/5)	14 (1/5) (14)	(0/5)	9.5 ± 2.1 (2/5) (8-11)	(0/5)	7 ± 0 (1/5) (7)	(0/5)	(0/5)	(0/5)	(0/5)

6.3.3 Serological responses

Only a few serum samples collected from mice during the experiment were tested for the presence of neutralizing antibodies (Appendix 3). Since no reference sera for LBV or MOKV are available the virus neutralizing antibody titer could not be converted into international units. Serum tested from mice that succumbed to disease, independent of route of inoculation or viral dose, were all seronegative (titers < 1:5), with four exceptions (Table 6.6). In all four these cases the incubation period was longer than observed in animals that tested seronegative. The gt 3 representative developed a high titer (1:625) of neutralizing antibodies before succumbing to the disease.

Table 6.6: Presence of neutralizing antibodies in mice that succumbed of lyssavirus infection

VIRUS ISOLATE	ROUTE OF INOCULATION	VIRAL DOSE	TAG NUMBER	INCUBATION PERIOD (DAYS)	TITRE OF NEUTRALIZING ANTIBODIES
LBVZIM1986	i.c.	LD_{50}	220	41	1:5 (Day 21, 28 and 35)
MOKVSA(252/97)	i.c.	1×10^{-2} LD_{50}	190	41	1:5 (Day 5 and 8) 1:25 (Day 14, 21 and 28) 1:625 (Day 35)
LagSA2003	i.m.	1×10^{3} LD_{50}	3463	21	1:25 (Day 14)
LagSA2003	i.m.	1×10^{6} LD_{50}	3463	18	1:25 (Day 14)

Antibody responses in the mice that survived varied (Table 6.7). Brain material tested with FAT from these animals was negative for lyssavirus antigens. Most mice that seroconverted was inoculated i.m. (Figure 6.5) and remained clinically healthy for 56 days. No animals seroconverted after oral inoculation.

Table 6.7: Presence of neutralizing antibodies in mice that survived lyssavirus infection after i.c. inoculation

VIRUS ISOLATE	VIRAL DOSE	TAG NUMBER	TITRE OF NEUTRALIZING ANTIBODIES
Mongoose2004	1×10^{-2} LD_{50}	350	1:25(Day 56)
LBVSEN1985	1×10^{-2} LD_{50}	8	1:625 (Day 56)
LagSA2003	LD_{50}	3472	1:625 (Day 56)
LagSA2003	1×10^{-2} LD_{50}	3462	1:25 (Day 56)
LagSA2003	1×10^{-2} LD_{50}	3461	1:25 (Day 56)
LagSA2003	1×10^{-2} LD_{50}	3459	1:25 (Day 56)
LagSA2003	1×10^{-2} LD_{50}	3458	1:25 (Day 56)

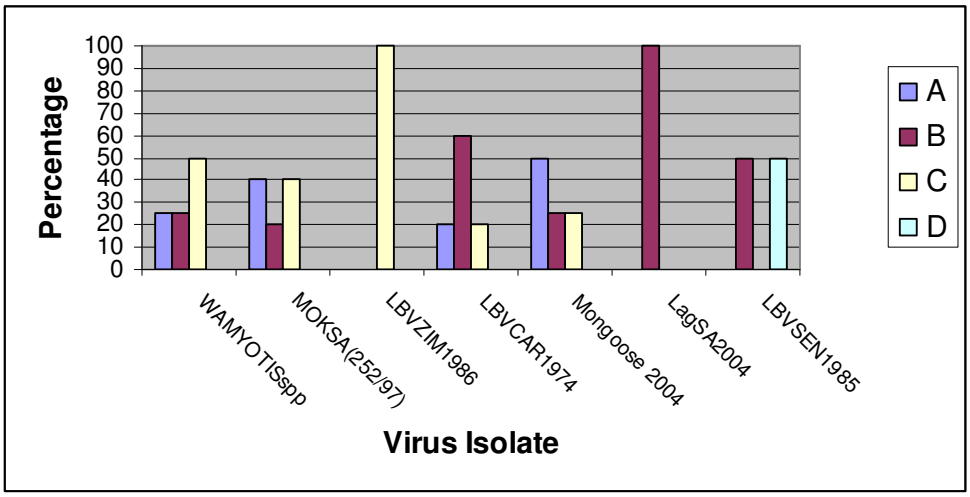


Figure 6.5: Presence of different titres of neutralizing antibodies in mice that survived lyssavirus infection after i.m. inoculation with 1 X 10^3 LD_{50}. A: 1:5; B: 1:25; C: 1:625 and D: 1:3125. Percentage indicates the number of mice that survived and demonstrated a specific titre of neutralizing antibodies.

6.3.4 Molecular determinants of pathogenesis

Multiple alignment of the P, M and G amino acids of different representatives of the lyssavirus genotypes was performed using ClustalW and regions previously implicated to play a role in pathogenesis were analysed. In Chapter V, gt 2 isolates analysed in this study were divided into three lineages (A-C) and lineage A was proposed to be a new lyssavirus genotype. In Section 6.3.2 differences in the pathogenicity of these gt 2 representatives (lineage A-C) were indicated. In this Section these differences observed (Section 6.3.2) will be correlated with molecular sequencing data of domains previously identified to be involved in lyssavirus pathogenesis.

i) LC8 dynein light chain binding domain on the phosphoprotein

The LC8 dynein light chain binding domain present on the phosphoprotein of lyssaviruses is indicated in Figure 6.6. This site was conserved between gt 1, 4, 5, 6, Aravan and Khujand virus but different in gt 2, 3, 7, Irkut and WCBV. Differences in pathogenicity between gt 1, 2 and 3 lyssavirus isolates analysed in this study may be related to differences in the binding site for the LC8 dynein light chain (Figure 6.6). This site is implicated in lyssavirus pathogenesis where it is involved in retrograde transport of virions (Poisson *et al*., 2001). Within gt 2 differences in pathogenicity

were observed and lineage A isolates (LBVSEN1985 and LBVAFR1999) demonstrated a higher mortality in mice when introduced i.m. compared to other gt 2 isolates (lineage C) (Figure 6.2). The dynein binding site was conserved in lineage C isolates but different for lineage A isolates and may contributed to the differences in pathogenicity observed.

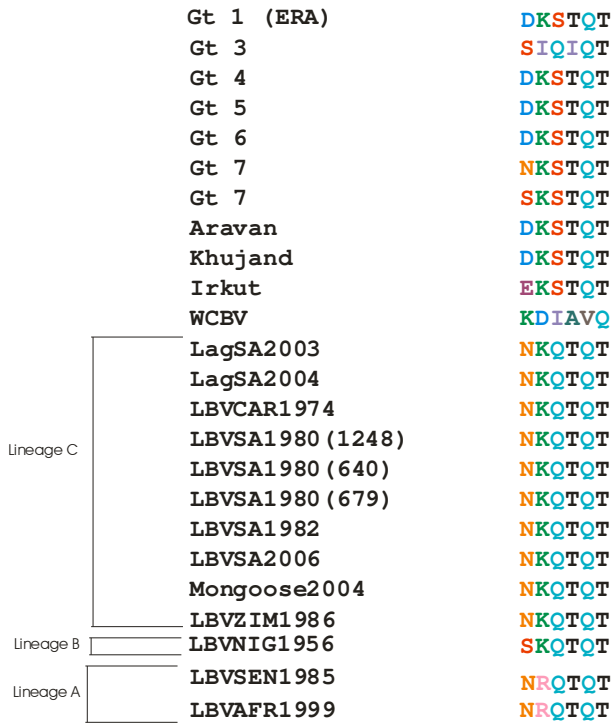


Figure 6.6: Aa 143-148 of the P protein of different representatives of the lyssavirus genus, indicating the dynein light chain binding site that may play a role in lyssavirus pathogenesis.

ii) The M protein

The M protein has previously been implicated in lyssavirus pathogenesis although specific pathogenicity regions have not yet been identified. A multiple alignment of the M protein of gt 2 representatives used in this study is indicated in Figure 6.7.

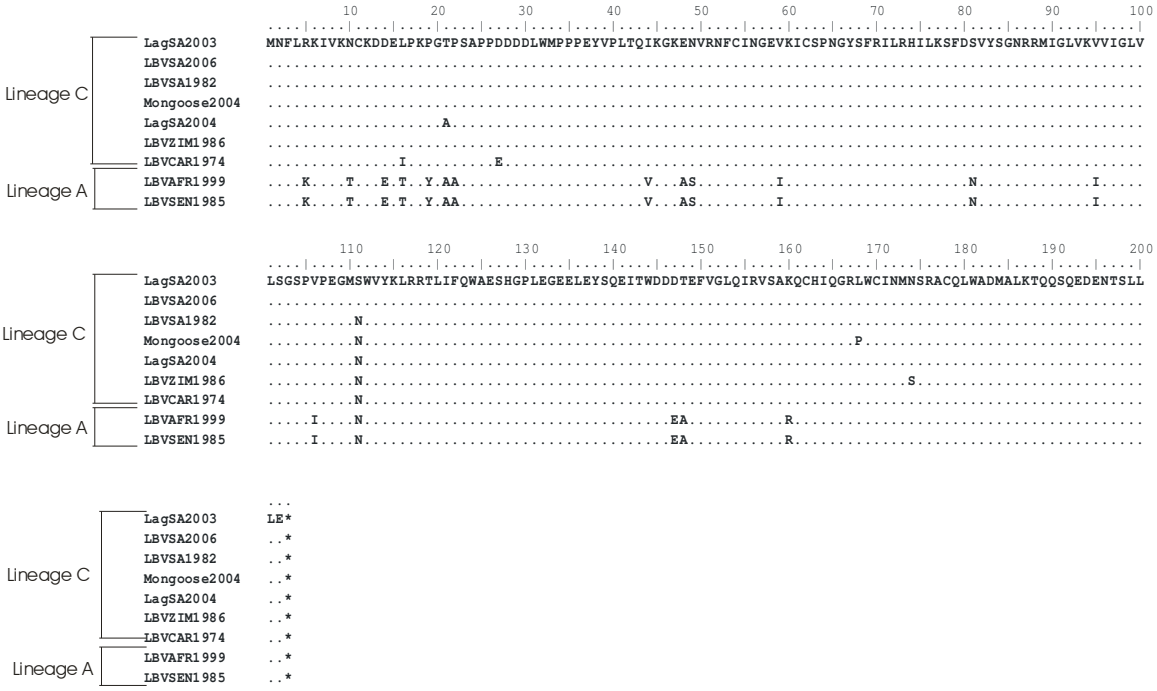


Figure 6.7: Multiple alignment of the matrixprotein of gt 2 representatives analysed in pathogenicity studies in a mouse model.

When introducing isolates via the i.c. or i.m. routes, lineage A isolates (LBVSEN1985 and LBVAFR1999) had a higher mortality in mice compared to other gt 2 isolates (Lineage C) (Figure 6.2). The highest mortality in lineage C (60%) was observed when the LagSA2004 isolate was introduced i.c. or i.m. compared to a 100% mortality when lineage A isolates were introduced. When analysing the M protein amino acid composition, lineage A isolates indicated a high amino acid diversity compared to other gt 2 isolates (Figure 6.7). Some of these amino acid changes may account for their increased pathogenicity. LagSA2004 has an amino acid change on position 21, A (alinine) instead of a T (threonine), compared to other lineage C isolates. This may account for this isolate's increased pathogenicity compared to other lineage C isolates. Lineage A isolates also have an A in this position instead of a T.

iii) Ectodomain of the glycoprotein

Aa 333 of the ectodomain of the glycoprotein has been implicated in pathogenic differences between lyssavirus genotypes (Badrane *et al*., 2001b) and this mutation together with a mutation in aa 330 has been associated with an even further reduction in pathogenesis (Coulon *et al*., 1998). Aa 330-333 was conserved in gt 1,4,

5, 6, 7, Aravan and Khujand virus and different in gt 2, 3 and WCBV except in a gt 4 isolate that indicate a Proline (P) instead of an R in position 333 (Figure 6.8). This change will need to be confirmed in future studies since this may have been a sequencing mistake. In gt 2 and 3 isolates, aa 333 has been replaced by an aspartic acid (D). With the exception of gt 2 members and WCBV, the aa at site 330 (K) is conserved within the lyssavirus genus. Isolates belonging to lineage B and C of gt 2 contained a Leucine (L) in position 330 but isolates belonging to lineage A contained a K in this position similar to all other lyssavirus genotypes. This may account for the increased pathogenicity of lineage A isolates.

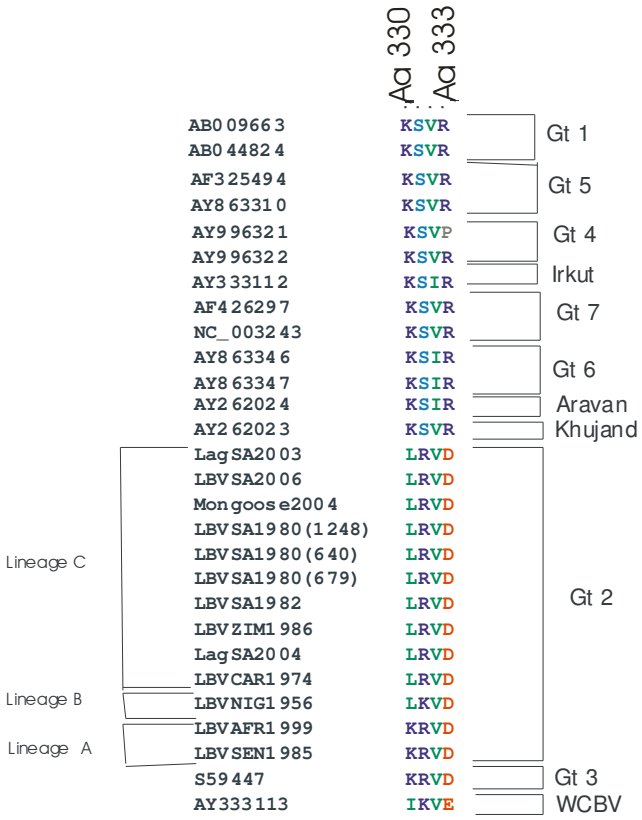


Figure 6.8: Multiple alignment indicating aa 330-330 of the ectodomain of the G protein of representatives of the lyssavirus genus.

Multiple alignment of glycoprotein sequences of gt 2 representatives used in this study is indicated in Figure 6.9. Lineage A isolates (LBVSEN1985 and LBVAFR1999) indicated a high sequence diversity in the glycoprotein compared to other gt 2 isolates. Lineage C isolates indicated almost no amino acid difference in the glycoprotein ectodomain associated with pathogenesis except on position 50 and 61 of the glycoprotein.

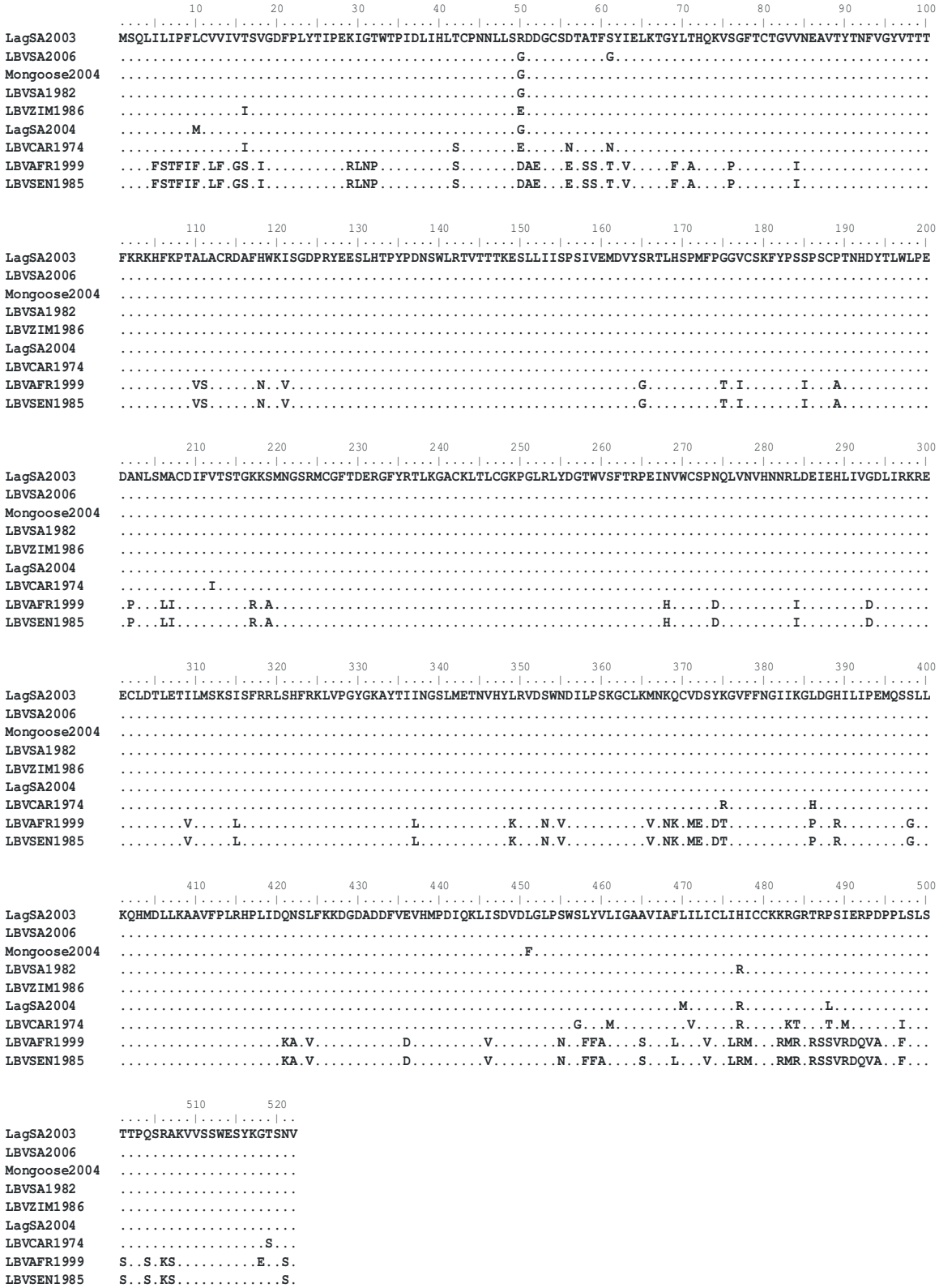


Figure 6.9: Multiple alignment, indicating differences in the G protein of gt 2 representatives analysed in pathogenicity studies.

6.4 Discussion

This study assessed the susceptibility of mice to different gt 2 isolates as compared to genotype 1 and 3 representatives. Infection dynamics of this experimental host may vary from infection in natural hosts. When virus isolates were inoculated intracerebrally they behaved similarly but differences were observed when inoculated intramuscularly. Previous studies suggested that phylogroup II viruses were not pathogenic to mice when introduced i.m. but only when introduced via the i.c. route. This led to a suggestion that phylogroup II viruses were less pathogenic and therefore less of a public health risk (Badrane *et al.*, 2001b). In this study the pathogenicity of several representatives of gt 2 was investigated and it was indicated that these viruses can cause mortality when introduced at high doses via the i.m. route into mice. When MOKV (gt 3) was inoculated i.m. at very high viral doses (1 x 10^6 LD_{50}) only 20% of the mice succumbed to disease compared to a gt 1 isolate where 100% of mice succumbed. Overall, gt 2 representatives had a lower pathogenicity compared to gt 1 except for lineage A isolates that had the same pathogenicity as gt 1 viruses. These two isolates (LBVSEN1985 and LBVAFR1999) have been shown to fulfil current criteria for their consideration as representatives of a new lyssavirus genotype (Chapter V) and their different pathogenic properties compared to other gt 2 isolates provide further support for this classification. Passage history of isolates may also change the properties of the virus. Where possible in this study, isolates were passaged only once in suckling mouse brain but for some isolates the passage history was unavailable or unknown.

Analysing aa differences of the P, M and G proteins of lyssavirus representatives revealed important molecular determinants that may be involved in pathogenesis. The LC8 dynein light chain binding domain is different between pathogenic (lineage A) and less pathogenic (lineage C) gt 2 representatives. This region has been implicated in retrograde transport (Poisson *et al.*, 2001) of virions and mutations in this region may lead to a reduction in spread of the virus. Analysis of this region also supported the classification of lineage A isolates as a new lyssavirus genotype, since this region is different for lineage A compared to other gt 2 representatives. Aa 333 on the ectodomain of the glycoprotein has previously been implicated in pathogenesis studies as an important domain for virulence (Badrane *et al.*, 2001b). A mutation in this region has been shown as the reason why phylogroup II viruses are less pathogenic. From results obtained in this study these observations are questionable since lineage A isolates indicated the same pathogenicity as a gt 1

representative and these isolates contain a mutation in aa 333 (R→D) similar to other phylogroup II representatives that had a low pathogenicity. A change in aa 330 of the glycoprotein ectodomain has previously been implicated in a further reduction in pathogenesis. Gt 2 representatives have a change in aa 333 and 330 rendering them even less pathogenic but lineage A isolates have the same aa in position 330 than other lyssavirus representatives including gt 1.

Our results suggest that certain virus isolates belonging to phylogroup II can have the same pathogenicity as gt 1 viruses (phylogroup I). The pathogenicity of phylogroup II has therefore been underestimated in previous studies (Badrane *et al*., 2001b). However, although indicated that phylogroup II viruses could not lead to mortality in mice when introduced via the i.m. route previous studies in other animal models indicated that MOKV (gt 3) could lead to mortality in shrews when introduced via the i.m. route. It has been shown that shrews can be experimentally infected with MOKV *via* several routes: subcutaneously (s.c.) behind the left foreleg, i.m. in the occipital muscle and orally (Kemp *et al*., 1973). A virus dose of less than 25 000 mouse LD_{50} inoculated s.c. was not infective and a dose of 30 000 LD_{50} produced a 100% fatality in shrews inoculated i.m., 66% mortality in shrews inoculated s.c. and 33% mortality in oral introduction. Four of the gt 2 isolates studied in this study also caused mortality in mice when introduced in high amounts via the oral route. Previous studies have shown that variation in pathogenicity can occur within a genotype depending on the animal model used as indicated when Mexican free-tailed bats were inoculated with gt 1 viruses using different routes of inoculation. These bats were relatively resistant (8 out of 46 bats succumbed) to peripheral inoculation with viruses isolated from other bat species (also gt 1) except when very large doses were administered (Baer and Bales, 1967). Lyssavirus pathogenesis may also differ depending on the virus variant e.g. raccoons infected with a raccoon RABV developed furious rabies whereas raccoons infected with a canine RABV developed paralytic rabies. Serological responses and viral exrection in saliva may also differ depending on the virus variant (Hill *et al*., 1993; Niezgoda *et al.,* 1997; Niezgoda *et al*., 1998). Results are therefore unpredictable in different species and conclusions about the pathogenicity of lyssavirus isolates depends largely on the type of animal model used in the experiment and the pathogenicity in another animal model cannot be predicted. Studies performed in an animal model such as a mouse model may provide an indication of virus infectivity but studies in the natural host will provide a better understanding.

Experimental infections of dogs and monkeys with phylogroup II viruses, LBV (Nigeria isolate) and MOKV (Nigeria isolate, 1968), were previously performed (Percy *et al*., 1973). Animals were inoculated via the i.c. and i.m. route using high amounts of virus (6.2 log10 LD_{50}/ml(LBV) and 7.2 log10 LD_{50}/ml(MOKV)). All animals inoculated i.c. died but following i.m. inoculation only one monkey inoculated with MOKV succumbed to disease 10 days after inoculation and virus was isolated from this animal. Central nervous system lesions were observed in dogs and monkeys inoculated i.m. with MOKV and LBV after they were euthanized and animals did not develop clinical signs. However, one monkey developed clinical signs after i.m. inoculation with LBV and survived up to 108 days whereafter it was euthanized. Virus isolation from this animal was unsuccessful. Both these phylogroup II viruses seemed to be less pathogenic when inoculated via the i.m. route. Serological responses were observed in animals inoculated i.m. but were insignificant in animals inoculated via the i.c. route. From this study it is evident that high amounts of phylogroup II viruses introduced via the i.m. route can lead to mortality in dogs. Unfortunately this study did not compare the pathogenicity of phylogroup I and II isolates and this will need to be investigated in future studies.

In this study the mean incubation period before mice succumbed of rabies seemed to be associated with the viral inoculation dose. A longer incubation period was observed when lower amounts of virus was introduced, irrespectively of the route of inoculation (i.m. or i.c.). Serological responses were observed in animals that succumbed of disease but this was associated with a long incubation period. In animals that survived, serological responses were observed in mice inoculated via the i.m. and the i.c. route but no serological responses were observed in animals inoculated via the oral route. There is still a lot to be discovered on rabies and rabies-related lyssavirus pathogenicity. The pathogenicity of phylogroup II viruses has previously been underestimated as indicated by the pathogenicity of lineage A and C viruses analysed in this study. These viruses are able to induce rabies and subsequent death when inoculated via the i.c. and the i.m. route and certain isolates in this group may be more virulent than others. The incidence of phylogroup II viruses should be assessed to determine the current situation. If the incidence of these viruses increases there may be more human contact with infective animals leading to a more significant public health threat.

CHAPTER VII

CONCLUSION

Before this study commenced, there were only twelve isolations of gt 2 (LBV) made throughout Africa. The molecular epidemiology of this virus was unknown and limited information on the pathogenicity was available. Previous studies indicated that this virus belongs to phylogroup II based upon genetic, immunological and pathogenic characteristics and the limited DNA sequencing information indicated a high sequence diversity in this group. Serological cross reactivity with other lyssavirus genotypes was found to be limited and therefore there current genotype 1 based vaccines do not offer cross protection against LBV infection. Human LBV cases have not been reported to date but surveillance for the African rabies-related lyssaviruses is poor and lyssavirus diagnostic methods used in most African laboratories cannot distinguish between different genotypes. The gt 2 viruses were also believed to be less pathogenic compared to phylogroup I viruses due to previous studies indicating that mice only succumb to disease when virus is introduced via the i.c. but not via the i.m. route. Phylogroup 1 viruses lead to mortality in mice when inoculated via both routes and phylogroup II viruses were therefore considered not to be a high public health risk.

After 12 years of no reports of LBV in South Africa, we commenced with a small scale passive surveillance programme in KwaZulu Natal, South Africa. In so doing, we were able to identify 6 new isolates over a period of four years. This emphasizes again that the incidence of LBV is underestimated due to no or poor surveillance. However, if more active surveillance for these viruses is instigated, the laboratories involved should also be able to apply diagnostic methods such as monoclonal antibody typing or genetic methods to distinguish between lyssavirus genotypes. Current methods used in most African laboratories can only identify the infectious agent as a lyssavirus. Two spill over events of LBV, one to a canine and one to a mongoose, have also been reported from South Africa during this study. This has been the first report of LBV from terrestrial wildlife and the first report of LBV infection of a canine in South Africa. It must therefore be emphasized that the lyssavirus involved in infection should be characterized to genotype level even in species previously associated with gt 1 lyssaviruses. In most of the LBV cases reported from South Africa there had been close contact with humans or other animals emphasising the public health and veterinary importance of this virus. It has also

been indicated in this study that the accurate identification of the host species involved is important and can be confirmed using DNA-based methods. Without accurate host-species identification it is impossible to make informed decisions about control and prevention of the disease.

The first report of the distribution pattern of LBV in naturally infected fruit bats, indicated that the viral tropism of gt 2 is not different to that of other lyssavirus genotypes. Virus moves to the brain via the CNS and is only disseminated to other organs when present in the brain. No support for non-productive lyssavirus infection was observed. Molecular epidemiological studies analysing the complete N, M, P and G gene and amino acids indicated a high sequence diversity in gt 2 isolates analysed. Further analysis of these isolates indicated that the LBVSEN1985 and LBVAFR1999 isolates, previously suggested to be part of gt 2, constitutes a new lyssavirus genotype based on genetic analysis of the N, P, M and G genes. This was supported by pathogenicity studies that showed that these two isolates are more pathogenic than other gt 2 isolates when introduced via the i.m. route. Previous suggestions that phylogroup II lyssaviruses are less pathogenic compared to phylogroup I viruses when introduced via the i.m. route in a mouse model were questioned. This study indicated that some phylogroup II isolates can lead to mortality in mice when introduced i.m. in high doses. Differences in pathogenicity observed between isolates analysed, have also been associated with aa changes in regions of the P, M and G proteins.

This study also suggested that the classification criteria for lyssaviruses should be revised. As more lyssavirus isolates are discovered and characterized on molecular level, the diversity of the lyssavirus genus will probably expand and intergenotypic and intragenotypic identities will overlap using the current criteria. This study has analysed complete N, P, M and G gene and protein sequences and the LBVSEN1985 and LBVAFR1999 isolates clustered together as a separate group in both NJ and MP phylogenetic analysis. Antigenic sites between these two isolates and other gt 2 isolates were different. A high level of susceptibility of mice to these isolates when introduced via the i.c. and i.m. routes was demonstrated and differences in molecular determinants involved in pathogenesis were different for lineage A isolates compared to other gt 2 virus isolates. All these combined characteristics strongly support these isolates as a new lyssavirus genotype perhaps specific to the West African geographical domain. To further investigate the taxonomy of lyssaviruses an effort should be made to obtain more sequencing data

of representatives of all lyssavirus genotypes to determine the diversity of the lyssavirus genus and to make informed decisions about lyssavirus classification.

An important shortcoming of neutralizing antibody detection assays (both FAVN and RRFIT) for African rabies-related lyssaviruses, is the lack of appropriate reference sera. As a result, interpretations regarding lyssavirus seropositivity in populations may be varied and controversial (Arguin *et al*., 2002; Reynes *et al*., 2004; Serra-Cobo *et al*., 2002). If serum titres could be converted to International Units (IU), a cut-off value of ≥0.5 IU/ml antibody would be considered positive by most laboratories – as an arbitrary standard. This standard is used by rabies reference laboratories as evidence for the induction of RABV neutralizing antibodies in humans, following rabies vaccination (WHO, 2005). Currently a reference serum is available for gt 1 only and therefore this conversion could not be performed for gt 2 in the present study. Our results are therefore only reported as the serum dilution at which a 50 % reduction in infectious centres was observed. Previous studies have applied a cut-off value of 50 % reduction in the infectious centres - compared to a positive control (Steece *et al*., 1989). In some studies this cut-off reduction value was taken as high as 90 % (Arguin *et al*., 2002). It has been shown in the case of rabies (Briggs *et al*., 1998) that the results of RFFIT and FAVN tests are not statistically diffirent – although this has not been clearly demonstrated for the rabies-related viruses. Modifications to the standard RFFIT test to accommodate smaller amounts of test sera, a typical limitation in the case of bat rabies-related virus have also not been validated. It is suggested that future studies to investigate the comparison of these two tests for the rabies-related lyssaviruses would be worthwhile and important towards a standardized serological assay for the measurement of neutralizing antibodies against rabies-related lyssaviruses.

When considering global mobility of humans and animals, gt 2 viruses can pose a threat to any country. The LBVAFR1999 was isolated in 1999 in a rabid bat imported illegally from Africa (as a pet) and developed rabies in France. The lyssavirus was identified as belonging to genotype 2 and 122 people needed to get post exposure prophylaxis due to exposure to this animal. The owner of the bat was bitten two times and also received post-exposure treatment. Information about the risk involved when handling bats and guidelines on how to minimize exposure should be communicated. Furthermore, additional surveillance among bat species in Africa is needed to establish more information about distribution, prevalence, genetic diversity and host species association of African lyssaviruses. These surveillance efforts should not

only include bats but also terrestrial wildlife and domestic cats and dogs. Certainly, studies of molecular epidemiology and pathogenicity should yield the type of information needed for balanced and informed decisions regarding the potential threat of these viruses to public and veterinary health.