SEQUENCE ANALYSIS OF THE SMALL (S) RNA SEGMENT OF VIRUSES IN THE GENUS ORTHOBUNYAVIRUS

Maizan Mohamed

A Thesis Submitted for the Degree of PhD at the University of St. Andrews



2007

Full metadata for this item is available in the St Andrews Digital Research Repository at: https://research-repository.st-andrews.ac.uk/

Please use this identifier to cite or link to this item: <u>http://hdl.handle.net/10023/434</u>

This item is protected by original copyright

This item is licensed under a <u>Creative Commons License</u>

SEQUENCE ANALYSIS OF THE SMALL (S) RNA SEGMENT OF VIRUSES IN THE GENUS ORTHOBUNYAVIRUS

MAIZAN MOHAMED

University of St. Andrews, UK

A thesis submitted for the degree of Doctor of Philosophy at the

University of St. Andrews

July, 2007

Abstract

Viruses in the genus *Orthobunyavirus* (family *Bunyaviridae*) are classified serologically into 18 serogroups. The viruses have a tripartite genome of negative sense RNA composed of large (L), medium (M) and small (S) segments. The L segment encodes the polymerase protein, the M segment encodes two glycoproteins, Gc and Gn, and a non-structural protein (NSm), and the S segment encodes nucleocapsid (N) and NSs proteins, in overlapping reading frames (ORF). The NSs proteins of Bunyamwera and California serogroup viruses have been shown to play a role in inhibiting host cell protein synthesis and preventing induction of interferon in infected cells.

To-date, viruses in only 4 serogroups: Bunyamwera, California, Group C and Simbu, have been studied intensively. Therefore, this study was conducted with the aim to sequence the S RNA segments of representative viruses in the other 14 orthobunyavirus serogroups, to analyse virus-encoded proteins synthesised in infected cells, and to investigate their ability to cause shutoff of host protein synthesis.

S RNA segment sequences were obtained from cloned RT-PCR products. They were compared with the available sequences and each other. Complete S RNA sequences of Anopheles A (ANAV) and Tacaiuma virus (TCMV) [Anopheles A serogroup], Anopheles B (ANBV) and Boraceia virus (BORV) [Anopheles B serogroup], Eretmapodites (E147V) and Nyando virus (NDV)[Nyando serogroup], Bwamba virus (BWAV) [Bwamba serogroup], M'Poko virus (MPOV) [Turlock serogroup], Tete (TETEV) and Batama virus (BMAV) [Tete serogroup], and Gamboa (GAMV) and San Juan 2441 virus (SJ244V) [Gamboa serogroup], and partial sequences of Patois virus (PATV) [Patois serogroup], Guama (GMAV) and Bertioga virus (BERV) [Guama serogroup], Capim virus (CAPV) [Capim serogroup] and Palestina virus (PLSV) [Minatitlan serogroup] were obtained. Complete S segment sequences revealed that viruses in the same serogroup have same length of N and NSs proteins, except for the viruses in Gamboa serogroup which were found to have two lengths of NSs protein. Viruses in 4 serogroups (Anopheles A, Anopheles B, Tete and Capim) were found not to encode an NSs ORF, presenting the first report of naturally isolated orthobunyaviruses without an NSs protein. Most of these viruses were found to have longer N proteins compared to those with NSs protein, with the largest N protein observed to date in TETEV and BMAV (258 amino acids). Other viruses

(EREV, NDV, GAMV, SJ2441V, BWAV and MPOV) were found to encode both N and NSs proteins in their S segment with the largest and smallest NSs protein detected to date in SJ2441V (137 amino acids) and MPOV (70 amino acids) respectively. The conserved CA rich motif in 5' non coding region (NCR) of Bunyamwera and California serogroups viruses was absent in BWAV and MPOV, while ANBV and BORV were found to have two copies of this motif. Repeated sequences, as observed previously in the 5' NCR of genomic-sense RNA of Lumbo virus (LUMV), were also detected in BWAV and TCMV S RNA segments.

Sequence comparisons and phylogenetic analyses of the sequences determined in this study were in agreement with previous serological classification of the viruses, except for BERV and TCMV. BERV, in the Guama serogroup, was found to have a closer relationship with CAPV compared to GMAV. However high sequence identities (>70%) were observed between these 3 viruses, suggesting that they are derived from the same ancestor. N protein and nucleotide sequence identities of TCMV with ANAV were only 53% and 59% respectively. However, Neighbour-Joining (NJ) plot based on complete N amino acid sequence and Maximum Parsimony (MP) plot based on partial N sequence supported previous serological classification which placed this virus in the same clade as ANAV.

This study first reports on the proteins synthesised by Bakau, Bwamba, Koongol, Gamboa, Minatitlan, Olifantsvlei and Tete serogroup viruses. Analysis of radio-labelled cell extracts revealed similar protein migration patterns for all the studied viruses compared with other viruses in the genus *Orthobunyavirus*. Shutoff of host cell protein synthesis, similar to that seen in Bunyamwera virus (BUNV)-infected cells was only observed in ACAV, BAKV, BWAV, CAPV, PAHV, PATV and WONV-infected cells. However, this shutoff was found not related to the presence of NSs protein. In general, viruses in the same serogroup were found to have almost same size of plaque and plaque-size did not correlate with the presence of NSs protein and the virulence of the virus in the mice.

In vitro transcription and translation (TnT) using rabbit reticulocyte and wheat germ lysate expression systems further confirmed the sequencing results that no NSs protein was expressed from S cDNA clones of ANAV, TCMV, ANBV, BORV, BMAV and TETEV. S RNA segments shutoff almost similar to BUNV-infected cells was observed in A549 cells

infected with TCMV, suggesting that TCMV might use a different mechanism to induce shutoff. No significant shutoff was observed in Hep2, Hep2/V and C6/36 cells infected with any of the viruses.

RT-PCR specific for IFN- β mRNA in 293 infected cells and IFN reporter gene assays revealed that TCMV was capable of counteracting IFN production similar to wt BUNV, whereas the other NSs minus viruses (ANAV, ANBV, BORV, TETEV and BMAV) were found to be capable of inducing IFN in infected cells. However, only low level of IFN- β mRNA and weak activation of the IFN- β promoter was detected in ANAV and BMAV-infected cells.

Declaration

(i) I, Maizan Mohamed, hereby certify that this thesis, which is approximately 27,000 words in length, has been written by me, that it is the record of work carried out by me and that it has not been submitted in any previous application for higher degree.

Date..... Signature of candidate.....

(ii) I was admitted as a research student in October 2003 as a candidate for the degree of Doctor of Philosophy in Molecular Virology; the higher study for which this is a record was carried out in the Faculty of Biomedical and Life Sciences at the University of Glasgow between 2003 to 2005 and Faculty of Science at University of St. Andrews between 2006 to 2007.

Date..... Signature of candidate.....

(iii) I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of Doctor of Philosophy in the University of St. Andrews and the candidate is qualified to submit this thesis in application for that degree.

Date..... Signature of supervisor.....

Unrestricted copyright declaration

In submitting this thesis to the University of St. Andrews I understand that I am giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. I also understand that the title and abstract will be published, and that a copy of the work may be made and supplied to any *bona fide* library or research worker.

Date..... Signature of candidate.....

Table of Contents

Page

List o	List of Tables 13				
List o	List of Figures 14				
Ackn	Acknowledgements 16				
List o	List of Abbreviations 17				
1.	Intro	duction			23
	1.1	The B	unyavirida	le	23
		1.1.1	Virion ch	naracteristics and properties	23
		1.1.2	Classific	ation of the viruses	32
	1.2	The O	orthobunya	<i>virus</i> genus	32
		1.2.1	Transmis	ssion of the virus	34
		1.2.2	Impact o	f Orthobunyavirus disease	34
		1.2.3	Epidemi	blogy	36
			1.2.3.1	Anopheles A serogroup	36
			1.2.3.2	Anopheles B serogroup	36
			1.2.3.3	Bakau serogroup	36
			1.2.3.4	Bunyamwera serogroup	37
			1.2.3.5	Bwamba serogroup	37
			1.2.3.6	Group C serogroup	37
			1.2.3.7	California serogroup	38
			1.2.3.8	Capim serogroup	38
			1.2.3.9	Gamboa serogroup	38
			1.2.3.10	Guama serogroup	38
			1.2.3.11	Koongol serogroup	39

		1.2.3.12	Minatitlan serogroup	39
		1.2.3.13	Nyando serogroup	39
		1.2.3.14	Olifantsvlei serogroup	39
		1.2.3.15	Patois serogroup	39
		1.2.3.16	Simbu serogroup	40
		1.2.3.17	Tete serogroup	40
		1.2.3.18	Turlock serogroup	40
	1.2.4	Genome orthobun	organisation and protein function of yaviruses	40
		1.2.4.1	L segment and protein	40
		1.2.4.2	M segment and protein	41
		1.2.4.3	S segment and proteins	42
			1.2.4.3.1 Nucleocapsid (N) protein	42
			1.2.4.3.2 The NSs protein	43
			1.2.4.3.3 Non-coding region (NCR)	44
	1.2.5	Virus rep	olication	45
		1.2.5.1	Attachment and entry	45
		1.2.5.2	Transcription and replication of RNA	47
		1.2.5.3	Morphogenesis and release of the virus from the cell	50
1.3	Viral e	evolution		50
1.4	Effect	of virus re	eplication in host cells	53
	1.4.1	Effects o	n host-cell metabolism	53
	1.4.2	Effects o	n host antiviral response	53
	1.4.3	Apoptos	is of the infected cells	56
	1.4.4	Persisten	t infection	57
	1.4.5	Defectiv	e interfering RNA	57

1.5	Phylog genus	genetic ar	nd sequence analyses of Orthobunyavirus	57
	1.5.1	Sequence	es and phylogenetic analysis of S segment	58
	1.5.2	Sequence	es and phylogenetic analysis of M segment	62
	1.5.3	Sequence	es and phylogenetic analysis of L segment	62
1.6	Aims	of the proj	ect	63
Mater	rials an	d Method	S	64
2.1	Mater	ials		64
	2.1.1	Media		64
		2.1.1.1	Cell culture media	64
		2.1.1.2	Plaque assay medium	64
	2.1.2	Cells		64
	2.1.3	Enzymes	8	65
	2.1.4	Reagents	s, chemicals and solutions	65
		2.1.4.1	Reagents and chemicals	65
		2.1.4.2	Oligonucleotides/ primers	66
		2.1.4.3	Reverse Transcription (RT) and Polymerase Chain Reaction (PCR)	66
		2.1.4.4	Competent cell preparation, transformation, and plasmid extraction	66
		2.1.4.5	Transfection of cultured cells	69
		2.1.4.6	RNA preparation	69
		2.1.4.7	<i>In vivo</i> protein labelling, polyacrylamide gel electrophoresis (PAGE) and <i>in vitro</i> transcription and translation (TnT)	69
	2.1.5	Plasmids	3	70
	2.1.6	Bacterial	strains	71
	2.1.7	Viruses		71

2.

Metho	ls		71	
2.2.1	Maintenance of	mammalian and insect cell lines	71	
2.2.2	Preparation of virus stock			
2.2.3	Plaque Assay		74	
2.2.4	Plaque purifica	tion of viruses	75	
2.2.5	Preparation of S	SDS-PAGE	75	
2.2.6	In vivo protein	labelling	76	
2.2.7	Preparation of v	virions for RNA extraction	76	
2.2.8	Preparation of v	viral nucleocapsid	76	
2.2.9	RNA Extraction	n	77	
2.2.10	RNA ligation		77	
2.2.11	Reverse Transcription-Polymerase Chain Reaction (RT-PCR)			
2.2.12	3'/5' RACE RT	T-PCR	78	
2.2.13	Modified 3'/5' RACE			
2.2.14	Rolling circle a	mplification-RACE (RCA-RACE)	80	
2.2.15	Cloning		81	
	2.2.15.1 Ligat	ion	81	
	2.2.15.2 Prepa	aration of competent E. coli cells	81	
	2.2.15.3 Trans	formation	81	
	2.2.15.4 Plasn	nid preparation	82	
	2.2.15.5 Restr plasm	iction endonuclease digestion of nid DNA	82	
2.2.16	Sequences anal	ysis and alignments	82	
2.2.17	In vitro Transci	ription and Translation (TnT) assay	83	
	2.2.17.1 Trans	scription of T7 transcription plasmid	83	
	2.2.17.2 Trans	cription and Translation reaction	84	

2.2

		2.2.18	Interferon a	ssay	84
			2.2.18.1 IF	N reporter gene assays	84
			2.	2.18.1.1 Dual Renilla luciferase assay system	84
			2.	2.18.1.2 IFN-specific RT-PCR	85
		2.2.19	Shutoff of infected cell	host cell protein synthesis in virus- ls	86
			2.2.19.1 In	vivo protein labeling at different time points	86
			2.2.19.2 L	uciferase expressing plasmid based assay	86
3.	Chara	acteriza	tion of Viral	Growth and Protein Profiles	87
	3.1	Introd	uction		87
	3.2	Result	s		88
		3.2.1	Virus propa	gation	88
		3.2.2	Plaque mor	phologies produced by viruses	88
		3.2.3	In vivo prote	ein labelling in Vero-E6 cells	99
		3.2.4	Shutoff of h	ost protein synthesis	99
	3.3	Discu	ssion		105
		3.3.1	Virus propa	gation	105
		3.3.2	Plaque char	acteristic of the viruses	106
		3.3.3	Protein synt	hesis in virus infected cells	106
		3.3.4	Shutoff of h	ost protein synthesis	107
4.	Seque	ence An	alysis of the	S RNA Segments of Orthobunyaviruses	108
	4.1	Introd	uction		108
	4.2	Result	s		108
		4.2.1	RT – PCR,	cloning and sequencing of the S segments	108
		4.2.2	'Architectur	re' of S segment	111
		4.2.3	Identificatio	on of potential NSs ORF in the S segments	130

		4.2.4	Sequence	s analysis of 3'/5' NCR	136
		4.2.5	N nucleot	ide and protein comparison	140
		4.2.6	NSs prote	ein comparison	145
		4.2.7	Phylogen	etic of S segments	145
		4.2.8	Partial m	and L sequences	151
	4.3	Discus	ssion		156
5.	Interf cause	eron R d by NS	esponse an Ss Minus C	nd Shutoff of Host Protein Synthesis Orthobunyaviruses	160
	5.1	Introd	uction		160
	5.2	Result	S		160
		5.2.1	<i>In vitro</i> tr	anscription and translation	160
		5.2.2	Shutoff of	f host protein synthesis	162
			5.2.2.1	Shutoff in A549 infected cells	162
			5.2.2.2	Shutoff in Hep2 infected cells	162
			5.2.2.3	Shutoff in Hep2/V cells infected with the viruses	162
		5.2.3	Shutoff or reporter b	of protein synthesis using a luciferase assed assay	165
			5.2.3.1	Shutoff in infected BHK-21 cells	165
			5.2.3.2	Shutoff in C6/36 infected cells	168
		5.2.4	Induction	of IFN- α/β in infected cells	168
			5.2.4.1	Reporter gene assay	168
			5.2.4.2	IFN – specific RT – PCR	170
	5.3	Discus	ssion		170
		5.3.1	<i>In vitro</i> tr	anscription and translation	173
		5.3.2	Shutoff of	f host protein synthesis	173
		5.3.3	Productio	n of interferon in infected cells	177

6.	General Discussion and Conclusion					
	6.1 Virus growth, protein synthesis and ability of the viruses to cause shutoff of host protein synthesis					
	6.2	Sequence analysis of the S RNA segments of orthobunyaviruses	179			
	6.3	Ability of NSs minus viruses in inhibiting the production of interferon and in causing the shutoff of host protein synthesis	182			
7.	Refe	rences	184			
8.	Appendix					

List of Tables

- Table 1.1Pattern of Bunyaviridae protein sizes (kDa).
- Table 2.1Sequences of the oligonucleotides used.
- Table 2.2Viruses used in this study.
- Table 3.1Propagation of the virus in Vero-E6, BHK-21 and
LLCMK2 cells.
- Table 3.2Summary of the plaque produced by the viruses in infected
Vero cells.
- Table 3.3Summary of the estimated size of protein synthesized by the
viruses and shutoff caused by the viruses in infected Vero
cells.
- Table 4.1 Primers used in the RT PCR.
- Table 4.2Summary of the positive-sense RNA sequences obtained
from the study and also from the prototype virus for each
published serogroups; Bunyamwera (BUNV), California
(CEV), Group C (ITQV) and Simbu (OROV).
- Table 4.3Nucleotide and amino acid sequence identity between
studied viruses and representative virus of Bunyamwera,
California, Group C and Simbu serogroups.
- Table 4.4NSs amino acid sequences identities of the viruses.

List of Figures

- Fig. 1.1 Schematic structure of an Orthobunyavirus virion.
- Fig. 1.2 Coding strategies of *Bunyaviridae* genome segments.
- Fig. 1.3 Terminal consensus sequences of the S, M and L genome segments of each genus of the family *Bunyaviridae*.
- Fig. 1.4 Complementary sequences and possible base-paired structures between the 3' and 5' termini of BUNV genomic RNA segments.
- Fig. 1.5 Antigenic relationships among viruses and serogroups within genus *Orthobunyavirus*.
- Fig. 1.6 Transmission cycle of La Crosse virus of *Orthobunyavirus* genus.
- Fig. 1.7 Replication cycle of viruses in the family *Bunyaviridae*.
- Fig. 1.8 Bunyavirus transcription and replication strategies.
- Fig. 1.9 Patterns of genome subunit exchange among members of the genus *orthobunyavirus*.
- Fig. 1.10 Mechanism of type I IFN induction, signaling and action.
- Fig. 1.11 Phylogenetic tree of aligned N ORF of members of the Bunyamwera, California, Group C and Simbu serogroups.
- Fig. 3.1 Plaques formed by the viruses in infected Vero cells.
- Fig. 3.2 Viral protein synthesis and shutoff of host protein synthesis of the viruses in Vero infected cells.
- Fig. 4.1 Nucleotide sequences of positive-sense S segments of the viruses.
- Fig. 4.2 Alignment of initiation codon of N and NSs proteins of the 5' region of positive-sense S segment to indicate their spacing.
- Fig. 4.3 Open reading frames of ANAV, ANBV, BORV, BMAV, CAPV, TCMV, TETEV and control virus BUNV using the DNAMAN programme (Lynnon biosoftware).
- Fig. 4.4 Alignment of complete 5' NCR of virus-sense S RNA segments.

- Fig. 4.5 Presence of repeated sequences in 5' NCR of virus-sense RNA.
- Fig. 4.6 Alignment of complete N protein of the viruses.
- Fig. 4.7 Alignment of NSs protein.
- Fig. 4.8 Phylogeny of the studied viruses based on complete N ORF amino acid sequence of ANAV, ANBV, BWAV, BORV, E147V, GAMV, MBOV, NDV, SJ2441V and TETEV, and partial N ORF of BERV, GMAV and PLSV.
- Fig. 4.9 Nucleotide and amino acid sequences of positive-sense of partial M (A) and L segments (B) of PATV and partial L of CAPV (C).
- Fig. 4.10 Position of L sequence obtained for CAPV and PATV (A), and M sequence obtained for PATV (B).
- Fig. 5.1 Proteins expressed from orthobunyavirus S segments in the *in vitro* TnT system using wheat germ lysate and rabbit reticulocyte lysate system.
- Fig. 5.2 Shutoff of host protein synthesis in infected A549 cells at 18 h pi.
- Fig. 5.3 Shutoff of host protein synthesis in infected Hep2 cells at 18 h pi.
- Fig. 5.4 Shutoff of host protein synthesis in infected Hep2/V cells at 18 h pi.
- Fig. 5.5 Shutoff of host cell protein synthesis in infected BHK-21 cells at 12 and 24 h pi from expression of pRL-SV40.
- Fig. 5.6 Inhibition of host cell protein synthesis in infected C6/36 cells from the expression of phRL-CMV at 18 h pi.
- Fig. 5.7 Activation of IFN- β promoter.
- Fig. 5.8 Detection of IFN- β and cellular γ -actin mRNA.

Acknowledgements

I would like to express my sincere thanks to my supervisor, Professor Richard Elliott for his advice, patience and enthusiasm throughout the years of work.

I would also like to acknowledge my assessors, Prof. Barklie Clements from Glasgow University and Prof. Rick Randall from St. Andrews University, for helpful discussions and comments of my work.

Thanks to everyone from RME group: Angela, Ping Li, Alain, Saleh, Anice, Carol, Xiaohong, Russell, Gjohn and Vincent for their cheerfulness, help and advice. My thank also goes to Dan Young, Monica and Bernie from Rick Randall's Group, Alex from Richard Iggo's Group and Pablo from Martin Ryan's Group for helping me in some of the technical work.

Gratitude also goes to the administration and support staff of BMS, St. Andrews University and Virology Section, Glasgow University.

Finally, special thanks to my husband, parent, children, brothers and sisters for their enduring patience, support and encouragement.

List of Abbreviations

AP	Anchor primer
APS	Ammonium persulphate
ATF	Activated transcription factor
ВНК	Baby hamster kidney
BSA	Bovine serum albumin
cDNA	Complementary deoxyribonucleic acid
CF	Complement fixation
CIP	Calf intestinal phosphatase
cpe	Cytopathic effect
CTD	C-terminal domain
DEPC	Diethyl pyrocarbonate
dH ₂ O	Deionised H ₂ O
DI	Defective interfering
DMEM	Dulbecco's modified Eagle's medium
DNA	Deoxyribonucleic acid
dsRNA	Double-strand RNA
E. coli	Escherichia coli
FBS	Foetal bovine serum
GMEM	Glasgow modified Eagle's medium
HI	Hemagglutination inhibition
IAP	Inhibitor of apoptosis
IFA	Immunofluorescent antibody
IFN	Interferon
IFNAR	IFN alpha receptor
IRF	IFN regulatory factor
ISRE	IFN-stimulated regulatory element
JAK	Janus kinase
LAR	Luciferase assay reagent
LB	Luria broth

MAbs	Monoclonal antibodies
MED	Mediator
MEM	Minimum essential medium
ML	Maximum-likelihood
MOI	Multiplicity of infection
NA	North America
NCR	Non coding region
NJ	Neighbor-joining
NS	Non-structural
nt	nucleotide
OAP	Oligo (dt) anchor primer
ORF	Open reading frame
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffer saline
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
PFU	Plaque forming unit
pi	Post-infection
PKR	Protein kinase R
PLB	Passive lysis buffer
RACE	Rapid amplification of cDNA ends
RCA	Rolling circle amplification
RNA	Ribonucleic acid
RNP	RNA polymerase
RT	Reverse transcription
SA	South America
SDS	Sodium dodecylsulphate
ssRNA	Single-strand RNA
STAT	Signal transducer and activator of transcription
TBE	Tris-Borate EDTA
TEMED	NNN,N,-tetramethylethylenediamine
TMD	Transmembrane domain
TnT	Transcription and translation
TPB	Tryptose phosphate broth

TSB	Transformation and storage buffer
VN	Viral neutralisation
VSP	Virus specific primer
VSPF	Virus specific primer forward
VSPR	Virus specific primer reverse

Abbreviations of virus names

ACAV	Acara virus
AINV	Aino virus
AKAV	Akabane virus
ANAV	Anopheles A virus
ANBV	Anopheles B virus
BAKV	Bakau virus
APEUV	Apeu virus
BATV	Batai virus
BERV	Bertioga virus
BMAV	Batama Virus
BORV	Boraceia virus
BOTV	Botambi virus
BUNV	Bunyamwera virus
BWAV	Bwamba virus
CAPV	Capim virus
CARV	Carapayu virus
CEV	California encephalitis virus
CMV	Cytomegalovirus
CVV	Cache Valley virus
DUGV	Dugbe virus
EREV	Eretmapodites virus
GAMV	Gamboa virus
GERV	Germiston virus
GMAV	Guama virus
GROV	Guaroa virus
ILEV	Ilesha virus
INKV	Inkoo virus
ITQV	Itaqui virus
JATV	Jatobal virus
JCV	Jamestown Canyon virus
JSV	Jerry Slough virus
KETV	Ketapang virus

KEYV	Keystone virus
KOOV	Koongol virus
KRIV	Kairi virus
LACV	La Crosse virus
LOKV	Lokern virus
LUMV	Lumbo virus
MADV	Madrid virus
MELV	Melao virus
MMLV	Murine leukaemia virus
MNTV	Minatitlan virus
MPOV	M'Poko virus
MTBV	Marituba virus
MURV	Murutucu virus
NDV	Nyando virus
NEPV	Nepuyo virus
NORV	Northway virus
OLIV	Olifantsvlei virus
ORIV	Oriboca virus
OROV	Oropouche virus
OSSAV	Ossa virus
PAHV	Pahayokee virus
PATV	Patois virus
PEAV	Peaton virus
PGAV	Pangola virus
PLSV	Palestina virus
RESV	Restan virus
RVFV	Rift Valley Fever virus
SHOV	Shokwe virus
SHUV	Shuni virus
SJ2441V	San Juan 2441 virus
SSHV	Snowshoe hare virus
SV40	Simian virus 40
TAHV	Tahyna virus
TCMV	Tacaiuma virus

TENV	Tensaw virus
TFV	Telok Forest virus
TRV	Tanjong Rabok virus
TURV	Turlock virus
TVTV	Trivittatus virus
UMBV	Umbre virus
UUKV	Uukuniemi virus
WONV	Wongol virus
WYOV	Wyeomyia virus
XINV	Xingu virus
ZELV	Zelga virus

1 Introduction

1.1 The Bunyaviridae

1.1.1 Virion characteristics and properties

The family *Bunyaviridae* is a large group of more than 350 viruses which are mainly arthropod-borne. The members of this family share characteristics such as particle morphology, molecular composition, mode of transmission, genome structure and coding strategies (Fenner, 1975). The viruses are spherical in shape with a diameter of 80-110 nm, with two surface glycoproteins (Gc and Gn) embedded in a lipid envelope. The envelope is usually derived from cellular Golgi membranes or rarely from surface membranes, and surrounds the tripartite single-stranded negative sense RNA genome. The three segments are designated large (L), medium (M) and small (S) (Bishop et al., 1980). The segments are present in virions as ribonucleoproteins (RNPs) and each of them consists of a single genomic RNA encapsidated by the viral nucleocapsid (N) and L proteins (Fig. 1.1) (Obijeski et al., 1976).

The buoyant densities of virions in sucrose and CsCl are 1.16-1.18 and 1.20-1.21 g/cm³ respectively, and virus particles are labile to heat, lipid solvents, detergents, formaldehyde, 70% ethanol and ultraviolet radiation (Bishop et al., 1980). The composition of virus particles was found to be 1-2% RNA, 58% protein, 20-30% lipid and 2-7% carbohydrates by weight (Obijeski and Murphy, 1977).

Members of the *Bunyaviridae* family are divided into five genera: *Orthobunyavirus, Hantavirus, Nairovirus, Phlebobovirus* and *Tospovirus*, based on their serological relationships and supported by biochemical analyses (Elliott, 1990). They are commonly referred as orthobunyaviruses, hantaviruses, nairoviruses, phleboviruses and tospoviruses, respectively (Elliott, 1997).

The lengths of the S, M and L segments vary among genera, ranging from 6.3-13 kb for the L segment, 3.5-5 kb for the M segment and 0.8-2.9 kb for the S segment. The sizes and coding strategies of proteins encoded by viruses with in each genus are similar (Fig. 1.2A, 1.2B, 1.2C and Table 1.1) (Elliott, 1990). The 3' and 5' terminal nucleotide sequences of



Fig. 1.1. Schematic structure of an Orthobunyavirus virion. The surface glycoproteins Gc () and Gn () are embedded in lipid bilayer as heterodimers. The three nucleocapsids are helical, circular and comprise one each of the unique ssRNAs (L, M and S) encapsidated by N proteiner) and associated with the L protein () to form ribonucleoproteins (RNP). The size of the virion is about 80-120 nm.

A. L segment



Fig. 1.2 Coding strategies of *Bunyaviridae* genome segments. Genomic RNAs are represented by purple boxes, black boxes indicate 3'/5' NCR, mRNAs are shown as green boxes, red boxes indicate host-derived primer sequence at 5' end, and arrowheads indicate truncated 3' end, and nt indicates nucleotides. Gene products, with their size in kilodaltons (kDa) are represented by light and dark blue boxes. Two examples of *Phlebovirus* M segments are given which differ with respect to the presence or absence of NSm (Adapted from Elliott, 1996).





C. S segment



RNA	GENUS				
Protein	Orthobunyavirus	Hantavirus	Nairovirus	Phlebovirus	Tospovirus
	<u>kDa</u>	<u>kDa</u>	kDa	<u>kDa</u>	<u>kDa</u>
L segment					
L	259-263	246-247	459	238-241	330-332
M segment					
Gn	29-41	68-76	30-45	50-70	52-58
Gc	108-120	52-58	72-84	55-75	72-78
NSm	15-18	none	78-85, 92-115	None or 78	34
L segment					
Ν	19-25	50-54	48-54	24-30	52
NSs	10-13	none	none	29-31	29

 Table 1.1
 Pattern of *Bunyaviridae* protein sizes (kDa) (Taken from Elliott, 2001).

the three RNA segments are conserved within a genus, but differ from those of other genera (Figure 1.3). They are base-paired, forming non-covalently closed and circular RNAs (Obijeski et al., 1976). In most of the orthobunyaviruses, the terminal 11 bases of S, M and L segment are complementary except a non-canonical pairing at the position 9 and - 9 (Fig. 1.4) (Elliott, 1990). The viral mRNAs are truncated at the 3' end compared to the genomic RNAs but they do not appear to be polyadenylated (Bishop et al., 1980).

In contrast to the other enveloped RNA viruses, bunyaviruses do not have an internal matrix protein, suggesting that morphogenesis of the virus involves the RNPs interacting directly with the cytoplasmic tails of the glycoproteins embedded in the Golgi membrane (Smith and Pifat, 1982). Viral RNA transcription is primed by host RNA sequences derived by a "cap-snatching" process which is similar to that of orthomyxoviruses (Kolakofsky and Hacker, 1991). Virus replication occurs in the host cell cytoplasm and virions mature by budding into intracytoplasmic vesicles from the internal membranes of the Golgi apparatus (Bishop et al., 1980; Schmaljohn and Dalrymple 1983; Elliott, 1990).

Bunyavirus particles are composed of four structural proteins: two external glycoproteins (Gc and Gn) encoded by the M segment, the N protein encoded by the S segment and the L transcriptase protein encoded by the L segment. Schematic representations of the genome, coding strategies and encoded products from a representative member of each genus are shown in Figures 1.2A, 1.2B and 1.2C. Viruses in Orthobunyavirus, Phlebovirus and Tospovirus genera also encode non-structural (NS) proteins such as NSm on the M segment and NSs on the S segment (Elliott et al., 1991). The N and NSs proteins of orthobunyaviruses are encoded in overlapping open reading frame (ORF) of the virion complementary S RNA, whereas the NSs protein of phleboviruses and tospoviruses, and NSm protein of tospoviruses are encoded in ambisense coding strategies (Figures 1.2B and C). Two NSm proteins about 70 and 85 kDa are encoded by Dugbe virus (DUGV) in the Nairovirus genus, which are precursors for the viral glycoproteins (Marriott et al., 1992). For phleboviruses, the NSm protein is encoded by M segment of Rift Valley Fever virus (RVFV) as part of the glycoprotein precursor, whereas Uukuniemi virus (UUKV) does not encode an NSm. The ORF of each segment is flanked by 3'/5' non-coding region (NCRs) sequences (Obijeski et al., 1980). No NS protein is encoded by the S and M segments of hantaviruses (Elliott et al., 1991).

Orthobunyavirus	3' UCAUCACAU 5' AGUAGUGUG
Hantavirus	3' AUCAUCAUCUG 5' UAGUAGUAUGC
Nairovirus	3' AGAGUUUCU 5' UCUCAAAGA
Phlebovirus	3' UGUGUUUC 5' ACACAAAG
Tospovirus	3' UCUCGUUA 5' AGAGCAAU

Fig. 1.3. Terminal consensus sequences of the S, M and L genome segments of each genus of the family *Bunyaviridae* (Elliott, 1996).



Fig. 1.4 Complementary sequences and possible base-paired structures between the 3' and 5' termini of BUNV genomic RNA segments. The terminal 11 nucleotides (left of black line) are conserved in all genome segments; red letters represent the next four nucleotides (only 3 in M segment) which are conserved for each segment within the orthobunyaviruses, and blue letters represent the non-canonical pairing at position 9 and -9 (Taken from Elliott et al., 1991). Except for hantaviruses which have no arthropod vector and are transmitted by rodent excretions, viruses in each genus have their specific biting arthropods for their transmission; orthobunyaviruses are transmitted by mainly mosquitoes or midges, nairoviruses mainly by ticks, phleboviruses mainly by sandflies or ticks, and tospoviruses are transmitted by thrips (Bishop and Shope, 1979; Pringle, 1996).

1.1.2 Classification of the viruses

Serological tests based on complement fixation (CF), hemagglutination inhibition (HI), indirect immunofluorescent antibody (IFA) and viral neutralisation (VN) were used to study the antigenic relationships among viruses in the *Bunyaviridae* family and their classification into genus or serogroup (Hunt and Calisher, 1979; Karabatsos, 1985). Many bunyaviruses are still uncharacterised and remain outside the existing generic structure (Zeller et al., 1989).

VN and HI assays are used to detect the relationships among the viral glycoproteins, while the CF assay is used to detect the relationships among the conserved N proteins (Beaty and Bishop, 1988). Viruses are grouped into a genus based on the CF test, while VN and HI tests are used to divide them into serogroups (Calisher, 1996). Because of the antigenic relationships between Bunyamwera virus (BUNV) and certain other groups of viruses, Casals and Whitman (1960) and Whitman and Shope (1962) suggested placing the viruses into the *Bunyamwera* supergroup. This supergroup was later known as *Bunyavirus* genus in the family of *Bunyaviridae* because of their similarity in molecular, morphogenetic, structure and mode of replication (Fig. 1.5) (Bishop et al., 1980). The International Committee for Taxonomy of Viruses later decided to rename this genus as *Orthobunyavirus* (Elliott et al., 2000). Based on varying degrees of their serological relationships, viruses in the *Orthobunyavirus, Phlebovirus* and *Nairovirus* are further divided into serogroups, complexes, subtypes, variants, varieties and strains.

1.2 The Orthobunyavirus genus

There are more than 170 viruses in this genus which are divided into 18 serogroups; Anopheles A, Anopheles B, Bakau, Bunyamwera, Bwamba, Capim, California, Gamboa,



Fig. 1.5. Antigenic relationships among viruses and serogroups within genus *Orthobunyavirus*. Anopheles A, Bakau and Nyando serogroups are not included in this figure (Taken from Calisher, 1988). Guama, Group C, Koongol, Minatitlan, Nyando, Olifanstlei, Patois, Simbu, Tete and Turlock, with a few additional as yet ungrouped viruses (Calisher, 1996).

1.2.1 Transmission of orthobunyaviruses

Members of this genus are capable of replicating alternately in vertebrates and several arthropod vectors. Their life cycle involves two types of vertebrate hosts, amplifier and dead end hosts (Fig. 1.6). For example in La Crosse virus (LACV) life cycle, chipmunks serve as an amplifier host where the infection is asymptomatic. During the viremic stage, infection of further mosquitoes occurs following ingestion of a blood meal. In this cycle, humans are considered a dead-end host as they are unlikely to transmit the virus back to mosquito and they became infected through the biting of infected mosquitoes (Borucki et al., 2002). Venereal and transovarial transmission in certain arthropod vectors has been reported for some members of the genus (Bishop et al., 1980; Thompson and Beaty, 1977). These two modes of infection are important for the maintenance of the virus in the mosquito population, especially during the winter season (Beaty and Bishop, 1988).

1.2.2 Impact of orthobunyavirus disease

Some members in the *Orthobunyavirus* genus are capable of causing disease in humans such as LACV, Cache Valley virus (CVV), Tahyna virus (TAHV) and Oropouche virus (OROV) (Elliott, 1997). BUNV, the prototype member of the family *Bunyaviridae* and the genus of *Orthobunyavirus*, causes febrile illness with headache, arthralgia and rash in humans in Sub-Saharan Africa (Nichol, 2001). In 1998, Garissa virus, a reassortant of BUNV and Batai virus (BATV) was isolated and found to be responsible for causing hemorrhagic fever outbreaks in Kenya and Somalia (Bowen et al., 2001). LACV and Jamestown Canyon virus (JCV) have been reported to be responsible for the majority of paediatric viral encephalitis cases in the United States (US) (Kappus et al., 1983). CVV has also been reported to be related to meningitis cases in humans in the US (Sexton et al., 1997; Campbell et al., 2006). OROV has been shown to cause an acute febrile dengue-like illness called Oropouche fever in Brazil, South America, Trinidad, Panama and Peru (Pinheiro et al., 1981; Saeed et al., 2000).



Fig. 1.6 Transmission cycle of La Crosse virus of *Orthobunyavirus* genus. This cycle involves two types of vertebrate hosts; small mammals as an amplifier-host and humans as dead-end host. Mosquitoes are the vector for this virus, which are capable of transovarial and venereal transmission of this virus (Taken from Borucki et al., 2002).
These viruses not only pose threat to humans. Some members in the Simbu serogroup could cause severe disease in animals. Akabane virus (AKAV) and Aino virus (AINV) were found to be associated with abortions, stillbirths, and congenital defects in cattle, sheep and goat (Inaba et al., 1975). These viruses are widely distributed in Southeast Asia, Australia, and East Asia leading to a devastating economic impact on the livestock in these countries. CVV which is found throughout the United States, Canada and Mexico, is associated with fetal death, stillbirths and multiple congenital malformations in sheep (Edwards et al., 1989)

1.2.3 Epidemiology

1.2.3.1 Anopheles A serogroup

With the exception of Tacaiuma virus (TCMV), which was also isolated from humans and primates, most members of this serogroup were isolated from various mosquitoes especially Anopheline and Culicine recovered in South and North America (Karabatsos, 1985). Data obtained from serology surveys indicate that some of the viruses in this serogroup cause natural infections in a range of species in Brazil such as livestock animals, birds and rodents. Serologic cross-reaction has been detected between Anopheles A and California serogroup viruses (Bishop, 1996).

1.2.3.2 Anopheles B serogroup

These viruses were obtained from mosquitoes collected in South America. Little is known about their vertebrate hosts (Karabatsos, 1985).

1.2.3.3 Bakau serogroup

Viruses in this serogroup were isolated from mosquitoes and ticks in Asia and Africa (Karabatsos, 1985). BAKV, Tanjong Rabok (TRV) and Telok Forest (TFV) viruses were isolated from monkeys. Serological surveys indicated that they also infect birds, bats, flying squirrels and rodents. Although antibody surveys have indicated that the viruses infect humans, no isolation has been made to date (Bishop, 1996).

1.2.3.4 Bunyamwera serogroup

Most of the viruses in this serogroup are transmitted by mosquitoes, except for Lokern (LOKV) and Main Drain (MDV), which have also been isolated from *Culicoides* (Karabatsos, 1985). Many of the viruses such as BUNV, Germiston (GERV), Guaroa (GROV), BATV, Ilesha (ILEV), Shokwe (SHOV), Wyeomyia (WYOV) and Xingu (XINV) have been isolated from human. Some of the viruses such as BUNV, GERV, GROV, BATV, ILEV, Tensaw (TENV) and WYOV are associated with human infections (Karabatsos, 1985). MDV was isolated from horses with encephalitis. WYOV has been reported in South and North America while BATV has been detected in Asia (Karabatsos, 1985). BUNV was first isolated from *Aedes* mosquitoes in Uganda in 1943 and also from viremic humans in Africa (Karabatsos, 1985). Antibodies to this virus have been detected in humans, primates, rodents, birds and livestock animals. CVV has been isolated from a variety of mosquitoes in North America and infection of sheep causes embryonic and fetal death, stillbirth, and multiple congenital malformations (Edwards et al., 1989; McConnell et al., 1987). Recently, this virus has also been associated with infection in humans (Sexton et al., 1997; Campbell et al., 2006).

1.2.3.5 Bwamba serogroup

Viruses in this serogroup are geographically restricted to Africa and most of them were isolated from humans with febrile illness (Karabatsos, 1985). Bwamba virus (BWAV) was first isolated in 1937 from an infected man in Bwamba, Uganda. Antibodies to these viruses have been detected in humans, livestock animals, and avian sera in Africa (Bishop and Shope, 1979). These viruses are serologically related to the California serogroup (Casals, 1963).

1.2.3.6 Group C serogroup

These viruses have been isolated from mosquitoes, rodents and marsupials in South and North America (Karabatsos, 1985). Some of the viruses such as Carapayu (CARV), Oriboca (ORIV), Itaqui (ITQV), Nepuyo (NEPV), Apeu (APEUV), Marituba (MTBV), Murutucu (MURV), Restan (RESV), Ossa (OSSAV) and Madrid (MADV) viruses have been associated with human disease such as self-limited and dengue-like illness including fever, headache, myalgia, nausea, vomiting and weakness (Nunes, 2005).

1.2.3.7 California serogroup

The prototype virus of this serogroup, California encephalitis (CEV) was isolated in 1943 from mosquitoes in North America (Hammon and Reeves, 1952). Viruses in this serogroup have a widespread distribution covering North and South America, Europe, Asia and Africa (Calisher, 1983). Besides mosquito vectors, these viruses have also been obtained from rodents and other animals (Karabatsos, 1985). Many of the viruses such as CEV, JCV, LACV, Inkoo (INKV), Snowshoe hare (SSHV), TAHV and Trivittatus (TVTV) cause infection in humans with encephalitic symptoms.

1.2.3.8 Capim serogroup

Culex mosquitoes serve as vectors for viruses in this serogroup. These viruses are associated with rodent hosts and a number of livestock animals, but no infection has been reported in humans. These viruses have been detected only in North and South America (Karabatsos, 1985).

1.2.3.9 Gamboa serogroup

These viruses have been isolated from Culicine mosquitoes collected in Central and South America (Karabatsos, 1985). Their natural vertebrate hosts are not known and none of the viruses infect humans (Bishop, 1996).

1.2.3.10 Guama serogroup

These viruses are restricted to South and North America. Mosquitoes of many different genera can act as vectors, and the viruses are associated with rodent and marsupial hosts (Karabatsos, 1985). They have also been isolated from birds, bats and a variety of livestock animals. Some of the viruses such as Catu (CATV) and Guama (GMAV) viruses have been associated with disease in humans (Vasconcelos et al., 2001).

1.2.3.11 Koongol serogroup

To date only two viruses were identified in this serogroup, Koongol (KOOV) and Wongol (WONV) viruses which were isolated from *Culex annulirostris* mosquitoes in Australia (Doherty et al., 1963). Although antibodies to these viruses were detected in cattle and other vertebrates such as marsupials and birds, no virus isolation has been made from these animals (Bishop and Shope, 1979).

1.2.3.12 Minatitlan serogroup

Only two viruses, Minatitlan (MNTV) and Palestina (PLSV) have been identified in this serogroup, which were isolated from *Culex* mosquitoes and hamsters in Mexico and Equador (Karabatsos, 1985). To date, no information is available on their vertebrate host and infection in humans.

1.2.3.13 Nyando serogroup

Viruses in this serogroup have been isolated from mosquitoes in Africa. Nyando virus (NDV) has been recovered from humans with a febrile illness in central Africa and antibodies to this virus have been detected in human sera in Kenya and Uganda (Karabatsos, 1985).

1.2.3.14 Olifantsvlei serogroup

These viruses are geographically restricted to Africa and were isolated from *Culex* mosquitoes. Little is known about their vertebrate hosts and these viruses have no association with disease in humans (Karabatsos, 1985).

1.2.3.15 Patois serogroup

These viruses are geographically restricted to Central, North and South America. They are vectored by mosquitoes and are associated with rodent hosts. Although serological surveys indicated that Patois (PATV) and Zelga (ZELV) viruses infect humans, to date, no isolation has been made from them (Karabatsos, 1985).

1.2.3.16 Simbu serogroup

These viruses are widely distributed in Asia, Australia, Africa, North and South America (Karabatsos, 1985). They are vectored by culicoids and mosquitoes. Some of the viruses have been recovered from birds and a number of vertebrate species including cattle and pigs. Some of them are human pathogens such as OROV and Shuni virus (SHUV), while AKAV and AINV are of significant veterinary importance (Karabatsos, 1985). OROV infects hundreds of human in Brazil. Serological survey results indicated that this virus can also infects monkeys, birds and rodents (Bishop, 1996).

1.2.3.17 Tete serogroup

Most of the viruses in this serogroup were isolated from birds and are vectored by Ixodid ticks or Culicoides species (Karabatsos, 1985). They have been isolated in Europe, Asia and Africa but none of them has been associated with human disease (Karabatsos, 1985).

1.2.3.18 Turlock serogroup

Viruses in this serogroup are vectored by Culicine mosquitoes from North and South America, Africa, Asia and Europe. Some of the viruses such as Turlock (TURV) and Umbre virus (UMBV) have been isolated from birds (Karabatsos, 1985).

1.2.4 Genome organisation and protein function of orthobunyaviruses

The genome of BUNV was the first virus in this family to be sequenced completely and has a size of 12294 nucleotides, of which 95.3% encodes amino acids (Elliott, 1989). The genome is richer in A + U than C + G residues (Elliott, 1990).

1.2.4.1 L segment and protein

Based on published sequences of complete L segment of Bunyamwera, California and Simbu serogroup viruses, their size lengths are in the range of 6.8-6.9 kb. This segment encodes the L protein (RNA dependent RNA polymerase) in a negative-sense coding strategy (Figure 1.2A). In virus-infected cells, only small amounts of this protein are detected (Elliott et al., 1991). To date, no other coding region beside polymerase ORF was reported in this segment (Elliott, 1989). Examination of the mRNAs synthesized by recombinant L protein revealed that they contain a 5' cap and host derived-primer sequence which is needed for transcription, suggesting that L protein has an endonuclease activity to mediate the 'cap-snatching' process (Jin and Elliott, 1993). Overall, little homology was observed between the L proteins of viruses in different genera. However, amino acid sequence alignment of LACV L protein with sequence of other polymerases from other members of *Bunyaviridae*, has identified the presence of conserved motifs containing a polymerase module common to all RNA dependent polymerase (Roberts et al.1995). It has been shown that the L protein of California serogroup viruses also plays a role in mouse neurovirulence and neurovasiveness, but the mechanism behind it is still unclear (Endres et al., 1991).

1.2.4.2 M segment and protein

Published nucleotide sequences of the M segments of orthobunyaviruses revealed that they are 4458-4534 nucleotides in length (Eshita and Bishop, 1984; Lees et al., 1986; Grady et al., 1987; Pardigon et al., 1988; Brockus et al., 1999; Campbell and Huang, 1999; Wang et al., 2001; Yanase et al., 2003). Similar to the N protein, the M sequences of the viruses from different serogroups exhibit limited sequence homology, while viruses in the same serogroup showed closer similarity, about 66% for Gn, 50% for NSm and 40% for Gc protein (Pringle, 1991).

The M segment of orthobunyaviruses encodes 3 proteins, two surface glycoproteins: Gc (108 to 125 kDa) and Gn (29 to 41 kDa), and a non-structural protein, NSm (15-18 kDa) in the form of a polyprotein precursor. The polyprotein has not been detected in infected cells, suggesting that it is co-translationally cleaved to give the mature proteins (Lappin et al., 1994). The gene order of the M segment of orthobunyaviruses is 5'-Gn-NSm-Gc-3'in the genome-complementary sense (Figure 1.2B) (Fazakerley et al., 1988, Fuller and Bishop, 1982; Nakitare and Elliott, 1993).

Alignment of the M segment sequences of Bunyamwera and California serogroup viruses identified three potential glycosylation sites that are relatively rich in cysteine, and N and C terminal hydrophobic domains which are conserved in Gc and Gn proteins of these viruses (Lees et al., 1986; Pardigon et al., 1988; Brockus and Grimstad, 1999). However, OROV was found to have only two glycosylation sites of which only that in Gc was conserved with the other two serogroups (Wang et al., 2001). These conserved regions are believed to involve in neutralising and protective epitopes (Wang et al., 1993). Najjar et al. (1985) have suggested that these protective epitopes are clustered within a single immunodominant antigenic site. The conserved regions in the Gn glycoprotein are found to contain type-specific antigenic determinants for hemagglutinating and neutralizing antibodies that are used to place the viruses into serogroups (Cheng et al., 2000). The Gn glycoprotein also contains the Golgi targeting and/or retention signals which are required for the Gc to interact with Gn to localize to Golgi compartment (Lappin et al, 1994; Bupp et al, 1996; Shi et.al., 2004). It has been shown that the N terminal domain of NSm is also contains some hydrophobic and non-hydrophobic domains that may be required for virus assembly and as an internal signal sequence for Gc glycoprotein (Shi et al., 2006).

The Gc is found to be responsible for fusion activity and also as a major determinant for viral attachment to mammalian cells (Pobjecky et al., 1986; Pekosz et al., 1995). Homologies with Sindbis virus E1 have suggested that LACV Gc is a class II viral fusion protein (Plassmeyer et al., 2005). Beside the L protein, the Gc protein is also shown to play a major role in determining the virus virulence (Gonzalez-Scarano et al., 1985; Elliott, 1990).

1.2.4.3 S segment and proteins

1.2.4.3.1 Nucleocapsid (N) protein

The N protein is the most abundant protein and first to be expressed in virus-infected cells. Alignment of nucleotide and amino acid sequences of three serogroup viruses of *Orthobunyavirus* genus (Bunyamwera, California and Simbu) revealed the presence of certain conserved regions in their N proteins (Akashi et al., 1984; Dunn et al., 1994; Bowen et al., 1995; Saeed et al., 2001). These regions have been suggested to be associated with the complement fixation antibodies that cross-react among the viruses in the same genus (reviewed by Calisher, 1996). Nucleotide and amino acid sequence comparisons of these viruses reveal no sequence identity between viruses in different genera but with at

least 40% identity between the viruses in the same genus (Dunn et al., 1994; Bowen et al., 1995; Saeed et al., 2001; Nunes et al., 2005). As mentioned in Section 1.1.1, N is used to encapsidate the genomic and antigenomic RNA to form RNPs (Jin and Elliott, 1991). The N protein of BUNV was shown to bind specifically to the 5' terminus of the S genome segment, which may contain the signal to initiate N encapsidation (Osborne and Elliott, 2000). Residues 17 to 20 at the amino terminus of N protein contain the motif FDPE conserved in almost all orthobunyaviruses and have been suggested to be structurally essential for N protein folding and/or stability (Leonard et al., 2005).

1.2.4.3.2 The NSs protein

For orthobunyaviruses, the NSs protein is translated from the same mRNA species as N protein using an alternative start codon in an overlapping reading frames +1 frame (Akashi and Bishop, 1983; Elliott, 1989). The size of NSs protein of orthobunyaviruses is between 10-13 kDa (Table 1.1). Most NSs proteins in Bunyamwera and California serogroup viruses are initiated with a double methionine whereas this feature is not shown by Group C and Simbu serogroup viruses. There is no sequence conservation between the NSs proteins of the viruses in different genera but some sequence identities were observed within the viruses in the same genus (Dunn et al., 1994).

A lot of studies have been carried out to determine the functions of the orthobunyavirus NSs protein. For BUNV, it was found to contribute to the viral pathogenesis by acting as an interferon antagonist that blocks the transcriptional activation of IFN- β (Weber et al., 2002). It is also able to inhibit host cell protein synthesis (Bridgen et al., 2001) and capable in delaying the early stage cell death by inhibiting IRF-3 mediated apoptosis (Kohl et al., 2003). In contrast to other IFN antagonists, NSs inhibits dsRNA-dependent IFN induction but has no effect on the dsRNA-activated PKR and RNase L systems (Streihtenfeld et al., 2003). Furthermore, the NSs protein of LACV was also found to be able to counteract RNA silencing directed against cellular and viral RNA (Soldan et al., 2005). The functions of NSs protein will be discussed in detail in sections 1.4.1, 1.4.2 and 1.4.3.

1.2.4.3.3 Non-coding region (NCR).

The ORF of each RNA segment of bunyaviruses are flanked by 3'/5' NCRs. Although the terminal 11 nucleotides of these NCRs are conserved between the three segments within a genus (Section 1.1.1), the internal regions are unique to each segment and largely nonconserved between different viruses (Haaster and Bishop, 1980). Mini-replicon systems based on reporter genes flanked by NCR sequences have been developed for BUNV (Weber et al., 2001; Kohl et al., 2006) and LACV (Blakqori et al., 2003), which demonstrated that the presence of NCRs alone are adequate to allow transcription, replication, encapsidation and packaging of mini-genome segments by viral proteins. By using this system, Kohl et al. (2004) have showed that the terminal 15 nt sequences of the 3'/5' NCR are required for BUNV S promoter to be functional. These complementary sequences are thought to be important in providing signals for recognition by the virusencoded polymerase and are involved in packaging of the viral genome (Elliott, 1990). Furthermore, the cooperation between 3' and 5' NCR sequences through base-pairing interaction is found to be required for BUNV RNA synthesis (Barr and Wertz, 2004). In addition, Barr and Wertz (2005) have showed that the non-canonical (G-U) at position 9 and -9 (Figure 1.4) is crucial for the signalling of BUNV transcription.

Recombinant BUNV with deletions in the 3'/5' NCR have shown that the 5' NCR is essential for viral growth, while the internal 3' NCR is important for the regulation of viral RNA synthesis (Lowen and Elliott, 2005). Furthermore, competition assays with a variety of viral RNAs have identified a region within the 5' terminus of the BUN S segment for which N had a high preference for binding to the 5' end of the S segment, suggesting that this site may constitute the signal for the initiation of encapsidation by N (Osborne and Elliott, 2000). Computer-assisted RNA folding models used by Kohl et al. (2006) revealed that a conserved sequence within nt 20-33 of the 3'/5' end of the genome segments was necessary for efficient transcription. The 5' NCRs of S segment genome also contain transcription termination signals which are responsible for the truncation of BUN S-segment mRNA (Elliott, 1990; Schmaljohn, 1996). The GU-rich region in the 5'NCR has been shown to play an essential role in directing this termination signals located in 5'NCR of BUNV: one at the nt 91(3' GUCGAC 5') which plays a major role in

termination signalling and the other one is a functionally independent termination signal at the nt 59 (3' UGUCG 5').

A recombinant virus called BUN MLM, in which the L segment open reading frame (ORF) is flanked by the M segment NCRs, was employed to investigate the segment-specific functions of the NCRs. In comparison to wt virus, BUN MLM virus was shown to be attenuated in cultured mammalian cells, had slower disease progression in mice, produced a smaller plaque size, expressed reduced levels of L mRNA and RNA polymerase protein, synthesized less L genomic and anti-genomic RNA, and had an increased particle-to-PFU ratio. The rescued of this BUN MLM mutant virus indicates that BUN NCRs have different efficiency in packaging of the viral genome (Lowen et al., 2005).

1.2.5 Virus replication

The process of bunyavirus replication involves attachment and penetration of the virus into the cell, primary transcription and translation, replication of the viral RNA, secondary transcription and translation, virus assembly and morphogenesis, and release of the virus from cells (Fig. 1.7).

1.2.5.1 Attachment and entry

Attachment is mediated by an interaction of the viral glycoproteins with unknown host receptor, while entry of the virus into the host cell and uncoating are thought to occur by receptor-mediated endocytosis of virions and fusion of viral membranes with endosomal membranes. The viral genomes and polymerase are released into the cytoplasm, where primary transcription is initiated (Schmaljohn and Hooper, 2001). The only cellular receptor identified is the β 3 integrin family which was used for cell attachment of some hantaviruses (Gavrilovskaya et al., 1999) but it is unlikely that the same receptor is used by other genera (Pekosz et al., 1995). Treatment of LAC virions with proteases (bromelain or pronase), which only degraded Gc, was found to abolish the infectivity of the virus to vertebrate host cells, suggesting that Gc is responsible in binding to vertebrate host cells (Kingsford and Hill, 1983), while Gn is shown to be responsible for the binding to mosquito cell lines and midgut cells (Ludwig et al., 1991). Exposing the endosomes at low



Fig. 1.7 Replication cycle of viruses in the family *Bunyaviridae*. Viruses enter the cell by receptor-mediated endocytosis (I). RNPs are released through fusion process of the viral envelope with the endosomal membrane (II). Three segments of genomic RNA are transcribed to mRNA by virion-associated polymerase; S and L mRNAs are translated by free ribosomes in the cytoplasm and M mRNAs by ER-bound ribosomes (III). Processing and glycosylation of envelope protein Gc and Gn (IV). Newly synthesized viral proteins mediate the replication of genomic RNAs to antigenomic RNAs, the templates to synthesize new genomic RNAs (V). Budding of the virus in the Golgi, assembly and release from the cells by endocytosis (VI)(Courtesy from Dr. Alain Kohl).

pH is thought to promote conformational change in Gc, causes fusion of the viral and cell membrane, suggesting a role of Gc in virus entry. Furthermore, treating the cells with ammonium chloride to prevent the acidification of endosomes inhibited infection by CEV (Hacker and Hardy, 1997). However, expression of Gc alone was found not to cause cell-cell fusion, suggesting that an interaction of Gc and Gn is needed for the membrane to be fused (Bupp et al., 1996).

1.2.5.2 Transcription and replication of RNA

Bunyavirus transcription involves the synthesis of mRNAs by virion-associated RNA polymerase or transcriptase which is complementary to genomic templates using host cellderived capped primers. Primary translation involves the translation of L and S segment mRNAs by free ribosomes and M segment mRNAs by membrane-bound ribosomes to yield the viral structural and non-structural proteins. Primary glycosylation and cotranslational cleavage of a precursor to yield envelope proteins (Gc and Gn) and NSm for some of the viruses occurs at the endoplasmic reticulum (ER) (Bishop, 1996). During this process, the L protein also copies vRNA into antigenomic RNA (Schmaljohn and Hooper, 2001). A model for transcription and replication of the family *Bunyaviridae* is presented in Figure 1.8.

Primary transcription of the vRNA to complementary mRNA is initiated by interaction of the virion-dependent RNA polymerase with the three viral RNAs (Bouloy and Hannoun, 1976; Bishop, 1996; Schmaljohn and Hooper, 2001). Studies on co-expression of viral proteins using minigenome RNAs have demonstrated that L and N proteins are necessary for the transcription process, implying that only RNPs and not free RNA act as a template (Dunn et al., 1995). In this transcription process, polymerase starts to synthesise mRNA at the 3' end of template with a capped primer and stop at nt 50-91 before the 5' end of the template, in response to mRNA termination signal (Barr et al., 2006).

The cap-snatching of the transcription process resulted in the presence of viral mRNAs that contain non-templated, host-derived capped primers (about 10 to 20 heterologous nucleotides) at the 5' terminus and are about 50-100 nucleotides shorter than full-length transcript but not polyadenylated at the 3' end (Eshita et al., 1985; Hacker et al., 1990; Jin and Elliott, 1993;). This cap-snatching mechanism was proven by the study of Patterson et



Fig. 1.8. Bunyavirus transcription and replication strategies. (A) from negative sense genome segments, and (B) from ambisense genome segments. Black boxes represent 5' cap structures. Solid blue boxes represent mRNAs, brown boxes represent nucleocapsid protein of genomic and antigenomic RNPs, pink circles represent L protein, red circles represent N protein and green and brown chain represent viral proteins.

al. (1984) which showed that purified LACV virions possess an endonuclease activity which specifically cleaves alfalfa mosaic virus RNA 4 containing a methylated cap group which could be selected by anti-cap antibodies (Patterson et al., 1990). Sequencing the 5' terminal region of the viral mRNAs revealed that these host-derived capped primers are rich in C and G residues and possess a U or a C adjacent to the viral sequence (Bouloy et al., 1990). Cleavage of the capped primers is mediated by endonuclease activity of the L protein (Jin and Elliott, 1993). However, unlike the influenza virus, this process occurs in the cell cytoplasm (Rossier et al., 1986) and is not affected by actinomycin D or α amanitin (Vezza et al., 1979).

Efficient transcription of bunyaviruses requires simultaneous translation (Bellocq et al., 1987; Bouloy, 1991). It has been shown *in vitro* that full length mRNAs could be synthesized only in the presence of rabbit reticulocyte lysate, which was active in translation (Bouloy, 1991). Blocking of translation resulted in the synthesis of incomplete viral transcripts, suggesting that the translation of the nascent bunyavirus mRNA is required to prevent premature termination of the transcription process (Bellocq and Kolakofsky 1987).

After primary transcription and translation, replication, which involves the synthesis of antigenomes to act as template for synthesis of further genomic strands, occurs (Kolakofsky and Hacker, 1991). However replication does not involve the 'cap-snatching' mechanism as no host primer sequences are observed at the 5'end of antigenomic RNA, suggesting that initiation is at the exact 5' end of the RNA and is primer independent (Bishop et al, 1983). Therefore a switch from transcription to replication is required. For bunyaviruses, the polymerase protein must first function as cap-dependent endonuclease to generate a primer for transcription and then switch to a process of independently initiating transcription to produce a full-length of complementary RNA (cRNA). Most likely some viral or host factor is required to initiate RNA replication and to suppress the transcription-termination signal responsible for the generation of truncated mRNA (Bishop, 1996). However, the mechanism involved has yet to be defined. It is likely that viral proteins such as L, N and NSs might be involved in this process since addition of translational inhibitors such as cycloheximide was found to prevent genome replication and secondary transcription (Vezza et al., 1979).

1.2.5.3 Morphogenesis and release of the virus from the cell

Maturation of Bunyavirus is usually occurs by budding at the smooth membrane vesicles in the Golgi complex (Lyons and Heyduk, 1973; Murphy et al., 1973). During this process, RNPs are also present at the same site, suggesting that budding is induced by a transmembrane recognition between the viral glycoproteins and the N protein (Smith and Pifat, 1982). The new virions are then released by exocytosis, which involves fusion of cytoplasmic vesicles with the plasma membrane (Matsuoka et al., 1991). This process was found to be inhibited by monensin, a monovalent ionophore (Cash, 1982).

The Golgi targeting and retention signal of BUNV glycoproteins was shown to reside in the transmembrane domain (TMD) of the Gn protein (Shi et al., 2004). The role of NSm in transport and Golgi retention are not known. However, Shi et al. (2006) have shown that domain I and part of domain II of the N-terminal region of NSm are required for virus assembly and hydrophobic domain V of C-terminal region may be function as an internal signal sequence for the Gc glycoprotein.

1.3 Virus evolution

Two major mechanisms are involved in the evolution of RNA viruses: genetic drift and genetic shift (Holland and Domingo, 1998). Genetic drift occurs through the accumulation of point mutations, deletions, duplications and inversions of viral RNA because of the RNA-dependent RNA polymerases lack the proofreading capabilities during the genome replication process. Genetic shift occurs through the reassortment of viral RNA segments resulted from a mixed infection in host cells (Beaty et al., 1985; Beaty and Calisher, 1991; Pringle, 1996; Elliott, 1996). Reassortment has been demonstrated to occur only between closely related viruses within the same serogroup but not with the viruses from different serogroups (Pringle, 1996).

Naturally occurring reassortant viruses have been reported to occur among Bunyamwera (Dunn et al., 1994; Bowen et al, 2001), Simbu (Akashi et al., 1997; Saeed et al., 2001), Group C (Nunes et al., 2005), California (Chandler et al., 1991) and Patois serogroup viruses (Ushijima et al., 1981). Studies of these reassortants have shown that they usually share the same L and S segments but different M RNA segment (Pringle, 1996). This

reassortment could confer beneficial traits on the progeny reassortants (Elliott, 1990), which sometimes could lead to the emergence of a variant with different pathogenicity and tropism (Elliott, 1996). Amongst orthobunyaviruses, Garissa virus, which caused haemorrhagic fever outbreaks in East Africa, was found to have the L and S segment sequences almost identical to those of BUNV and the M segment sequence closely related to BATV, an orthobunyavirus first detected in Malaysia and which has not been isolated from humans (Briese et al., 2006).

The chances of dual infection and natural reassortment of orthobunyaviruses are enhanced in the arthropod vector, especially in ticks because they feed on different hosts at different life stages (Bishop, 1996). This may be one of the main reasons contributing to the existence of so many virus serotypes, subtypes, variants and varieties of these viruses (Bishop, 1996). In the case of transovarial transmission, the viruses may persist through generations of infected arthropods without being transmitted to a vertebrate host and this can promote the emergence of a new strain of viruses via both genetic drift and genetic shift (Beaty and Bishop, 1988; Calisher, 1988).

Pringle (1996) has suggested that there are different gene pools within the different serogroups in the genus Orthobunyavirus (Figure 1.9). In this figure, he summarised the outcome of *in vitro* heterologous recombination experiments which involved the crossing of ts mutants and determining the phenotypes and genotypes of non-ts reassortments. In this study, all the six viruses in the California serogroup were able to exchange genome segments. However in the Bunyamwera serogroup viruses, only five out of eight were able to exchange segments. The three viruses: MDV, Kairi virus (KRIV) and GROV are genetically isolated and less able to reassort, suggesting that the pattern of restriction within the Bunyamwera serogroup is in agreement with their serological relationships (Hunt and Calisher, 1979); the more divergent their antigenic relationship, the greater the degree of restriction for the virus to reassort. It also illustrates that the tendency of the viruses to undergo reassortment is not related to the known geographical range of these viruses. For example in MAGV which is thought to be limited to South America, can exchange genome segments with African, Eurasian and North American Bunyamwera serogroup viruses, suggesting that there is no genetic barrier for these viruses to reassort (Iroegbu and Pringle, 1981). In contrast, Northway (NORV), MDV and KRIV, which are reported only in North America cannot reassort among themselves, suggesting that

	Africa	Eurasia	North America	South America
Bunyamwera	Batai Bunyamwera Germiston	Batai	Northway	Maguari
Serogroup			Main Drain Kairi	Guaroa
California Encephalitis Serogroup	Lumbo Tahyna	Tahyna	La Crosse Snowshoe Hare Trivittatus	
Simbu Serogroup	Sathuperi	Sathuperi		

GEOGRAPHIC RANGE

Fig. 1.9. Patterns of genome subunit exchange among members of the genus *Orthobunyavirus*. The large rectangles enclose serologically related viruses. Viruses which can exchange genome sub-units by genetic reassortment and can be regarded as a common gene pool are contained within the same heavily lined box. Viruses in different boxes cannot exchange genome sub-units although they belong to the same serogroup. The viruses are arranged vertically by geographic range; all except Batai, Tahyna, and Sathuperi are restricted to a single continent (Taken from Pringle, 1991) restricted host-vector pairing has led to the genetic isolation and stability of the virus (Calisher, 1988).

Based on CF tests, GROV was previously included in the Bunyamwera serogroup virus. However, with NT tests, it exhibits some reactivity similar to California serogroup viruses (Whitman and Shope, 1962), and therefore is considered as a bridging virus or reassortant virus between Bunyamwera and California serogroups (Bowen et al., 1995; Pringle, 1996). However, sequencing of both S (Dunn et al., 1994) and M segment (Briese et al., 2004) has showed that it has closer relationship to Bunyamwera than California serogroup, therefore confirming that GROV is a bridging virus between these two serogroups.

1.4 Effect of virus replication in host cells

1.4.1 Effects on host-cell metabolism

Shutoff of host protein synthesis has been noted in mammalian cells infected with orthobunyaviruses (Lazdins and Holmes, 1979; McPhee and Westaway, 1981; Bridgen et al., 2001; Blakqori and Weber, 2005) and some phleboviruses (Ikegami et al., 2006). It has been suggested that the shutoff of host protein synthesis caused by LACV in mammalian cells could be due to mRNA instability produced by the virus, most probably mediated by the endonuclease activity of the viral transcriptase during the cap-snatching process (Raju and Kolakofsky, 1988). A drastic reduction in shutoff has been observed in cells infected with BUNdelNSs, a mutant BUNV lacking the NSs protein, suggesting that NSs plays an important role in the shutoff of host protein synthesis (Bridgen et al., 2001). No shutoff was observed in cells infected with hantaviruses and nairoviruses which do not encode the NSs protein in their S segment (Elliott et al., 1984; Watret et al., 1985). Shutoff is important in decreasing competition between the virus and cell for cellular factors such as transcription and translation components (Lyles, 2000).

1.4.2 Effects on host antiviral response

Production of interferon (IFN) in infected cells is the first line of host defence against virus infection. Infected cells secrete α/β IFN which stimulates the neighbouring cells to

synthesize antiviral factors, thereby limiting the spread of virus infection (Weber et al., 2002). Induction and activation of specific host molecules by IFN blocks virus infection at several levels, including transcription, translation and RNA degradation. Viruses have evolved several ways for evading the IFN responses; sequestering double-stranded RNA or inhibition of activation of the double-stranded RNA-dependent protein kinase R (PKR), NF-kB, and other IFN regulatory factors (IRF), e.g., IRF-1 and IRF-3, and inhibition of IFN signalling at different levels (signalling of IFN receptors, JAK/STAT activation, and signalling of p48 and ISGF3 transcriptional factors) (Liu et al., 2005). The onset of IFN- α/β response in virus-infected cells occurs on viral entry and release of viral components, including ds-RNA intermediates (Figure. 1.10). These viral components activate the transcription factors IRF-3, IRF7 and NF-kB, and activated transcription factor 2 (ATF-2). Secreted IFN- α/β binds to the IFN alpha receptor (IFNAR) on the surface of both infected and neighbouring cells, resulting in activation of Janus kinase (JAK) and signal transducer and activator of transcription (STAT) pathway and transcription of numerous genes from promoters containing IFN-stimulated regulatory elements (ISRE), resulting in the induction of antiviral state (Munoz-Jordan et al., 2003).

For orthobunyaviruses, the NSs protein of BUNV and LACV has been found to suppress induction of IFN- α/β , inhibit the action of small interfering RNAs and interfere with hostcell apoptosis signalling pathway. Mutant viruses, BUNdelNSs and rLACVdelNSs, which do not express NSs protein, were found to be strong inducers of IFN- α/β . IFN induction by BUNdelNSs correlated with the activation of NF-kB and the IFN transcription factor IRF-3 by the virally-produced double-stranded RNA (dsRNA). In IFN-nonresponsive cells and mice, both BUNdelNSs and wt BUNV were found to replicate almost to the same level. However in IFN-competent systems, wild-type BUNV grew more efficiently than BUNdelNSs, confirming that the NSs protein is an IFN antagonist that blocks the transcriptional activation of IFN-alpha/beta (Weber et al., 2002) and has a role in contributing to viral pathogenesis (Bridgen et al., 2001).

The NSs protein of BUNV inhibits IFN- β gene expression in the mammalian host by dysregulating the phosphorylation at serine 2 in the heptapeptide repeat (YSPTSPS) of the C-terminal domain (CTD) of RNA polymerase (RNAP) II, which would prevent the elongation step of transcription without disturbing initiation of transcription. Interestingly, no interference with CTD phosphorylation of RNAP II was observed in insect cells



Fig. 1.10. Mechanism of type I IFN induction, signaling and action. (A) ds RNA (by-product of virus replication) activates the transcription factors NF-kB and IRF-3. These actions are required for full activation of the IFN-β promoter. IRF-3 is phosphorylated by the kinases IKK and TBK-1 which are activated by the RNA-sensing complex of RIG-I, MDA5 and IPS-1/MAVS. A second signaling pathway involves endosomal TLR-3 and TRIF. (B) Newly synthesized IFN- β binds to the type I IFN receptor (IFNAR) and activates the expression of numerous ISGs via the JAK/STAT pathway. IRF-7 amplifies the IFN response by inducing the expression of several IFN- β subtypes. SOCS and PIAS are negative regulators of the JAKSTAT pathway. Mx, ISG20, OAS and PKR are proteins with antiviral activity (Taken from Haller et al., 2006).

(Thomas et al, 2004). The ability to inhibit both host transcription and the interferon response is linked to the interaction of BUNV NSs protein with the MED8 component of Mediator, a protein complex necessary for host mRNA production. The interacting domain on NSs was mapped to the C-terminal region, which contains 3 amino acids (LPS motif) conserved among orthobunyavirus NSs proteins. A recombinant virus in which the interacting domain in NSs was deleted had strongly reduced ability to inhibit host protein expression and was unable to inhibit the interferon response (Leonard et al., 2006).

1.4.3 Apoptosis of the infected cells

Apoptosis is the process by which abnormally behaving cells are induced to die to cause minimum disruption to neighboring cells. This mechanism is necessary in clearance or removal of infected, cancerous or damaged cells. If there is no clearance, the cell will die necrotically, releasing harmful intracellular components. Successful viral replication requires not only the efficient production and spread of progeny, but also evasion of host defense mechanisms that limit replication by killing infected cells. In addition to inducing immune and inflammatory responses, infection by most viruses triggers apoptosis or programmed cell death of the infected cell (Aigner, 2002).

Some viruses seem to use apoptosis as a mechanism of cell killing and virus spread. The NSs protein of California serogroup viruses has been shown to have some sequence similarity to Reaper, a proapoptotic protein from *Drosophila*. Although NSs protein is lacking in the Reaper N-terminal motif critical for Inhibitor of Apoptosis (IAP) inhibition, they do retain other functions of Reaper at their conserved C-terminal regions. Similar to Reaper, NSs protein of California serogroup viruses also induces mitochondrial cytochrome c release and caspase activation in cell-free extracts and promotes neuronal apoptosis and mortality in a mouse model. Independent of caspase activation, NSs proteins of these viruses also share with Reaper the ability to directly inhibit cellular protein translation (Colon-Ramos et al., 2003)

In mammalian cells, the NSs protein of BUNV has been found to delay cell death in the early stages of BUNV infection by inhibiting IRF-3-mediated apoptosis. BUNdelNSs, which does not express NSs was found to induce apoptosis more rapidly than wt BUNV. Screening for apoptosis pathways revealed that both BUNV and BUNdelNSs could

activate the proapoptotic transcription factor IRF-3, but only BUNV was able to suppress signaling downstream of IRF-3 (Kohl et al., 2003). However in contrast with BUNV NSs protein, the NSs protein of LACV was found to exhibit a strong proapoptotic response (Blakqori and Weber, 2005).

1.4.4 Persistent infection

Infection of mammalian cells by orthobunyaviruses is generally cytolytic while infection of mosquito cells e.g. *Aedes albopictus* C6/36 cells is asymptomatic and persistent (Rossier et al., 1988; Scallan and Elliott, 1992). Persistent infection resulted in generation of viruses with variable genetic and phenotypic characteristics (Elliott and Wilkie, 1986). Reduction of viral RNA and protein synthesized, and the presence of defective RNAs in infected cells were suggested to contribute to maintenance of persistence (Newton et al., 1981).

1.4.5 Defective interfering RNA

Passages at high multiplicity of infection (e.g. moi 10) of BUNV in insect and mammalian cells have been reported to produce defective interfering (DI) particles (Kascsak and Lyons, 1978; Newton et al., 1981). In general, the presence of DI particle has been shown to interfere with the virus growth and contributed to the establishment of persistent infection (Giachetti and Holland, 1989). DI RNAs which have deletions in L were described before in BUNV (Patel and Elliott, 1992). However these deletions were found still in-frame, allowing the translation of truncated L polypeptides to occur, suggesting that the signals necessary for replication, transcription and translation are still retained and this interference only occurred at certain steps in the replication process. The presence of defective L RNA segments and resultant DI particles were also reported in GERV (Cunningham and Szilagyi, 1987).

1.5 Phylogenetic and sequence analyses of Orthobunyavirus genus

Because of the high error rate of RNA polymerase during RNA replication, RNA viruses have been shown continually to evolve rapidly in vivo and in vitro (Holland et al., 1982). For orthobunyaviruses, analyses of oligonucleotide fingerprints of a number of LACV isolates from various ecological niches have showed that every isolate has a different L, M and S genome sequence, suggesting that genetic drift plays a major role in orthobunyavirus evolution (El Said et al., 1979; Klimas et al., 1981).

Although serological techniques have mostly been used to characterise the viruses in *Bunyaviridae* family, there are yet many isolates which cannot be characterised using these methods and remain as unclassified viruses. Based on numbers and genetic diversity of these viruses, it is very difficult to get a panel of monoclonal antibodies (MAbs) that could recognise all epitopes of the viruses (Kingsford, 1991). Therefore other techniques are needed especially for new virus isolates.

Gene sequencing is found to be a rapid and accurate method for characterisation of viruses. Furthermore, there is also a need to understand at the molecular level the basis of genetic diversity and its direct result on antigenic relationship, host susceptibility, pathogenesis, transmission by vector, etc (Kingsford, 1991). In addition, sequencing data will also provide us information on the gene functions of the viral proteins in replication and protein recognition by the immune system (Elliott, 1990).

1.5.1 Sequences and phylogenetic analysis of S segment

Data on the presently available S segment sequences of Bunyamwera, California, Group C and Simbu serogroup viruses clearly show that the patterns of sequence relationships are mostly in agreement to the previous serological classification (Dunn et al., 1994; Bowen et al., 1995; Saeed et al., 2001; Nunes et al., 2005).

Comparison of the sequences of the S segments of 34 viruses in four serogroups of the *Orthobunyavirus* genus (Bunyamwera, California, Group C and Simbu) revealed that the lengths of the segments are between 850 to 1077 nucleotides, with variable lengths of 3'/5' NCR (Dunn et al., 1994; Bowen et al., 1995; Saeed et al., 2001; Nunes et al., 2005). The GU-rich motif in the 3'NCR of the positive-sense RNA is conserved in Bunyamwera, Group C and California serogroup viruses but this motif was not present in Simbu serogroup viruses (Dunn et al., 1994; Bowen et al., 1995; Nunes et al., 2005). The 3' NCR of positive-sense RNA of Bunyamwera, California, Group C and Simbu serogroup viruses are always longer than 5'NCR. The longest 3'NCR was observed in Lumbo virus

(LUMV), which was found to have a duplicated sequence including two GU-rich motifs which are separated by 79 nt (Dunn et al., 1994). Nucleotide alignment of the terminal 3'/5' NCR of positive-sense S RNA sequences of California serogroup viruses revealed that the first 33 nt of 5' terminus and the first 22 nt of the 3' terminus were highly conserved between viruses within this serogroup (Bowen et al., 1995).

The encoded N proteins are 233 amino acid residues for Bunyamwera and Simbu serogroups, except for OROV and Jatobal virus (JATV), which have 231 amino acids, 234 amino acids for Group C and 235 amino acids for California serogroup. The NSs proteins of the orthobunyaviruses are found to be more variable, 91 to 96 amino acids for Simbu serogroup, 83 to 109 amino acids for Bunyamwera serogroup, 92 or 97 amino acids for California serogroup and 91 or 98 for Group C serogroup, with a lower sequence identity and fewer conserved residues between viruses in different serogroups (Dunn et al., 1994; Bowen et al., 1995; Saeed et al., 2001; Nunes et al., 2005).

In general, viruses within the same serogroup exhibit close similarity between each other (Pringle, 1991). N amino acid sequence identity of the viruses in the same serogroup is found to be more than 62% identity, and between serogroups is more than 40% (Dunn et al., 1994; Bowen et al., 1995; Saeed et al., 2001; Nunes et al., 2005), with Group C viruses closer to Simbu serogroup and Bunyamwera serogroup viruses closer to California serogroup. Viruses in Group C and Simbu serogroups were found to display larger divergence in the N sequence (more than 40% divergence) compared to Bunyamwera and California serogroup viruses (Saeed et al., 2001; Nunes et al., 2005). The possible reasons for these divergences might be that these viruses existed earlier than Bunyamwera and California serogroup viruses, and they have broader geographical distribution and vector association compared to Bunyamwera and California serogroup viruses (Saeed et al., 2001).

Alignments of the N and NSs protein sequences of the viruses in Bunyamwera, California and Simbu serogroups revealed that conserved amino acid motifs of the N protein are located at amino acid residues 39 to 43, 90 to 116 and 125 to 165. These regions might be responsible in inducing the CF antibodies that cross react between the viruses in the genus *Orthobunyavirus* and might be functional domains that are responsible in the interaction of N protein with L protein or viral RNA (Dunn et al., 1994; Bowen et al., 1995). Alignments

of the NSs protein of Bunyamwera and California serogroup viruses identified the variation in length to be at the carboxy terminus, and the presence of a conserved amino acid motif LPS near this terminus. However this LPS motif was not observed in GROV, Group C and Simbu serogroup viruses. The NSs protein was found to be more divergent than the N protein (Dunn et al., 1994).

Phylogenetic analysis of Bunyamwera and California serogroup viruses by Bowen et al. (1995) led the authors to conclude that the trend in orthobunyaviruses evolution is toward a smaller NSs protein by truncation at the carboxy terminus. However, GROV is found not to fit this pattern, with its NSs protein being shortened at both amino and carboxy termini, resulting a very short NSs (83 residues) and low identity with other members of the Bunyamwera serogroup. Shortening of NSs has been suggested as an evolutionary advantage to the S segment of the virus by freeing additional ORF codons from the constraint of encoding both N and NSs proteins (Bowen et al., 1995).

Based on phylogenetic analyses of N ORF of viruses in Bunyamwera, California, Group C and Simbu serogroups (Fig. 1.11), each serogroup was divided into several clades. For example, viruses in Bunyamwera serogroup are divided into four clades; BUN, GER, GRO and KRI (Dunn et al., 1994), California serogroup into three clades; MEL, CE and TVT (Bowen et al., 1995), Group C serogroup into three clades; I, II, III with MADV representing a distinct lineage (Nunes et al., 2005) and Simbu serogroup into five clades; I, II, III, IV and V (Saeed et al., 2001).

Viruses in Bunyamwera and California serogroups are found to have 16 nt between the spacing of N and NSs AUG initiation codon, while Group C and Simbu serogroups are found to have 19 nt (Dunn et al., 1994; Bowen et al., 1995; Saeed et al., 2001; Nunes et al., 2005). However, GROV is found not to follow these spacing with 31 nt (Dunn et al., 1994). With the exception of GROV, GERV, AINV and Group C viruses, most of the viruses in California (Bowen *et al.*, 1995), and Bunyamwera (Dunn et al., 1994) serogroups and certain OROV strains in Simbu serogroup (Saeed et al., 2001) start their NSs ORF with two successive ATG codons.



Fig. 1.11. Phylogenetic tree of aligned N ORF of members of the Bunyamwera, California, Group C and Simbu serogroups. The tree was constructed by the Neighbor-Joining (NJ) method. Numbers adjacent to each branch represent bootstrap support calculated for 100 replicates. The viruses in each serogroup were divided into several complexes: Bunyamwera serogroup into BUN, KRI, GRO and MD complexes, California serogroup into CE, MEL and TVT complexes; Group C into 3 groups (I, II and III), with the exception of MADV which was outside the group and Simbu serogroup into 5 groups (I, II, III, V and IV) (Generated using references from Dunn et al., 1994; Bowen et al., 1995; Nunes et al., 2005; Saeed et al., 2001)

1.5.2 Sequences and phylogenetic analysis of M segment

Comparison of the M segment sequences of BUNV (Lees et al., 1986), LACV (Grady et al., 1987) and GERV (Pardigon et al., 1988) with SSHV (Eshita and Bishop, 1984) revealed an overall identity of 60% for Gc, 50% for NSm and 40% for Gn with the conserved sequence of –Lys-Ser-Leu-Arg-Ala/Val-Ala-Arg- at the carboxy terminal of Gc (Elliott 1990). Phylogenetic analyses of M sequences of viruses in the California serogroup (Campbell and Huang, 1999) and some viruses in Simbu serogroup (Wang et al., 2001) were found to be consistent with the previously reported phylogeny based on S RNA sequences (Bowen et al., 1995 and Saeed et al., 2001). However, the phylogenetic tree of AKAV, AINV and PEAV in the Simbu serogroup based on the M ORF sequences revealed that PEAV first forms a cluster with AKAV, and is distant from AINV, while the tree based on N ORF sequences showed closer relationships of Peaton virus (PEAV) with AKAV and AINV. This suggested that RNA segment reassortment had occurred among the ancestors of these viruses (Yanase et al., 2003). Based on complete S and partial M segment sequences of Group C viruses, reassortments were also observed in these viruses (Nunes et al., 2005).

1.5.3 Sequences and phylogenetic analysis of L segment

To date, only limited sequence data are available for the L segments of orthobunyaviruses. Comparisons of these sequences data indicate that generally little similarity was observed between viruses in the different genera. However, phylogenetic analysis of these viruses was found to be correlated with their current serological classification (Elliott, 1996). Comparison of the L protein sequences of BUNV with LACV and SSHV reveal that 46 amino acid residues of the amino-terminal show strong identity (63%) (Elliott,1989). Phylogenetic analyses of the complete L sequences of nine members of four genera revealed that orthobunyaviruses are more related to TSWV in *Tospovirus* genus than the other genera (Roberts et al, 1995). Sequence comparison of L amino acid sequences of OROV, BUNV and LACV showed that BUNV and LACV are more closely related to each other than to OROV.

Sequence data of the viruses from other serogroups in *Orthobunyavirus* genus are needed. Therefore, in this study, I made an attempt to sequence viruses in the other 14 serogroups to give fuller understanding of their phylogenetic relationships and to provide a greater insight of their evolution.

1.6 Aims of the project

The main objective of this project was to sequence the S segments of the viruses in the unstudied 14 serogroups of the *Orthobunyavirus* genus and to compare these sequences between themselves and with the published sequences of other orthobunyaviruses. The S segment was chosen in this study because it displays relatively large sequence divergence (Dunn *et al.*, 1994), and also because of its size, being the smallest (less than 1000 nucleotides) among the three segments. I hope in this study to get a clearer picture of the phylogenetic relationships and evolution of the viruses in all eighteen serogroups of viruses in the *Orthobunyavirus* genus.

The second objective was to better characterize the viruses in terms of their ability to grow in different cell lines, analysis of viral-encoded proteins synthesised in infected cells and investigation of the function of the NSs protein in inhibiting the production of interferon and causing shutoff of host protein synthesis.

2 Materials and Methods

2.1 Materials

2.1.1 Media

2.1.1.1 Cell culture media

Foetal bovine serum (FBS), Dulbecco modified Eagle's medium (DMEM), Glasgow's modified Eagle's medium (GMEM), Liebowitz (L15) medium, and L-glutamine were purchased from Invitrogen. 1x trypsin (0.025%) and versene (0.02% EDTA) were purchased from Sigma.

2.1.1.2 Plaque assay medium

Minimum essential medium 2x (MEM) (Gibco BRL) supplemented with 2% FBS and 2 mM L-glutamine,

2.1.2 Cells

A549 is a cell line derived from human alveolar basal epithelium. This cell was obtained from MRC Virology Unit, University of Glasgow.

293 is a cell line derived from human embryo kidney (Graham et al., 1977), which was obtained from Prof. Rick Randall.

BHK-21 clone 13 (Macpherson and Stoker, 1962), a cell line derived from baby hamster kidney.

C6/36 cells (Igarashi, 1978) were derived from the mosquito Aedes albopictus

Hep2 and Hep2/V cells were obtained from Prof Rick Randall. Hep2 is a human negroid cervical carcinoma cell line and Hep2/SV5-V is a Hep2 cell expressing the V protein of Simian virus 5 (SV5) (Andrejeva et al., 2002).

LLCMK2 (ATCC CCL 7) is a monkey kidney cell line.

Vero-E6 (ATCC no. CRL-1586) is an African green monkey kidney cell line.

BHK-21 cells were maintained in GMEM supplemented with 10% FBS, 2 mM Lglutamine and 10% tryptose phosphate broth (TPB) (Gibco BRL). A549, Vero-E6, 293, Hep2, Hep2/V and LLCMK2 cells were maintained in DMEM containing 10% FBS and 2 mM L-glutamine. C6/36 cells were maintained in L-15 medium supplemented with 10% FBS, 2 mM L-glutamine and 10% TPB.

Except for C6/36 cells which were incubated at 28° C in a tightly sealed container, the other cells were grown at 37° C in the presence of 5% CO₂ in a humidified incubator.

2.1.3 Enzymes

Restriction endonuclease enzymes, T4 RNA and DNA ligases, and calf intestinal phosphatase (CIP) were purchased from Promega. All the reactions using these enzymes were carried out according to the manufacturer's instructions.

2.1.4 Reagents, chemicals and solutions

2.1.4.1 Reagents and chemicals.

All chemicals and reagents as listed below were purchased from BDH Chemical Ltd, Fisher Scientific or Sigma Chemicals except as stated.

DNA loading buffer: 0.25% (w/v) bromophenol blue; 20% (w/v) Ficoll 400; 0.1 M Na₂EDTA, pH8.0; 1% SDS; 0.25% (w/v) xylene cyanol;

Ampicillin, agarose HSA, and DNA size markers were purchased from Research Biolabs.

Ethidium bromide and nuclease-free water were purchased from Promega.

10x Tris-borate EDTA (TBE): 1.0 M Tris; 0.9 M boric acid; 0.01 M EDTA was obtained from Invitrogen.

Multi-purpose agarose was obtained from Roche.

Penicillin/Streptomycin: 10 000 U/ml penicillin and 10 000 µg/ml streptomycin sulphate in 0.85% saline solution was purchased from Gibco BRL.

Phosphate buffered saline (PBS): 170 mM NaCl; 3.4 mM KCl; 10 mM Na₂HPO₄; 1.8 mM KH₂PO₄, pH 7.4; 0.68 mM CaCl₂; 0.49 mM MgCl₂.

lysis buffer (0.15 M NaCl; 0.05 M Tris HCl, pH7.5; 0.6% NP₄₀).

2.1.4.2 Oligonucleotides/ primers

Oligonucleotides were purchased from Sigma-Genosys. Sequences of the oligonucleotides used are listed in Table 2.1.

2.1.4.3 Reverse Transcription (RT) and Polymerase Chain Reaction (PCR)

Access RT-PCR one step system, *Taq* DNA polymerase, 10 mM deoxynucleotide triphosphate (dNTPs), Avian myeloblastosis virus (AMV) reverse transcriptase, and recombinant ribonuclease inhibitor (RNasin) were purchased from Promega.

5'/3' RACE (rapid amplification of cDNA ends) (2nd generation) kit was obtained from Roche.

2.1.4.4 Competent cell preparation, transformation, and plasmid extraction

Ampicillin (Penbritin): 50 mg/ml

L- broth (LB): 10 g NaCl; 10 g bactopeptone; 5 g yeast extract per litre.

Oligonucleotide	Sequences (5'-3')	Reference
BUN S+	AGTAGTGTACTCCAC	Dunn et al.,
		1994
BUN S-	AGTAGTGTGCTCCAC	Dunn et al.,
		1994
S13+	GTCACAGTAGTGTACYCC	This study
S13-	CTGACAGTAGTGTGCYCC	This study
S25F	GTCACAGTAGTGTACTCCACWDNAA	This study
S400F	TCCCAGGAACAGAAATGTT	This study
S400R	AACATTTCTGTTCCTGGGA	This study
MPO400PR	AGAAACATTTCTGCACCWG	This study
S350F	TCATCAACCCAATTGCTGA	This study
ORO2S	AGTAGTGTGCTCCACAT	Saeed et al.,
		2001
BWAPR	CTGTATATGCCAATTGCTAGTGG	This study
BERPF	CTGCTGCAGCCAGAGCCTTCCT	This study
BUNCAL2	CCCCTACCACCCACCC	Bowen et al.,
		2001
CAPPF	ATCGTGAACCCGATTGCTG	This study
CAPPR	AGTGAACGTAGCAGTGCCAA	This study
PGTR	AAAAACATCTCTGTGCCAGG	This study
GMAPF	GTCTGCTGCAGCCAGAAACTTCT	This study
GMAPR	GCTCKRTARATTCCRATTGC	This study
NIFFF	TTAAGGATTTTCTTCCTCAATGC	This study
NNHR	ACTTGTTGTGGTTATT	This study
PATNNHR	CAACCATAACCCAAGACTGG	This study
PATPF	CTCTGTAGATCCCGATGGCAAGTGG	This study
PLSPF	TCATAAACCCGATTGCAGCA	This study
SJPF	CTGCAAGAATGATTGAGTTCTTC	This study
Oligo (dT)	GACCACGCGTATCGATGTCGACT	RACE kit
anchor primer	TTTTTTTTTTTTTTV	(Roche)
(OAP)		
PCR anchor	GACCACGCGTATCGATGTCGAC	RACE kit

primer (AP)		(Roche)
DT88	GAAGAGAAGGTGGAAATGGCGTTTTGG[3d_A]	Li et al.,
		2005
DT89	CCAAAACGCCATTTCCACCTTCTCTTC	Li et al.,
		2005
Exo-oligo	NNnnN	Polidoros et
random hexamer		al., 2006
InvR	TGGACATAAGCCGTAGCATG	Polidoros et
		al., 2006
5' Xba/BUN S+	TCTAGAAGTAGTGTACTCCAC	This study
5'Kpn/BUN S	GGTACCAGTAGTGTGCTCCAC	This study
IFN-βF	GACGCCGCATTGACCATCTA	Spiegel et
		al., 2005
IFN-βR	CCTTAGGATTTCCACTCTGACT	Spiegel et
		al., 2005
human γ-actin F	GCCGGTCGCAATGGAAGAAGA	Spiegel et
		al., 2005
human γ-actin R	CATGGCCGGGGTGTTGAAGGTC	Spiegel et
		al., 2005

Table 2.1. Sequences of the oligonucleotides used. [3d_A] indicates endocypin- blocked adapter primer, D stands for A/G/T, K for G/T, N for A/C/G/T, R for A/G, V for A/C/G, W for A/T and Y for C/T.

LB agar: L-broth plus 1.5% (w/v) agarose.

transformation and storage buffer (TSB): L-broth + 10% PEG (3350 Mwt); 5% DMSO; 10 mM MgCl₂; 10 mM MgSO₄.

TSB glucose: 0.36% (w/v) glucose in TSB.

QIAquick gel extraction kit, Mini and Midi plasmid preparation kits were obtained from Qiagen.

5-bromo-4-chloro-3-indolyl-b-D-galactopyranoside (X-gal) and isopropyl-1-thio-β-D-galactoside (IPTG) were obtained from Invitrogen.

2.1.4.5 Transfection of cultured cells.

Lipofectamine 2000 was purchased from Invitrogen

Luciferase reporter gene assay kit was obtained from Promega.

2.1.4.6 RNA preparation

TRIzol reagent (a mono-phasic solution of phenol and guanidine isothiocyanate) was obtained from Invitrogen.

diethyl pyrocarbonate (DEPC)-treated dH₂O: DEPC was added to double distilled water to 0.01%, mixed and incubated overnight at room temperature before autoclaving.

2.1.4.7 *In vivo* protein labelling, polyacrylamide gel electrophoresis (PAGE) and *in vitro* transcription and translation (TnT).

Acrylamide/bis-acrylamide stock solution: 30% (w/v) acrylamide; 0.8% bis-acrylamide, final ratio 37.5:1.

ammonium persulphate (APS) was obtained from BioRad.

DMEM without methionine solution was obtained from Gibco BRL. gel fixing buffer: 50% (v/v) methanol; 10% (v/v) acetic acid; 40% dH₂O.

sodium dodecylsulphate (SDS): 10% (w/v) in dH₂O.

TEMED (N,N,N',N' tetramethylethylenediamine) was purchased from BioRad.

10x Tris-glycine buffer: 2.4 M Tris; 1.9 M glycine.

Running buffer: 1x Tris-glycine buffer; 10% SDS.

Protein dissociation mix: 125mM Tris, pH 6.8; 4% (w/v) SDS; 20% (v/v) glycerol; 0.1% (w/v) bromophenol blue; 10% (w/v) 2-mercaptoethanol.

Resolving gel buffer (RGB): 59.0 g Tris; 4.0 g SDS in 1000 ml of dH₂O, pH 8.9.

Stacking gel buffer (SGB): 59.0 g Tris; 4.0 g SDS in 1000 ml of dH₂O; pH 6.7

NUPAGE 6-12 % gradient pre-cast polyacrylamide gel was purchased from Invitrogen.

Transcription and Translation (TnT) kit was obtained from Promega.

2.1.5 Plasmids

pGEMT easy vector was purchased together with pGEMT easy cloning kit from Promega.

pRL-SV40 contained the *Renilla* luciferase gene under the control of the constitutive Simian virus 40 (SV-40) promoter and phRL-CMV contained the *Renilla* luciferase gene under the control of the cytomegalovirus (CMV) immediate-early promoter. These plasmids were purchased from Promega.

pBluescript2 KS plus was obtained from Stratagene.

pT7ribo and pT7riboBUNS (Dunn et al., 1995) containing full coding sequence of BUNV S segment and IFN- β reporter plasmid, pIF Δ (-125)lucter (King and Goodbourn, 1994) were supplied by Prof.R.M Elliott.

2.1.6 Bacterial strains

Two bacterial strains were used for growth and maintenance of the plasmids:

E.coli DH5 α : Φ 80d *lacZ*, Δ M15, *rec*A1, *gyr*A96, *thi*-1, *hsd*R17 (r_k-,m_k+), *rel*A1, *supE*44, *rel*A1, *deo*R, Δ (*lacZYA-argF*) U169, *pho*A.

E.coli JM109: *end*A1, *rec*A1, *gyr*A96, *thi*-1, *hsd*R17 (r_k -, m_k +), *rel*A1, *sup*E44, Δ (*lac*-*pro*AB),[F', *tra*D36, *pro*AB,*lac*/^qZ Δ M15].

2.1.7 Viruses

Two viruses of each serogroup were used in this study (Table 2.2). The viruses were obtained from Dr. R.E. Shope, University of Texas Medical Branch, Galveston, Texas. Most of the viruses were in the form of freeze-dried infected mouse brain homogenates.

2.2 METHODS

2.2.1 Maintenance of mammalian and insect cell lines

Mammalian cell lines (A549, BHK-21, Vero-E6, Hep2 and Hep2/V) were maintained in medium-(80 cm²) or large-sized (175 cm²) tissue culture flasks and were split every three to four days. To passage the cells, the monolayers were rinsed twice with versene solution. Then, 1 ml or 2 ml of 1x trypsin solution diluted in versene was added to each medium or large flask, respectively. After incubation at 37°C for 5 min, cells were resuspended in 3 ml growth medium by vigorously pipetting up and down. Additional growth medium was added to the cell suspension to give a final volume of 5 ml for a medium flask and 10 ml
Serogroup	Virus	Virus	Passage cell	Isolation	Date of	Country of	Vector	Isolation
		strain	and no. of	reference	freeze	origin		year
			passage	no.	drying virus			
			(pass)		stock			
Anopheles	Anopheles	-	Mouse Pass	-	23/12/1984	South	Mosquito	1940
А	A (ANAV)		9			America		
						(SA)		
	Tacaiuma	BeAn73	-	TVP-7995	21/7/2001	SA	Mosquito+	1955
	(TCMV)						Culicoides	
Anopheles B	Anopheles B	-	Mouse Pass	-	15/2/1968	SA	Mosquito	1940
	(ANBV)		179					
	Boraceia	SP Ar395	Mouse Pass	-	6/12/1965	SA	Mosquito +	1962
	(BORV)		4				Culicoides	
Bakau	Bakau	MM2322	Mouse Pass	-	9/5/1980	Asia	Mosquito	1956
	(BAKV)	5	8					
			Vero Pass 1					
	Ketapang	MM2549	Mouse Pass	R16179	15/5/1970	Asia	Mosquito	-
	(KETV)		5					
Bwamba	Bwamba	-	Mouse Pass	-	3/80	Africa	Mosquito	-
	(BWAV)		86					
			Vero Pass 2					
	Pangola	-	Mouse Pass	-	17/3/1971	Africa	Mosquito	-
	(PGAV)		9					
Conim	Acoro	BeAn276	Smb Pass 3	-	26/6/1966	SA	Mosquito	1961
Capini	Acara	39	5110 1 455 5		20/0/1900	North	mosquito	1901
	(ACAV)					America		
						(NA)		
	Capim	-	Mouse Pass	-	8/2/1994	SA	Mosquito	1958
	(CAPV)		16					
	()		Vero Pass 1					
Gamboa	Gamboa	GML4385	-	TVP2657	18/12/1990	SA	Mosquito	-
	(GAMV)	24						
	San Juan	78V2441	Vero Pass 4	-	4/12/1981	NA	Mosquito	-
	(SJ2441V)							
Guama	Bertioga	-	Mouse Pass	-	4/1/1966	SA	Mosquito	-
	(BFRV)		4					
	(BERV)							
	Guama	BeAn277	Mouse Pass	-	13/1/1992	SA	Mosquito	-
	(GMAV)		5			NA		
			Vero Pass 1		00/14/10/5	A		40.40
Koongol	Koongol	MRM31	Mouse Pass	-	23/11/1965	Australia	Mosquito	1960
	(KOOV)		1					
	Wongal	MRM168	Mouse Pass	-	6/6/1966	Australia	Mosquito	1960
	(WONV)		9					

Minatitlan	Minatitlan	70U39	Mouse Pass	-	25/4/1974	NA	Mosquito	1967
	(MNTV)		4					
	Palestina	76/V1565	Pass sm 3	-	31/8/1982	SA	Mosquito	1976
	(PLSV)		Vero 2					
Nyando	Eretmapodit	-	Mouse Pass	-	3/10/1966	Africa	Mosquito	-
	es		7					
	(E147V)							
	Nyando	Mp401	Mouse Pass	-	15/3/1971	Africa	Mosquito	-
	(NDV)		7					
Olifantsvlei	Botambi	BA937	Mouse Pass	-	8/3/1970	Africa	Mosquito	1968
	(BOTV)		3					
			N D		20/7/10/1			10/2
	Olifantsvlei	SA 4.:5122	Mouse Pass	-	28/7/1964	Africa	Mosquito	1963
	(OLIV)	AI3133	5					
Patois	Pahayokee	FE3 52F	Mouse Pass	-	14/9/1968	NA	Mosquito	1964
	(PAHV)		4					
	Patois	63A49	Mouse Pass	-	8/3/70	NA	Mosquito	1961
	(PATV)		3					
Tete	Batama	DAK	Mouse Pass	-	15/12/1984	Africa	Tick	1970
	(BMAV)	AnB1292	8					
	Tete	SA	Pass sm 7.	-	12/7/1984	Africa	Tick	1959
	(TETEV)	An3518	Vero 1					
Turlock	M'Poko	BA365	Mouse Pass	-	22/8/1980	Africa	Mosquito	-
	(MPOV)		11					
	Turlock	-	-	MP847-32	9/2/1959	NA,SA	Mosquito	-
	(TURV)							

Table 2.2.Viruses used in this study. The information of passage history was providedby Dr. R.E. Shope. Country of origin, year of isolation and vector of these viruses areretrieved from International Catalogue of Arboviruses, 1985.

for a large flask. One ml of this cell suspension was used to seed a large flask containing 50 ml growth medium and 0.5 ml for a medium flask containing 25 ml growth medium.

The insect cell line C6/36 was maintained in medium size flasks and split once a week. The cells were washed twice with growth medium, then 2 ml of medium was added to the flask. The side of the flask was hit vigorously to dislodge the cells. An additional 3 ml of the medium was added to the flask to give a final volume of 5 ml. 0.5 ml of the cell suspension was seeded in a medium-sized flask containing 25 ml of L-15 growth medium.

2.2.2 Preparation of virus stock

A 70% confluent monolayer of Vero-E6 or BHK-21 cells in a 25 cm² flask was infected with 100 μ l of the original virus stock and incubated at 37°C for 1 h with gentle agitation every 15 minutes. After the absorption, the medium was discarded and 5 ml of DMEM containing 2% FBS was added. The infected cells were further incubated for up to 5 days at 33°C until 70% cytopathic effect (cpe) was observed. The culture supernatants were transferred to 15 ml Falcon tubes and cell debris were then removed by centrifuging at 3,000 rpm for 5 min in a swinging bucket rotor. The supernatant was aliquoted in 1 ml volumes and frozen at -70°C.

If no cpe was observed, the culture supernatants were harvested at 7 days post infection (pi) and were further passaged until a good cpe was obtained.

2.2.3 Plaque Assay

The titre of the virus stocks was determined by plaque formation. Virus-infected supernatant was serially diluted from 10^{-1} to 10^{-6} dilution in PBS containing 2% FBS. A 200 µl volume of each dilution (from dilution 10^{-2} to 10^{-6}) was inoculated to each well of a 6-well tissue culture plate containing confluent Vero-E6 or BHK-21 cells. The last well was used to inoculate 200 µl PBS supplemented with 2% FBS as a control well. The cells were then incubated at 37°C for 1 h with rocking every 15 min. The inoculum was removed and the infected cells were then overlaid with 2 ml equal volume of 1.2% agarose (pre-warmed at 55°C) and 2x MEM containing 2% FBS (pre-warmed at 45°C). The plates

were incubated at 33° C for 4-5 days. After incubation, the plates were fixed with 2 ml of 4% (v/v) formaldehyde and incubated at room temperature for at least 3 hours. The overlay was removed and the cells were then stained with Giemsa's for 5-10 min. The stained cells were then washed with the tap water until the clear plaques were obvious.

2.2.4 Plaque purification of viruses

This assay was conducted as above, except, rather than staining the infected cells with Giemsa, monolayers were stained with 1 ml neutral red (diluted 1/20 in PBS) per dish at 37°C for at least 4 h without prior fixing. The plaques were then picked using 1 ml blue micropipette tip, placed into 1 ml growth medium, and vortexed to release the virus particle from the agarose. Half was used to inoculate 25 cm² flasks of Vero-E6 cells. The cells were then incubated at 37°C and the virus supernatants were harvested after 70% cpe was observed. The harvested supernatants were used as a stock to infect 80 cm² flask of Vero-E6 cells to get a working stock.

2.2.5 Preparation of SDS-PAGE

Ten ml of 15% mini resolving gels were used in this study. The gel was prepared as below; 7.2 ml of (30%) acrylamide: bis solution was added with 3 ml RGB, 1.8 ml water, 100 μ l ammonium persulphate, and 10 μ l TEMED. This resolving gel was poured in-between the assembled glass plate. A space equivalent to two comb-depths was left for the stacking gel and 70% isopropanol was layered onto the resolving gel to create a smooth surface. Once the gel had polymerised, the isopropanol was removed and the gel was rinsed twice with distilled water.

A 6 ml stacking gel was prepared which containing 1 ml 30% acrylamide:bis solution, 1.5 ml SGB, 3.5 ml water, 55 μ l ammonium persulphate, and 10 μ l TEMED. The gel mixture was then poured on-top of the resolving gel, the comb of 0.75 mm was inserted and the gel was allowed for polymerisation.

2.2.6 In vivo protein labelling

A 35 mm-diameter dish of monolayer Vero-E6 cells was infected with 1 PFU/cell of virus and incubated at 37°C for 24 h. The cells were then starved of methionine by incubating them in 1 ml DMEM without methionine for 30 minutes at 37°C. After labelling the cells with 30 μ Ci [³⁵S] methionine for 2 h at 37°C, the medium was removed, washed once with cold PBS and the cells were lysed in 200 μ l protein dissociation mix by scraping them using a rubber policeman and then passing through a 23g needle several times. Fifteen μ l of the cell lysate was used to load each well of the mini gel. The gel was run for 1.5 h at 100 V in 1x Tris-glycine buffer using the Bio-Rad Mini–Protean III apparatus. The gel was then fixed with 20% methanol containing 10% acetic acid for 1 h at room temperature and transferred to distilled water for about 30 min. The gel was placed onto Whatman filter paper, covered with cling film and dried for 50 min at 80°C. The gel was exposed to X-Omat S x-ray film overnight at room temperature.

2.2.7 Preparation of virions for RNA extraction

A 175 cm² flask of Vero-E6 cells were infected with 5 PFU/cell virus, incubated for 24-48 h until 70% cpe was observed. The supernatants were harvested and centrifuged at 3,000 rpm for 5 min to remove cell debris. The virus particles in the supernatant were collected by centrifugation at 26,000 rpm for 2 h in an AH629 rotor at 4° C. The pellet was resuspended in 1 ml TRIzol reagent for RNA extraction.

2.2.8 Preparation of viral nucleocapsid

This method was based on Leppet et al. (1979). A 175 cm² flask of Vero-E6 cells was infected with 5 PFU/cell virus, incubated for 24-48 h until 50% of the cpe was observed. The cells were then scraped into the tissue culture medium and centrifuged at 3,000 rpm for 5 min. The cell pellet was washed with PBS, re-spun for another 5 min and resuspended in 2 ml cold lysis buffer. After incubation on ice for 5 min, the suspension was vortexed for 2 min and centrifuged at 8,000 rpm for 5 min at 4°C in a refrigerated microcentrifuge. The supernatant was transferred to a new tube and 25 μ l 0.5 M EDTA was added to each 2 ml cell extract. CsCl gradients comprised 6 ml of 20-40% (w/w) CsCl in 25 mM Tris-HCl (pH7.5); 2 mM EDTA (TE), overlaid with 2 ml of 5% (w/v) sucrose

in 50 mM NaCl; 25 mM Tris-HCl (pH7.5); 2 mM EDTA (NTE). Gradients were overlaid with the above cell extract and centrifuged for at least 16 h at 32,000 rpm at 12°C. The visible band of the nucleocapsid around the middle of the gradient was harvested with a syringe, diluted at least 3-fold with NTE buffer, and the nucleocapsids were pelleted at 40,000 rpm for 2 h at 4°C in TST rotor. The pellet was dissolved in 1 ml TRIzol and the viral RNA was extracted.

2.2.9 RNA Extraction

Total RNA was extracted from infected Vero-E6 cells or virions or nucleocapsids using TRizol, essentially as described by the manufacturer. Total RNA extracted from infected 35 mm-diameter dishes of Vero-E6 cells at 24 h p.i. was prepared by first removing the culture medium, resuspending the monolayer with 1 ml TRIzol, and transferring to a 1.5 ml Eppendorf tube. For RNA extracted from virions and nucleocapsids, the pellet was resuspended in 1 ml of TRIzol reagent. The TRIzol and the virus mixture was pipetted up and down, and incubated at room temperature for 5 min. After the incubation, 200 μ l of chloroform was added and the tube was inverted a few times, then the mixture was further incubated at room temperature for 3 min. After centrifugation at 13,000 rpm for 15 min, the upper phase was transferred to a new 1.5 ml tube and 500 μ l isopropanol was added. The mixture was then incubated at room temperature for 10 min followed by centrifugation at 13,000 rpm for 10 min. The supernatant was removed and the pellet was washed with 1 ml cold 70% ethanol. After centrifugation at 10,000 rpm for 5 min at 4°C, the pellet was air-dried and resuspended in 30 μ l DEPC-water or nuclease-free water. The extracted RNA was stored at -20°C until use.

2.2.10 RNA ligation

Aliquots of 1-5 μ g of TRIzol extracted RNA were heated at 75°C for 5 min to remove secondary structure, followed by ligation in 25 μ l reaction volume containing 2.5 μ l 10x reaction buffer (500 mM Tris-HCl, pH 7.5; 100 mM MgCl₂; 50 mM DTT; 10 mM ATP), 20 U of RNasin and 20 U of T4 RNA ligase. The reaction mixture was incubated at 37°C for 2 h followed by 65°C for 15 min to inactivate the enzyme. Ten μ l of the ligated RNA was used for cDNA synthesis.

2.2.11 Reverse Transcription-Polymerase Chain Reaction (RT-PCR)

The one step Access RT-PCR system from Promega was used for amplification of full and partial length S segment cDNA, using primers that correspond to the highly conserved termini of the genome RNAs published for Bunyamwera, California, Group C and Simbu serogroup viruses. RT-PCR were carried out in a 50 µl reaction mixture containing 1-5 µg of viral RNA, 20 pmol of BUN S+ forward primer , 20 pmol of BUN S- reverse primer (Dunn et al., 1994), 1x RT-PCR buffer [20 mM Tris-HCl, pH 8.4; 50 mM KCl; 1 mg/ml bovine serum albumin(BSA); 2.5 mM MgCl₂], 10 U of RNasin, 0.2 mM of dNTPs, 2 U of *Thermus flavus (Tfl)* DNA polymerase and 5 U AMV reverse transcriptase. RT-PCR was performed for 60 min at 45°C, followed by 35 cycles consisting of 94°C for 45 s, 54°C for 45 s and 68°C for 1.5 min and a final extension at 72°C for 10 min.

For amplification of partial cDNA, RT-PCR was conducted as above but using primers S400F (this study) and BUNCAL2 (Bowen et al., 2001).

The PCR products were analysed by agarose gel electrophoresis. Ten μ l of PCR product was mixed with 2 μ l of 6x loading dye and electrophoresed on 1% agarose gel containing 0.5 μ g/ml ethidium bromide prepared in 1x TBE buffer. Electrophoresis was at 100 V for 30 min. The product was observed and photographed using a UV transluminator at 365nm.

2.2.12 3'/5' RACE RT-PCR

3'/5' RACE RT-PCR was conducted using the kit from Roche, following the manufacturer's instruction. In the 5' RACE, 1-5 μ g of virion RNA was reverse-transcribed to cDNA in a reaction of 20 μ l containing cDNA synthesis buffer, 0.2 mM dNTPs, 12.5 μ M specific primer 1F, 25 U Transcriptor reverse transcriptase and nuclease-free water. The cDNA product was further purified using the PCR purification kit from the same company and 19 μ l of the cDNA was then subjected to oligo(A) tailing in a 25 μ l reaction containing 2.5 μ l 10x reaction buffer, 2.5 μ l 2 mM dATP. After the denaturation at 94°C for 3 min and quick chilling on ice for 4 min, 10 U terminal transferase was added to the reaction. The reaction was incubated at 37°C for 20 min and stopped by heating at 70°C for 10 min. Five μ l of the tailed cDNA product was used in a first PCR reaction in 50 μ l reaction containing 1x PCR reaction buffer, 2.5 U *Taq* DNA polymerase, 12.5 μ m second

specific primer 2F and 37.5 μ m oligo (dT) anchored primer (OAP), 0.2 mM dNTPs, 1.5 mM MgCl₂, and nuclease-free water. The PCR was performed for 2 min at 94°C, followed by 35 cycles of 94°C for 15 s, 55°C for 30 s and 72°C for 40 s, followed by incubation at 72°C for 7 min.

If no band was observed, one μ l of 10-fold diluted first PCR product was then subjected to the nested PCR reaction using PCR anchor primer (AP) and third specific primer 3F. For the 3' RACE, the extracted RNA was firstly tailed with A residues using poly (A) RNA polymerase (Epicentre) and reverse transcription was conducted using the OAP primer at 50°C for 1 h. The cDNA was then treated with 2 U of RNase H at 37°C for 30 min and further purified using Qiagen PCR purification kit. The first PCR was conducted on 5 μ l of the purified cDNA using OAP primer and specific primer 1R. If the first PCR was not observed, 1 μ l of the first PCR was further subjected to nested PCR using AP primer and specific primer 2R. The products were examined in 1% agarose gel electrophoresis and purified by QIAquick gel purification kit (Qiagen). The products were either sequenced directly, or, if the band was weak, it was cloned into pGEMT vector.

2.2.13 Modified 3'/5' RACE

This method was conducted as described by Li et al. (2005). For 5' RACE, 1 µg RNA extracted from nucleocapsid was incubated at 65°C for 5 min, followed by RT reaction to synthesize cDNA. The reaction was conducted in 20 µl volume containing 20 pmol virus specific (VSP) forward-1, 5 U AMV reverse transcriptase enzyme, 1x RT buffer (50 mM Tris-HCl; 75 mM KCl; 10 mM MgCl₂), 20 U of RNasin and 0.2 mM of dNTPs (Promega). The reaction mixture was incubated at 42°C for 60 min, followed by inactivation at 85°C for 5 min and treatment with 2 U of RNaseH (Invitrogen) at 37°C for 30 min. After purification using the QIAquick PCR purification kit (Qiagen), the cDNA was ligated with a 3' end cordecypin-blocked adapter DT88 using T4 RNA ligase (Promega). The resulting adapter-ligated cDNA was then amplified using primers DT89, which is complementary to primer DT88, and VSP forward-2 primer.

The cDNA amplification was conducted in a 50 μ l reaction containing PCR buffer (10 mM Tris-HCl; 50 mM KCl; 0.1% Triton X-100), 1.5 mM MgCl₂, 2 U *Taq* DNA polymerase, 0.2 mM dNTPs, 20 pmol of each primer and 1 μ l of a 1:10 diluted cDNA mix. The PCR

was performed for 35 cycles of 94°C for 20 s, 55°C for 20 s and 72°C for 30 s, and a final 7 min extension at 72°C. Five μ l of the PCR products were electrophoresed on a 1% TBE agarose gel. If no band was observed, a secondary PCR was carried out using DT89 and VSP forward-3 primer, using essentially similar conditions as the primary PCR, except replacing the cDNA template with 1 μ l of a 1:100 diluted primary PCR product.

For 3' RACE, adapter ligation was carried out first before cDNA synthesis as follows:

1 μ g RNA, 20 pmol of DT88, 2 μ l of 10x RNA ligase buffer, 20 U of T4 RNA ligase, and nuclease-free water to a final volume of 20 μ l. The reaction mixtures were mixed and incubated at 37°C for 1 h. A 2 μ l aliquot was taken out to make cDNA using 20 pmol of primer DT89. RT, PCR and RNase treatment were performed as for 5' RACE above except primer DT89 was used with VSP reverse-1 in primary PCR and VSP reverse-2 in secondary PCR.

2.2.14 Rolling circle amplification-RACE (RCA-RACE)

The reaction was conducted as described by Polidoros et al. (2006). Total RNA was extracted from nucleocapsids. First strand cDNA was synthesized using 3 μ g RNA in a reaction containing 20 pmol (VSPF1) for genomic-sense RNA and VSPR1 for antigenomic-sense RNA, 0.2 mM dNTPs, 10 mM DTT, 1x RT buffer, and 200 U Moloney murine leukaemia virus (MMLV) reverse transcriptase (Promega). The reaction was incubated at 42°C for 1 h, followed by enzyme inactivation at 70°C for 15 min. After the RNaseH treatment, the reaction was purified using QIAquick PCR purification kit (Qiagen).

Half of the purified cDNA was then circularised using 100 U CircLigase, 1x reaction buffer (Epicentre Biotechnologies), and 50 μ M ATP at 60°C for 1 h, followed by inactivation of the enzyme at 80°C for 10 min and purification using the QIAquick PCR purification kit (Qiagen).

RCA reactions were performed in 50 μ l volume containing a 10 μ l aliquot of the circularised cDNA, 1 mM dNTPs, 200 μ g/ml BSA, 1x Φ 29 DNA polymerase reaction buffer (New England Biolabs), 10 U Φ 29 DNA polymerase, and 10 μ M random hexamers modified by the addition of phosphothioate linkages on the 3' end to make primer resistant

to the $\Phi 29$ exonuclease activity. Control reaction containing all reagents except $\Phi 29$ DNA polymerase was also performed. The reaction was incubated at 30°C for 21 h followed by heat inactivation at 60°C for 10 min.

 $0.5 \ \mu$ l of neat and 10^{-1} dilution of RCA mixture were then used as a template in inverse PCR amplification. The PCR was performed as above but using primers VSPF2 for genomic RNA or VSPR2 for antigenomic RNA, and InvR. The cycling parameters for PCR were 94°C for 3 min, followed by 35 cycles of 94°C for 30 s, 56°C for 45 s and 72°C for 1.5 min, and a final extension step at 72°C for 10 min.

2.2.15 Cloning

2.2.15.1 Ligation

The PCR product was eluted and purified using QIAquick gel extraction kit following the procedures recommended by the manufacturer (Qiagen). The DNA was cloned into pGEMT easy vector. Ligation was performed in 12 μ l reaction mixtures containing 4 μ l of fresh PCR product (less than 2 days), 6 μ l 2x ligation buffer, 1 μ l of vector DNA and 1 μ l (2 U) of T4 DNA ligase. The ligation mixture was incubated at 4°C overnight or 37°C for 1 h.

2.2.15.2 Preparation of competent *E. coli* cells

One colony of *E.coli* JM109 or DH5 α was inoculated into 10 ml LB broth without ampicillin and incubated overnight at 37°C with shaking. One ml of the culture broth was then subcultured into 100 ml LB broth and incubated with shaking for 2 h at 37°C. The culture was centrifuged at 3,000 rpm for 5 min and the bacterial pellet was resuspended in 2 ml TSB with 5% DMSO and was stored at 4°C for up to 2 days.

2.2.15.3 Transformation

The ligation mixtures were briefly spun down and placed on ice. Five μ l of ligation mixture were added to 200 μ l of competent cells and gently mixed by flicking the side of the tube. The tube was incubated on ice for 30 min and 800 μ l of the TSB glucose with 5%

DMSO were added to the tube. The mixture was then incubated at 37°C for 1 h and centrifuged for 1 min at 6,000 rpm. Eight hundred μ l of the medium were removed and the remaining 200 μ l were used to resuspend the bacteria pellet. The bacterial suspension was then spread onto LB agar plates containing ampicillin (50 μ g/ml), IPTG (0.1 mM) and X-gal (20 μ g/ml), and incubated at 37°C overnight. White colonies most probably contain the inserts were selected for plasmid preparation.

2.2.15.4 Plasmid preparation

Small scale plasmid preparation was carried out using Qiagen spin miniprep kit according to the manufacturer's protocol. The plasmid DNA was extracted from 3 ml of bacterial culture grown overnight in LB broth containing 50 μ g/ml ampicillin.

2.2.15.5 Restriction endonuclease digestion of plasmid DNA

Three μ l of the small scale plasmid preparation were digested with the restriction enzyme in a 10 μ l reaction containing 1 μ l 10x reaction buffer, 5-10 U restriction enzyme and distilled water to a final volume of 10 μ l. The reaction was incubated at the appropriate temperature for 1-2 h. Two μ l of loading dye was added to the digestion mixture and electrophoresed on 1% agarose gel containing ethidium bromide. The gel was then visualised on a UV transluminator.

2.2.16 Sequences analysis and alignments

Automated DNA sequence determination was performed by University of Dundee. Three independent plasmids containing the expected insert were sequenced in both directions using M13F and M13R primers.

The sequences were analysed using Bioedit software and Sequence Navigator of the ABI system, and aligned using the ClustalX (NCBI) and ClustalW program (EMBL-EBL). Phylogenetic analyses were generated by Neighbor-joining (NJ) (Saitou and Nei, 1987) and Maximum Parsimony (MP) analysis methods, implemented with the Mega 2.1 software packages (Kumar et al., 2000).

Accession numbers of the S segment sequences used are as follows:

Bunyamwera serogroup viruses; Maguari (P16606), Cache Valley (CAA51845), Main Drain (CAA51853), Kairi (CAA51849), Batai (CAA51843), Guaroa (CAA51847), Bunyamwera (AAL37356).

California serogroup viruses; California encephalitis(AAC54051), Jameston Canyon (AAC54049), Keystone (AAC54053), La Crosse (P04873), Tahyna (CAA92805), Trivittatus (U12803), Melao (U12802), Snowshoe hare (J02390), Lumbo (X73468). Group C serogroup viruses; Caraparu BeAn3994 (ABA54067), Itaqui BeAn12797 (ABA54073),Vinces 75V-807(ABA54087), Murutucu BeAn 974(ABA54065), Restan TRVL 51144(ABA54083), Marituba BeAn 15 (ABA 54069), Nepuyo TRVL 18462 (ABA 54071), Gumbo Limbo FE-371H (ABA 54081), Madrid BT 4075 (ABA 54085).

Simbu serogroup viruses; Oropouche (AAU29359), Aino (P12414), Akabane (BAE98058), Tinaroo (AB000819), Yaba-7 (AF362392), Douglas (AF362393), Sabo(AF362396), Buttonwillow (AF362398), Mermet (AF362399), Peaton (AF362401), Sathuperi (AF362403), Shuni (AF362405), Facey's Paddock (AF362400), Ingwavuma (AF362395).

Accession numbers of the M segment sequences used are as follows: Bunyamwera virus (NC001926), La Crosse virus (U18979) and Oropouche virus (NC005775)

Accession numbers of the L segment sequences used are as follows:

Bunyamwera virus (NC001925), La Crosse virus (AF525489) and Oropouche virus (AF484424)

2.2.17 In vitro Transcription and Translation (TnT) assay

2.2.17.1 Transcription of T7 transcription plasmid

The full length S segment cDNAs from the pGEMT easy vector were amplified using forward primer 5' Xba/BUN S+ and reverse primer 5'Kpn/BUN S-. The PCR products were then digested with Kpn I and Xba I, excised from the gel and ligated to the

pBluescript2 SK+ vector (previously cut with Xba I and Kpn I, and treated with CIP to remove the 5' phosphate from the vector). The ligation mixture was transformed into *E.coli* JM109. The plasmid was extracted using the miniprep kit column and the insert was checked by restriction enzymes, followed by checking the orientation by sequencing. Plasmids containing inserts in the correct orientation were then linearised by digestion with Kpn I and were used in the TnT reaction.

2.2.17.2 Transcription and Translation reaction.

The TnT system was used to express proteins from plasmid *in vitro*. Reactions were performed according to the manufacturer's condition. A 25 μ l reaction volume was used which containing 12.5 μ l TnT rabbit reticulocyte lysate or wheat germ lysate; 1 μ l TnT reaction buffer; 0.5 μ l (10 U) T7 RNA polymerase; 0.5 μ l (1 μ M) amino acid mixture, minus methionine; 0.5 μ l (10 U) RNAsin; 1 μ l (10 μ Ci) [³⁵S] Methionine; 1 μ g plasmid with insert, and top up to final volume with nuclease-free water.

The reaction was incubated at 30°C for 90 min and 13 μ l of the reaction were analysed by electrophoresis in NUPAGE 6-12% gradient Pre-cast polyacrylamide gel.

2.2.18 Interferon assay

2.2.18.1 IFN reporter gene assays

2.2.18.1.1 Dual Renilla luciferase assay system

1 µg of the IFN- β reporter plasmid pIF Δ (-125) lucter and 0.1 µg/µl of control plasmid pRL-SV 40 or phRL-CMV diluted with 200 µl OptiMEM were added slowly to 200 µl OptiMEM containing 3 µl Lipofectamine 2000 preincubated at room temperature for 5 min. The mixture was then incubated at room temperature for 20 min and 400 µl of the mixtures were added slowly to a confluent monolayer A549 cells. The cells were then incubated for at least 4 h at 37°C with gentle shaking for every 30 min. The OptiMEM was removed and the cells were infected with 1 PFU/cell virus. After incubating for 1 h at 37°C, 2 ml of growth medium was added to the cells and further incubated for 18 h at

37°C. Luciferase assays were performed using the dual *Renilla* Luciferase Assay System purchased from Promega. Briefly, the cells in 35-mm diameter dishes were lysed by removing the culture medium, washing the monolayer once with PBS, adding 200 μ l 1x Passive lysis buffer (PLB) and incubating on an orbital shaker for 15 min at room temperature. The lysate was collected into 1.5 ml tube and cell debris was pelleted by brief centrifugation. 20 μ l of the cell lysate was added to 100 μ l of Luciferase Assay Reagent (LAR) 11 solution and the luminosity of the firefly and *Renilla* luciferase activities in the sample was read using a luminometer. One hundred μ l of Stop & Glo reagent was added in between the two readings. If the reading of luciferase activity was too high (>9999), the lysates were further diluted with 1x PLB. The firefly luciferase activities were normalised to the corresponding *Renilla* luciferase activities to determine the level of IFN induction.

2.2.18.1.2 IFN-specific RT-PCR

Confluent 293 cells were infected with 1 PFU/cell virus for 18 h and total infected cell RNA was extracted using TRIzol reagent. A two step RT-PCR was used in this experiment. Prior to the RT reaction, 1 μ g of RNA was first treated with 2 U DNase I (Promega) by incubating at 37°C for 30 min, followed by enzyme inactivation at 65°C for 10 min. The RNA was then reverse-transcribed to cDNA in a 20 μ l reaction containing 100 ng of random hexanucleotides (Amersham), 5 U AMV reverse transcriptase enzyme, RT buffer (50 mM Tris-HCl; 75 mM KCl; 10 mM MgCl₂), 20 U RNasin and 0.2 mM of dNTPs. The mixture was incubated at 42°C for 1 h, followed by 65°C for 10 min.

The resulting cDNA was amplified in 50 μ l reaction volume containing PCR buffer (10 mM Tris-HCl; 50 mM KCl; 0.1% Triton X-100), 1.5 mM MgCl₂, 2 U *Taq* DNA polymerase, 0.2 mM dNTPs and 20 pmol forward and reverse primers. The primers used to amplify IFN- β mRNA were IFN- β F and IFN- β R primers, and human γ -actin F and human γ -actin R for human γ -actin mRNA (Spiegel et al., 2005). The PCR was performed for 35 cycles consisting of 94°C for 45 s, 58°C for 45 s and 72° C for 1min. Five μ l of the PCR products were electrophoresed on a 1% TBE agarose gel.

2.2.19 Shutoff of host cell protein synthesis in virus-infected cells

2.2.19.1 *In vivo* protein labeling at different time points

This experiment was conducted as described previously (Section 2.2.6), except that Veroinfected cells were labelled with 30 μ Ci [³⁵S] methionine at 12 and 24 h, and 36 h pi for cell-infected with ANAV, TCMV, ANBV, BORV, TETEV and BMAV. After labelling the cells for 2 h at 37°C, the medium were removed, washed once with cold PBS and 200 μ l protein dissociation mix was added. 15 μ l of the cell lysate were analysed by SDS-PAGE.

For ANAV, TCMV, ANBV, BORV, TETEV and BMAV infected cells, the cell lysate was analysed by electrophoresis in NUPAGE 6-12% gradient pre-cast polyacrylamide gel.

2.2.19.2 Luciferase expressing plasmid based assay

0.1 μ g control plasmid pRL-SV 40 or phRL-CMV was transfected into subconfluent BHK-21 or C6/36 cells respectively, using 3 μ l of Lipofectamine 2000 diluted in 400 μ l of OptiMEM. The cells were incubated for at least 5 h at 37°C with gentle shaking every 30 min. The OptiMEM was removed and the cells were infected with 1 PFU/cell virus, incubated for 1 h at 37°C, and then 2 ml of growth medium was added. The cells were incubated for another 18 h at 37°C. The cells were then washed once with PBS, 200 μ l 1x PLB was added to the cells, and further incubated for 15 min at room temperature with shaking. Twenty μ l of the lysate was added to 100 μ l of LAR 11 solution and the luminosity of the firefly and *Renilla* luciferase activities of the sample was then taken using the luminometer.

3 Characterization of Viral Growth and Protein Profiles

3.1 Introduction

Viruses in family *Bunyaviridae* are divided into genus and serogroup based on their serological relationships, and morphological and biochemical characteristics (Elliott, 1990). Biochemical analyses of representatives of 10 serogroups viruses of genus *Orthobunyavirus*: Anopheles A, Anopheles B, Bunyamwera, California, Capim, Group C, Guama, Patois, Simbu and Turlock serogroups revealed that the molecular weight of their proteins are estimated to be 259 kDa for L, 108-120 kDa for Gc, 29-41 kDa for Gn and 15-18 kDa for NSm, 19-25 kDa for N and 10-13 kDa for NSs (Gentsch et al., 1977; White, 1975; Klimas et al., 1981; McPhee and Westaway, 1981; Ushijima et al., 1980; Ushijima et al., 1981; Short et al., 1982; Elliott, 1985; McPhee and Della-Porta, 1988). To date, no biochemical data on Bakau, Bwamba, Koongol, Gamboa, Minatitlan, Nyando, Olifantsvlei and Tete serogroup viruses are available.

Shutoff of host cell RNA and protein synthesis during the infection of orthobunyaviruses in the mammalian cells have been reported before (Pennington et al., 1977; Lazdins and Holmes, 1979; McPhee and Westaway, 1981; Short et al., 1982; McPhee and Della-Porta, 1988; Frugulhetti and Rebello, 1989; Bridgen et al., 2001; Blakqori et al., 2005). The NSs protein of BUNV and LACV were shown to be responsible in causing this shutoff (Bridgen et al., 2001; Blakqori et al., 2005).

This study was undertaken with the aims to grow the viruses in different cell culture systems (Vero-E6, BHK-21 and LLCMK2), to determine the pattern of viral proteins synthesized in infected cells, and to investigate the ability of the virus to cause shutoff of host protein synthesis in infected cells.

3.2 Results

3.2.1 Virus propagation

Most of the viruses received were as freeze dried, infected mouse brain homogenates (Table 2.2). Therefore the viruses had first to be reconstituted and then attempts were made to grow them in different cell lines until a good cpe was observed. For some of the viruses (BAKV, KETV, ACAV, PGAV, SJ2441V, KOOV, WONV, MNTV, PLSV, OLIV, PAHV, PATV, MPOV and TURV), longer adaptation to the cell culture systems (more than 5 passages) were needed. These could have been due to the low viability of the frozen stock of the viruses. Once they were adapted, growth occurred more rapidly with earlier appearance of cpe. Using the cell culture systems (Vero-E6, BHK-21 and LLCMK2), I managed to grow 27 of the 28 viruses received. I was unable to grow BOTV from the Olifantsvlei serogroup. An attempt to grow this virus in mosquito Aedes albopictus (C6/36) cells was also unsuccessful. There were also difficulties in obtaining high virus titers (more than 10⁶ PFU/ml) for some viruses (ANAV, ANBV, BORV, TETEV, BMAV, SJ2441V, PGAV and KETV). Therefore I decided to use low moi of virus (1 PFU/cell) for all my experiments as it was not possible to do high moi (5 PFU/cell) with these stocks to allow synchronous infection of all cells. The results of the virus propagation studies are summarised in Table 3.1. Some of the viruses were plaque purified in Vero-E6 cells. The titers of the virus stocks obtained using Vero-E6 cells are also recorded.

3.2.2 Plaque morphologies produced by the viruses

This assay was performed to examine the size of plaque produced by the different viruses in Vero-E6 cells. All the viruses used in this study produced cpe in this cell line (Table 3.1), thus, plaque production in these cells was also expected. In general, viruses in the same serogroup were found to produce almost similar sizes of plaques. Although some of the viruses were plaque purified, mixed sizes of plaques were still observed in Vero-E6 cells. With the exception of TCMV, most of the viruses isolated from human (BWAV, PGAV, GMAV and NDV) and primate (BAKV) were found to produce clear plaques that were either medium or large in size (Figure 3.1 and Table 3.2). However some viruses such as KOOV, WONV, BERV and KETV, which were isolated from mosquitoes, also produced large sized plaques. The plaque assay results are summarised in Table 3.2.

SEROGROUP	VIRUS		VIRUS TITER IN		
		Vero-E6	BHK-21	LLCMK2	VERO CELLS PFU/ml
1. Anopheles A	Ana A (ANAV)	cpe (P3)	cpe (P3)	cpe (P3)	1.0 x 10 ⁶
	Tacauima (TCMV)	cpe (P3)	cpe (P3)	cpe (P3)	2.0 x 10 ⁶
2.Anopheles B	Ana B (ANBV)	cpe (P3)	cpe (P3)	cpe (P3)	1.0 x 10 ⁶
	Boraceia (BORV)	cpe (P3)	cpe (P3)	cpe (P3)	1.0 x 10 ⁶
3.Bakau	Bakau (BAKV)	cpe (P5)	cpe (P5)	cpe (P5)	4.5 x 10 ⁶
	Ketapang (KETV)	cpe (P5)	cpe (P5)	cpe (P5)	4.0 x 10 ⁵
4.Bwamba	Bwamba (BWAV)	cpe (P3)	cpe (P3)	cpe (P3)	1.0×10^7
	Pangola (PGAV)	cpe (P7)	cpe (P7)	cpe (P7)	1.0 x 10 ⁵
5.Capim	Acara (ACAV)	cpe (P5)	cpe (P5)	cpe (P5)	5.0 x 10 ⁶
	Capim (CAPV)	cpe (P4)	cpe (P4)	cpe (P4)	3.5 x 10 ⁶
6.Gamboa	Gamboa (GAMV)	cpe (P3)	cpe (P3)	cpe (P3)	3.0 x 10 ⁶
	San Juan (SJ2441V)	cpe (P5)	cpe (P5)	cpe (P5)	1.0 x 10 ⁵
7.Guama	Bertioga (BERV)	cpe (P3)	cpe (P3)	cpe (P3)	1.0×10^7
	Guama (GMAV)	cpe (P3)	cpe (P3)	cpe (P3)	6.5×10^7
8.Koongol	Koongol (KOOV)	cpe (P5)	cpe (P5)	cpe (P5)	4.0×10^7
	Wongol (WONV)	cpe (P5)	cpe (P5)	cpe (P5)	6.0 x 10 ⁶
9. Minatitlan	Minatitlan (MNTV)	cpe (P5)	cpe (P5)	cpe (P5)	2.0×10^6
	Palestina (PLSV)	cpe (P5)	cpe (P5)	cpe (P5)	1.5×10^7
10.Nyando	Eretmapodites (E147V)	cpe (P3)	cpe (P3)	cpe (P3)	5.0 x 10 ⁶
	Nyando (NDVV)	cpe (P3)	cpe (P3)	cpe (P3)	2.0×10^7
11. Olifantsvlei	Botambi (BOTV)	No cpe (>P10)	No cpe (>P10)	No cpe (>P10)	ND
	Olifantsvlei (OLIV)	cpe (P5)	cpe (P5)	cpe (P5)	7.0 x 10 ⁷
12.Patois	Pahayokee (PAHV)	cpe (P5)	cpe (P5)	cpe (P5)	6.5 x 10 ⁶
	Patois (PATV)	cpe (P5)	cpe (P5)	cpe (P5)	6.5 x 10 ⁶
13.Tete	Batama (BMAV)	cpe (P3)	cpe (P3)	cpe (P3)	1.0 x 10 ⁶
	Tete (TETEV)	cpe (P3)	Cpe (P3)	cpe (P3)	2.0 x 10 ⁶
14.Turlock	M'Poko (MPOV)	cpe (P5)	cpe (P5)	cpe (P5)	2.0 x 10 ⁷
	Turlock (TURV)	cpe (P5)	cpe (P5)	cpe (P5)	$1.0 \ge 10^7$

Table 3.1 Propagation of the virus in Vero-E6, BHK-21 and LLCMK2 cells. P indicates passage level; cpe indicates cytopathic effect; ND indicates not done and > indicates more than.





BUNV

BUNdelNSs







Fig.3.1 Plaques formed by the viruses in infected Vero cells after 6 days pi. The plaques were fixed for more than 3 h with 4% formaldehyde solution and stained with Giemsa solution for 5-10 min.





ANBV

BORV





TETEV

BMAV





CAPV







BWAV

PGAV





GMAV







KOOV

WONV





TURV







BAKV

KETV







PLSV









SJ2441V





E147V







PAHV

PATV





OLIV

MOCK

VIRUSES	SEROGROUP	PLAQUE SIZE	PLAQUE TYPE	HOST ASSOCIATION	PLAQUE CLONED IN VERO CELLS
1. Anopheles A	Ana A	Small	Turbid	NR	Yes
	(ANAV) Tacauima	Small	Turbid	Human,	Yes
	(TCMV)			Primate	
2.Anopheles B	Ana B (ANBV)	Small	Turbid	NR	Yes
	Boraceia (BORV)	Small	Turbid	NR	Yes
3.Bakau	Bakau (BAKV)	Medium and Large	Clear	Primate	No
	Ketapang (KETV)	Medium	Clear	NR	No
4.Bwamba	Bwamba (BWAV)	Small, Medium and Large	Clear	Human	Yes
	Pangola (PGAV)	Large	Clear	Human	Yes
5.Capim	Acara (ACAV)	Medium	Clear	Rodent	No
	Capim (CAPV)	Medium and large	Turbid	Rodent	No
6.Gamboa	Gamboa (GAMV)	Large	Turbid	NR	Yes
	San Juan (SJ2441V)	Medium and Large	Turbid	NR	Yes
7.Guama	Bertioga (BERV)	Small, Medium and Large	Clear	NR	No
	Guama (GMAV)	Small, medium and Large	Clear	Human	No
8.Koongol	Koongol (KOOV)	Large	Clear	NR	No
	Wongol (WONV)	Large	Clear	NR	No
9. Minatitlan	Minatitlan (MNT)	Small and medium	Turbid	NR	No
	Palestina (PLSV)	Small and Medium	Turbid	NR	No
10.Nyando	Eretmapodites (E147V)	Medium	Turbid	NR	Yes
	Nyando (NDVV)	Medium and Large	Clear	Human	Yes
11.Olifantsvlei	Botambi (BOTV)	ND	ND	NR	ND
	Olifantsvlei (OLIV)	Medium	Turbid	NR	No
12.Patois	Pahayokee (PAHV)	Medium	Turbid	Rodent	Yes
	Patois (PATV)	Medium	Turbid	Rodent	Yes
13.Tete	Batama (BMAV)	Medium	Turbid	Bird	Yes
	Tete (TETE)	Medium	Turbid	Bird	Yes
14. Turlock	M'Poko	Medium	Clear	NR	Yes
	(MPOV) Turlock	Medium	Clear	Bird	Yes
	(TURV)				

Table 3.2 Summary of the plaques produced by the viruses in infected Vero cells. ND represents not done; NR represents not reported. Plaque sizes; big (3-5 mm), medium (2-4 mm) and small (<2 mm).

3.2.3 In vivo protein labeling in Vero-E6 cells

To analyse the proteins synthesized by the viruses, Vero-E6 cells were infected with 1 PFU/cell of each virus, and labeled with [³⁵S] methionine at 12, 24 and up to 36 h pi for ANAV, TCMV, ANBV, BORV, TETEV and BMAV. The cell lysates were analysed by SDS PAGE. Analysis of the gels (Figure 3.2) revealed that all the viruses had a typical orthobunyavirus protein profile, with L, Gc and N easily identified; the Gn, NSm and NSs protein were only clearly observed for some of the viruses. However in cells infected with BMAV, TETEV, PLSV, MNTV and PGAV, L and Gc proteins were not clearly observed, where lack of shutoff prevented identification of these proteins. The N protein was the most abundant protein observed in virus infected cells. Different sizes of N protein were observed, for instance, E147V, PGAV, OLIV, KOOV, PATV, PAHV, TURV, KETV, GMAV and NDV have N protein similar in size to BUNV (22 kDa), ANAV, TCMV, ANBV, BORV, TETEV, BMAV, PLSV, MNTV, ACAV, CAPV, BWAV, MPOV, BAKV, BERV, GAMV and SJ2441V were found to have N protein larger than that of BUNV, with the largest N protein observed in TETEV and BMAV (26 kDa), while WONV has the smallest N protein among the studied viruses (20 kDa) (Table 3.3). Not much variation in the size of L protein was seen, Gc protein showed more variation in size with the smallest in KOOV, GAMV, SJ2441-infected cells (<110 kDa). The presence of different sizes of NSm and NSs were also observed in some of the infected cells profiles. The suspected NSs protein bands were only observed in cells infected with ACAV, BERV, E147V, WONV, PATV, PAHV, SJ2441V, TURV, NDV, KETV, GAMV, GMAV and BUNV with the largest NSs protein observed in GAMV and SJ2441V-infected cells (14 kDa). A strange protein profile was observed in cells infected with ACAV, this could be due to the presence of mixed viruses in this extract since it was not from the plaque purified culture. A number of protein bands were also observed in some of the virus infected extracts but not in mock-infected cells. It is unclear whether they are virus or hostcoded proteins.

3.2.4 Shutoff of host protein synthesis

Previously, BUNV was shown capable of producing shutoff of host protein synthesis in infected Vero cells, but not with its mutant virus that lacks the NSs protein, BUNdelNSs (Bridgen et al., 2001). Inspection of the gels in Figure 3.2 show that after 24 h pi, shutoff



Fig. 3.2 Viral protein synthesis and shutoff of host protein synthesis of the viruses in infected Vero-E6 cells. The cells were infected with 1 PFU/cell virus, labelled at 12, 24 and up to 36 h pi for ANAV, TCMV, ANBV, BORV, TETEV and BMAV using 30 μ Ci [³⁵S] methionine and analysed by (A) NUPAGE 6-12 % gradient gel (Invitrogen) and (B) 15 % polyacrylamide gel electrophoresis. Blue arrows indicate unknown protein bands that were not observed in mock-infected cells. Position of molecular weight standards are indicated at the right.











VIRUSES	SEROGROUP	PROTEIN SIZE						LEVEL OF SHUTOFF
		L kDa	Gc kDa	Gn kDa	N kDa	NSm kDa	NSs kDa	-
1. Anopheles A	Ana A (ANAV)	259	110	32	24	NO	NO	++
	Tacauima (TCMV)	>259	110	32	23	NO	NO	++
2.Anopheles B	Ana B (ANBV)	>259	110	32	24	NO	NO	++
	Boraceia (BORV)	259	110	32	24	18	NO	++
3.Bakau	Bakau (BAKV)	259	110	32	23	18	NO	+++
	Ketapang (KETV)	259	110	32	22	18	11	-
4.Bwamba	Bwamba (BWAV)	259	110	NO	23	18	NO	+++
	Pangola (PGAV)	NO	NO	NO	22	18	NO	-
5.Capim	Acara (ACAV)	259	110	28	23	17	12	++
	Capim (CAPV)	259	110	NO	23	18	NO	+++
6.Gamboa	Gamboa (GAMV)	259	<110	32	23	18	14	++
	San Juan (SI2441V)	259	<110	NO	23	18	14	++
7.Guama	Bertioga (BERV)	259	110	32	23	18	11	++
	Guama (GMAV)	259	110	32	22	18	11	++
8.Koongol	Koongol (KOOV)	259	<110	NO	22	18	NO	++
	Wongol (WONV)	259	110	30	20	17	13	+++
9. Minatitlan	Minatitlan (MNT)	NO	NO	NO	23	17	NO	-
	Palestina (PLSV)	NO	NO	NO	23	18	NO	-
10.Nyando	Eretmapodites (E147V)	259	110	NO	22	18	11	++
	Nyando (NDVV)	259	110	32	22	18	11	++
11.Olifantsvlei	Olifantsvlei (OLIV)	259	110	NO	22	17	NO	++
12.Patois	Pahayokee (PAHV)	259	110	NO	22	18	11	+++
	Patois (PATV)	259	110	NO	22	18	11	+++
13.Tete	Batama (BMAV)	NO	NO	NO	26	NO	NO	-
	Tete (TETE)	NO	NO	NO	26	NO	NO	-
14. Turlock	M'Poko (MPOV)	259	110	32	23	18	NO	++
	Turlock (TURV)	259	110	32	22	18	11	++
Bunyamwera	BUNV	259	110	32	22	18	11	+++

Table 3.3 Summary of the estimated size of protein synthesized by the viruses and shutoff caused by the viruses in infected Vero cells. NO represents not obvious, +++ indicates drastic (same as shutoff caused by BUNV), ++ indicates moderate shutoff (same as shutoff caused by BUNV), and – indicates no shutoff.

of almost similar as in BUNV-infected cells was seen in cells infected with CAPV, BWAV, WONV, PAHV, PATV and BAKV, whereas lesser shutoff similar to BUNdelNSs infected cells was observed in cells infected with ACAV, ANAV, TCMV, ANBV, BORV, E147V, NDV, OLIV, KOOV, GAMV, SJ2441V, TURV, MPOV, SJ2441V, BERV and GMAV. No shutoff was observed in TETEV, BMAV, PLSV, MNTV, PGAV and KETV-infected cells (Figure 3.2)

3.3 Discussion

3.3.1 Virus propagation

Some of the viruses were received as mouse brain homogenate form and had been kept frozen for more than 40 years. Furthermore some had not been passaged in cultured mammalian cells before (Table 2.2), therefore a problem was encountered in growing and getting a good cpe of the viruses in these cells. Some of them needed more than 5 passages in different cell types. Since some of the viruses were isolated from mosquito, I also tried to grow OLIV, BOTV, SJ2441V and KETV in *Aedes albopictus* C6/36 cells for at least 3 passages before propagating them in mammalian cell cultures. All viruses except BOTV generated cpe. Perhaps BOTV was already dead since it has been kept frozen for more than 40 years, or alternatively, was unable to replicate in mammalian cells used. RT-PCR also gave a negative result with this virus and therefore it was excluded from further experiments.

It was very difficult to produce high titre stocks of some viruses. This problem delayed me to do further experiments. To overcome this problem, a few steps were taken including pelleting the virus infected supernatant at high speed (26,000 rpm for 2 h) to concentrate the virus. However these attempts were unsuccessful. According to Lazdins and Holmes (1979), this could be due to the fragility of bunyaviruses, giving rise to great losses during concentration of the virus by centrifugation. The other step is growing the virus from the isolated plaques. However, this method was also failed for some of the viruses. The production of virus defective interfering particles after repeating passage the virus at high multiplicity might also contributed to this failure (Obijeski et al., 1976).

3.3.2 Plaque characteristic of the viruses

The plaque-forming properties are used as a marker to distinguish attenuated and virulent strains of some viruses such as Rhinovirus (Dougles et al., 1974), Newcastle disease virus (Singh et al., 1970) and Measles virus (Matumoto, 1966). Although mutant virus BUNdelNSs that lacking the NSs protein has been shown to exhibit a smaller plaque size, and less pathogenic in mice compared to wt BUNV (Bridgen et al., 2001), for the other orthobunyaviruses, the presence of NSs protein and the pathogenicity of the virus have been shown not to correlate with the plaque size. This was as previously reported in MA-104 embryonic rhesus monkey kidney cells infected with GROV which encodes NSs protein in its S segment, two sizes of plaques were produced in the infected cells and both were found to have same pathogenicity in suckling mice (Tauraso, 1969). This confirms that the size of the plaque is not always correlated with the presence of NSs protein and the virulence of the virus. Mixed sizes of plaques were also observed with some of the viruses used in this study.

3.3.3 Protein synthesis in virus infected cells

Analysis of protein profiles revealed that all the viruses used in this study have proteins of almost similar migration pattern to those of other viruses in the *Orthobunyavirus* genus. Therefore these results support the serological classifications, which include the viruses in *Orthobunyavirus* genus. In this study, NSs protein was only detected in certain viruses. Most probably that these viruses do not have NSs protein in their S segment, the protein is too short or too little to be detected by SDS-PAGE.

The results obtained with ANAV, CAPV, and BORV were in agreement with the results obtained by Ushijima et al (1980) and Klimas et al (1981), where the N proteins were found to be bigger than BUNV N protein. A similar result was obtained by Ushijima et al (1981) who found the N protein of PAHV was almost similar in size to that of BUNV. The result obtained for TURV in Turlock serogroup was found in contrast with the result of Klimas et al. (1981), where bigger N protein size was reported (24 kDa) compared to this study (22 kDa). Therefore, sequencing of N protein is needed to confirm this result. However my study revealed that MPOV, which is also in the Turlock serogroup possess an N protein larger than BUNV N protein (23 kDa).

In agreement with the previous result by Ushijima et al. (1980), who found that not all of the viruses in the same serogroup have the same size of N protein. For example viruses in Bwamba (BWAV and PGAV), Bakau (BAKV and KETV), and Koongol (KOOV and WONV) serogroups had different sizes of N protein. Considering these, it is not possible to group the viruses to the same serogroup based on their protein profiles alone. Therefore, serological and sequencing data are needed to support this classification.

3.3.4 Shutoff of host protein synthesis

BUNV and LACV have been shown to cause a drastic shutoff of host protein synthesis in infected BHK and Vero cells (Lazdins and Holmes, 1979; McPhee and Westaway, 1981; Short et al., 1982; Bridgen et al., 2001; Blakqori et al., 2004) but it depends on multiplicity of infection of the virus (Lazdins and Holmes, 1979). However, no shutoff was observed in Vero cells infected with Uukuniemi virus in the Phlebovirus genus (Pattersson, 1974), in hantavirus infected cells (Elliott et al., 1984) nor in some nairovirus-infected cells (Watret et al., 1985). My study revealed that no shutoff was observed in cells infected with TETEV, BMAV, PLSV, MNTV, PGAV and KETV. It is a possible that these viruses replicated more slowly and might need longer time to cause shutoff of host protein synthesis. It has been shown previously that a mutant virus of BUNV, BUNdelNSs which is lacking the NSs protein, caused shutoff later than BUNV (Bridgen et al., 2001). No NSs protein band was observed in the cells infected with TETEV, BMAV, PLSV, MNTV, PGAV and KETV, leading to a preliminary conclusion that these viruses are similar to BUNdelNSs and do not produce an NSs protein. However these speculations needed to be confirmed by the sequencing results. Some shutoff was still observed in cells infected with BUNdelNSs, suggesting that beside the presence of NSs protein, other factor such as capsnatching process by L protein might also plays a role in causing this shutoff. However the shutoff caused by this process is minimal compared to the shutoff caused by NSs protein (Hart, 2004). Besides that, shutoff also depends on the multiplicity of infection, where the shutoff caused by 10 PFU/cell virus was found to be less than the shutoff caused by 1 PFU/cell virus due to the presence of autointerference (McPhee and Westaway, 1981).
4 Sequence Analysis of the S RNA Segments of Orthobunyaviruses

4.1.1 Introduction

As discussed in Chapter 1, the genus *Orthobunyavirus* contains more than 170 viruses that are classified serologically into 18 serogroups (Calisher, 1991). To-date, molecular data and full length sequences are only available for viruses in 4 serogroups: Bunyamwera, California, Group C and Simbu. Thus, sequences of viruses in the other serogroups are needed to give us a fuller understanding of their phylogenetic relationships and a greater insight into their evolutions. Therefore, in this chapter, I aimed to sequence the S RNA segments of representatives of the other 14 serogroups with a view to rationalise their classification and to understand their evolutionary relationships.

4.2 Results

4.2.1 RT-PCR, cloning and sequencing of the S segments

Twenty seven viruses which represent 14 serogroups viruses were grown up as listed in Table 2.2. Due to the unavailability of a viable isolate, BOTV was excluded from this study. RT-PCR, 3'/5' rapid amplification cDNA ends (RACE), modified RACE, Rolling Circle cDNA Amplification (RCA)-RACE and cloning were conducted as described in the "Materials and Methods" section. The sequences of the primers used are as listed in Table 2.1. The steps and primers used to get the sequence of each virus S segment are summarised in Table 4.1.

To synthesize S segment cDNA, RNA extracted either from virus infected cells or from nucleocapsid RNA was used for RACE, and the RNAs extracted from virions were used for RT-PCR. Initially, primers BUN S+/S- (Dunn et al., 1994) based on the 15 terminal nucleotides of the published S segment sequences of Bunyamwera, California, Group C and Simbu serogroups viruses were used to amplify cDNA. These primers were thought to be conserved in all orthobunyaviruses. However, only complete S segment sequences of

SEROGROUP	VIRUS	PRIMER	PRODUCT	FINAL
			SIZE (bp)	PRODUCT (bp)
Anopheles	Ana A	BUN S+/BUNCAL2	884	947
А	(ANAV)	S25F/S13-	947	(Complete)
	Tacaiuma	BUN S+/BUNCAL2	910	960
	(TCMV)	S25F /S13-	960	(Complete)
Anopheles	Ana B	BUN S+/BUNCAL2	884	991
В	(ANBV)	S13+/ORO2S	991	(Complete)
	Boraceia	BUN S+/ BUNCAL2	909	982
	(BORV)	S13+/ORO2S	982	(Complete)
Bakau	Bakau	NP	NP	NP
	(BAKV)			
	Ketapang	NP	NP	NP
	(KETV)			
Bwamba	Bwamba	BUN S+/BUN S-	868	1096
	(BWAV)	BWAPR/ORO2S	578	(Complete)
	Pangola	NP	NP	NP
	(PGAV)			
Capim	Acara	NP	NP	NP
	(ACAV)			
	Capim	S400R/PGTR	357	991
	(CAPV)	CAPPF/CAPPF	545	(Partial)
		CAPPF/CAPPR on	235	
		ligated RNA		
Gamboa	Gamboa	BUN S+/BUNCAL2	1170	1170
	(GAMV)			(Partial)
	San Juan	BUN S+/BUNCAL2	1179	1250
	(SJ2441V)	SJPF/S13-	498	(Complete)
Guama	Bertioga	S350F/BUNCAL2	632	695
	(BERV)	S400F/BERPF/RACE	420	(Partial)
		OAP and AP		
	Guama	S400F/BUNCAL2	510	621
	(GMAV)	GMAPF/GMAPR on	427	(Partial)
		ligated RNA		
Koongol	Koongol	NP	NP	NP
	(KOOV)			
	Wongol	NP	NP	NP
	(WONV)			

Minatitlan	Minatitlan	NP	NP	NP
	(MNTV)			
	Palestina	NIFFF/PGTR	305	662
	(PLSV)	PLSPF/NNHR on	543	(Partial)
		Ligated RNA		
Nyando	Eretmapodites	BUN S+/BUN S-	902	902
	(E147V)			(Complete)
	Nyando	BUN S+/BUN S-	941	941
	(NDV)			(Complete)
Olifantsvlei	Olifantsvlei	NP	NP	NP
	(OLIV)			
Patois	Pahayokee	NP	NP	NP
	(PAHV)			
	Patois	S400F/S13-	501	768
	(PATV)	PATPR/PATNNHR	286	(Partial)
		/RACE OAP and		
		RACE AP		
Tete	Tete	BUN S+/BUNCAL2	947	1022
	(TETEV)	BUN S+/S13-	1022	(Complete)
	Batama	BUN S+/BUNCAL2	944	1021
	(BMAV)	BUN S+/S13-	1021	(Complete)
Turlock	M'Poko	BUN S+/BUN S-	951	1075
	(MPOV)	MPO400PR+S13+	578	(Complete)
	Turlock	NP	NP	NP
	(TURV)			

Table 4.1 Primers used in the RT-PCR. The sequences and the location of the primers were shown in Table 2.1 and Fig. 2.1. F and + stand for forward ; R and – stand for reverse ; P for partial; bp for base pair; NP stands for no product obtained, RACE for rapid amplification cDNA ends, OAP stands for Oligo (dT) anchor primer and AP stands for anchor primer.

viruses in the Nyando serogroup, NDV and E147V, were obtained with these primers. Therefore, combinations of published and designed primers for the partial and complete sequences were used for the other viruses. Table 4.1 summarised the primers used and PCR products obtained using these primers. Besides NDV and E147V, complete sequences of the S segment were also obtained for ANAV, TCMV, ANB, BORV, TETEV, BMAV, SJ2441V, NDV, E147V, BWAV and MPOV, while partial sequences of CAPV, GAMV, PLSV, PATV, BERV and GMAV were obtained. No products were obtained from ACAV, BAKV, KETV, MNTV, OLIV, KOOV, PAHV, PGAV, TURV and WONV RNAs. The 3'/5' RACE, modified 3'/5' RACE, RCA-RACE and RT-PCR on ligated RNA were conducted using internal primers that were designed based on the partial sequences obtained initially and RACE primers. Even though I failed to get the 3' and 5' end sequences using these methods, a product shorter than an expected length was obtained. It is likely that annealing of RACE primers to the truncated RNA was greater than annealing to complete RNA.

The PCR products were cloned into pGEMT easy vector. All the sequences obtained represent the consensus from sequencing three independent plasmids containing the insert of each S segment, using M13 forward and reverse primers. For 3'/5' RACE products only two clones were available for some of the viruses. Low numbers of nucleotide variations (less than 5 nucleotides) were recorded between clones of the same segment. The nucleotide present in the majority of the clones was accepted for compiling the consensus sequence. Direct sequencing of PCR product was found to produce unsatisfactory sequencing results.

4.2.2 'Architecture' of S segment

The nucleotide sequence data obtained were manipulated and analysed using the Lasergene program (DNAStar inc). A summary of the sequences is shown in Table 4.2. Similar to those of the published viruses, different lengths of S segment were also observed either between viruses within the same serogroup or between different serogroups (Fig. 4.1 & Table 4.2). SJ2441V in the Gamboa serogroup was found to have the longest S segment to date (1250 nucleotides) while the shortest S segment was noted in E147V (902 nucleotides). However the shortest S segment to date is reported for OROV in the Simbu serogroup (754 nucleotides) (Saeed et al., 2001). Similar to previous reports, the S RNAs

A. S sequence with NSs protein

(i) SJ2441V-complete sequence

1	AGTA	GTGT	ACT	CCAC	CAGI	TAAZ	ACTZ	ATTC	CATZ	AAZ	ATCI	TAC		TC	TTC	CAAZ	AGTT	TTT	CAA
61	ACCT	GGAG	GTG	CTTT	TTC	GAG	ACO	ACG	AAG	CAF	AGAT	GTO	TA7	TGZ	TCT	TT7	ACGI	TAT	CGA
											N	4 5	5 1	1 I) I	. s	د ۱	7 E	D
														м	I	F	т	Y	S
121	CGAT	GATA	CGT	CGGC	CAAC	CTC	TAC	СТС	TTZ	TG	ATCO	CAAZ	GAG	GAGZ	ATZ	ATGI	TA	ATT	TAT
	D	D T	S	A	S	S	S	S	Y	D	Р	K	R	E	Y	D	K	F	I
	т м	I	R	R	Q	Α	L	Α	L	М	I	Q	R	Е	N	М	I	N	L
181	TAGT	ACTT	ACG	GGCI	GCI	AC	TAC	TCC	CAGO	CAAZ	ACAT	TAT	ATC	TTT	TCT	CTT	rgc <i>i</i>	ATGO	CACG
	S	т ү	G	Q	Q	L	т	Р	Α	N	I	к	С	F	L	L	н	Α	R
	l V	L	т	G	S	N	L	L	Q	Q	т	L	N	v	F	S	С	М	н
241	CACG	GCCA	AGG	CTCC	STCI	TTC	CAAZ	AGGZ	CCC	CGGC	CGGI	ACC	CAC	SAAC	GCI	TAC	CTTT	TGC	SAAC
	т	A K	A	R	L	S	к	D	Р	Α	Е	R	R	т	L	т	F	G	т
	A R	P	R	L	v	F	Q	R	т	R	R	N	Α	Е	R	L	L	L	E
301	TTGG	AAGG	TTG	AGG	TG	TAT	ATAZ	ACCE	TT	ATCO	GACO	SAGO	CGG	TCC	CAA	TAC	GG	GCI	AGGA
	W	кν	E	v	v	N	N	н	Y	R	R	Α	Α	Р	I	т	v	Q	D
	L G	R	L	R	L	L	I	т	I	I	D	Е	Р	L	Q	L	R	С	R
361	CCAT	GATC	TTA	CCTT	GCI	ACCO	SCCI	TTTC	CGG	GATZ	ATA1	TGC	CTI	GT	ACTO	STCI	rgc <i>i</i>	AGGC	CAG
	н	D L	т	L	н	R	L	S	G	Y	I	Α	L	Y	С	L	Q	Α	S
	т м	I	L	Ρ	С	т	Α	F	Р	D	I	L	Р	С	т	v	С	R	Р
421	TGTG	GAGA	ATG	CGCI	GTC	TTT	rcgz	GCC	GAA	CAC	GATO	CAAC	CAGI	CAT	CAZ	ACCO	GA	СТС	TGA
	v	E N	A	Q	S	F	Ε	R	I	R	S	т	v	I	N	Р	I	S	E
	V W	R	М	R	s	L	S	S	Е	s	D	Q	Q	s	s	т	R	S	L
										5	Stop	NSs	5						
481	GTCC	ATGG	GGA'	FCA	ACTO	GCI	AGCI	TGO	GAGC	CAA'	[GA]	GT	ATC1	TTC	CAAC	CCC	TAC	TGC	FTGC
	S	M G	I	N	W	Q	н	G	Α	М	М	Y	L	S	т	L	Р	G	Α
	S P	W	G	S	т	G	S	М	Е	Q	*								
541	AGAA	ATGT	TTC	FAA C	CAG	AGT 7	rccz	AGC1	TCT	CCC	CACI	TAGO	CATI	TGC	TG	GG	GAC	SAG	TAAA
	E	M F	L	т	E	F	Q	L	L	Р	L	Α	F	Α	v	v	R	v	ĸ
601	AAAG	AACC	TGG	CGAC	CCC	CAGI	AAC	GGI	TAP	GAZ	\AG'	'AC'	TAAG	SAC	GAC	SATZ	ATG2	AGGC	GCAG
	ĸ	N L	A	S	Р	E	т	v	K	K	v	L	R	Q	R	Y	Е	G	R
661	ACTT	CCGG	GAG	AATO	GAI	GAZ	ACTO	CAGA	GC1	GGI	\AA /	AGGI	CAP	AGC	TGC	CAA	CAC	CAC	SAGT
	L	P G	E	W	М	N	S	E	L	E	K	v	K	Α	Α	I	т	R	v
721	GGAG	TCGT	TGC	ACAZ	GAC	TTT	TG	CGG	GAG	GAC	CTGC	CAAC	GAAT	GAI	TGF	AGTT	CTT	CAC	STGC
	E	S L	H	K	т	F	v	G	v	т	A	R	М	I	E	F	F	S	Α
									St	iopl	1								
781	TCTA	GGCA	TTA	ATCI	TGF	ATT?	GAC	SAAG	GAT!	AGZ	ACC	CAG	ATCF	TT7	TTT	GAZ	ATC	TTT	TAA
	L	G	I I	N I	- I) 1	L F	ξF	٤ ٦	•									
841	TATT	TGAC	AAT	[TTT	TAT	'AC'	CTT	TTF	AAZ	ACTT	CTTT	GTZ	TAT	ATA	ACTT	TTZ	ATA	TAT	TTAT
901	TTTA	TCAA	TAT	FTA	ACCZ	GTO	TAF	ACCZ	TAT	TAT	rccz	AGT1	'AA'	CCZ	TAT	TAT	AGG2	CAT	ATA
961	TATA	TTTT	TAT!	FAA C	ATT	'ATZ	ATC	ATA	TAT	TATC	CATZ	ATCO	CAGI	'AA'	TATO	TAT	ACCO	CATZ	ACTA
1021	TCCT	GATT	TAT	CAAT	TTT	TAT	TAT	TAT	'AT	CCZ	ACAT	CAT	TTF	GTT	TTZ	AGCI	AAA	ATT1	GGT
1081	AAGG	ACTA	ACA	ATAC	TAT	ACAZ	ATT1	TAT	AGG	TAC	STTI	TAC	TAC	TAZ	ACAC	CATT	'AG	ACTZ	GTT
1141	AGGA	CCCA	AAA	GCAG	CTC	TCI	ATC	GGG	GGG	TGG	TCO	GGGG	CAA	CTT	TGC	CAP	ATT?	CTT	CAA
1201	ATCA	TTTT	AGT	AGA	TAT	ATC	CAAZ	ACGZ	ATC	TTT	TAC	TGG	AGC	CAC	ACTZ	ACT			

Fig. 4.1 Nucleotide sequences of positive-sense S segments of the viruses. * indicates stop codon for N protein; * stop codon for NSs protein; red bold letters represent N protein and green letters represent NSs protein. The nucleotides were translated to amino acid using DNAMAN version 4.0 program (Lynnon Biosoft)

(ii)GAMV-incomplete sequences

1	AGTAGTGTACTCCACAGTAAACTATTCTTAAAAATACTACAAAGTTTTGCAAACCTGGAG													
	StartN StartNSs													
61	ATCCTCTTTGAGTCCGACGAAGCAAG <mark>ATG</mark> TCTGACTCTCTATACGTATTTGACGATGATT													
	MI													
	M S D S L Y V F D D D S													
121	CGACGGCAAGCTCTAGCACTTATGATCCAACGCAGGAATACAATGTCTTTATTAGTACTC													
	R R Q A L A L M I Q R R N T M S L L V L													
	TASSSTYDPTQEYNVFISTH													
181	ACGGGGACCAGCTTACTCCAGCAAACATTAAATGTTTCTTCCTGCATGCA													
	ΤGΤSLLQQTLNVSSCMΗVQΡ													
	G D Q L T P A N I K C F F L H A R T A K													
241	AGGCTCGTCTGGCAAAGGACCCGGCGGAACGCAGAACGCTTACTTTTGGATCTTGGAAGG													
	R L V W Q R T R R N A E R L L L D L G R													
	A R L A K D P A E R R T L T F G S W K V													
301	TTGAGGTTGTTAATAACCATTACAGAAGAGCCACTCCGGTTACAGTGCAGGACCATGATA													
	L R L L I T I T E E P L R L Q C R T M I													
	E V V N N H Y R R A T P V T V Q D H D N													
361	ATACCTTGCACCGCATTTCCGGATATCTTGCCCTGTACTGCTTGCAGACCAGCTCCCAGA													
	I P C T A F P D I L P C T A C R P A P R													
	T L H R I S G Y L A L Y C L Q T S S Q N													
421	ATAAGCAGAACTTTGAGCGAATCAAGTCAGTCGTCATCAACCCGATCTCTGAGTCCATGG													
	I S R T L S E S S Q S S S T R S L S P W													
	K Q N F E R I K S V V I N P I S E S M G													
	StopNSs													
481	GGATCAACTGGCAGCATGGAGCAATGATCTATCTGTCTACGCTTCCCGGAGCAGAAATGT													
	GSTGSMEQ*													
	INWQHGAMIYLSTLPGAEMF													
541	TTCTCACTGAGTTTCAGCTGTTCCCACTTGCATTTGCAGTTGTAAGAGTGAAAAAGAACC													
	L T E F Q L F P L A F A V V R V K K N L													
601	TGGCAAGCCCTGAGACAGTGAAGAAAGTGCTGAGACAGAGATATGACGGCAGATTGCCTG													
	A S P E T V K K V L R Q R Y D G R L P G													
661	GAGAATGGATGAATTCAGAACTAGAAAAGGTTAAGACCGCAATCACAAAGGTTGAAGCAC													
	E W M N S E L E K V K T A I T K V E A L													
721	TTCACAAGACCTTTGTTGGAGTTACTGCCAGAATGATTGAGTTCTTCACTGCTCTTGGAA													
	H K T F V G V T A R M I E F F T A L G I													
	StopN													
781	TCAACCTTGACATGAGACGTTAATTCAATTTTGAATTTCCAGTACAATTTTACTTTTATT													
	NLDMRR*													
841	TCCAATTTTTATTTTGCTTTATATATATTTATTACTTGACTATATTTTATTATTATTATTCTTA													
901	TTTTATCCATATATAACCAATATATATATCTATATATCCCAATATGTAATCCATATACCAA													
961	TTATATATGTATCGATAGTTAATCATTATTATATCCAGTTTATTTA													
1021	ͲͲϹϪϪͲͲͲͲϾͲͲͲϾϹͲΑͲΆͲͲϤϾϹΑΑͳΑͲͲϹϪϪͲΆϾΑΑͲͲϪͲϹϪϪϪͲͲͲϪϾϪͲϪϪϾϪϪϾͲϪ													
1081	ΑCΑΑΤΑCTAACATTTTTTTTTTTTTTTTTTTTTTTTTTTT													
1141	AACAGCTGTCAAATGGGTGGGTGGTAGGG													

4.1.1 S segment sequences of Gamboa serogroup viruses.

(i)NDV-complete sequence

AGI	[AG]	rgro	SCTO	CCAC	CAA	FTG	AAT	CT7	ATC	ATT	CAC	ATA	GTT	ATC	ACA'	TAC	AAG	IGA'	ICTT
			St	art	τN					Sta	rtN	Ss							
CTO	SAC	FTG	rt <mark>a</mark> t	CGTO	CTG	AGA!	rcgo	CTT	CTT2	ATG	ATG:	rcci	AGT	CTA	CTG	CCC	AAA	ATG	GATT
			М	S	E	I	Α	F	Y	D	v	Q	S	т	Α	Q	N	G	F
										М	М	S	s	L	L	Р	к	м	D
TGF	ATCO	CTG	ATC	AC	AATZ	ACT	rggo	CAT	CTA2	AAG	CTG	CTG	CAG	GAG	CAG	GGC!	FTA	ATA	FTGT
D	Р	D	Q	Q	Y	L	Α	F	ĸ	Α	Α	Α	G	Α	G	L	N	I	v
L	I	L	I	N	N	т	W	н	L	к	L	L	Q	Е	Q	G	L	I	L
TTC	CCGC	стсо	GA	CTT	ICT:	rcc:	rca/	ATGO	ccc	GGA	AAG	CCA	AAG	ATC	AAC	ICT(стс	GTA	GACC
S	A	R	I	F	F	L	N	Α	R	K	A	K	D	Q	L	S	R	R	Р
F	Р	L	G	s	s	S	S	м	Р	G	к	Р	к	I	N	s	L	v	D
AGZ	AGCO	CAAZ	AGAT	TAT	ACC	CTA2	AAT	TGG	GCAG	CAT	GGT	CGG	rggi	AGG	TGG'	TCA	ATA	ACC	ATTT
Е	Р	к	I	N	L	к	F	G	т	W	S	v	Е	v	v	N	N	н	F
Q	s	Q	R	L	т	L	N	L	А	н	G	R	W	R	W	s	I	т	I
TCF	AGG	GAAZ	ACCO	GGG	ACAZ	ATG:	TTA:	CAC	GCG	ACA	CGG	ATC	TTA	CAA'	rcci	ACA	GAC!	IGT	CAGG
0	G	N	R	D	N	v	I	S	D	т	D	L	т	I	н	R	L	S	G
F	к	Е	т	G	т	м	L	s	А	т	R	I	L	0	s	т	D	С	0
	5	Stor	oNS	5										~					-
ATZ	ATA	FAG	CAC	GAT	TTA:	TTC:	rggi	ACCZ	ATA	ATC	TTG	CAG	GCA	ATA	CTG	IGG!	гтси	AGT	CTGG
Y	I	A	R	F	I	L	D	0	Y	L	Α	G	N	т	v	v	0	S	G
D	I	*						-									-		
		יסממ	TTC	AGA	TTG:	TAT	ATCO	CTAT	TGG	CTG	AGT	CCA	ATG	GAA'	TAA	AGT	GGA	GCG	CAGG
GAT	rccz																		
GA1		L	0	I	v	N	Р	I	Α	E	S	N	G	I	K	W	S	Α	G
GAT I TGC	rcc <i>i</i> Q CAGI	L	Q FCT2	I	V FGTO	N CCT	P FTT:	I rcco	A CAG	E GAA(S CTG2	N AAA'	G TGT	I TTC	K TGG2	W AAA	S AAT	A FCA	G ATTT
GAT I TGC A	CCZ Q CAGZ E	L AGAT	Q FCT2 Y		V TGT(S	N CCT: F	P TTT: F	I ICCO P	A CAG(G	E GAA(T	S CTG2 E	N AAA' M	G IGT: F	I FTC L	K TGG2 E	W AAA K	S AAT: F	A FCAZ N	G ATTT F
GAT I TGC A TTZ	CAGI	L AGAI I CATI	Q FCT2 Y FGG0		V TGT(S TTG(N CCT: F GAA:	P TTT: F TCT2	I ICCO P ACCO	A CAG G GAG	E GAA T TGA	S CTG2 E AGA(N AAA' M GAG(G TGT' F GGA'	I ITC L IGA	K TGG2 E TGG2	W AAAA K ATG	S AAT F CAC	A FCA N N	G ATTT F ATCT
GAT TGC A TTZ Y	CAGA E ATCO	L AGAI I CATI	Q FCT2 Y FGG0		V TGTO S TTGO G	N CCT: F GAA: I	P TTT: F TCT2 Y	I ICCO P ACCO R	A CAGO G GAG! V	E GAA T TGA K	S CTG2 E AGA(R	N AAA' M GAG(G	G TGT' F GGA' M	I ITC L IGA M	K TGG2 E TGG2 D	W AAAA K ATG A	S AAT F CAC	A FCAX N AATX Y	G ATTT F ATCT L
GAT TGC A TTZ Y AAZ	CAGI E ATCO P	L AGAT I CATT L AAGO	Q FCT2 Y FGGC A CAT1	I ACCI L CAAI I IGAG	V FGT(S FTG(G GAC)	N CCT F GAA I AGA	P TTT F TCT2 Y GAT2		A CAGO G GAG' V GAA	E GAA T IGA K ATC	S CTG2 E AGA(R FAA(N AAA M GAG G CTG	G TGT' F GGA' M CAG	I ITC: L IGA: M ACC	K TGG2 E TGG2 D AGT(W AAAA K ATG A GGA	S AAT F CAC Q IGC	A FCA N AAT Y AAT	G ATTT F ATCT L CAAA
GAT TGC A TTZ Y AAZ K	ICCA Q CAGA E ATCO P AAAA K	L AGAJ I CATJ L AAGO	Q ICTA Y IGGO A CATI	I ACCI L CAAI I IGAO R	V IGTO S ITTGO G GAC2 O	N CCT F GAA I AGA(R	P FTTT: F FCT2 Y GAT2 Y	I ICCO P ACCO R ACGO	A CAGO G GAG SAG V SAAI	E GAA(T IGA) K ATC! L	S CTG2 E AGA(R TAA(T	N AAA' GAG(G CTG(A	G IGT F GGA GGA CAG	I ITC: L IGA: M ACC: O	K IGG2 IGG2 D AGT0 W	W AAAA K ATG A GGA M	S AAT F CAC Q IGC	A ICA N AAT Y AAT S	G ATTT F ATCT L CAAA K
GAT I TGC A TTZ Y AAZ K ATC	CAGA E ATCO P AAAA K CAGA	L AGAI I CATI L AAG(A ATG2	Q FCTA Y FGGC A CATI L ATGI	I ACCI L CAAI I IGAO R FGAO	V FGT(S FTG(G GAC/ Q FGA(N CCT F GAA I AGA R GGG	P FTTT F FCTZ Y GATZ Y CAA	I ICCO P ACCO R ACGO G IAGO	A G G GAG SAG SAA SAA N CAG	E GAA T IGA I K ATC L I ATC	S CTG2 E AGA0 R TAA0 T TGG2	N AAA M GAG G CTG A AGA2	G IGT: F GGA: M CAGI D AGT:	I ITC: L IGA: M ACC: Q IGC:	K TGG2 TGG2 TGG2 NGT0 W AGT0	W AAAA K ATG A GGA GGA GGG	S AAT F CAC Q IGC Q GCA	A FCA N AAT Y AAT S AAT	G ATTT F ATCT L CAAA K CAGG
GAT I TGC A TTZ Y AAZ K ATC	CAGA E ATCO P AAAA K CAGA	L AGAI I CATI L AAGO A ATGZ	Q ICTA IGGO A CATI L ATGI	I ACC: L CAA: I IGAC R ICT: L	V IGT(S ITG(G GAC) Q IGA(R	N CCT F GAA I AGA(R GGG(A	P TTT: F TCT2 Y GAT2 Y CAA: I	I ICCO P ACCO R ACGO G IAGO	A G G G G G A G G A G C A G C A G C A G C A G V V V V V V V V V V V V V V V V V V	E GAA(T IGA I ATC L IAT	S CTGJ E AGAG R TAAG T TGGJ E	N AAA GAGO G CTGO A AGAX	G IGT F GGA M CAG D AGT	I ITC: L IGA: M ACC: Q IGC: O	K IGG2 IGG2 D AGT0 W AGT0 W	W AAAA K ATG A GGA GGG G GGG	S AAT F CAC Q FGC Q GCA K	A FCA N AAT AAT S AAT S AAT S	G ATTT F ATCT L CAAA K CAGG G
GAT I TGC A TTZ Y AAZ K ATC S	CAGA E ATCO P AAAA K CAGA D	L AGAT I CATT L AAGO A ATGZ	Q TCTA Y TGGC A CATI L ATGI V	I ACCI L CAAI I IGAO R ICTI L	V IGTO S ITTGO G GAC2 Q IGAO R	N CCT F GAA I AGA R GGG A	P F TCT2 Y GAT2 Y CAAS I	I ICCO P ACCO R ACCO G IACO A	A CAGO G GAG' CAG' V CAG' V	E GAA(T IGA) K ATC: L IAT: L	S CTG2 E AGA0 R TAA0 T TGG2 E	N AAA G G G CTG CTG A AGA K	G IGT F GGA GGA CAG D AGT L	I ITC: L IGA: M ACC: Q IGC: Q	K TGGJ TGGJ D AGT(W AGT(W	W AAAA K ATG(A GGA GGG(G Storesting Storesting	S AAT F CAC Q TGC GCA SCA K	A ICAI N AATI Y AAT(S AAT(S	G ATTT F ATCT L CAAA K CAGG G
GAT I TGC A TTZ Y AAZ K ATC S	CAGA E ATCO P AAAA K CAGA	L AGAI I CATI L AAGO A ATGI D	Q FCT# FGG0 A CATT L ATGI V	I ACCI L CAAI I IGAC R ICTI L	V FGT(S FTG(G GAC/ Q FGA(R R	N CCT F SAA I AGA(R GGG(A SGG(A	P FTTT: F FCTZ Y GATZ Y CAA: I	I ICCO P ACCO R ACGO G IAGO A	A G G G G G G G G G G G G G G G G G G G	E GAAG T IGAX K ATC' L IAT' L	S CTG2 E AGAO R FAAO T GG2 E	N AAAA M GAGO G CTGO A CTGO K	G IGT F GGA M CAG D AGT L	I FTC: L FGA: M ACC: Q FGC: Q FGC: TCT'	K TGG2 TGG2 D AGT0 W AGT0 W	W AAAA K ATGO A GGA GGGO G G Sto	S AAT F CAC Q IGC IGC Q GCA K S OPN	A TCA N AAT Y AAT S AAT S	G ATTT F ATCT L CAAA K CAGG G
GAT TGC A TTZ Y AAZ K ATC S CCT	CAGA E ATCO P AAAA K CAGA D	L AGAI I CATI L AAG(A ATG/ D GCG/ E	Q FCT2 FGGC A CAT3 L ATG3 V AAGC A	I ACC: L CAA: I GAC R FCT: L CAGC	V FGTC S FTGC G GAC Q FGAC R CTAC R	N F GAA I AGAO R GGGO A GGGO A	P FTTT: F ICTI Y GATI Y CAA: I AGT: F	I ICCO P ACCO R ACGO G IAGO A ICC! I.	A G G G G G G G G G G G G G G G G G G G	E GAAG T IGAJ K ATC: L IAT: L GAAJ	S CTG2 E AGA0 R TAA0 T TGG2 E AGT2 F	N AAAA G G CTG CTG A AGAI K TTG G	G IGT F GGA D CAG D AGT L GTA	I TTC: L TGA: M ACC: Q TGC: Q TGC: L	K TGGJ D AGT(W AGT(W TAA!	W AAAA K ATG A GGA GGG G G St TTT *	S AAT F CAC Q IGC IGC IGC IGC IGC IGC IGC IGC IGC IGC	A TCA N AAT AAT S AAT S AAT S	G ATTT F ATCT L CAAA K CAGG G CTAT
GAT TGC A TTZ Y AAZ K ATC S CCT L TTC	CAGI E ATCO P AAAI CAGI D TGAO	L AGA1 I CAT1 L AAG(A ATG2 D GCG2 E	Q FCT2 FGGC A CAT1 L ATG1 V AAGC A AGC	I ACC: L CAA: I GAC R ICT: L CAGC A	V IGTO S ITGO G GACI Q IGAO R CTAO R GGT	N F SAA: I AGA(R SGG(A SGG(A SGC) C TTG	P TTT: F TCTI Y GATI Y CAA: I AGT: F	I ICCO P ACCO R ACGO IAGO A ICCC L	A G G G G G G CAG CAG CAG CAG CAG CAG CA	E GAAG T GAA ATC: L TAT: L GAA GAA	S CTG2 E AGAG R TAAG T TGG2 E AGT F CAA	N AAAA G G CTG CTG A AGAA K ITG G	G IGT F GGA CAG D AGT L GTA I GTA	I TGA I ACC Q IGC I CT L ATT	K TGGJ TGGJ AGT(W AGT(W TAA TAA	W AAAA K ATG A GGA G GGG Sto FTT X AAAA	S AAT CAC Q IGC IGC Q GCA K SOPN AGA	A TCA N AAT AAT S AAT S AAT S AAT S	G ATTT F ATCT L CAAA K CAGG G CTAT
GAT TGC A TTZ Y AAZ K ATC S CCT L TTC TCZ	CCC Q CAGA E ATCO P AAAA K CAGA D CAGA CAGA	LAGAI I CATI LAAGO AATGA B GCGA E AATI	Q FCT2 FGGC A CAT3 L ATG3 V AAGC A FTA2	I ACCI L CAAI I GAAI CAGO A ATTO	V IGTO G GACI Q IGAO R CTAO R GGT GGT	N CCT F GAA I AGA(R GGG(A GGCI Q TTG GGG	P TTT: F TCTI Y GATI Y CAA: I AGT: F TACI	I CCCC P ACCCC R ACCGC G IACCC L CCC I CCC I CCC I CCC I CCCC I CCCC I CCCC I I CCCCC I I I CCCCC I I I I I I I I I I I I I	A CAGO G GAGO V SAAN N CAGO V CAGO CAGO G G CATO	E GAAG T IGAJ K ATC' L IAT' GAAJ K IAT(GGGG	S CTG2 E AGA0 TAA0 T TGG2 E AGT2 F CAA2 GAA0	N AAA G G CTG CTG A AGAI K ITG G ITG CAC	G TGT F GGA CAGA D AGT L GTA I GCT	I IGA IGA IGA IGC IGC IGC I ICT I ATT	K TGGJ TGGJ TGGJ M AGT W AGT W TAA T TCAJ	W AAAA K ATGO A GGA G GGGO G Sto Sto I TTTI * AAAA	S AAT F CAC Q IGC IGC Q GCA K S OPN AGA GAC	A FCA N AAT AAT S AAT S AAT S AAT S	G ATTT F ATCT L CAAA K CAGG G CTAT CGGG IAGA
	AGY TGZ D L TTC S F AGZ Q F F ATZ Y D	TGATCO D P L I TTCCGO S A F P AGAGCO E P Q S TCAAGO Q G F K S ATATAT Y I D I	TGATCCTGA D P D L I L TTCCGCTCC S A R F P L AGAGCCAAK Q S Q TCAAGGAAA Q G N F K E Stop ATATATAGC Y I A D I *	AGTAGTGTGTGTG St CTGACTTGTTAT M TGATCCTGATCA D P D Q L I L I TTCCGCTCGGAT S A R I F P L G AGAGCCAAAGAT E P K I Q S Q R TCAAGGAAACCC Q G N R F K E T StopNSS ATATATAGCCAC Y I A R D I *	AGTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG	AGTAGTGTGTGTGTGTCCACAAA StartN CTGACTTGTTATGTCTGZ M S E TGATCCTGATCAACAATZ D P D Q Q Y L I L I N N TTCCGCTCGGATCTTCT S A R I F F F P L G S S AGAGCCAAAGATTAACCZ E P K I N L Q S Q R L T TCAAGGAAACCGGGACAZ Q G N R D N F K E T G T StopNSS ATATATAGCCAGATTTAZ Y I A R F I D I *	AGTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG	AGTAGTGTGTGTCTCCACAATTGAATT StartN CTGACTTGTTATGTCTGAGATCGC M S E I A TGATCCTGATCAACAATACTTGGC D P D Q Q Y L A L I L I N N T W TTCCGCTCGGATCTTCTTCCTCAA S A R I F F L N F P L G S S S S AGAGCCAAAGATTAACCTTAAATT E P K I N L K F Q S Q R L T L N TCAAGGAAACCGGGACAATGTTAT Q G N R D N V I F K E T G T M L StopNSS ATATATAGCCAGATTTATTCTGGA Y I A R F I L D D I *	AGTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG	AGTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGGAGATCGGCTTTTTZ M S E I A F Y TGATCCTGATCAACAATACTTGGGATCGCTTTTAT M S E I A F Y TGATCCTGATCAACAATACTTGGCATTTAX D P D Q Q Y L A F Y TGATCCTGGATCAACAATACTTGGCATTTAX D P D Q Q Y L A F Y TGATCCTGGATCAACAATACTTGGCAATGCTCAATGCCGCGATCTTCTTCCTCCAATGCCAATGCCCGAAGACCGGGACAATGTTAATCGCAACAATGTGACCAAAGATTAACCCTAAAATTTGGCAA E P L G S S S M P AGAGCCAAAGGATAACCTGGACAATGTTATCAGCGA E P K I N L K F G T Q S Q R L T L N L A TCAAGGAAACCGGGACAATGTTATCTGGCACAATGTTATCAGCGG Q G N R D N V I S D F K E T G N N L	AGTAGTGTGTGTGTGTGTGTGTGTGAGATTGATTGTCTAGAGTTGTCTTATGTStartNStartNStartNStartNStartNMTGACTGGTTGTATGTCTGAGATCGCTTTTTATAGGDPDQYLFKIAFFKINVLARINVIARISSATLNVISSATLDTLNVISSARITLDT	AGTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG	AGTAGTGTGTGTGTGTGTGAGAATTGAATTGAATTGAA	AGTAGTGTGTGTGTGTGTGTGTGTGAGATTGAATTGTATGATG	AGTAGTGTGTGTGTGTGTGTGTGTGAGATTGAATTGTATGATTGTCTAGAGTGTGCAGTGTGTATGGTGTGTGT	AGTAGTGTGTGTGTGTCAAATTGAATTGTATGAATTGTTATGAATTGTTATGATGTCCAGTATATGTTATGTGTGTCAATAGTTATGATGTCCAGTCTACTAGA M S S I A F Y D V Q S T A M S E I A F Y D V Q S T A M M S S L L T M M S S L L TGATCCTGATCAACAATACTTGGCATTTAAAGCTGCAGGAGGCAGGGCAGGACGGTCCTGGATCTTCTTCCTCAATGCCCGGAAAGCCAAAGATCAACC S A A G A G A G A G A G A G A G A G A G A G C G C C C Q Z Z A G R G C <td>AGTAGTGTGTGTGTGTGAGATTGAATTGTATCACATAGTTATCACATAGTTATGACGTGTGTATGTCTAGGATCGCCAGAGTGAGT</td> <td>AGTAGTGTGTGTGTGTGTGTGTGAGATTGAATTCAATTC</td> <td>AGTAGTGTGTGTGTGTGAATTGAATTGTATCACATAGTGAT StartN StartNSs CTGACTTGTTATGTCTGAGATCGCTTTTTATGATGTCCAGTCTACTGCCCAAAATGG M S E I A F Y D V Q S T A Q N G M M S S L L P K M TGATCCTGATCAACAATACTTGGCATTTAAAGCTGCTGCAGGAGCAGGGCTTAATAT D P D Q Q Y L A F K A A G A G L N I L I L N N T W H L K L Q E Q G L I I N N I I I N I N I I I N I N I I I I N N I I I I I I I I I I I I I I I I I<!--</td--></td>	AGTAGTGTGTGTGTGTGAGATTGAATTGTATCACATAGTTATCACATAGTTATGACGTGTGTATGTCTAGGATCGCCAGAGTGAGT	AGTAGTGTGTGTGTGTGTGTGTGAGATTGAATTCAATTC	AGTAGTGTGTGTGTGTGAATTGAATTGTATCACATAGTGAT StartN StartNSs CTGACTTGTTATGTCTGAGATCGCTTTTTATGATGTCCAGTCTACTGCCCAAAATGG M S E I A F Y D V Q S T A Q N G M M S S L L P K M TGATCCTGATCAACAATACTTGGCATTTAAAGCTGCTGCAGGAGCAGGGCTTAATAT D P D Q Q Y L A F K A A G A G L N I L I L N N T W H L K L Q E Q G L I I N N I I I N I N I I I N I N I I I I N N I I I I I I I I I I I I I I I I I </td

(ii)EREV-complete sequence

	StartN	
1	AGTAGTGTGCTCCACAATTGAATTCTATTTTGAACTGCTATTAAATGTCGGAACTGGTG	
	MSELV	
	StartNSs	
61	TTCTATGATGTCGAGCCAACTGCCCAAAATGGATTTGATCCTGATAAGCAGTATGTGGCA	A
	FYDVEPTAONGFDPDKOYVA	
	M M S S O L P K M D L T L T S S M W F	н
121	TTTAAAGCTTCAGCTGGAGCAGGGCTTAACATTGTTTCCGCTAGGATCTTCCTTC	r
	F K A S A G A G I. N T V S A R T F F I. N	•
		v
181	GCCAGGAAAGCCAAAGATCAACTCGCTCGTAGACCAGAGCCGAAGGTTGGTCTTAAATTT	Г
	A R K A K D O L A R R P E P K V G L K F	
	PGKPKINSLVDOSRRLVLNI	ι.
241	GGAACATGGCAGGTGGAAGTGGTCAATAACCATTTTCAAGGAAACCGGGACAATCCTATC	2
	G T W O V E V V N N H F O G N R D N P I	
	EHGRWKWSTTTFKETGTTLS	5
	StopNSs	1
301	GGCGACTCAGATCTCACAATGCATAGACTCTCAGGATATATAGCCCCGGTATATCTTAGAT	г
	G D S D L T M H R L S G Y T A R Y T L D	-
	ATOISOCIDSODI*	
361	CAATATCTAGCTGGGAATAGTGTTGCACAAGCTGGCATCCAGCTTCAGATCATCAATCCA	A
	Q Y L A G N S V A Q A G I Q L Q I I N P	
421	ATAGCAGAATCAAATGGAATAAAGTGGAGTGCTGGTGCAGAGGTGTACCTATCTTTTTT	з
	I A E S N G I K W S A G A E V Y L S F F	
481	CCAGGCACAGAAATGTTCTTAGAAAAATTTAATTTTTACCCCTTGGCAATTGGAATATAC	з
	P G T E M F L E K F N F Y P L A I G I Y	
541	AGAGTAAAAAAAGGGATGATGGAAGCCCAATTCTTGAAGAAATCTCTCAGACAAAGATAT	г
	R V K K G M M E A O F L K K S L R O R Y	
601	GGCCAAATGACAGCAGATCAATGGATGCAGACAAAATCGGATGATGTGATGAGGGCAGT	A
	G O M T A D O W M O T K S D D V M R A V	
661	GCAGTTCTTGAAAAATTGAGTTGGGGGAAGATCAGGTCTGAGTGAAGCTGCCCGTCAATTC	з
	A V L E K L S W G R S G L S E A A R O F	
	StopN	
721	CTGGGCCGATTCGGGATTGTTATCTAATTCATAAAATATACGCCTATTCGACTCCAAAT	г
	LGRFGIVI*	
781	CTAAAATATTTAATTTGACACCAAAAGACCTTCGGGTCAAAGCCGACAGCTTAATGGGTG	3
841	GGTGGTAGGGGAAACATAGCTTCAAAAGCATTTTTTCAAAATTCATTGTGGAGCACACAT	A
901	СТ	-

4.1.2 S segment sequences of Nyando serogroup viruses.

(i)BWAV-complete sequence

1	AGTAGTGTACTCCACTTAAGTACTAAAGCTATTAATTGAGGACATCGACTAATAAAAGCA														
	StartN StartNSs														
61	ACTTTAAGTAGAGGCACTATTATGTCCGATTTGATCTTCTATGATGTCGACTCGGTTAAT														
	MSDLIFYDVDSVN														
	M M S T R L M														
121	GCCAATGAGTTTGACCCTGATACAGGATACCTGGATTTTAAAAAATAACTATCCAGGGGCA														
	A N E F D P D T G Y L D F K N N Y P G A														
	P M S L T L I Q D T W I L K I T I Q G H														
181	CTCAATACAAATACCGCTAGGACATTCTTCCTCAATGCCGCAAAGGCCAAGAATGTGCTC														
	LNTNTARTFFLNAAKAKNVL														
	S I Q I P L G H S S S M P Q R P R M C S														
241	CGCAATAAACCTGACAAGAAGGTCAATCCTAAATTTGGAAACTGGGAGGTGGAGGTTGTC														
	R N K P D K K V N P K F G N W E V E V V														
	AINLTRRSILNLETGRWRLS														
301	AATAATCATTTTCCTGGAAACAGGAACAACCCAATTGGTAAAGATGATCTTACCCTCCAC														
	N N H F P G N R N N P I G K D D L T L H														
	I I I F L E T G T T Q L V K M I L P S T														
	StopNSs														
361	AGAATTTCAGGATATTTAGCAAGATGGATTTTGGAGGAATATAAGAGAGATGATGAATCA														
	R I S G Y L A R W I L E E Y K R D D E S														
	EFQDI*														
421	GAAAAGGAAATCATTGAAAGCACAGTTATTAATCCAATAGCTGAATCAAATGGGATTAGA														
	E K E I I E S T V I N P I A E S N G I R														
481	TGGAGTGATGGAGCAGAGATATATCTATCATTTTTCCCTGGCACAGAGATGTTTTTGGAA														
	W S D G A E I Y L S F F P G T E M F L E														
541	CCATTCAAATTCTATCCACTAGCAATTGGCATATACAGAGTCAAACATAAGATGATGGAT														
	P F K F Y P L A I G I Y R V K H K M M D														
601	GCACAGTTCCTGAAAAAAGCCTTGCGGCAGAGATATGGGAAAATGACAGCTGAAAAATGG														
	A Q F L K K A L R Q R Y G K M T A E K W														
661	ATGTCAACAAAAGTGAAGACTATAGGTGAGGCTGTTAGAAATGTAGAAAAGCTAAAGTGG														
	M S T K V K T I G E A V R N V E K L K W														
721	GGCAAGGGTGGTTTGAGCGATGCTGCAAGAAACTTCCTAAAGAAGTTTGACATTGCAATG														
	G K G G L S D A A R N F L K K F D I A M														
	StopN														
781	ATATAACCAGCCATATGCAAATATTTAGCTATCACTAAAACATATAACCAATATCAATAA														
	T *														
841															
901	AATAATAAAAAAATATAGGGTTAGGGTAAAAAACAAAAAACAAAAAACACAAAAAAAA														
961	AAAACACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA														
1021	CAGTGTTCTCATGGCCGATATTAATCTTTGCTTTGCCTAATGGAAATAATTAAAGGTACTT														
1081	AGTGGAGCACCACTACT														

4.1.3 S sequence of Bwamba serogroup virus.

(i)MPOV-complete sequence

1	AGTAGTGTACTCCACGTGTCTACTTTGGTAAAGTCTTGCTAAAATCTTGCAAATAAAAACTA														
61	GCTATCAAATTAAACTGTCAATCGCAATTCTGGGGAGCTTTGAACAATATCCAGGATGTC														
	StartNSs M S														
121	AGACCTTGTGCTCACATTTTCAGACTCCGACGATGTCAGTCGAAGTACTTATAACCCAGG														
	DLVLTFSDSDDVSRSTYNPG														
181															
101															
241															
211															
301															
301															
	TVETTELCCCDWITTTCIVE														
	LKFILEHSSSRWLITISLVE														
261	StopNSs														
301															
	A Q K E V A D S D L T L H K V S G I L A														
401															
421															
401															
481	AAGCAGAATCATCAACCCAATTTCTGCTAAGATGGGCTTCCAATGGAGTGTTGGGCCAGA														
	S K I I N P I S A K M G F Q W S V G P E														
541	GATTTACCTTTCTACTCTGCCTGGAACTGAAATGTTTCTGGATGCATTCAAAATGTATCC														
c 0 1	I I L S T L P G T E M F L D A F K M Y P														
601	TTTAGCATTCATCCTGCTTAGAGTAAAGAGGGGGTGAAATCAAGATTGACATGGCAAAAAA														
<i>c.c.</i>															
661	AGCTATGCGTCAGAGGTATGGTGACAAAGAGGCTTCAACCTGGCTTGAGGAAGAGGTCGA														
	A M R Q R Y G D K E A S T W L E E E V D														
721	CACTGTAAAAAGTGCAATCAAGACAGTTGAGAAGCTGAAGCCTACTCTCACTGGACTTGC														
	TVKSAIKTVEKLKPTLTGLA														
	StopN														
781	TGCAAGCATGACCAAGTTCCTCCAGGAGCTTGGTATTTCTAAA TAA ATTTCATATTCCAA														
	A S M T K F L Q E L G I S K *														
841	ATTAATATTTCATGTGTTATCTCCCTATTATATCCAATATATTATCCAATTTATCCATATA														
901	TAACCAATTTATAATCTTTTAGTTATCTTAAGTTTATTAATCATACTAATAATTTAGTTT														
961	TAAGTTTTCATCTAATCTTCATCTAATCATCCCCCATCCCATCCAATAACAGCGCAATA														
1021	AATCAATGTATTCAAGTTTGCATTAATTAATAGTAGTCGTGGAGCACACTACT														

4.1.4 S segment sequence of Turlock serogroup virus.

B S sequence without NSs protein

(i)ANAV-complete sequence

	StartN	
1	AGTAGTGTACTCCACATAAAGTCAAACTATCGCTTGTTGTAAAAAACATTCAAAATGTCA	C
	M S S	5
61	CCATGGATTATATGTTTGACGATCCTGGTCATATCCAGATCATCAACTATATACCAGAG	3
	M D Y M F D D P G H I Q I I N Y I P E H	2
121	AGGAATATATGGCATTTTGTGCAAAATATCAAGATCAACTGACAGTAGAGAACATTAGGA	ł
	EYMAFCAKYQDQLTVENIRI	E.
181	TTTTCTTCCTTAATGCAGGCAAATTCAAGAAAATGTTGAGGACAAATCCAAAATTGAAAA	3
	FFLNAGKFKKMLRTNPKLKV	7
241	TTCAAGTCAAATTTGGCACGCTTCAACTTGAGATCATCAATACTCACAATCCAGCAAATC	2
	Q V K F G T L Q L E I I N T H N P A N H	Ł
301	GTGAAATTCAATTGTCATCAAGTGACCTAACATTACATAGAGTCTCTGCCTATTTGGCT	ł
	EIQLSSSDLTLHRVSAYLAH	S.
361	GAAAATCCCTTCGACTTTGGGAAGCCCTAACATTGGATCAAAAAGAAGAATTCAAGCTAA	ł
	K S L R L W E A L T L D Q K E E F K L H	C
421	AGATCATCATGCCATTAGCAGAAAAAGCTGGAGTCACTTGGGATAATGCCGGATCTGTTG	3
	IIMPLAEKAGVTWDNAGSVI	2
481	AGATGTACCTAGGTTTTGCTCCAGGTGCTGAGTTCTTCTTAAAAGATTTCAAATTTTATC	2
	MYLGFAPGAEFFLKDFKFYI	2
541	CACTTGCAATTGGCATTGCTAGAGTTAAGAAAGGACTGATGAGCCCTGATTTCTTATCA	7
c 0.1	L A I G I A R V K K G L M S P D F L S H	C
601	AAACTCTGAGACAGAGATATGGAGGAATACCACCAGCTGACTGGATGGCTATTCACAAA	
c c 1	T L R Q R Y G G I P P A D W M A I H K S	5
66T	CAGCCGTGAGGACCGCCGTTGAATTTGTTGACAAATTCCCACTAATGAGATCCAATGCT	
701		2
/21		1
		4
701		
101		•
841		P
901		•
301	CAGIIGAATATICIAAAATTATGIACITIATGTGGAGCACACTACT	

(ii)TCMV-complete sequence

	StartN													
1	AGTAGTGTACTCCACTAGAAGCAAAGTTTGTAAAAGTTTAAAATGTCAGGAATAGATTTC													
	MSGIDF													
61	ATCTACGATGATCCTGGTAGGATCCAGCAATCAGACTTTGTTCCTAGAGAGGAGGAGTATGCT													
	IYDDPGRIQQSDFVPREEYA													
121	GTTTTCTGCTCAAAGTTTGCACTTGTCTTAATATTGCAAAACATACGGATTTTTTTT													
	V F C S K F A L V L I L Q N I R I F F L													
181	AATGCAGGGAAACTAAAAGCTGCACTTAAATTATCTCCCTAAAAGGAAAGTAAAGGCAAAA													
	N A G K L K A A L K L S P K R K V K A K													
241	TTTGGAACTCTGGAAGTGGAGATTACAAATACGCATAACAGGGCATTTGCTGACTCAAAA													
	FGTLEVEITNTHNRAFADSK													
301	TTGGAAAAAGGTGACATCACACTTCACAGACTTTCAGGATTTCTTGCAAAAAAGATTTTG													
	LEKGDITLHRLSGFLAKKIL													
361	GATATATACAATGCTGGCTCGTTTGATCTTCAAACACAGATTAAGGCTGAAATCATCATC													
	DIYNAGSFDLQTQIKAEIII													
421	CCATTAGCTGAGGTTGCTGGTGTCTCATGGGCTAATTCTACGCCAGAAATTTACCTAAGC													
	PLAEVAGVSWANSTPEIYLS													
481	TTTTGCCCAGGTTCAGAATTCTTTCTAGCAGATTTCAAATTTTATCCACTAGCAATTGGC													
	F C P G S E F F L A D F K F Y P L A I G													
541	ATAGCTAGGGTGAAAAAGGATCTAATGAAGCCAGAGTTTCTGGCAAAATTACTGAGGCAA													
	IARVKKDLMKPEFLAKLLRQ													
601	AGATATGGCAATATGACACCTGGAGAATGGATGAAAACACATTCCGTCTCAGTGCAGTCT													
	RYGNMTPGEWMKTHSVSVQS													
661	GCTATTTCTGAAATAGAGAAATACCCACTCATGAAGGTGAATTCCCGTCCTCATGTTGAG													
	A I S E I E K Y P L M K V N S R P H V E													
721	GAGTTCCTAAAGTCTCTAGGTCTATCTGCCCACTATATAAATGGGGGCTCTAAGAGGGTAA													
	E F L K S L G L S A H Y I N G A L R G *													
781	GCATAGCAAATACAGTCTGAATATCTAACACTAACATTAAATCACATATACTATAAACAC													
841	TAACATTAAATCACATACACTATAAGCACTACATACACTAAAATACTAACACAAGGGTGG													
901	GTGGTTGGGGACTGCATCTTATTAGCTGTTAGTTTTTGTCTTGATGTGGAGTACACTACT													

4.1.5 S segment sequence of Anopheles A serogroup viruses.

(i)ANBV-complete sequence

$\begin{array}{cccccccccccccccccccccccccccccccccccc$		StartN
M A S Q V D F A F 61 TTGAAGACACAGGAAATATCACTCAGTCTAGACTTTATTCCTGATGTGGGGTATACTGCAT E D T G N I T Q S D F I P D V G Y T A F 121 TCTGCCTGGGCAAGACAGCACATCTGTCTTTGAGAATATCAAGATATTCTTCTCCTAAATG C L G K T A H L S L E N I K I F F L N A 181 CTGGCAAATTAAAGCAGCAGATGAAAACGTGCTCAAAAACTAAAATAAAAGCAAAGTTTG G K L K Q Q M K T C S K T K I K A K F G 241 GCACGTTAGAAATTGAACTGATAATAGAAACTTAGAAGTCTCACTTC T L E I E L V N T L N R S L G Q V S L Q 301 AGCCGAATGATGTGGACACTCCACAGGGCATCCGCATATTGGCAAGAAAGCTCTGAAC P N D V T L H R A S A Y L A R K A L E L 361 TTTACCGGGAGGGCCAAGCTGATTTCCAAGCCGGAAGTCCACTTGGTT A E V A G V Q F K P E V P P E L Y I G F 481 TTGCCCTGGTGGTGGTGGTGTCAATTCAAGCCGGAAGTGCCACCGGCATATGGAG A P G A E F F M K L F K C Y P L A I A V 541 TTGTTGGTGTAAAGAAAGGGCAAATGCCAGCGGAATGCCACTAGGGCCTACGTCG A P G A E F F M K L F K C Y P L A I A V 541 TTGTTGGTGGTAAAGAAAGGGCAAATGCAGCCTGAATGCCGCTAAGGAGCCTACGTCAG A P G A E F F M K L F K C Y P L A I A V 541 TTGTTGGTGGTAAAGAAAGGGCAAATGCAGCCTGAATGCCGCCAACAGAGCCTACGTCAC V R V K K G Q M Q P E F L A K S L R Q R 661 GTTATGCTGGGGCTCCACCAGCAGGTGGATTCCTAGGTGCTCACGGCAATTAGGCAG Y A G A P P A E W M S T N A A A I K A A 661 CCACTGCCAATGGTGGAGAATACCCTATGCAAGGTTGCCACCAGCAATTAGGCAG Y A G A P P A E W M S T N A A A I K A A 661 CCACTGCCAATGGTGGAGAAATACCCTATGCAAGGTTGCCACCAGCAATTAAGGCAG Y A G A P P A E W M S T N A A A I K A A 661 CCACTGCCAATGGTGGAGAAATACCCTATGAAGGAGGTTGCCTCAAGGCCTCACATTGAAGGAGA Y A G A P P A E W M S T N A A A I K A A 661 CCACTGCCAATGGTGGAGAAATACCCTATGCAAGGTTGCGTCTAAGGCCTCACATTGAAGTAAGGCAGG Y A G A P P A E W M S T N A A A I K A A 7 A M V E K Y P M M K V A S R P H I E K StopN 721 AGCTTCTGTCCGAACTTGGACATTGGACACTATTATATTGAAACCAATTATAACCAATTGGAGCATATTGGGGGCGTTGGTGGGGGCTTCAAACAGCATATATAACCATATTATATATA	1	AGTAGTGTACTCCACATTATAGTAATAAAGTCAAG <mark>ATG</mark> GCGTCTCAAGTTGATTTTGCTT
61 TTGAAGACACAGGAAATATCACTCAGTCTGACTTTATTCCTGATGTGGGGTATACTGCAT E D T G N I T Q S D F I P D V G Y T A F 121 TCTGCCTGGGCAAGACAGCACACCTCTTGTGTGAGAATATCAAGATATTCATCTTCTTCATAATG C L G K T A H L S L E N I K I F F L N A 181 CTGGCAAATTAAAGCAGCAGAGACACACTCTTGTGGGCTCAAAAACCTAAAAACTAAAAACTAAAAAAGCAAAGCTCATTG G K L K Q M K T C S K T K A K F G 241 GCACGTATGGAAATGGACAGCACAGGCAGCGGCATAGAGGCAAGCTAAGGCAAAGCTCCACATGCAAGACTCCACACAGGCAATGGAGAAAGCTCATCCACTCACACAGAGAAAGCTCTCACTCA		M A S Q V D F A F
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	61	TTGAAGACACAGGAAATATCACTCAGTCTGACTTTATTCCTGATGTGGGGGTATACTGCAT
 121 TCTGCCTGGGCAAGACAGCACATCTGTCTTTGAGAATATCCAAGATATTCTTCCTAAATG C L G K T A H L S L E N I K I F F L N A 181 CTGGCAAATTAAAGCAGCAGATGAAAACGGCTCAAAAACTAAAATAAAAGCAAAGTTG G K L K Q Q M K T C S K T K I K A K F G 241 GCACGTTAGAAATGAACTAGTAAAACAATGAAACCTTAGGACAAGCTTCACTTC T L E I E L V N T L N R S L G Q V S L Q 301 AGCCGAATGAGACACTCCACAGGGCATCCGCATATTGGACAAGCTCTGAAC P N D V T L H R A S A Y L A R K A L E L 361 TTTACCGGGAGGGCCAGCTGATTCCAAGCAGGCCATGGAGGATCAATTGTATGCCTC Y R E G Q A D F Q A A M R D Q F V M P L 421 TTGCTGAGGTTGCTGGTGTTCAATCAAGCGGGAAGTGCCACCAGAGCTTTATATTGGTT A E V A G V Q F K P E V P P E L Y I G F 481 TTGCCCCTGGTGGCGAATTCTTCATGAAGCGCCTGAAGTGCCACCGCAAGTGCAATTGCAG A P G A E F F M K L F K C Y P L A I A V 541 TTGCTCGTGAAAAGGGCAAATGCAGCCTGAATTCCTAGGTGCTGCAGCTAACTGCAGC V R V K K G Q M Q P E F L A K S L R Q R 601 GTTATGCTGGGGCTCCACCAGCAGAGTGGAGGATGATCCTAGGCGCAACTGCAGCG A P P A E W M S T N A A A I K A A 661 CCACTGCAATGGGGAGAAATACCCTATGAAGGTTGCGTCTAGGCCTACCACATTGAAA T A M V E K Y P M M K V A S R P H I E K 721 AGCTTCTGTCCGAACTTGGACATTCCAGCTCCACATATACGGAGCCTACCACATTGAAA T A M V E K Y P M M K V A S R P H I E K 731 CCACAGTGCTCCACATTGGACATTCCAGCTCCCACATATAACGAGAGG A L S E L G L S A P T I Q Q A L T K * 741 CCACAGTGCTTCATCTCATACTGTAAATAGATATCCATATTAATCGGAACCATATAAGT 741 CCACAGTGCTTCATCTCATACTGTAAATAGATATCCATATTAATCGGAACCATATATAGT 741 CCACAGTGCTTCATCTCATACTTATATATATATATATAACAGCATATGCGTGGGGCGCTCCAACAGCATATAAGT 741 CCACAGTGCTTCATCTCATACTTATATATATATATATATA		E D T G N I T Q S D F I P D V G Y T A F
$ \begin{array}{ccccc} C & L & G & K & T & A & H & L & S & L & E & N & I & K & I & F & F & L & N & A \\ 181 & CTGGCAAATTAAAGCAGCAGATGAAAACGTGCTCAAAAACTAAAAATAAAAGCAAAGTTTG \\ G & K & L & K & Q & Q & M & K & T & C & S & K & T & K & I & K & A & K & F & G \\ 241 & GCACGTTAGAAATTGAACTAGTAAATACATTGAATCGAAGCTTAGGACAAGTCTCACTTC \\ T & L & E & I & E & L & V & N & T & L & N & R & S & L & G & Q & V & S & L & Q \\ 301 & AGCCGAATGATGTGACACTCCACAGGGCATCCCGCAATTTGGCAAGAAAAGCTCTTGAAC \\ P & N & D & V & T & L & H & R & A & S & A & Y & L & A & R & K & A & L & E & L \\ 361 & TTTACCGGGAGGGCCAAGCTGATTTCCAAGCAGCCATGAGAGAATCAATTTGTTATGCCTC \\ & Y & R & E & G & Q & A & D & F & Q & A & A & M & R & D & Q & F & V & M & P & L \\ 421 & TTGCTGAGGTTGCTGGTGTTCCAATCAAGCCGGAAGTGCCACCAGAGCTTTAATTGGTT \\ & A & E & V & A & G & V & Q & F & K & P & E & V & P & P & E & L & Y & I & G & F \\ 481 & TTGCCCCTGGTGCGGAATTCTCATGAAGCTCTTCAAGGTGCTACCCGCTAGGAATTGCAG \\ & A & P & G & A & E & F & F & M & K & L & F & K & C & Y & P & L & A & I & A & V \\ 541 & TTGTTCGTGTGAAAAAAGGGGCAAATGCAGCCTGAATTCCTAGGTAAGAGCCTACGTCAAC \\ & V & R & V & K & K & G & Q & M & Q & P & E & F & L & A & K & S & L & R & Q & R \\ 601 & GTTATGCTGGGGCTCCACCAGCAGAGTGGATGTCCACAAATGCTGCAGCAATTAAGGCAG \\ & Y & A & G & A & P & P & A & E & W & M & S & T & N & A & A & I & K & A & A \\ 661 & CCACTGCAATGGTGGAAATACCCTATGATGAAGAAGGTTGCGTCTAGGCCTCACAATTAAGGCAG \\ & Y & A & G & A & P & P & A & E & W & M & S & T & N & A & A & I & K & A & A \\ 661 & CCACTGCAATGGTGGAAATACCCTATGATGAAGAGGTTGCGTCTAGGCCTCACAATTAAGGCAG \\ & Y & A & G & A & P & P & A & E & W & M & S & T & N & A & A & I & K & A & A \\ 661 & CCACTGCAATGGTGGAAATACCCTATGGACTATCAGCTCCCACTATCAGGCTCAGGCCTACGACAATTAAGGCAGGC$	121	TCTGCCTGGGCAAGACAGCACATCTGTCTCTTGAGAATATCAAGATATTCTTCCTAAATG
181CTGGCAAATTAAAGCAGCAGATGAAAACGTGCTCCAAAAACTAAAATAAAAGCAAAGTTTG G K L K Q Q M K T C S K T K I K A K F G241GCACGTTAGAAATTGAACTAGTAAATACATTGAATCGAAGCTTAGGACAAGTCTCACTTC T L E I E L V N T L N R S L G Q V S L Q301AGCCGAATGAGACACGTGCACCACAGGGCATCCGCATATTGGCAAGAAAAGCTCTGGAAC P N D V T L H R A S A Y L A R K A L E L361TTTACCGGGAGGGCCAAGCTGATTTCCAAGCAGCCATGAGAGAACAAGCTCTGGACC Y R E G Q A D F Q A A M R D Q F V M P L421TTGCTGAGGTTGCTGGTGTCAATTCAAGCCGGAAGTGCCACCAGAGCTTTATATTGGTT A E V A G V Q F K P E V P P E L Y I G F481TTGCCCCGGGGGCGCGAAGTGCTCATGAGACGCCGAAGTGCCACCAGGCCTACGCAATTGCAG A P G A E F F M K L F K C Y P L A I A V541TTGTTCGTGTAAAAGAAAGGGCAAATGCAGCCTGAATTCCTAGGCACAATGCAGCCTACGTCAAC V R V K K G Q M Q P E F L A K S L R Q R601GTTATGCTGGGGCTCCACCAGCAGAGTGGATGTCCACAAATGCTGCAGCAATTAAGGCAG Y A G A P P A E W M S T N A A A I K A A661CCACTGCAATGGTGGAGAATACCCTATGATGAAGAAGGCTCCACATTAAGGCCTACGTCACAATGGTGGAGAATACCCTATGATGAAGGATGCGCCTCACATTAAGGCAG T A M V E K Y P M M K V A S R P H I E K StopN721AGCTTCTGTCCGAACTTGGACTATCAAGCTCCCACAATTCCAAGCAGGCTTAAACCAAGTGAG G L L S E L G L S A P T I Q Q A L T K *781CCACAGTGGCTCACATATATATATATATATAATAGGGGGGGG		C L G K T A H L S L E N I K I F F L N A
$ \begin{array}{c} \mathbf{G} \mathbf{K} \mathbf{L} \mathbf{K} \mathbf{Q} \mathbf{Q} \mathbf{M} \mathbf{K} \mathbf{T} \mathbf{C} \mathbf{S} \mathbf{K} \mathbf{T} \mathbf{K} \mathbf{I} \mathbf{K} \mathbf{A} \mathbf{K} \mathbf{F} \mathbf{G} \\ \mathbf{G} \\ \mathbf{G} \\ \mathbf{C} \\ \mathbf{C} \\ \mathbf{C} \\ \mathbf{C} \\ \mathbf{C} \\ \mathbf{A} \\ \mathbf{C} \\ \mathbf$	181	CTGGCAAATTAAAGCAGCAGATGAAAACGTGCTCAAAAAACTAAAATAAAAGCAAAGTTTG
241GCACGTTAGAAATTGAACTAGTAAATACATTGAATCGAAGCTTAGGACAAGTCTCACTTC T L E I E L V N T L N R S L G Q V S L Q301AGCCGAATGATGTGACACTCCACAGGGCATCCGCATATTTGGCAAGAAAAGCTCTTGAAC P N D V T L H R A S A Y L A R K A L E L361TTTACCGGGAGGCCAAGCTGATTCCAAGCAGCACATGAGAGAACAATTTGTTATGCCTC Y R E G Q A D F Q A A M R D Q F V M P L421TTGCTGAGGTGCTGGTGTCAATCAAGCCGGAAGTGCCACCAGAGCTTTATATGGTT A E V A G V Q F K P E V P P E L Y I G F481TTGCCCCTGGTGCGGAATTCTTCATGAAGCCGCAATGCCACCAGAGCTAATTGGAG A P G A E F F M K L F K C Y P L A I A V541TTGTTCGTGTAAAGAAAGGGCAAATGCAGCCTGAATTCCTAGCTAG		G K L K Q Q M K T C S K T K I K A K F G
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	241	GCACGTTAGAAATTGAACTAGTAAATACATTGAATCGAAGCTTAGGACAAGTCTCACTTC
$\begin{array}{rclcrcrcrcrcrcrcrcrcrcrcrcrcrcrcrcrcrcr$		T L E I E L V N T L N R S L G Q V S L Q
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	301	AGCCGAATGATGTGACACTCCACAGGGCATCCGCATATTTGGCAAGAAAAGCTCTTGAAC
$\begin{array}{rrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrr$		PNDVTLHRASAYLARKALEL
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	361	TTTACCGGGAGGGCCAAGCTGATTTCCAAGCAGCCATGAGAGATCAATTTGTTATGCCTC
421 TTGCTGAGGTTGCTGGTGTTCAATTCAAGCCGGAAGTGCCACCAGAGCTTTATATTGGTT A E V A G V Q F K P E V P P E L Y I G F 481 TTGCCCCTGGTGCGGAATTCTTCATGAAGCTCTTCAAGTGCTACCCGCTAGCAATTGCAG A P G A E F F M K L F K C Y P L A I A V 541 TTGTTCGTGTAAAGAAAGGGCAAATGCAGCCTGAATTCCTAGCTAAGAGCCTACGTCAC V R V K K G Q M Q P E F L A K S L R Q R 601 GTTATGCTGGGGCTCCACCAGCAGAGTGGATGTCCCACAAATGCTGCAGCAATTAAGGCAG Y A G A P P A E W M S T N A A A I K A A 661 CCACTGCAATGGTGGAGAATACCCTATGATGAAGGTTGCGTCTAGGCCTCACATTGAAA T A M V E K Y P M M K V A S R P H I E K StopN 721 AGCTTCTGTCCGAACTTGGACTATCAGCTCCCACTATTAATCGGAACCATATAGT 41 CCACAGTGCTTCATCATTGATGAAATAGCATATTAATCGGAACCATATAAGT 841 CCACAGTGCTTCATCATTCAATATATATATAGATATCCATATTAATCGGAACCATATAAGT 841 CAATATATACCATATATATATATATAGGGGGTGGTGGGGGCTTCAAACAGCATAAC 901 ATATGGGTGGGTGGTGGGGACACCACATATT		Y R E G Q A D F Q A A M R D Q F V M P L
A E V A G V Q F K P E V P P E L Y I G F TTGCCCCTGGTGCGGAATTCTTCATGAAGCTCTTCAAGTGCTACCCGCTAGCAATTGCAG A P G A E F F M K L F K C Y P L A I A V 541 TTGTTCGTGTAAAGAAAGGAGCAAATGCAGCCTGAATTCCTAGCTAAGAGCCTAAGGACCTACGTCAAC V R V K K G Q M Q P E F L A K S L R Q R 601 GTTATGCTGGGGCTCCACCAGCAGGAGTGGCAGTGTCCACAATGCTAGCAGCAATTGAGAGG Y A G A P P A E W M S T N A A A I K A A 661 CCACTGCAATGGTGGAGAAATACCCTATGATGAAGGTTGCGTCTAGGCCTCACATTGAAA T A M V E K Y P M M K V A S R P H I E K StopN 721 AGCTTCTGTCCGAACTTGGACTATCAGCTCCCACTATCCAGCAGTGAG L L S E L G L S A P T I Q Q A L T K * 781 CCACAGTGCTTCATCATTCATATATATAGATATCCATATTAATCGGAACCATATAGT 841 CAATATATAACCATATATATATATATAGGGGGTGGTGGGGGCTTCAAACAGCATAAC 901 ATATGGGTGGGTGGTGGGGGACAGCACTACT	421	TTGCTGAGGTTGCTGGTGTTCAATTCAAGCCGGAAGTGCCACCAGAGCTTTATATTGGTT
$\begin{array}{rcl} 481 & & \mbox{trgcccctggtgcggaattcttcatgaagctcttcaagtgctacccgctagcaattgcag} \\ A & P & G & A & E & F & F & M & K & L & F & K & C & Y & P & L & A & I & A & V \\ \hline 541 & & \mbox{trgctgtgtgaagaaggggcaatgccagattgcccgaattcctaggcctaagagcctacgctaacgcctacgctaag} \\ V & R & V & K & K & G & Q & M & Q & P & E & F & L & A & K & S & L & R & Q & R \\ \hline 601 & & \mbox{gtatgggggcaatgcccacagcggatgtcccacaatgctgcagcaattaaggcag} \\ Y & A & G & A & P & P & A & E & W & M & S & T & N & A & A & I & K & A \\ \hline 661 & & \mbox{ccactgcaatggtggaaatacccctatgatgaagatgtcccacaatggcgcattaaggcag} \\ Y & A & G & A & P & P & A & E & W & M & S & T & N & A & A & I & K & A \\ \hline 661 & & \mbox{ccactggtggagaatacccctatgatgaagatgtcccacaatggcgcctccacattgaaa} \\ T & A & M & V & E & K & Y & P & M & M & K & V & A & S & R & P & H & I & E & K \\ \hline & & \mbox{stopN} \\ 721 & & \mbox{agcttctggcccgaacttggactatcaggagctgcaggcgctttaaccatagtgag} \\ L & L & S & E & L & G & L & S & A & P & T & I & Q & Q & A & L & T & K & * \\ 781 & & ccacagtggtggggggggggggggggggggggggggggg$		A E V A G V Q F K P E V P P E L Y I G F
A P G A E F F M K L F K C Y P L A I A V TTGTTCGTGTAAAGAAAGGGCAAATGCAGCCTGAATTCCTAGCTAAGAGCCTACGTCAAC V R V K K G Q M Q P E F L A K S L R Q R 601 GTTATGCTGGGGCTCCACCAGCAGAGTGGAGTGTCCCACAATGCTGCAGCAATAAGGCAG Y A G A P P A E W M S T N A A A I K A A 661 CCACTGCAATGGTGGAGAAATACCCTATGATGAAGGTGCGTCTAGGCCTCACATTGAAA T A M V E K Y P M M K V A S R P H I E K StopN 721 AGCTTCTGTCCGAACTTGGACTATCAGCTCCCACTATCCAGCAGGCTTTAACCAAGTGAG L L S E L G L S A P T I Q Q A L T K * 781 CCACAGGTGCTTCATCATGTAAATAGATATCCATATTAATCGGAACCATATATAGT 841 CAATATAAACCATATATATATATAATAGGGGTGGTGGTGGGGCTTCAAACAGCATAAC 901 ATATGGGTGGGTGGTTGGGCACACTACT	481	TTGCCCCTGGTGCGGAATTCTTCATGAAGCTCTTCAAGTGCTACCCGCTAGCAATTGCAG
541 TTGTTCGTGTAAAGAAAGGGCAAATGCAGCCTGAATTCCTAGCTAAGAGCCTACGTCAAC V R V K K G Q M Q P E F L A K S L R Q R Q P E F L A K S L R Q R G R V K K G Q R Q P E F L A K S L R Q R G A P P A E W M S T N A A I K A <		A P G A E F F M K L F K C Y P L A I A V
V R V K K G Q M Q P E F L A K S L R Q R GTTATGCTGGGGGCTCCACCAGCAGAGTGGATGTCCACAAATGCTGCAGCAATTAAGGCAG Y A G A P P A E W M S T N A A A I K A A 661 CCACTGCAATGGTGGAGAAATACCCTATGATGAAGGTTGCGTCTAGGCCTCACATTGAAA T A M V E K Y P M M K V A S R P H I E K StopN 721 AGCTTCTGTCCGAACTTGGACTATCAGGCTCCACATCCAGGGGCTTAACCAAGTGAG L L S E L G L S A P T I Q Q A L T K * 781 CCACAGTGCTTCATCATACTGTAAATAGATATCCATATTAATCGGAACCATATATAGT 841 CAATATATAACCATATTATATATATAGGGGGTGGGGGGGG	541	TTGTTCGTGTAAAGAAAGGGCAAATGCAGCCTGAATTCCTAGCTAAGAGCCTACGTCAAC
 601 GTTATGCTGGGGCTCCACCAGCAGAGTGGATGTCCACAAATGCTGCAGCAATTAAGGCAG Y A G A P P A E W M S T N A A A I K A A 661 CCACTGCAATGGTGGAGAAATACCCTATGATGAAGGTTGCGTCTAGGCCTCACATTGAAA T A M V E K Y P M M K V A S R P H I E K StopN 721 AGCTTCTGTCCGAACTTGGACTATCAGGCTCCCACTATCCAGGCGGTGTAACCAAGTGAG L L S E L G L S A P T I Q Q A L T K * 781 CCACAGTGCTTCATCATGTGAAATAGATATCCATATTAATCGGAACCATATAAGT 841 CAATATAAACCATATATATATATATATAGGGGGGGGGGG		V R V K K G Q M Q P E F L A K S L R Q R
Y A G A P P A E W M S T N A A A I K A A CCACTGCAATGGTGGAGAAATACCCTATGATGAAGGTTGCGTCTAGGCCTCACATTGAAA T A M V E K Y P M M K V A S R P H I E K StopN 721 AGCTTCTGTCCGAACTTGGACTATCAGGTCCCACCTATCCAGGCAGG	601	GTTATGCTGGGGGCTCCACCAGCAGAGTGGATGTCCACAAATGCTGCAGCAATTAAGGCAG
 661 CCACTGCAATGGTGGAGAAATACCCTATGATGAAGGTTGCGTCTAGGCCTCACATTGAAA T A M V E K Y P M M K V A S R P H I E K StopN 721 AGCTTCTGTCCGAACTTGGACTATCAGCTCCCACTATCCAGCAGGCTTTAACCAAGTGAG L L S E L G L S A P T I Q Q A L T K * 781 CCACAGTGCTTCATACTGTAAATAGATATCCATATTAATCGGAACCATATTAAGT 841 CAATATATAACCATATATATATATATATAATAGGGGGGGG		Y A G A P P A E W M S T N A A A I K A A
T A M V E K Y P M M K V A S R P H I E K StopN721AGCTTCTGTCCGAACTTGGACTATCAGCTCCCACTATCCAGCAGGCTTTAACCAAGTGAG L L S E L G L S A P T I Q Q A L T K *781CCACAGTGCTTCATACTGTAAATAGATATCCATATTAATCGGAACCATATTAAGT 841641CAATATATAACCATATATATATATATAATAGGGGGGGGGG	661	CCACTGCAATGGTGGAGAAATACCCTATGATGAAGGTTGCGTCTAGGCCTCACATTGAAA
StopN 721 AGCTTCTGTCCGAACTTGGACTATCAGCTCCCACTATCCAGCAGGCTTTAACCAAGTGAG L L S E L G L S A P T I Q Q A L T K * 781 CCACAGTGCTTCATACTGTAAATAGATATCCATATTAATCGGAACCATATATAGT 841 CAATATAAACCATATATATATATATAGGGGGGGGGGGGG		TAMVEKYPMMKVASRPHIEK
721 AGCTTCTGTCCGAACTTGGACTATCAGCTCCCACTATCCAGCAGGCTTTAACCAAGTGAG L L S E L G L S A P T I Q Q A L T K * 781 CCACAGTGCTTCATCTCATACTGTAAATAGATATCCATATTAATCGGAACCATATATAGT 841 CAATATATAAACCATATATATATATATAATAGGGGGTGGTGGTGGGGGGCTTCAAACAGCATAAC 901 ATATGGGTGGGTGGTTGGGGGACAGCAGCAGCAGCTGTTATATTGACATATGCTTCGCTG 961 TAATATAATCTTTTATGTGGAGCACACTACT		StopN
L L S E L G L S A P T I Q Q A L T K * 781 CCACAGTGCTTCATCTCATACTGTAAATAGATATCCATATTAATCGGAACCATATATAGT 841 CAATATATAACCATATATATATATATAGGGGGGGGGGGG	721	AGCTTCTGTCCGAACTTGGACTATCAGCTCCCACTATCCAGCAGGCTTTAACCAAGTGAG
781 CCACAGTGCTTCATCTCATACTGTAAATAGATATCCATATTAATCGGAACCATATATAGT 841 CAATATATAAACCATATATATATATAATAGGGGGTGGTGGTGGGGGGCTTCAAACAGCATAAC 901 ATATGGGTGGGTGGTTGGGGGACAGCAACTGTTATATTCAAAATTGACATATGCTTCGCTG 961 TAATATAATCTTTTATGTGGAGCACCACTACT		LLSELGLSAPTIQQALTK*
841 CAATATATAACCATATATATATATAATAGGGGGTGGTGGTGGGGGGCTTCAAACAGCATAAC 901 ATATGGGTGGGTGGTTGGGGGACAGCAACTGTTATATTCAAAATTGACATATGCTTCGCTG 961 TAATATAATCTTTTATGTGGAGCACCACTACT	781	CCACAGTGCTTCATCTCATACTGTAAATAGATATCCATATTAATCGGAACCATATATAGT
901 ATATGGGTGGGTGGTTGGGGACAGCAACTGTTATATTCAAAATTGACATATGCTTCGCTG 961 TAATATAATCTTTTATGTGGAGCACCACTACT	841	CAATATATAACCATATATATATATAATAGGGGGTGGTGGTGGGGGCTTCAAACAGCATAAC
961 TAATATAATCTTTTATGTGGAGCACACTACT	901	ATATGGGTGGGTGGTTGGGGACAGCAACTGTTATATTCAAAATTGACATATGCTTCGCTG
	961	TAATATAATCTTTTATGTGGAGCACACTACT

(ii)BORV-complete sequence

	StartN																			
1	AGTA	GTG	TGC'	TCC	ACA	TCA	AAG	TAA'	TAA	AGT	CAC	AAT	GTC	GTC	TAC	TGT	GGA	TTT	CGC	AT
												М	S	S	т	V	D	F	Α	F
61	TTGA	AGA	CAC	AGG	GAA	TGT	CAC	TCA	GTC'	TGA	CTT	TGT	TCC.	AGA	TGT	GGG	GTA	TAC	TAC	СТ
	L	K	т	Q	G	М	S	L	S	L	т	L	F	Q	М	W	G	I	L	Ρ
121	TCTG	CAT	rgg	GAG	GTC	CAA	TCA	ACT	GTC	ACT	GGA	GAA	TAT	CAA	AAT	ATT	СТТ	CGT	CAA	TG
	С	I	G	R	S	N	Q	L	S	L	Е	N	I	к	I	F	F	v	N	Α
181	CTGG	AAA	GCT	AAA	AAT	GCA	AAT	GAA	AAG	CAG	TCT'	TAA	ACC.	AAG	AAT	CAA	AGC	TAA	ATT	TG
	G	K	L	к	М	Q	М	к	S	S	L	к	Р	R	I	к	Α	к	F	G
241	GAAC	ATT	AGA	AAT	TGA	AAT	GGT	AAA	CAC	ATT.	AAA'	TAA	GGC	TTT	GGG	CCA	AAT	AAG	ССТ	GC
	т	L	Е	I	Е	М	v	N	т	L	N	к	Α	L	G	Q	I	S	L	Q
301	AGCC	TAC	TGA	CCT	TAC	TCT	GCA	TAG	GGC	ATC	TGC	ATT	TTT	GGC	TAG	AAG	AGC	TCT	GGA	GA
	P	т	D	L	т	L	н	R	Α	S	A	F	L	Α	R	R	Α	L	Е	М
361	TGTA	CCA	IGG	GGG.	AGC	TGC	AGA	TTT	CCA	GGC.	AGC	TAT	GAG	GGA	CCA	ATT	TGT	GAT	GCC	TC
	Y	н	G	G	Α	Α	D	F	Q	Α	Α	М	R	D	Q	F	v	М	Р	L
421	TAGC	TGA	AGT	TTC	CGG	CGT	CCA	ATT	CAA	GCC	TGA	GGT	TCC	TCC	AGA	GCT	TTA	CAT	CGG	AT
	A	Е	v	S	G	v	Q	F	к	Ρ	Е	v	Ρ	Ρ	Е	L	Y	I	G	F
481	TTGC	TCC	CGG	AGC.	AGA	ATT	CTT	CAT	GAA	GCT.	ATT	CAA	ATG	CTA	CCC	TCT	TGC	CAT	TGC	TG
	A	Ρ	G	Α	Е	F	F	М	к	L	F	к	С	Y	Ρ	L	Α	I	Α	v
541	TGGT	TAG	AGT	GAA	AAA	AGG	GCA	GAT	GCA	ACC	CGA	ATT	CTC.	AGC	TAA	GTG	CCT	AAG	GCA	GC
	v	R	v	к	к	G	Q	М	Q	Ρ	Е	F	S	Α	к	С	L	R	Q	R
601	GGTT	TGG	TGG	AGC	ACC	TCC	AGC	TGA	ATG	GAT	GAC	CAC	TAA	TGC	TCC	TGC	AAT	CAA	GAT	TG
	F	G	G	Α	Ρ	Ρ	Α	Е	W	М	т	т	N	Α	Ρ	Α	I	к	I	Α
661	CAAC	TGC	AAT	GGT.	AGA	GAA	GTT	CCC	AAT	GAT	GAG	AAC	TGC.	ATC	CAG	GCC	TCA	TGT	TGA	GA
	т	Α	М	v	Е	к	F	Ρ	М	М	R	т	Α	S	R	Ρ	н	v	Е	к
																			St	opN
721	AGCT	TCT	CTC	AGA	GCT	GGG	CCT	TTC:	TAC'	TCC.	AAC	TAT	TCA	GAG	AGC	TCT	TAC	CAA	GTA	AG
	L	L	S	Е	L	G	L	S	т	Р	т	I	Q	R	Α	L	т	к	*	
781	CATA	ACC	TTC	ACA	TAT	ATC	TAG	CAG	TTA	GTC	TAT	ATA	ACC.	ATT.	ATA	ATC	CAT	ATT	TAG	AG
841	GTTT	TGG	GTA	GAT	ATT	AGG	GGT	CAG	GGG	GGT	GGG	ACT	GGG	CAG	CAA	GAG	АТА	TGG	GTG	GG
901	TGGT	TGG	GGA	CAG	CAT	TCA	AAA	CTA	CTG	CAA	ATA'	TAC	ATA	TGA	GAT	GCA	GAA	AAT	ACA	AT
961	CTAT	AAA	GTG	GAG	TAC	ACT	ACT													

4.1.6 S segment sequence of Anopheles B serogroup viruses.

(i)TETEV-complete sequence

1	AGTAGTGTACTCCACTGGATACAAAATCGTTAATACTGAGAATTTATAGATTGCCAATAG
61	
01	CAAAICAIACAAAAIGCCCAAAGGAGGAAAAAGCGAICCIGAACCAICCAI
101	
121	TGCCAGTGCCGCATTCGAAGATGATGATGAGATCGATTATCATTTTACAGCAGGGGATGATGG
1.01	
181	AGGTGCTCCATTCAATCCCATGACTGCTTATAAAAGGTTCATGGAAGTCCATGGCAAGGA
	GAPFNPMTAYKRFMEVHGKE
241	ACTAAACCTGCCCAATATCAAGGTCTTTTTTTTTTGAAGGCTCGCCAGGCAAAGGAGATAAT
	LNLPNIKVFFLKARQAKEIM
301	GCGGTCCAAGGCAAAGTCAGAAATGACTTTTACTTTCGGATCTCTGACTCTGACCTTCAA
	R S K A K S E M T F T F G S L T L T F K
361	GAACACCCATCATCCATCTAATCGTCACCTGGCTGTTGAGCAGGACGACTTGACTATCAA
	N T H H P S N R H L A V E Q D D L T I N
421	CAGAGTCACCGGATTTATGGCTCATGCTATTTTGCTGACTCATAGAGACCCAAAGCATAA
	RVTGFMAHAILLTHRDPKHK
481	GGATGCAGTTGAAAAAACTATCATTAATCCAATTGCGGATTCAAAAGGTGTGACCTGGAA
	DAVEKTIINPIADSKGVTWK
541	AGAAAGGGGCAAATGTATATCTTGTTTCTTCCCTGGAACTGAGATGTTCATGCTCGAGTT
	E R G K C I S C F F P G T E M F M L E F
601	CAAGTTCTTCCCGCTTGCAGTTGGATTAGCTCGCTGCCATAAGGAAAAAATGGACACCGA
	K F F P L A V G L A R C H K E K M D T E
661	ATACTTGAAGAAGCCAATGCGCCAGATGCTCACTGATGGAACAAAGGCACAGGTTTGGCT
	Y L K K P M R Q M L T D G T K A Q V W L
721	AGGTGCTAAGATTGAAGAGATCAGGAAGGCCTACAAGGTTTGCATGAACCTTAAATTTGT
	G A K I E E I R K A Y K V C M N L K F V
781	CAAGGCTGGGTTCTCTGAGGCTGCAAGAGAGTTCCTAAAGGAATTTGGATTGGACCAGGA
	KAGFSEAAREFLKEFGLDOD
	StopN
841	TAAGAACTAAACAATTCTCTTTTTAAGAGATAGGGGGTTGGTATTTATATACCAACTCCGAC
•	K N *
901	ΔΨ <u></u> Τ Τ Τ Τ Τ Τ Τ Τ Τ Τ Τ Τ Τ Τ Τ Τ Τ Τ Τ
961	ΔCTCΔTTCCΔTTCTTCTTCTTCTTCTCTCTTTCTCTCTTTΔTΔΔΔCTCGΔCTACACAC
1021	
	C1

(ii) BMAV-complete sequence

1 AGTAGTGTACTCCACAAAATCCATAAAATCGTTAATACTGAGAATTTATAGATTGCATATAG StartN 61 CAAACCATACAAAATGTCCCAAGGTGAAGAGGGGGGGGGG		
StartN 61 CAAACCATACAAAAATGTCCAAACAAAAAAAAGGGTAGATGAACCACAAAAAAAGGGTCCAACCAA	1	AGTAGTGTACTCCACAAAATACAAAATCGTTAATACTGAGAATTTATAGATTGCATATAG
 61 CAAACCATACAAAATGTCCAAAGTAAAAAGAGGTGAGTCAGAACCATCAATCA		StartN
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	61	CAAACCATACAAAATGTCCAAAGTAAAAAGAGGTGAGTCAGAACCATCAATCA
121 AGCTAGTGCCGCATTTGCTGAAGATGAGATGAGATTGATT		M S K V K R G E S E P S I S V A
ASAFAEDEIDYHFSAGDDG181AGGTGCTCCATTCAACCCTATGGCTGCATACAAAGAGTTCATGGAAACTCATGGCAAGGA GAPFNPMAAYKEFMETHGKD241TTTAACTGTCACCAACACAAAGGGTCTTTTTTTTGGAAGGCCCCCAGGCTAAAGAAATTAT LTNNRNVFFLKARQAKEIM301GAGGTCTAAAGCCAAATCTGAAATGGACTTTCACTTTTGGCAGCCCTGGCGCCCAGCCTTCAA RRAKSEMTFTFGSLTLTFKKEIMAAYXILTTFKKSKAKSSKAKSSKAKSSKAKSSKKKEKMTTFKKSKAKSKAKSSKAKSSKAKSSKAKSSKAKSSKAKSSKAKSSKAKSKAKKKSKAA <t< td=""><td>121</td><td>AGCTAGTGCCGCATTTGCTGAAGATGAGATTGATTACCATTTCTCAGCAGGGGATGATGG</td></t<>	121	AGCTAGTGCCGCATTTGCTGAAGATGAGATTGATTACCATTTCTCAGCAGGGGATGATGG
 181 AGGTGCTCCATTCAACCCTATGGCTGCATACAAAGAGTTCATGGAAACTCATGGCAAGGA G A P F N P M A A Y K E F M E T H G K D 241 TTTAACTGTCACCAACATAAGAGTCTTTTTCTTGAAGGCTCGCCAGGCTAAGAAATTAT L T V T N I R V F F L K A R Q A K E I M 301 GAGGTCTAAAGCCAAATCTGAAATGACTTTCACTTTTGGCAGCCTAACCCTGACCTTCAA R S K A K S E M T F T F G S L T L T F K 361 GAACACTCATCACCATCTAATGGCATTGGCCATGGCCAGGATGATTGGACTATCAA N T H H P S N R H L T V D Q D D L T I N 421 TAGGGCCACTGGGTTTATGGCCTACGCAATCCTCCTGACCAGGAGGGGGGGG		A S A A F A E D E I D Y H F S A G D D G
G A P F N P M A A Y K E F M E T H G K D 241 TTTAACTGTCACCAACATAAGAGTCTTTTTCTTGAAGGCTCGCCAGGCTAAAGAAATTAT L T V T N I R V F F L K A R Q A K E I M 301 GAGGGTCTAAAGCCAAATCTGACATTGACATTTGACTATTTGGCAGCCTAAAGCCAACCCTGCAAATGGCCAGGGGGGGG	181	AGGTGCTCCATTCAACCCTATGGCTGCATACAAAGAGTTCATGGAAACTCATGGCAAGGA
241 TTTAACTGTCACCAACATAAGAGTCTTTTTCTTGAAGGCTCGCCAGGCTAAAGAAATTAT L T V T N I R V F F L K A R Q A K E I M 301 GAGGTCTAAAGCCAAATCTGAAATGACTTTCACTTTTGGCAGCCTAACCCTGACCTGACCTGACCTAACCCTGACCACGACCTTCAA R S K A K S L T L T F K A K E I M T F F F G S L T L T F K A K S L T L T F K A K S L T L T F K S L T L T L T L T L T L T I N N T H P S N N L T I N N L T I N N L T I N N		G A P F N P M A A Y K E F M E T H G K D
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	241	TTTAACTGTCACCAACATAAGAGTCTTTTTCTTGAAGGCTCGCCAGGCTAAAGAAATTAT
301 GAGGTCTAAAGCCAAATCTGAAATGACTTTCACTTTTGGCAGCCTAACCCTGACCTTCAA R S K A K S E M T F T F G S L T L T F K 361 GAACACTCATCACCCATCTAATCGGCATTTGACCAGGATGACCAGGATGATTTGACTATCAA N T H H P S N R H L T V D Q D D L T I N 421 TAGGGCCACTGGGTTATGGCCTACGCAATCCTCTGGACACCAAAGGGAGTGAAAGGACCACTCATCAATCCTATCTGAGCTGCAAATGGAAAGGAGGGGGAGGTGAAAGGACCAGGAAA R A T G N K I L L T H R E P K S 421 TAGGGCCACTGAGATGACACACCACCACCACCACACTCAACCTCAACCCAGAAAGGGCGTGACAAAGGGGGGGG		L T V T N I R V F F L K A R Q A K E I M
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	301	GAGGTCTAAAGCCAAATCTGAAATGACTTTCACTTTTGGCAGCCTAACCCTGACCTTCAA
361 GAACACTCATCACCCATCTAATCGGCATTTGACCGTTGACCAGGATGATTTGACTATCAA N T H H P S N R H L T V D Q D L T I N 421 TAGGGCCACTGGGTTTATGGCCTACGCAATCCTCCTGACTCATAGGAGCCTAAAAGCAA R A T G F M A Y A I L L T H R E P K S K 481 GGATGCAGTTGAAAAGACCATCATCATCATCCATCTGAATCCTGGAATGAGGGTGTTACCTGGAA D A V E K K G V W K S K 541 GGAGGGAGCAAATATCTATCTGAGGTTCTTCCCTGGGCAGATGATGAACGACATGATCATGGTTGATT E G A N I Y L S F F P G T E M F M L E F 541 GGAGGGAGCAAATATCTATCTGGCTGTGGCTGGCTGGCAGATGCCACAGAAAGATGGCACATGA G T L K F F G T E M L L E F		R S K A K S E M T F T F G S L T L T F K
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	361	GAACACTCATCACCCATCTAATCGGCATTTGACCGTTGACCAGGATGATTTGACTATCAA
421 TAGGGCCACTGGGTTTATGGCCTACGCAATCCTCCTGACTCATAGGGAGCCTAAAAGCAA R A T G F M A Y A I L L T H R E P K S K 481 GGATGCAGTTGAAAAGACCATCATCAATCCTATTGCAGAATCAAAGGGTGTTACCTGGAA D A V E K T I N P I A E S K G V T W K 541 GGAGGGAGCAAATATCTATCTGGGCTCTGGCTCGGGCAGACTGAAATGTTCATGGACTGGAATGCCACAAAAGATGGACACTGA E G A N L E F P G T E M F M L E F 601 CAGGTTCTTCCCACTAGCTGGTCGGCTCGGCTGGCTAGACGCACAAAGGACTGGAACAAGGCCCAAGGAAGG		N T H H P S N R H L T V D Q D D L T I N
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	421	TAGGGCCACTGGGTTTATGGCCTACGCAATCCTCCTGACTCATAGGGAGCCTAAAAGCAA
 481 GGATGCAGTTGAAAAGACCATCATCATCCTATTGCAGAATCAAAGGGTGTTACCTGGAA D A V E K T I I I N P I A E S K G V T W K 541 GGAGGGAGCAAATATCTATCTGAGCTTCTTCCCTGGGACTGAATGTTCATGCTTGAATT E G A N I Y L S F F P G T E M F M L E F 601 CAGGTTCTTCCACTAGCTGTTGGTCTGGCTAGATGCCACAAAGAAAAGATGGACACTGA R F F P L A V G L A R C H K E K M D T E 661 ATACCTGAAGAAGCCCATGCGCTCAAATGTTGACTGATGGAACAAAGGCTCAAGTCGGCT Y L K K P M R Q M L T D G T K A Q V W L 721 CGGTGCCAAGATTGAAGAGACCAGGAAGCACTAGGGAGGCATATAAGGTTGATGAACCATGAATTGTG G A K I E E I R K A Y K V C M N L K F V 781 TAAAGCTGGTTTCCTGAGGCTGCTAGGGAGGAGGAGGATGGAATTCGGATTAGACCAGGA K A G F S E A A R E F L K E F G L D Q D StopN 		R A T G F M A Y A I L L T H R E P K S K
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	481	GGATGCAGTTGAAAAGACCATCATCAATCCTATTGCAGAATCAAAGGGTGTTACCTGGAA
541 GGAGGGAGCAAATATCTATCTGGGCTTCTTCCCTGGGACTGAATGTTCATGCTTGAATT E G A N I Y L S F F P G T E M F M L E F 601 CAGGTTCTTTCCACTAGCTGTTGGTCTGGCTAGATGCCACAAAGAAAG		D A V E K T I I N P I A E S K G V T W K
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	541	GGAGGGAGCAAATATCTATCTGAGCTTCTTCCCTGGGACTGAAATGTTCATGCTTGAATT
601 CAGGTTCTTTCCACTAGCTGTTGGTCTGGTCTGGCTAGATGCCACAAGAAAAGATGGACACTGA R F F P L A V G L A R C H K E K M D T E 661 ATACCTGAAGAAGCCCATGCGTCAAATGTTGACTGATGGACAAAAGGCTCAAAGGCTCAAAGGCCTCAAAGGCCTGAGAGAGA		EGANIYLSFFPGTEMFMLEF
R F F P L A V G L A R C H K E K M D T E 661 ATACCTGAAGAAGCCCATGCGTCAAATGTTGACTGATGGAACAAAGGCTCAAGTCTGGCT Y L K K P M R Q M L T D G T K A Q V W L 721 CGGTGCCAAGATGAAGAGATCAGGAAGGCATATAAGGTTGTATGAACCTTAAATTGT G A K I E E I R K A V V W L K F V 781 TAAAGCTGGTTTCTCTGAGGCTGCTAGGGGAGTCCTTAAAGAATTCGGATTAGACCAGGA K K F S C M N L K F V 781 TAAAGCTGGTTTCTCTGAGGCTGCTGATGGGGGAGTCCTTAAAGAATTCGGATTAGACCAGGAA K K F S L A R F L K L D Q D 781 TAAAGCTGAAACTAAATAATTCTCTTTTAAGAGATAGGGGGTTGGTATATTTACCAACCCCGGAAACAGGAAACTAAAATTCTCTTTTTAAGAGGATAGGGGGTTGGTATAATTTACCAACCCCGA	601	CAGGTTCTTTCCACTAGCTGTTGGTCTGGCTAGATGCCACAAAGAAAG
661 ATACCTGAAGAAGCCCATGCGTCAAATGTTGACTGATGGAACAAAGGCTCAAGTCTGGCT Y L K K P M R Q M L T D G T K A Q V W L 721 CGGTGCCAAGATGAAGAGATCAGGAAGGCATATAAGGTTGTATGAACCTTAAATTTGT G A K I E E I R K A V V W L 781 G A K I E E I R K A V C M N L K F V 781 TAAAGCTGGTTTCTCTGAGGCTGCTAGGGGAGTTCCTTAAAGAATTCGGATTAGACCAGGA K A G F S E A R E F L K D Q D 781 TAAAGCTGATTGTTCTCTGAGGCTGCTCATAGGGGGAGTCCTTAAAGAATTCGGATTAGACCAGGA K A G F S E A R E F L L D Q D StopN S41 CAGAAACTAAATAATTCTCTTTTAAGAGATAGGGGGTTGGTATATTTTACAACCCCGACAT </td <td></td> <td>R F F P L A V G L A R C H K E K M D T E</td>		R F F P L A V G L A R C H K E K M D T E
Y L K K P M R Q M L T D G T K A Q V W L CGGTGCCAAGATTGAAGAGATCAGGAAGGCATATAAGGTTTGTATGAACCTTAAATTTGT G A K I E E I R K A Y K V C M N L K F V 781 TAAAGCTGGTTTCTCTGAGGCTGCTAGGGAGTTCCTTAAAGAATTCGGATTAGACCAGGA K A G F S E A A R E F L K E F G L D Q D StopN 841 CAGAAACTAAATAATTCTCTTTTAAGAGATAGGGGTTGGTATATTTACCAACCCCGACAT	661	ATACCTGAAGAAGCCCATGCGTCAAATGTTGACTGATGGAACAAAGGCTCAAGTCTGGCT
721 CGGTGCCAAGATTGAAGAGATCAGGAAGGCATATAAGGTTTGTATGAACCTTAAATTTGT G A K I E E I R K A Y K V C M N L K F V 781 TAAAGCTGGTTTCTCTGAGGCTGCTAGGGAGTTCCTTAAAGAATTCGGATTAGACCAGGA K A G F S E A A R E F L K E F G L D Q D StopN 841		Y L K K P M R Q M L T D G T K A Q V W L
G A K I E I R K A Y K V C M N L K F V 781 TAAAGCTGGTTTCTCTGAGGCTGCTAGGGAGTTCCTTAAAGAATTCGGATTAGACCAGGA K A G F S E A R E F L K E F G L D Q D StopN 841 CAGAAACTAAATAATTCTCTTTTTAAGGAGTAGGGGTTGGTATATTTACCAACCCCGACAT CAGAAACTAAATAATTCTCTTTTAAGAGATAGGGGTTGGTATATTTACCAACCCCGACAT	721	CGGTGCCAAGATTGAAGAGATCAGGAAGGCATATAAGGTTTGTATGAACCTTAAATTTGT
 781 TAAAGCTGGTTTCTCTGAGGCTGCTAGGGAGTTCCTTAAAGAATTCGGATTAGACCAGGA K A G F S E A A R E F L K E F G L D Q D StopN 841 CAGAAACTAAATAATTCTCTTTTAAGAGATAGGGGTTGGTATATTTACCAACCCCGACAT 		G A K I E E I R K A Y K V C M N L K F V
K A G F S E A A R E F L K E F G L D Q D StopN 841 CAGAAACTAAATAATTCTCTTTTTAAGAGATAGGGGTTGGTATATTTACCAACCCCGACAT	781	TAAAGCTGGTTTCTCTGAGGCTGCTAGGGAGTTCCTTAAAGAATTCGGATTAGACCAGGA
StopN 841 CAGAAACTAAATAATTCTCTTTTAAGAGATAGGGGTTGGTATATTTACCAACCCCGACAT		K A G F S E A A R E F L K E F G L D O D
841 CAGAAACTAAATAATTCTCTTTTTAAGAGATAGGGGTTGGTATATTTACCAACCCCGACAT		StopN
	841	CAGAAACTAAATAATTCTCTTTTTAAGAGATAGGGGTTGGTATATTTACCAACCCCGACAT
R N *		RN*
901 GTAAACAGCTAAATTAAGCTGCAATATGTGGTGGTGGTTGGGGCAAGATTCAGATCAAC	901	GTAAACAGCTAAATTAAGCTGCAATATGTGGTGGGTGGTTGGGGGCAAGATTCAGATCAAC
961 TCATTACAATCTTTTTCCTCATCATTAGTTCTTATGTATCCAGTGGAGCACACACTA	961	TCATTACAATCTTTTTCCTCATCATTAGTTCTATTCTTATGTATCCAGTGGAGCACACTA
1021 CT	1021	СТ

4.1.7 S segment sequence of Tete serogroup viruses.

(i)CAPV-incomplete sequence

	StartN
1	AGTAGTGAACTTCGTAGGAAGTTAAAATCATTCCAGAGAATAAATCAAT <mark>ATG</mark> GGAGAAGT
	MGEV
61	TGCTAATTTTGAATTTGGTGAGCGTTTGATAGCAGAAGGGAGGAACCCTTTTGTACCTGA
	ANFEFGERLIAEGRNPFVPD
121	TGTAGCCTATGCTCGGTTCATCAGAGAGCACGTACCTATGCTTACGCCTGGGAATATCAG
	V A Y A R F I R E H V P M L T P G N I R
181	GATCTTTTTCCTGCGTGCATATGATGCAAAGCAGAAACTTAAGACCACAACAGCTCGGAC
	I F F L R A Y D A K Q K L K T T T A R T
241	TGCCAACCTTAAATTTGGCACTGCTACGTTCACTGTGAAAAATAACCACAATGAGCGAAA
	ANLKFGTATFTVKNNHNERN
301	CGCCAATATTGAGGTTGAAGCTGAAGACCTGACTTTGCATCGGATCTCTGGATTCCTTGC
	A N I E V E A E D L T L H R I S G F L A
361	AAAGTTCGTGAGAGAGTCTATGGTGGACGAGACTGCATCTAAGATGATCAGGGACACAAT
	K F V R E S M V D E T A S K M I R D T I
421	CGTGAACCCGATTGCTGAAAGTCTAGGTATAACCTGGGAAAATGGGGATGATGTATATCT
	V N P I A E S L G I T W E N G D D V Y I
481	CTCATTCTTCCCTGGCACAGAGATGTTCTTGGACACCTTCCACATGCTACCTCTGGCTAT
	S F F P G T E M F L D T F H M L P L A I
541	TGGAATCTACAGAGTCCAGCAAAAGCAGATGAAGGCTGAGTTCCTGAAGAAGCACCTGCG
	GIYRVQQKQMKAEFLKKHLR
601	CCAACAATATGGAGGACTTGCTGCTTCCCAGTGGATGTCAGAAAGGAAAGAAGAAGATGTCAA
	Q Q Y G G L A A S Q W M S E R K E D V K
661	GCTTGCCTTAACATTAATTTCCAGGCTGCCCTGGGGAAAAGCTGGGCTCTCTGCTGCAGC
	LALTLIS RLPWGKAGLSAAA
	StopN
721	TAGAAATTTCCTAGCTGACTTCGGTATAACCTTT <mark>TAA</mark> TATCTCCCAGCTTTTTGTCTTGT
	RNFLADFGITF*
781	CCCCCAAATCCATTTTAATCAAAATTTAAAACCGAAATTAATGATCCAATGGTAAGTTCA
841	AATGGTAAGTTCTAATGGTAAGTCAATGGTAAGTTTCAAAAACCTAAAATTCAAAATTTTG
901	AAATTAAAATAAGTAAATAATTCGCCAAATGGCAAAACAATAACAGCAAAAGGGGTGGCG
961	ATAAGGGGGAACACTGTCACAGTAGGGTACT

4.1.8 S segment sequence of Capim serogroup virus.

(i)PLSV-incomplete sequence

	Sta	irtl	I .																	
1	AT	GAA'	rga	CCC	GAT	GCC	CTC	TTT	TGA	GTT'	TAA	GGA	GGA	GTC	CGC	TGC	AGT	AGG	GAG	TAAT
	М	N	D	Ρ	М	Ρ	S	F	Е	F	К	Е	Е	S	A	Α	v	G	S	N
61	CC!	TTT:	TAT'	TCC	GGA	FGA	AGC	CTA'	rggi	AGT'	TTT	CGT	TAA	CGA	AAA	CCA	GGC	CCA	ATT	GACC
	Ρ	F	I	Ρ	D	Е	Α	Y	G	v	F	v	N	Е	N	Q	Α	Q	L	т
121	AT	CCC	GAA'	TAT	TAG	GAT'	TTT(CTT	CCT	GCG	AGC	GTA	TGA	CGC	CAA	ACT	TAA	ACT	CAA	GAGA
	I	Ρ	N	I	R	I	F	F	L	R	Α	Y	D	Α	к	L	к	L	к	R
181	AC	GGA	AGC	AAG	GTC	IGC	GAA	TCT	CAC	ATT	TGG	AAC	TAT	GAA	GCT	GCA	TAT	AAT	GAA	TAAC
	т	Е	Α	R	S	Α	N	L	т	F	G	т	М	к	L	н	I	Μ	N	N
241	CAO	CAA	CAA	CAA	GTT	FGC	AAA	CAC	TGA	ACT	CGA	ACC	CCA!	IGA	гст	GAC	ССТ	CCA	CAG	AGTG
	н	N	N	К	F	Α	N	т	Е	L	Е	Ρ	н	D	L	т	L	н	R	v
301	TC	rgg	CTT	TTT	GGC	AAG	ATT	CCT	GAA	AGA	GGC	TAT	GGC'	FAA	CAC	TGC	AAC	CCG	CAT	GGAC
	S	G	F	L	Α	R	F	L	К	Е	Α	М	Α	N	т	Α	т	R	М	D
361	AT	CCA	CAA	GAA	CAT	CAT	AAA	CCC	GAT'	rgc/	AGC	AGC	TCT	GAA	ACT	GGA	ATG	GAC	AGT	TGGA
	I	н	к	N	I	I	N	Ρ	I	Α	Α	Α	L	к	L	Е	W	т	v	G
421	GA	CGA	CAT	CTA	TCT	CTC	ATT	CTT	CCC!	rgg	CGC	TGA	GAT	GTT(CAT	GGA	CAG	ATT	CAA	AATG
	D	D	I	Y	L	S	F	F	Ρ	G	Α	Е	М	F	М	D	R	F	K	М
481	TT	CCC	CCT	GGC	CAT	FGC	AAT	CTT	CAG	AGT	CCA	ACA	GAA	GCA	GAT	GGA	CCC	AGA	GTT	TCTA
	F	Ρ	L	Α	I	Α	I	F	R	v	Q	Q	К	Q	М	D	Ρ	Е	F	L
541	AA	GAA	ACA	CCT	GCG	CCA	GTA	CTG	CGGi	AGC	CAT	CCC	ATC	CTC	FCA	GTG	GAT	GCT	GAC	TACG
	К	к	н	L	R	Q	Y	С	G	Α	I	Ρ	S	S	Q	W	М	L	т	т
601	AA'	TTG	СТТЛ	ACT	GAG	CGA	ATA	TTT	CAA	ACA	CTA	CAT	FGC	ATT	ACA	TTG	СТТ	TAA	TTG	GTCA
	N	С	L	L	S	Е	Y	F	к	н	Y	I	Α	L	н	С	F	N	W	S
661	AA																			

Patois serogroup

(i)PATV-incomplete

1	AAGC	ACC	AAG	AGC	TAC	CATI	GCC	CTT	TCO	TGF	TTG	GAG	AGI	TGA	GGI	TGT	CAA	CAA	CCA	TA
	Α	Ρ	R	A	т	L	R	F	R	D	W	R	v	Е	v	v	N	N	Н	N
61	ACCC	AAG	ACT	GGG	TCA	AGI	TGF	AG	TCT	ACC	GAC	AGA	ACI	TAC	TCI	ACA	TCG	CAT	ATC	AG
	Р	R	L	G	Q	v	Е	v	L	Р	т	Е	L	т	L	н	R	I	S	G
121	GTTT	CAT	TGC	AAG	ACA	ACC1	TCT	TG	AA	TGC	CAA	TGG	AGI	TGA	CCA	GGC	CCC	CAT	'AAT	тG
	F	I	Α	R	н	L	L	E	I	Α	N	G	v	D	Q	A	R	I	I	D
181	ATAT	GCA	AGA	GCA	GAI	TGC	TAP	TCO	CATT	rggc	CACI	CAG	CAA	GGG	GAAT	CAC	СТС	GAG	TGA	TG
	М	Q	Е	Q	I	A	N	Р	L	Α	L	S	K	G	I	т	W	S	D	G
241	GAGC	AAC	CAT	ATA	CTI	GTC	CAT1	CTT	rcco	CAGO	SAAC	AGA	AAT	GTI	TCI	'AGA	TCI	GTI	'CAA	GT
	A	т	I	Y	L	S	F	F	Р	G	т	Е	М	F	L	D	L	F	к	F
301	TCTA	CCC	ACT	TGC	CAI	CGG	GAI	CTI	ACAG	GAGI	TCA	GAT	CGG	AGA	CAI	GGA	TGC	TGA	ATA	тс
	Y	Ρ	L	A	I	G	I	Y	R	v	Q	I	G	D	М	D	A	Е	Y	L
361	TAAA	GAA	AGC	TTT	AAG	ACA	ACA	ATZ	ATGO	TGC	TCI	GAA	ACC		TGA	ATG	GGGC	:AA1	CCT	GA
	K	К	A	L	R	Q	Q	Y	G	Α	L	K	Р	N	Е	W	A	I	L	K
421	АААА	AGA	TGC	AGT	CAA	AAC	AGC	TG	TATO	CTCF	TGI	'GGC	AGC	ACI	CAA	ATG	GGGG	ATC	CAA	AG
	K	D	A	v	к	т	Α	v	S	н	v	Α	Α	L	K	W	G	S	к	G
																	St	opN	I	
481	GTTT	CTC	TCA	GGC	TGC	CACA	AGA	CCI	TCT	CAF	AGA	ATT	TGG	GAI	TAP	GGI	CTA	ATT	TTG	ТΑ
	F	S	Q	A	Α	Q	D	L	L	K	Е	F	G	I	K	v	*			
541	AATA	TTT	CAC	GTA	GGA	AAC	ATA	GG	TAT	ATT]	AGG	TTT	AAG	TTA	ATC	GTA	AGI	CAA	TGG	TA
601	GGTT	TAA	ATC	AAT	TGI		TAT	TAT	ATT	AGG	GTA	ATT	TTA	ATT	AGG	GTI	AGG	GCA	AAA	AC
661																				
••-	AGCA	TCA	GTT	ATT	TTA	GCA	GTI	'AA'	rggo	GTG	GGT	GGI	'AGG	GGA	ACA	AAA	CTA	ATC	TGA	ΤТ

4.1.10 S segment sequence of Patois serogroup virus.

(i)BERV-incomplete sequence

1	TCAT	CAA	CCC	AAT	TGC	TCA	GTC	AGI	AGG	GAI	AAC	ATG	GGA	AAA	CGG	TGA	TGA	TGT	CTA	CC
	I	N	Ρ	I	Α	Q	S	v	G	I	т	W	Е	N	G	D	D	v	Y	L
61	TTTC	CTT	TTT	ccc	AGG	AAC	AGA	AAT	GTI	TCT	GGA	AAA	ATT	CAA	GCT	GTT	GCC	TCT	GGC	AA
	S	F	F	Р	G	т	Е	М	F	L	Е	к	F	к	L	L	Р	L	Α	R
121	TCGG	AAT	СТА	TCG	TGI	TCA	GCA	AAA	ACA	CAT	GGA	TGC	TGA	CTT	CCT	TAA	AAA	ACA	CCT	cc
	I	G	I	Y	R	v	Q	Q	Κ	н	М	D	A	D	F	L	к	Κ	Н	L
181	GTCA	GCA	АТА	TGG	GGA	CAT	ACC	TGC	TTC	ACA	GTG	GAT	GCT	GAC	AAT	GAA	AGA	AGA	TGT	CA
	Q	Q	Y	G	D	I	Ρ	Α	S	Q	W	Μ	L	т	М	к	Е	D	v	K
241	AAAG	TGC	ACT	GAC	TTT	GAT	TTC	CAA	GCI	'ACC	TTG	GGG	CAA	AGC	AGG	CAT	GTC	TGC	TGC	AG
	S	Α	L	т	L	I	S	к	L	Ρ	W	G	к	Α	G	М	S	Α	Α	A
												St	opN	1						
301	CCAG	AGC	СТТ	ССТ	TGC	TGA	CTI	TGG	AAT	CAC	AAT	'G <mark>TA</mark>	AAA	ACC	CAC	CCT	ATT	TGT	AAG	TA
	R	Α	F	L	Α	D	F	G	I	R	М	*								
361	GATT	AGA	TTC	TAT	AAT	TTA	GTI	TAT	GAA	ATA	ACC	ATA	GGA	TTT	ACA	GTT	TAG	TTT	GTA	AG
421	GTAA	CTA	TAG	TAT	AGG	ATA	AGI	'GA'I		TAA	TGG	TAA	GAA	AAT	GAT	ACA	AAA	TGG	TAT	TA
481	ATGG	TAA	GGT	TTT	TAT	TTT	TTT	'ACC	TTA	AAT	TAC	AAA	TTA	ATT	GAT	TCC	CCC	CCT	ACC	CA
541	TATT	TTA	ATT	AAA	ATC	CGC	TCC	CAC	CGC	TCC	CTC	:ccc	AAA	TAC	CCA	CAT	CCC	AAA	TTC	ΤG
601	AAGC	AGT	CCA	TAA	AAG	GGT	TGG	GTG	GTI	'GGG	GAA	CGA	GAC	AAA	CGA	ATT	TGT	TGT	GTT	АΤ
661	AATC	TTT	ACA	TTC	TTT	'AAC	AAC	TTA	AAA	GGG	TCT	'G								

(ii)GMAV-incomplete sequence

1	CCAG	GAAC	CAGA	AAT	GTT	TTT	rgga	AGA	CTI	CAC	AAT	GAI	cco	GTI	GGC	CAAT	TGG	GAAT	TT	
	P G	т	Е	Μ	F	L	Е	D	F	R	М	I	Р	L	Α	I	G	I	Y	
61	ACAG	AGTT	CAA	CAG	GAAG	CAC	ATG	AAA	GCA	GAA	TTC	сто		AAG	CAC	ATC	CGC	CAG	CAZ	٩T
	R	v	Q	Q	к	Q	М	к	Α	Е	F	L	к	к	н	М	R	Q	Q	Y
121	ATGG	AGAC	ATG	CCJ	GCF	AGC	CAG	TGG	ATC	ACT	CTA	AAA		GCA	GAJ	GTT	CAC	JAAT	GC	ГТ
	G	D	М	Р	Α	S	Q	W	М	т	L	к	к	Α	D	v	Q	N	Α	ь
181	TGAC	ATT	GTG	TCT	TAAC	CTT	rccc	TGG	ACA	AGA	GCI	GGF	TTC	TCT	GCI	GCZ	GCC	AGA	AA	СТ
	т	L	v	S	к	L	Р	W	т	R	Α	G	L	S	Α	A	Α	R	N	F
									Sto	pN										
241	TCTT	GGCI	GAA	TTT	GGG	CATC	CACC	TTG	TAA	TCI	CAC	TCC	CCI	TTG	сто	ATA	AGGT	AAC	AG	ГТ
	L	Α	Е	F	G	I	т	L	*											
301	TCCG	ATA	TAT	TAT	ACCO	CATA	AGGA	ccc	:AA1	GGI	AAG	TTT		TGA	TAZ	GTT	TTA	AAT	'GG'	٢A
361	AGTC	TCAZ	TGG	TAF	AGCI	CTF	ATG	GTA	AGI	TTT		TCF	AAC	TGT	ACA	CTZ	AAA	\TTT	CTZ	١A
421	ATTC	TAAZ	ATTC	CAF		ATT Z	AAT	TCO	ATT	TAA	ATT	AGC	CTT	GAT	TGG	CAC		CAT		٩A
481	CAGC	TTA		ATC	GGT	TGG	GTG	GTI	GGG	GAA	CAT		AGG	TGC	AGC	CAT	TAT	TCT	TC	٢A
541	GTTC	ATA	GTI	TAT	AGA	ATCO	GCGI		TGC	ATA	TTA	TAT	TCT	GTA	TTC	TTC	CATA	TTT	TAT	٢G
601	TAAC	TTGO	TAC	GAZ	GAZ	ACAC	СТ													

4.1.11 S segment sequence of Guama serogroup virus.

SEROGROUP	VIRUS		NUCLI	EOTIDE		AMIN	O ACID	REFERENCE		
		Total	5'NCR	3'NCR	A + U					
					(%)	Ν	NSs			
Anopheles	Ana A	947	53	156	63	245	No NSs	This study		
А	(ANAV)									
	Tacaiuma	960	42	181	62	245	No NSs	This study		
	(TCMV)									
Anopheles	Ana B	991	35	213	57	247	No NSs	This study		
В	(ANBV)									
	Boraceia	982	35	203	56	247	No NSs	This study		
	(BORV)									
Bwamba	Bwamba	1096	81	310	65	234	92	This study		
	(BWAV)									
Capim	Capim	991	ICS	ICS	ICS	235	No NSs	This study		
	(CAPV)	(P)								
Gamboa	Gamboa	1170	86	ICS	62	238	130	This study		
	(GAMV)	(P)								
	San Juan	1250	94	439	62	238	137	This study		
	(SJ2441V)									
Guama	Bertioga	695	ICS	ICS	ICS	111 (P)	ICS	This study		
	(BERV)	(P)								
	Guama	621	ICS	ICS	ICS	88 (P)	ICS	This study		
	(GMAV)	(P)								
Minatitlan	Palestina	662	ICS	ICS	ICS	220 (P)	ICS	This study		
	(PLSV)	(P)								
Nyando	Eretmapodites	902	45	155	58	233	92	This study		
	(E147V)									
	Nyando	941	70	170	58	233	92	This study		
	(NDV)									
Patois	Pahayokee	768	ICS	ICS	ICS	176 (P)	ICS	This study		
	(PAHV)	(P)								
Tete	Tete	1022	73	172	58	258	No NSs	This study		
	(TETEV)									
	Batama	1021	73	171	58	258	No NSs	This study		
	(BMAV)									
Turlock	M'poko	1075	117	247	61	236	79	This study		
	(MPOV)									
Bunyamwera	BUNV	961	85	174	58	233	101	Elliott et al.,		
								1989		

California	California	978	78	192	58	235	97	Bowen et
	Encephalitis							al.,1995
	(CEV)							
Group C	Itaqui (ITQV)	920	76	139	54	234	98	Nunes et al., 2005
Simbu	Oropouche (OROV)	754	44	14	53	231	91	Saeed et al., 2000

Table 4.2 Summary of the positive-sense RNA sequences obtained from this study and also from the prototype virus for each published serogroups; Bunyamwera (BUNV), California (CEV), Group C (ITQV) and Simbu (OROV). P indicates partial; ICS – incomplete sequences; NCR-non coding region; REF represents reference; F and + represent forward; A-adenine and U-Uridine

of these viruses were also found to have more A and U residues (more than 55%) compared to G and C residues.

The S segments of viruses in Bunyamwera, California, Group C and Simbu serogroups encode N and NSs proteins in overlapping reading frames. Similar organisation was observed in NDV, E147V, BWAV, MPOV, SJ2441V and GAMV S sequences (Figure 4.1 and Table 4.2), whereas, no potential NSs ORF was detected in ANAV, TCMV, ANBV, BORV, BMAV, TETEV and CAPV. Except for CAPV, all these viruses were found to have longer N proteins compared to those viruses that encode an NSs protein (Table 4.2).

4.2.3 Identification of potential NSs ORF in the S segments

The S segments of Bunyamwera, California, Group C and Simbu serogroup viruses encode the NSs protein in the +1 frame with respect to N. There are 16 nucleotides difference between the N and NSs initiation codons of Bunyamwera and California serogroup viruses, and 19 nucleotides different for viruses in Group C and Simbu serogroups. This study detected the NSs ORF only in Bwamba, Turlock, Nyando and Gamboa serogroup viruses but not in Anopheles A, Anopheles B and Tete serogroup viruses. To determine the possible initiation codon of NSs protein, the initiation codons for N and NSs were aligned without gap using ClustalW program (Figure 4.2). The alignment showed that only the NSs initiation codons of NDV, E147V and BWAV had similar spacing as BUNV and CEV, whereas the spacing between the ATG of N and NSs protein of SJ2441V, GAMV and MPOV were found to be 4, 25 and 34 nucleotides respectively. The complete sequences of the S segments were subjected to analysis by the FRAMES program using DNAMAN software (Figure 4.3). From this analysis, a short ORF was detected in BORV, CAPV and TCMV. However they were considered too short to encode for NSs protein due to the presence of multiple stop codons in the frames. Different initiation codons were present at the same relative spacing as BUNV and CEV in BMAV (GTG) and CAPV (TTG), and at the same relative spacing as OROV and ITQV in ANAV (TTG), TCMV (ACG), ANBV (TTG) and BORV (TTG), and TETEV (ATC).

Previously, it has been shown the presence of six conserved nucleotides between the N and NSs initiation codon of viruses in Bunyamwera and California serogroups (Figure 4.2) (Dunn et al, 1994). However, only five of these nucleotides were present in OROV and

MPOV	ATG TCA GA CC T T G TGC T CACATTTTCAGACTCCGACG ATG TCAGTCGAAGTACTTA56
SJ2441V	ATG TCTA ATG ATCTTTACGTATTCGACGATGATACGTCGGCAAGCTCTAGCTCTTA56
GAMV	ATG TCT GA CTCTCTA T ACGTATTTGACG ATG ATTCGACGGCAAGCTCTAGCACTTA56
OROV	ATG TCA GA GT T CATT TT CAACG ATG TACCACAGCGGACTACATCTACATT50
ITQV	ATG TCA GA GT T CATT TT CAACG ATG TACCCTCATTCACAGACAATACATT50
CEV	ATG TCG GA CT T G G TT TT CT ATG ATGTCGCATCCACAGGTGCAAATGGATT50
BUNV	ATG ATT GA GT T G G AA TT TC ATG ATGTCGCTGCTAACACCAGCAGTACTTT50
NDVV	ATG TCG GA AC T G G TG TT CT ATG ATGTCGAGCCAACTGCCCAAAATGGATT50
E147V	ATG TCT GA GA T C G CT TT TT ATG ATGTCCAGTCTACTGCCCAAAATGGATT50
BWAV	ATG TCC GA TT T GATC TT CT ATG ATGTCGACTCGGTTAATGCCAATGAGTT50
ANAV	ATG TCATCCA T G G AT T ATATGT TTG ACGATCCTGGTCATATCCAGATCATCAACTA56
TCMV	ATG TCA G GAA T A G AT TT CATCT ACG ATGATCCTGGTAGGATCCAGCAATCAGACT TTG -56
ANBV	ATGGCGTCTCAAGTTGATTTGCTTTTGAAGACACAGGAAATATCACTCAGTCTGACTT59
BORV	ATGTCGTCTACTGTGGATTTCGCATTTGAAGACACAGGGAATGTCACTCAGTCTGACTT59
TETEV	ATG CCCA A AGGA G GAAAAAGCG ATC CTGAACCATCCATCAACGTTGTTGC50
BMAV	ATG TCCA A AG T AAAAAGAG GTG AGTCAGAACCATCAATCAGCGTTGCAGC50
CAPV	ATG GGA GA AG T T G CTAATT TTG AATTTGGTGAGCGTTTGATAGCAGAAGG50

Fig. 4.2 Alignment of initiation codon of N and NSs proteins of the 5'region of positive sense S segment to indicate their spacing. Red bold letters denote the initiation codon of N protein; blue bold letters denote initiation codon of NSs protein; brown bold letter denotes possible initiation codon of NSs protein; green bold letters indicates conserved nucleotide in-between the spacing relative to BUNV and CEV and dotted lines indicate gaps inserted to maximize identity.

A.ANAV



B.TCMV



Fig. 4.3 Open reading frames of ANAV, ANBV, BORV, BMAV, CAPV, TCMV, TETEV and control virus BUNV using the DNAMAN programme (Lynnon biosoftware). The red boxes represent the N protein ORF, blue boxes represent the NSs protein ORF and green boxes represent the potential NSs protein ORF. The upper vertical lines indicate the start codons and lower vertical lines are for stop codons.

C.ANBV



D.BORV



E. TETEV



F. BMAV



G. CAPV



BUNV



ITQV. In this study, only viruses in NDV serogroup (E147V and NDV) have all the six conserved nucleotides; BWAV has five, MPOV has 4, GAMV has three and SJ2441V has any of them. However, some of these conserved nucleotides were also present in the NSs minus viruses, for example; ANBV and TCMV has five, while other viruses (ANAV, BORV, TETEV, BMAV and CAPV) have four or less of these nucleotides. Therefore, it is very difficult to relate the presence of these conserved nucleotides with the present of NSs protein. Most of the viruses in Bunyamwera and California serogroup start their NSs protein with double methionine. However in this study, only BWAV, E147V and NDV followed this pattern.

4.2.4 Sequences analysis of 3'/5' NCR

In agreement to previous findings with viruses in Bunyamwera, California, Group C and Simbu serogroups, alignments of 5' and 3' NCR of the positive-sense RNA of the viruses used in this study revealed that the 3' NCR is longer than 5'NCR. The longest 3' NCR was observed in SJ2441V (439 nucleotides), while the longest 5'NCR was detected in MPOV (117 nucleotides). These regions were found richer in A and U compared to C and G. The nucleotide sequences of 3'/5' NCR were largely non-conserved between the serogroup viruses, except in the first 15 nucleotides of the terminal sequences (Figure 4.4). This is as expected since these are the primer sequences used to obtain the products. However, some conserved nucleotide sequences were observed within the viruses in the same serogroup. As reported in Bunyamwera and California serogroup viruses, the non-canonical nucleotide pairing (G and U) at position 9 of the 3'/5' terminal sequences was also observed in ANAV, ANBV, BMAV, BWAV, BORV and MPOV. However, some of the viruses such as TCMV, TETEV, SJ2441V, E147V and NDV were found to have the same nucleotide sequence at these positions (either both A or G), suggesting that the same primers had annealed to their both ends. As observed in OROV (Figure 4.4), the CA-rich motif in the 5' NCR of virus-sense S segments downstream of the predicted mRNA termination site was also not observed in BWAV and MPOV. However ANBV and BORV were found to have two copies of this motif that are separated by 19 nucleotides. An extensive GAU/AU/GU-rich motif was observed in the 5'NCR of most of the studied viruses especially in BWAV, MPOV and SJ2441V with more than 74% of A and U residues. Using the DNAMAN software, repeated sequences as observed in LUMV (Dunn et al., 1994) were also detected in BWAV and TCMV (Figure 4.5).

OROV		
MPOV		
CEV		
LI4/V		
NDVV		
1 TQV		
TETEV		
BMAV		
SJ2441V	CUUGGUCUAGUAAAAAAACUUUACAAAAUUAUAAACUGUUAAAAAAUAAUGAGAAAAUUUUG	60
BUNV		
ANBV		
BORV		
ANAV		
BWAV		
TCMV		
OROV		
MPOV		
CEV		
E147V		
NDVV		
TTOV		
±±⊻v TETEV		
IEIEV		
BMAV		
SJ2441V	AAAAACAUAUAUAUGAAAAUAUAAUAUAAAAAUAGUUAUAAAAUUGGUCACAUUGGUAUAA	120
BUNV		
ANBV		
BORV		
ANAV		
DMAN		
BWAV		
TCMV		
OROV		
MPOV		
CEV		
E1470		
EI4/V		
NDVV		
ITQV		
TETEV		
BMAV		
SJ2441V	UAGGUCAAUUAGGUAUAUAUCCUGUAUAUAUAUAAAAAUAAUUCUAAUAUAGUUAUAUAU	180
BUNV		
ANRV		
DODY		
BORV		
ANAV		
BWAV	GGUCGGUAUACGUUUAUAAAUCGAUAGUGAUUUUGUAUAUUGGUUAUA	48
TCMV		
OROV		
MPOV	UAAAGUAUAAGGUUUAAUUAUAAAGUACACAA	32
CEV		
LL4 / V		
NUVV		
ITQV		
TETEV		
BMAV		
SJ2441V	AGUAUAGGUCAUUAUACAUUGGGUAUGAUAGGACUAAAUAGUUAAAAAAUAAUAUAAUAA	240
BUNV		20
ANRV		1 ⊑
		1 J
DUKV	GUAUUG-GAAGU	ΤŢ
ANAV		
BWAV	GUUAUUGGUAUCCUG UAUAUAUUUUAUUUUAUAUUUUAUGUUUUUUUUUU	108
TCMV	CGUAU	5
OROV		
MPOV	UAGAGG AUAAUAUAGGUUAUAUAAUA GG UUAAAUAGGUAUAUAUU GG UUAAAUAUUA GAA	92
CEV	CGAGGCCGUCCGUUUAACAAAUCCUGAUUUAAAAAUUUUAUUAACCUAUAAAGUUUUAA	58
E147V	CUUUCGAUAAAGUCUUAAAUUAACCAAACAUGUGCAUA	38
NDVV		40
TTOV		20
++V TETEV		29 00
ILILV DMAN		20
BMAV	UUAUUAAGAGAAAAUU CUCUAUCCCCA	21

SJ2441V	GUGUAGUAAAUCAAAAUCGUUUUAAACCAUUCCUGAUUGUUAUGAUUGUUAAAAUAUCCA 300
BUNV	CGUCCCUACGUAAAAAUUAGCCCGAUUUCAGUA 33
ANBV	AGUAUGACAUUUAUCUAU-AGGUAUAAUUAG CCUUGGUAUAUAUCAGUUAUAUAUUGGUA 74
BORV	-GUAUAUAGAUCGUCAAUCAGAUAUAUUGGUA-AUAUUAGGUAUAAAUCUCCA 62
ANAV	AAGUACUA-AUGUUAGUUA-AAAUAGGUU 27
BWAV	AUAUUUUUAUUAUUUUU-UAUAUCCCAAUCCCAUUUUUGUUUUUGUUUUUUGUGUUUU 165
TCMV	-CGUUUAUGUCAGACUUAUAGAUUGUGAUUGUAAUUUAGUGUAUAUGAUAUUUGUGAUUG 64
OROV	
MPOV	AAUCAAUAGAAUUCAAAUAAUUAGUAUGAUUAUUAAAUCAAAAUUCAAAAGUAGAUUA 150
CEV	UCAAUUAACCGAUUGAUUUUCCCAAGGAUCCCCGGGUGUUGGUAUCGUCGAUUUAC-CCA 117
E147V	AUAGUUAACCGAUAAAGUUUUCUGGAAGCCCAGUGUUAUAGUCGUCGACCCAC-CCA 94
NDVV	UAAAUUAAACUGUGGUUUUCUGGAAGCCCAGUUUCGGCUGUCGAAUUAC-CCA 92
ITQV	AGAGUUCUUCGGUUAUGUUCUGGGGUCGUAUUGC-CCA 66
TETEV	ACC AUAAAUAUAUGGUUG AGGCUGUACAUUUG-UCGAAUUCAUUCGACAUUAU ACUCCA 84
BMAV	ACC AUA-UAAAUGGUUGGGGCUGUACAUUUG -UC GAUUUAAUU CGACGUUAU ACACCA 83
SJ2441V	UCAAAAUCAUGAUUGUGUAAUCUGAUCAAUCCUGGGUUUUCGUCGACAGU-UAC-CCA 356
BUNV	GACAAAAUUAAACCGAUUUUCCCAACAAA GUU-GGGUGUUUUAUUGU CGACG AAC-CCA 90
ANBV	-UA-UAUAUAUAUUAUCCCCCACCACCACCACCACCACGAAGUUUGUCGUAUUGUA-UAC-CCA 128
BORV	-AAACCCAUCUAUAAUCCCCCAGUCCCCCCCCCCCGUCGUUCUCUA-UAC-CCA 117
ANAV	-CAAUUAGCGUAUAUUAGUGUUUAUAAUUA-GUGAUUGUAAAUAUUAAAACCC-CCA 81
BWAV	UUUGUGUGUUUUUUGUGUUUUUUUUGUGUGUUGUGU
TCMV	UAAUUUAGUGUAUGUGAUAUUCGUGAUGUAUGUG-AUUUUAUG-AUUGUGUUCCCA 118
OROV	
MPOV	GAAGUAGAUUAGUAGGGGGUAGGGUAGGUUAUUGUCGCGUUAUUUAGUUACAUAAGUUC 210
CEV	CCCACCAUCCCCUGUCGUU-UUAUUAUCUUCGAUUUA-GUAAUUAAAUUAUU 167
E147V	CCAUCCCCUUGUGGUUUAU-AGAUUAACUCAGAUCUA-ACAAUUUAUUUUUU 143
NDVV	CCCACCAUCCCCUUUGUAU-CGAAGUUUUC-GUAAAAAA 129
ITQV	CCCACCAGCCCCUGUGCUU-GUAAUAAU CCACACUAAAGUGAUUAAUUU 114
TETEV	CCCACCAACCCCGUUCCAAGUCUAGUUGAGUAACGUUAGAAUAAGAAGUAAUAAGAA 141
BMAV	CCCACCAACCCCGUUCUAAGUCUAGUUGAGUAAUGUUAGAAAAAGGAGUAGUAAUCA 140
SJ2441V	CCCACCAGCCCCGUUGAAACGGUUAAAGAAGUUUAGUAAAAUCAUCUAAUUUA 409
BUNV	CCCACCAACCCCUGUCUUUCUGUCGCCCGAUUUAGUUGUAAUA-UAACAAUUA 142
ANBV	CCCACCAACCCCUGUCGUU-GACAAU-AUAAGUUUUAACUGUAUACGAAGCGACAU 182
BORV	CCCACCAACCCCUGUCGUA-AGUUUUGAUGACGUUUAUAUGUAUACUCUACGUCUU 172
ANAV	CCCACCAACCCCUGUCGUAGAUUCAUUAG-UCAACU-UAUAAGAUUU 126
BWAV TCMV	UCUUUCAUCACUUGUCACAAGAGUACCGGCUAUAAUUAGAAACGAACCGAUUACCUU 278 CCCACCAACCCCUGACGUAGAAUAAUCGACAAUCAAAAACAGAACUACACCUC 171
OROV	ACCUCGUGUGAUGA 14
MPOV	AAACGUAAUUAAUUAUCAUCAGCACCUCGUGUGAUGA 247
CEV	UUCCAUAAGUCACCUCGUGUGAUGA 192
E147V	AGUUAUAAGUAACACCUCGUGUGAUGA 170
NDVV	-GUUUUAAGUAACACCUCGUGUGAUGA 155
ITQV	AGAUCUUAACCACCUCGUGUGAUGA 139
TETEV	AGACAAGAAUAUAUUUCACCUCAUGUGAUGA 172
BMAV	AGAUAAGAAUACAUAGGUCACCUCGUGUGAUGA 173
SJ2441V	GUUUGCUU <mark>ACAAA</mark> UGC <mark>A</mark> CCUCGUGUGAUGA 439
BUNV	CCAUAA-AAUUCAAAAUCCACCUCGUGUGAUGA 174
ANBV	U-AUAUUAGAAAAUACACCUCGUGUGAUGA 211
BORV	UUAUGUUAGAUAUUUCACCUCAUGUGAUGA 202
ANAV	UAAUACAUGAAA-UACACCUCGUGUGAUGA 155
BWAV	UAUUAAUUUCCAUGAAUCACCUCGUGAGAUGA 310
TCMV	AUGUGAUGA 180

Fig. 4.4 Alignment of complete 5' NCR of virus-sense S RNA segments. Dotted lines indicate gaps inserted to maximize identity, green letters represent the CA-rich region RNA, blue bold letters represent GU rich motif and red bold letter represent the G/A/U-rich region.

BWAV

61 CUGUAUAUAUUUAUUUAUAUAUUUUAU 87
103 UUGUAUAUAUUUUUAUUAUUUUUAUAUUUUUAUA 129
151 UGUUUUUGUGUUUUUUGU 170
171 GUGUUUUUGUGUUUUUUU 189
TCMV
27 UUGUGAUUGUAAUUUAGUGUAUAUGAUAU 56
57 UUGUGAUUGUAAUUUAGUGUAUGUGAUAU 85

Fig. 4.5 Presence of repeated sequences in 5' NCR of virus-sense RNA. Red letters indicate the duplicated sequences

4.2.5 N nucleotide and protein comparison

Viruses in the Bunyamwera, California, Group C and Simbu serogroups were reported to have the same length of N protein, 233, 235, 234 and 233 amino acids respectively. As seen in Table 4.2, the same number of amino acids was also observed in the N proteins of viruses within the same serogroup; ANAV and TCMV (245), ANBV and BORV (247), GAMV and SJ2441 (238), NDV and E147V (233), and TETEV and BMAV (258). The N protein of Tete serogroup viruses was the longest reported to date among orthobunyaviruses. Amino acid sequence alignment of the N proteins of these viruses (Figures 4.6) revealed that the extensions of the N protein at the carboxyl terminus was also observed for these viruses. Furthermore, 8 amino acids were apparently inserted after amino acid 23 of their amino terminus. In agreement with previous reports, amino acid sequence alignments of the presence of some conserved motifs in the N protein of these viruses.

The nucleotide and predicted amino acid sequences of the N protein coding region of complete and partial S sequences of the viruses obtained in this study, together with the published sequences for representatives of other serogroups, (BUNV, CEV, ITQV and OROV) were compared (Table 4.3). Except for ANAV and TCMV in the Anopheles A serogroup, whose S segments have only 59% nucleotide sequence identity and 53% amino acid identity between the N proteins, the S segments and encoded N protein of the viruses within the same serogroup were found to have nucleotide and amino acid sequence identity greater than 70%. N sequence identity between viruses in different serogroups was as low as 24% for amino acid sequences and 5% for nucleotide sequences, which were observed between TCMV with TETEV and BMAV, and between ANBV with GAMV and OROV (Table 4.3). This low value could be due to the more serogroups (12 serogroups) used in these comparisons compared to only four in the previous comparisons. However, greater sequence identity was also observed between CAPV with viruses in Guama serogroup (BERV and GMAV), with 80% and 77%, and 73% and 74% for amino acid and nucleotide identity respectively, while BERV and GMAV exhibited 73% identity for both nucleotide and amino acid sequences, indicating that BERV and GMAV are closer to CAPV than to each other. BWAV was found to exhibit high sequence identity with CEV with nucleotide and amino acid sequence identity of 70% and 68% respectively.

California, Group C and Simbu serogroups. Values in the top right half represent nucleotide sequence identity and values in the bottom left half represent amino acid sequence identity. Red numbers indicate the sequence identity of the viruses within the same Nucleotide and amino acid sequence identity between studied viruses and representative virus of Bunyamwera, serogroup and blue numbers indicate identity value of possible 'bridging' virus between the two serogroups. Table 4.3

	ANAV	TCMV	ANBV	BORV	BWAV	CAPV	GAMIV	S/2441V	BERV	GMAV	PLSV	AQN	E147V	PATV	BMAV	TETEV	MBOV	BUNV	CEV	VQTI	OROV
ANAV		53	8	45	8	34	34	34	44	42	33	41	8	35	29	38	38	36	39	35	37
TCMV	59		51	铃	6	36	36	34	42	台	35	42	41	35	28	28	36	35	38	29	32
ANBV	ß	8		82	33	31	31	31	33	35	31	35	8	34	25	25	Я	38	34	8	32
BORV	31	53	15	×	32	31	29	31	31	35	32	33	33	36	36	24	32	26	32	31	32
BWAV	я	ą	16	я		4	8	8	8	8	8	0	8	R	31	8	đ	Ą	8	ਲ	4
CAPV	30	23	15	42	47		35	37	80	3	51	43	4	49	37	35	37	42	44	42	37
GAMV	16	15	2	18	38	21		88	39	36	33	37	36	39	31	31	43	35	37	30	33
SJ244IV	13	13	15	13	33	29	83	*	d	36	31	8	33	38	30	30	42	35	36	33	34
BERV	39	39	28	24	62	73	45	46	***	73	37	64	59	54	45	43	45	54	54	39	43
GMAV	4	33	36	33	62	74	57	99	73		37	59	55	23	43	42	42	54	51	39	43
PLSV	15	31	8	IJ	40	57	17	м	4	4		ষ	89	8	8	8	8	8	35	8	8
NDV	38	45	17	各	65	47	43	37	61	50	41	*	85	왕	34	34	38	47	60	39	46
E147V	20	42	33	41	67	47	42	\$	64	62	8	80		51	36	36	38	4	60	8	42
PATV	32	20	37	16	58	各	4	41	59	57	35	56	64		38	35	38	4	48	35	38
BMAV	17	7	ន	00	R	23	12	8	8	8	8	8	8	8		88	8	କ୍ଷ	36	8	8
TETEV	17	5	6	00	37	<u>56</u>	12	8	58	<u></u> 26	30	16	16	30	8		8	33	34	26	8
MBOV	6	25	11	43	45	29	53	33	44	31	38	37	37	59	41	83		39	38	30	32
BUNV	21	17	18	S	39	38	20	31	61	63	27	\$	55	32	29	32	27	*	42	46	9
CEV	4	4	13	18	R	\$	31	8	62	8	ঞ	8	3	ð	8	18	ą	8	*	ð	\$
ITQV	12	12	10	14	54	25	10	21	33	56	12	8	54	31	15	25	6	46	48		51
OROV	22	17	7	14	55	28	34	46	52	55	13	60	8	33	26	25	41	41	51	28	***

Akabane Oropouche Maguari Cache_Valley Batai Bunvamwera Main_Drain Kairi California_encephalitis Tahvna La Crosse Jameston_Canyon Keystone BWAV E147V NDV CAPV ANA_BV BORV TCMV Anopheles_A GAMV SJ2441V MPOV TETEV BMAV Caraparu_BeAn3994 Itaqui_BeAn12797 Vinces Murutucu BeAn974 Restan_TRVL51144 Marituba_BeAn Nepuyo_TRVL18462 Gumbo LimboFE-371H Madrid_BT4075 Aino Akabane Oropouche Maguari Cache_Valley Batai Bunyamwera Main Drain Kairi California_encephalitis Tahyna La Crosse Jameston_Canyon Kevstone BWAV E147V NDV CAPV ANA BV BORV TCMV Anopheles_A GAMV SJ2441V MPOV TETEV BMAV

Caraparu_BeAn3994

Itaqui_BeAn12797

Murutucu_BeAn974

Restan_TRVL51144

Gumbo_LimboFE-371H

Marituba_BeAn Nepuyo_TRVL18462

Madrid_BT4075

Vinces

Aino

SAYGSQLVSAVVRVFYINAAKVKAYLSRSSELLVFLTLGGVNLAVLNTHF 77 SAYGRQLVSAVVRVFYINAAKVKAYLSRSSELLVFLTLGGVNLAVLNTHF 77 SVYGQQLVAAVVRVFYINAAKVKAYLRRSRELLVFLTLGGVNLAVLNTHF 77 SDYGSQLESAIVRVFYINAAKVKAYLRRSSELLVSLTLGGVNLAVLNTHF 77 SAYGSQLESAIVRVFYINAAKVKAYLRRSSELLVSLTLGGVNLAVLNTHF 77 PAYGSELESSIVRVFYINAAKVKAYLRRSSELLVLLTLGGVNLAVLNTHF 77 PAYGSELESAIVRVFYINAAKVKAYLRRSSELLVSLTLGGVNLAVLNTHF 77 PTHGSELESSTVRIFYINAAKVKAYLRRSAEQLVSLTLGGVNLAVLNTHF 77 PRNGRVLEATIVRIFYINAAKVKAYLHRSSEQLVLLTLGGVNLAVLNTHF 77 AKHGAQLNFDTVRFFFLNQKKAKMVLSKTAQPSVDLTFGGIKFTLVNNHF 78 SKYGQQLNFTVARVFFLNQKKAKMVLHKTPQPSVDLTFAGVKFTVVNNHF 78 ARYGQVLNAGVVRVFFLNQKKAKDVLRKTSRPMVDLTFGGVQFAMVNNHF 77 RIHTTGLSYDHIRVLYIKGREIKTSLTKRSEWEVTLNLGGWKVAVFNTNF 77 RIHTTGLSYDHIRIFYIKGREIKTSLTKRSEWEVTLNLGGWKVTVFNTNF 77 RIYTTGLSYDNIRIFYIKGREIKTSLSKRSEWEVTLNLGGWKVTVFNTNF 77 RLYTTGLSYDHIRIFYIKGREIKTSLAKRSEWEVTLNLGGWKVTVFNTNF 77 RVHTTGLSYDHIRIFYIKGREIKTSLSKRSEWEVTLNLGGWKVAVFNTNF 77 RRHTTGLNYDHIRIFFLNGKKAKDTLSKRSETTITLNFGGWKIPVVNTHF 77 AKHGESINLAAVRIFFLNAAKAKAALSRKPERKANPKFGEWQVEVVNNHF 77 IKHGEAINLHSVRIFFLNAAKAKAALARKPERKASPKFGEWQVEVVNYHF 77 VKNAESLNLAAVRIFFLNAAKAKAALSRKPERKANPKFGEWOVEVINNHF 77 ADHGESINLSAVRIFFLNAAKAKAALARKPERKATPKFGEWQVEIVNNHF 77 ANHGESISLSTVRIFFLNAAKAKAALTRKPERKATPKFGEWQVEIVNNHF 77 NNYPGALNTNTARTFFLNAAKAKNVLRNKPDKKVNPKFGNWEVEVVNNHF 77 AAAGAGLNIVSARIFFLNARKAKDQLSRRPEPKINLKFGTWSVEVVNNHF 77 ASAGAGLNIVSARIFFLNARKAKDQLARRPEPKVGLKFGTWQVEVVNNHF 77 REHVPMLTPGNIRIFFLRAYDAKQKLKTTTARTANLKFGTATFTVKNNHN 81 LGKTAHLSLENIKIFFLNAGKLKQQMKTCSKTKIKAKFGTLEIELVNTLN 80 IGRSNOLSLENIKIFFVNAGKLKMOMKSSLKPRIKAKFGTLEIEMVNTLN 80 SKFALVLILQNIRIFFLNAGKLKAALKLSPKRKVKAKFGTLEVEITNTHN 79 AKYQDQLTVENIRIFFLNAGKFKKMLRTNPKLKVQVKFGTLQLEIINTHN 79 STHGDQLTPANIKCFFLHARTAKARLAKDPAERRTLTFGSWKVEVVNNHY 79 STYGQQLTPANIKCFLLHARTAKARLSKDPAERRTLTFGTWKVEVVNNHY 79 DTYREHLTVDNIRIFFLKAAEAKAHMAKVNNEVVEVHFGTLVLPLVNNHK 79 EVHGKELNLPNIKVFFLKARQAKEIMRSKAKSEMTFTFGSLTLTFKNTHH 100 ETHGKDLTVTNIRVFFLKARQAKEIMRSKAKSEMTFTFGSLTLTFKNTHH 100 1 1 1 . * : 1 :.

. .

-----VASFTDNTFDPEAAFLRFE 27 -----VPSFTDNTFDPEAAFLRFE 27 -----EPSFTDNTFDPEAAFLRFE 27 -----VASFTDNTFDPEAAYLRFE 27 -----VASFTDNTFDPEAAYLRFE 27 -----VASFTDNTFDPEAAYVRFK 27 -----EASFTDNTFDPEAAYVRFK 27 -----ETSFADNTFNPEAAFVAFK 27 ------VPSFADNTFNPEAQYVTFI 27 -----VPQRNLATFNPEVGYVAFI 28 -----VPQRNAATFNPDAGYVAFI 28 -----VPORTTSTFDPEAAYVAFE 27 -----VAANTSSTFDPEIAYVNFK 27 -----VAANTSSTFDPEIAYVNFK 27 -----VAANTSSTFDPEVAYINFK 27 -----VAANTSSTFDPEVAYINFK 27 -----VAANSSSTFDPEVAYASFK 27 -----VTANTSSTFDPEAGYAAFK 27 -----VASTGANGFDPDAGYVDFC 27 -----VASTGANGFDPDAGYVDFC 27 -----VASTGANGFDPDAGYMDFC 27 -----VASTGANGFDPDAGFVAFM 27 -----VASTGANGFDPDAGYVAFM 27 -----VDSVNANEFDPDTGYLDFK 27 -----VQSTAQNGFDPDQQYLAFK 27 -----VEPTAQNGFDPDKQVAFK 27 -----LIAEGRNPFVPDVAYARFI 31 -----TGNITQSDFIPDVGYTAFC 30 -----TGNVTQSDFVPDVGYTTFC 30 -----PGRIQQSDFVPREEYAVFC 29 -----PGHIQIINYIPEEEYMAFC 29 -----DSTASSSTYDPTQEYNVFI 29 -----DTSASSSSYDPKREYDKFI 29 -----SDDVSRSTYNPGEEYDTFV 29 MPKGGKSDPEPSINVVASAAFEDDEIDYHFTAGDDGGAPFNPMTAYKRFM 50 MSKVKRGESEPSISVAASAAFAEDEIDYHFSAGDDGGAPFNPMAAYKEFM 50 : * : :

Akabane Oropouche Maguari Cache_Valley Batai Bunyamwera Main Drain Kairi California_encephalitis Tahyna La Crosse Jameston_Canyon Keystone BWAV E147V NDV CAPV ANA_BV BORV TCMV Anopheles_A GAMV SJ2441V MPOV TETEV BMAV Caraparu_BeAn3994 Itaqui_BeAn12797 Vinces Murutucu BeAn974 Restan_TRVL51144 Marituba_BeAn Nepuyo_TRVL18462 Gumbo LimboFE-371H Madrid BT4075 Aino Akabane Oropouche Maguari Cache_Valley Batai Bunyamwera Main Drain Kairi California_encephalitis Tahyna La Crosse Jameston_Canyon Keystone BWAV E147V NDV CAPV ANA_BV BORV TCMV Anopheles_A GAMV SJ2441V MPOV TETEV BMAV Caraparu_BeAn3994 Itaqui_BeAn12797

Caraparu_BeAn3994

Itaqui_BeAn12797

Murutucu BeAn974 Restan_TRVL51144

Nepuyo_TRVL18462

Gumbo_LimboFE-371H

Marituba_BeAn

Madrid BT4075

Vinces

Aino

AESKGVTWK--EGANIYLSFF**PGTEMFMLEF**RFFPLAVGLARCHKEKMDT 195 :*:*::: * **.. : : NCLEKAIWQRYM-GIGADQWITKETVAILTRFEDVEKLDWARTAFTSAARE 224 NCLEKAIWQRYM-GIGADQWITKETVAILTRFEDVEKLDWARTAFTSAARE 224

AEENGITWR--DGVEMYLAFL**PGAEMFMSTF**EFYPLTIGMYRVRRWSMDP 174 AEENGITWR--DGVEMYLAFLPGAEMFMSTFEFYPLTIGMYRVRRWSMDP 174 AEENGITWR--DGVEMYLAFLPGAEMFMSTFEFYPLTIGMYRVRRWSMDP 174 AEENGITWS--DGVGMYLAFL**PGAEMFMSTF**TFYPLTIGMYRVRRWSMDP 174 AEENGITWS--DGVGMYLAFL**PGAEMFMSTF**AFYPLTIGMYRVRRWSMDP 174 AEENGITWT--DGVAMYLAFL**PGAEMFMSTF**KFFPLTIGMYRVRRWSMDP 174 AEENGITWT--DGVAMYLAFL**PGAEMFMSTF**TFFPLTIGMYRVRRWSMDP 174 AEENGITWC--DGVAMYLAFL**PGAEMFMSTF**AFFPLTIGMYRVRRWTMDP 174 AEKNGITWC--DGVTMYLAFLPGSEMFMSTFRFFPLTIGMYKVQRWGMDP 174 AEVKGCTWT--EGLTMYLGFAPGAEMFLETFEFYPLVIDMHRVLKDGMDV 174 AEVKGCTWS--DGYAMYLGFAPGAEMFLETFEFYPLVIDMHRVIKDGMDV 174 AEVKGCTWN--DGDAMYLGFAAGAEMFLQTFTFFPLVIEMHRVLKDGMDV 172 AEKNGITWA--DGEEVYLSFF**PGSEMFLGTF**KFYPLAIGIYKVQKKEMEP 174 AEKNGITWS--DGEEVYLSFF**PGSEMFLGTF**KFYPLAIGIYKVQRKEMEP 174 AEKNGITWA--DGEEVYLSFF**PGSEMFLGTF**RFYPLAIGIYKVQRKEMEP 174 AEKNGITWS--DGEEVYLSFF**PGSEMFLGTF**RFYPLAIGIYKVQRKEMEP 174 AEKNGITWA--DGEEVYLSFF**PGSEMFLGIF**KFYPLAIGIYKVORKEMEP 174 AASHGITWD--DGEEVYLSFF**PGSEMYLTTF**KFYPLAIGIYKVQRKLMDP 174 AESNGIRWE--NGAEIYLAFF**PGTEMFLETF**KFYPLTIGIYRVKNGMMDS 175 AESNGIRWD--NGAEIYLAFF**PGTEMFLETF**NFYPLTIGIYRVKQGMMDP 175 AESNGVGWD--SGPEIYLSFF**PGTEMFLETF**KFYPLTIGIHRVKOGMMDP 175 AESNGIHWN--NGPEIYLSFF**PGTEMFLEIF**KFYPLTIGIYRVKHGMMDP 175 AESNGIRWG--NGVEIYLSFF**PGTEMFLELF**KFYPLTIGIYRVKHGMMDA 175 AESNGIRWS--DGAEIYLSFF**PGTEMFLEPF**KFYPLAIGIYRVKHKMMDA 174 AESNGIKWS--AGAEIYLSFF**PGTEMFLEKF**NFYPLAIGIYRVKRGMMDA 174 AESNGIKWS--AGAEVYLSFF**PGTEMFLEKF**NFYPLAIGIYRVKKGMMEA 174 AESLGITWE--NGDDVYLSFF**PGTEMFLDTF**HMLPLAIGIYRVOOKOMKA 176 AEVAGVQFKPEVPPELYIGFAPGAEFFMKLFKCYPLAIAVVRVKKGQMQP 179 AEVSGVQFKPEVPPELYIGFAPGAEFFMKLFKCYPLAIAVVRVKKGQMQP 179 AEVAGVSWANSTP-EIYLSFCPGSEFFLADFKFYPLAIGIARVKKDLMKP 177 AEKAGVTWDNAGSVEMYLGFA**PGAEFFLKDF**KFYPLAIGIARVKKGLMSP 178 SESMGINWQ--HGAMIYLSTLPGAEMFLTEFQLFPLAFAVVRVKKNLASP 175 SESMGINWQ--HGAMMYLSTLPGAEMFLTEFQLLPLAFAVVRVKKNLASP 175 SAKMGFOWS--VGPEIYLSTLPGTEMFLDAFKMYPLAFILLRVKRGEIKI 177 ADSKGVTWK--ERGKCISCFF**PGTEMFMLEF**KFFPLAVGLARCHKEKMDT 195

PGTQSIPLPDYGLTLHRISGYLARWALYLIRTNPESDIAL-VRSTFVVPL 126 PGTQSIPLPDYGLTLHRISGYLTRWALYLIKTNPESDIAL-VRSTFVVPL 126 PGTOSIPLPDYGLTIHRISGYFSRWALYLIATNPESDIAL-VRSTFVVPL 126 PGTQSIPLPDYG**LTIHRLSGYLSR**WALYLIRPSPESEIGL-VRSTFIVPL 126 PGTQSIPLPDYG**LTLHRLSGYLAR**WALYLIRPSPESDIGL-VRSTFIVPL 126 PGTQSIPLPDYG**LTIHRLSGYLSR**WALYLIYIIPESEIDL-VRTTFIVPL 126 PGTQSIPLPDYGLTIHRLSGYLARWALYLICPIPDSDIDL-VRTTFIVPL 126 PGTOSIPLPDYGLTIHRLSGYLARWALYLICPSPDSSIGL-VRATFVVPL 126 PGNOSIPLPDYGLTIHRVSGYLARWALYLIAPSPYSDRDI-VRTTFIVPI 126 PQYTANPVPDTALTLHRLSGYLAKWVADQCKTN-QIKLAE-AMEKIVMPL 126 PQYTANPVSDTAFTLHRISGYLARWVAEQCKAN-QIKFAE-AAATIVMPL 126 PQFQSNPVPDNGLTLHRLSGYLARWAFTQMRS--PIKQAE-FRATVVVPL 124 PGNRNSPVPDDGLTLHRLSGFLARYLLEKILKVSDPEKLI-IKSKIINPL 126 PGNRNSPVPDDGLTLHRLSGYLARYLLEKILKVSDPEKVI-IKSKIINPL 126 PGNRNSPVPDDGLTLHRLSGFLARYLLEKILKVSDPEKLI-IKSKIVNPL 126 PGNRNSPVPDDGLTLHRLSGYLARYLLEKILKVSDPEKLI-IKSKIINPL 126 PGNRNSPVPDDGLTLHRLSGFLARYLLEKILKVSEPEKLL-IKSKIINPL 126 LENRNMSVPDDGLTLHRVSGYLARYLLDRVYSAGEPEKLK-IKTTIINPI 126 PANRNNPIGNNDLTIHRISGYLARWVLEQYKENEDESQRELVKTTVINPI 127 PGNRNNPIDNNDLTIHRLYGYLARWVLEQFKENEDAAQRELIKTTVINPI 127 PGNRNNPIGNNDLTIHRLSGYLARWVLDQYNENDDESQHELIRTTIINPI 127 SGNRNNPIGNNDLTIHRLSGYLARWVLEHFNSDDDESQRELIRSTIINPI 127 PGNRNNPIGNNDLTLHRISGYLARWVLEHFGEGEDESQKELIKSTVINPI 127 PGNRNNPIGKDDLTLHRISGYLARWILEEYKR-DDESEKEIIESTVINPI 126 QGNRDNVISDTDLTIHRLSGYIARFILDQYLAGNTVVQSG-IQLQIVNPI 126 QGNRDNPIGDSDLTMHRLSGYIARYILDQYLAGNSVAQAG-IQLQIINPI 126 ERNANIEVEAEDLTLHRISGFLAKFVRESMVDETASKMIR---DTIVNPI 128 RSLGQVSLQPND**VTLHRASAYLAR**KALELYREGQADFQAA-MRDQFVMPL 129 KALGQISLQPTDLTLHRASAFLARRALEMYHGGAADFQAA-MRDQFVMPL 129 RAFADSKLEKGD**ITLHRLSGFLAK**KILDIYNAGSFDLQTQ-IKAEIIIPL 128 PANREIQLSSSDLTLHRVSAYLARKSLRLWEALTLDQKEE-FKLKIIMPL 128 RRATPVTVQDHDNTLHRISGYLALYCLQTSSQNKQNFER--IKSVVINPI 127 RRAAPITVQDHD**LTLHRLSGYIAL**YCLQASVENAQSFER--IRSTVINPI 129 PGRAQREVADSDLTLHRVSGYLAKFVLELYKKLRKSSEREAITSRIINPI 129 PSNRHLAVEQDDLTINRVTGFMAHAILLTHRDPKHKDAVE---KTIINPI 147 PSNRHLTVDQDDLTINRATGFMAYAILLTHREPKSKDAVE---KTIINPI 147 :*::* .::: . * • 10
Vinces Murutucu_BeAn974 Restan_TRVL51144 Marituba_BeAn Nepuyo_TRVL18462 Gumbo_LimboFE-371H Madrid_BT4075 Aino Akabane Oropouche Maguari Cache_Valley Batai Bunyamwera Main_Drain Kairi California_encephalitis Tahyna La_Crosse Jameston_Canyon Kevstone BWAV E147V NDV CAPV ANA_BV BORV TCMV Anopheles A GAMV SJ2441V MPOV TETEV BMAV Caraparu_BeAn3994 Itaqui_BeAn12797 Vinces Murutucu BeAn974 Restan_TRVL51144 Marituba BeAn Nepuyo_TRVL18462 Gumbo_LimboFE-371H Madrid_BT4075 Aino Akabane Oropouche Maguari Cache_Valley Batai Bunyamwera Main_Drain Kairi California_encephalitis Tahyna La_Crosse Jameston_Canyon Kevstone BWAV E147V NDV CAPV ANA_BV BORV TCMV Anopheles A GAMV SJ2441V MPOV TETEV BMAV

FVSRFGIRLA	234
FVSRFGIRLA	234
FVIQFGIRLA	234
FVSRFGIRLA	234
FVSSFGIRLA	234
FVSQFGIKLA	234
FVSQFGIKLA	234
FVSKFGIRLA	234
FVAQFGIKLA	234
FLAQFGINI	233
FLAQFGIQI	233
FLRQFGIGI	231
FLAKFGINM	233
FLAKFGISM	233
FLAKFGISM	233
FLAKFGISM	233
FLAKFGINM	233
FLNKFGINI	233
FLQKFGIKLP	235
FLQKFGIRLP	235
FLQKFGIRLP	235
FLIKFGVKLP	235
FLAKFGVRLP	235
FLKKFDIAMI	234
FLGKFGILI	233
FLGRFGIVI	233
FLADFGITF	235
LLSELGLSAPTIQQALTK	247
LLSELGLSTPTIQRALTK	247
FLKSLGLSAHYINGALRG	245
FLVDLGLSRATVASMKR-	245
FFTALGINLDMRR	238
FFSALGINLDLRR	238
FLQELGISK	236
FLKEFGLDQDKN	258
FLKEFGLDQDKN	258
: :*	

Fig. 4.6 Alignment of complete N protein of the viruses. * indicate 100% conserved; brown letters indicate conserved motif; dotted lines indicate gaps inserted to maximize identity; : denotes more than 90% conserved; . denotes more than 50% conserved.

NCLEKAIWQRYM-GIGADQWITKETVAILTRFDAVEKLDWARTAFTSAARE 224 NCLEKAIWQRYM-GIGADQWITKETVAILTRFEDVEKLDWAKTAFTSSARE 224 NCLEKAIWORYM-GIGADOWITKETLAILTREEDVEKLDWAKTAFTSAARE 224 NCLEKAIWQRYM-GIGADDWITKETVTILARFEDVEKLDWARTAFTSSARE 224 NCLEKAIWQRYM-GIGADDWITKETVTILARFEDVEKLDWARTAFTSSARE 224 QCLEKAIWQRYM-GIGADEWITKETVSILARFEDVEKLDWAKTAFTSAARE 224 KCLEKAIWQRYA-GIGADTWIQKEYADINARFSVVESLDWARTAFTSTARE 224 NFMRKVLRORYG-TLTAEOWMTOKIDAVRAAFNAVGOLSWAKSGFSPAARA 224 NFMRKVLRQRYG-QLTAEEWMTSKLDAVKAAFSSVAQISWAKSGFSPAARA 224 NFMKKVLRQRYG-QKTAEQWMREEIVAVRAAFEAVGTLAWARTGFSPAGRD 222 KYLEKTMRQRYM-GLEAATWTVSKVNEVQAALTVVSGLGWKKTNVSAAARE 224 KYLEKTMRORYM-GLEASTWTISKVNEVQAALTVVSGLGWKKTNVSAAARE 224 KYLEKTMRQRYM-GLEASTWTVSKLNEVQSALTVVSGLGWKKTNVSSAARE 224 KYLEKTMRQRYM-GLEASTWTVSKLNEVQSALTVVSGLGWKKTNVSSAARE 224 KYLEKTMRQKYM-NMDAATWTVTQVSEVQAALTVVSGLGWKKTNVSAAARE 224 KYLEKTMRORYM-NLDASOWTOKHFSDVNSALTVVAGLGWKKANVSIAAKD 224 QYLKKALRQRYG-SLTAEKWMSQKTGMIAKSLKEVEQLKWGRGGLSDTART 225 QYLKKALRQRYG-SLTADKWMSQKTTAIAKSLKDVEQLKWGRGGLSDTART 225 QYLKKALRQRYG-TLTADKWMSQKVAAIAKSLKDVEQLKWGKGGLSDTAKT 225 OYLKKALRORYG-TLTAEKWMAOKTVLIAKSLKDVEOLKWGRGGLSDAART 225 OYLKKALRORYG-TLTADKWMAOKTSMITKSLKDVEOLKWGKGGLSDTARA 225 QFLKKALRQRYG-KMTAEKWMSTKVKTIGEAVRNVEKLKWGKGGLSDAARN 224 QYLKKALRQRYG-NLTADQWMQSKSDDVLRAIAVLEKLQWGKSGLSEAARQ 224 QFLKKSLRQRYG-QMTADQWMQTKSDDVMRAVAVLEKLSWGRSGLSEAARQ 224 EFLKKHLROOYG-GLAASOWMSERKEDVKLALTLISRLPWGKAGLSAAARN 226 EFLAKSLRQRYA-GAPPAEWMSTNAAAIKAATAMVEKYPMMKVASRPHIEK 229 EFSAKCLRQRFG-GAPPAEWMTTNAPAIKIATAMVEKFPMMRTASRPHVEK 229 EFLAKLLRQRYG-NMTPGEWMKTHSVSVQSAISEIEKYPLMKVNSRPHVEE 227 DFLSKTLRORYG-GIPPADWMAIHKSAVRTAVEFVDKFPLMRSNASSHVEO 228 ETVKKVLRORYD-GRLPGEWMNSELEKVKTAITKVEALHKTFVGVTARMIE 225 ETVKKVLRQRYE-GRLPGEWMNSELEKVKAAITRVESLHKTFVGVTARMIE 225 DMAKKAMRQRYG-DKEASTWLEEEVDTVKSAIKTVEKLKPTLTGLAASMTK 227 EYLKKPMROMLTDGTKAQVWLGAKIEEIRKAYKVCMNLKFVKAGFSEAARE 246 EYLKKPMRQMLTDGTKAQVWLGAKIEEIRKAYKVCMNLKFVKAGFSEAARE 246 :*:: * : . .

4.2.6 NSs protein comparison

NSs amino acid sequences of NDV, E147V, BWAV, GAMV, SJ4441V and MPOV were aligned with the published NSs sequences of BUNV, CEV, ITQV and OROV (Fig. 4.7). This alignment showed that the NSs proteins vary in size especially at the carboxy terminus. The largest NSs protein was observed in SJ2441V (137 amino acids) in Gamboa serogroup and the smallest (79 amino acids) in MPOV in Turlock serogroup. These are the largest and the smallest NSs protein reported to date. As observed with GROV NSs protein, the relative truncation of MPOV was at both termini. In agreement with previous results (Dunn et al., 1994; Bowen et al., 1995), NSs proteins reported in this study were more divergent than the corresponding N proteins with no conserved amino acid motif observed. The conserved LPS motif at the carboxy terminus in most of NSs protein of Bunyamwera and California serogroup viruses was only observed in BWAV NSs protein. This motif was also absent in ITQV, OROV and GROV NSs proteins (Figure 4.7). The sequence identity of the NSs protein among viruses within the same serogroup was more than 80% and between serogroups were as low as 4% (between ITQV and BWAV) (Table 4.4). In contrast to viruses in Bunyamwera, California, Group C and Simbu serogroup, only viruses in Gamboa serogroup were found to have two lengths of NSs protein; 130 (GAMV) and 137 (SJ2441V) amino acids respectively. This could be due to the limited number of viruses used for each serogroup in this study.

4.2.7 Phylogenetic analysis of S segments

To establish genetic relationships among the studied viruses with other orthobunyaviruses, phylogenetic trees were constructed from the alignments of complete N amino acid sequences of ANAV, ANBV, BMAV, BORV, BWAV, CAPV, E147V, GAMV, MPOV, TCMV, NDV, TETEV and SJ2441V (Figure 4.8A) using NJ analysis, and partial N ORF of ANAV, ANBV, BMAV, BORV, BWAV, CAPV, E147V, GAMV, MPOV, TCMV, NDV, TETEV and SJ2441V, BERV, GMAV, PLSV and PATV using NJ and MP analysis methods (Figure 4.8B and 4.8C). For NJ analysis, a distance matrix was calculated from the aligned sequences using the Kimura two parameter formulas (Kimura, 1980) and NJ tree was computed. The reliability of the inferred tree was tested by bootstrap analyses using 100 replicates data sets generated from the original sequence alignment (Felsenstein, 1993) and a bootstrap consensus tree was generated. For MP analysis, phylogenetic tree

NDVNSs	MMSSQLPKMDLILISSMWHLKLQLEQGLTLFPLGSSSSMPGKPKINSLVDQSRRLVLNLE	60
E147VNSs	MMSSLLPKMDLILINNTWHLKLLQEQGLILFPLGSSSSMPGKPKINSLVDQSQRLTLNLA	60
BWANSs	MMSTRLMPMSLTLIQDTWILKITIQGHSIQIPLGHSSSMPQRPRMCSAINLTRRSILNLE	60
CEVNSs	MMSHPQVQMDLILMQGMWTSVLNMENQLTLLQLGSSSSMPQRPRLLSRVSQRGKLILNLA	60
BUNVNSs	MMSLLTPAVLLTQRSHTLTLSVSTPLGLVMTTYESSTLKDARLKLVSQKEVNGKLHLTLG	60
GAMVNSs	MIRRQALALMIQRRNTMSLLVLTGTSLLQQTLNVSSCMHVQPRLVWQRTRRNAERLLLD	59
SJ2441VNSs	MIFTYSTMIRRQALALMIQRENMINLLVLTGSNLLQQTLNVFSCMHARPRLVFQRTRRNAERLLLE	66
OROVNSs	MYHNGLHLHLIRRQHMWHLKLDTDKCSMLVLLESSSSTKRRPKMSYVRHRGPWLTLLLV	59
ITQVNSs	MYPHSQTIHLTLKLLSLDSKVHTAGNWFPPSYESSTSTRRRSKLTSVGQVNSLFFLLLV	59
MPOVNSs	MSVEVLITQERNTIRLWIHTENISQWITLESSSSRLQRPKLTWRRLTTRSLKFILE	56
	:: : : : : : : : : : *	
NDVNSs	HGRWKWSITIFKETGTILSATQISQCIDSQDI	92
E14/VNSs	HGRWRWSITIFKETGTMLSATRILQSTDCQDI	92
BWANSs	TGRWRLSIIIFLETGTTQLVKMI LPS TEFQDI	92
CEVNSs	SGRWRLSIIIFQQTGTIQLVTTI LPS TASQDTLPDGS	97
BUNVNSs	AGRLLYIIRIFLATGTTQFLTMV LPS TASVDSLPGTYLRRC	101
GAMVNSs	LGRLRLLITITEEPLRLQCRTMIIPCTAFPDILPCTACRPAPRISRTLSESSQSSSTRSLSPWGST	125
SJ2441VNSs	LGRLRLLITIIDEPLQLRCRTMILPCTAFPDILPCTVCRPVWRMRSLSSESDQQSSTRSLSPWGST	125
OROVNSs	GSNLQWLITISHSSSRIQCRTTVLPCTVCQDT	91
ITQVNSs	VLTLPYLTHISQEPSPFHCLTTDSRCTAYRVTSPDGPCI	98
MPOVNSs	HSSSRWLITISLVEPSVRLRTLT	-79
	*	
NDVNSs		
E147VNSs		
BWANSs		
CEVNSs		
BUNVNSs		
GAMVNSs	GSMEO 130	
SJ2441VNSs	GSMEO 130	
OROVNSs		
TTOVNSs		
MPOVNSs		

Fig. 4.7 Alignment of NSs protein. * indicate 100% conserved; dotted lines indicate gaps inserted to maximize identity; red bold letters indicate LPS conserved motif : denotes more than 90% conserved.

JS BWAV EI	Δ	7 50	V 28 2	V 7 2	45 8	IV 31 2	7 31 3	51 5	4	V 25 3
ITV GAN			7	4 29	1 23	6 8 0	2 30	2 30	5 19	1 29
ΔN						_			_	_
MBOV					22	29	25	25	15	30
NDV						23	29	51	13	30
SI2441V							30	8	8	8
BUNV								90	25	25
CEV									18	34
идті										31
OROV										

Table 4.4 NSs amino acid sequences identities of the viruses. Bold values represent sequence identity of the viruses within the same serogroup



A. Phylogenetic tree using Neighbor-joining based on complete amino acid N sequence



B. Phylogenetic tree using Neighbor-joining based on partial amino acid N sequence.



C. Phylogenetic tree using Maximum Parsimony method (Mega 2.1 software) based on partial amino acid N sequence.

Fig. 4.8 Phylogeny of the studied viruses based on complete N ORF amino acid sequence of ANAV, ANBV, BWAV, BORV, E147V, GAMV, MBOV, NDV, SJ2441V and TETEV, and partial N ORF of BERV, GMAV and PLSV. Viruses in the Bunyamwera, California, Group C and Simbu serogroups were used as a control group. The sequences were aligned using the ClustalX (NCBI) and ClustalW program (EMBL-EBL). Phylogenetic analyses were generated by Neighbor-joining (NJ) based on complete N amino acid sequences (A) and partial N sequences (B) and Maximum Parsimony (MP) based on partial N sequences (C). Numbers adjacent to each branch represent the percentage bootstrap support calculated for 100 replicates. All the trees yielding almost identical topologies, except with TCMV. Blue letters indicate the phylogeny results that not correlated with serological grouping and red letters indicate serogroups used in this study. Scale bar represent 10% amino acid sequence divergence. Accession numbers of the control viruses are listed in 2.2.11.

was obtained by the heuristic search using equal weighting of all changes. Weighting schemes of 3:1 and 10:1 were used and gaps were treated alternatively as missing data. The reliability of the inferred tree was tested by bootstrap test performed on 100 replicates data sets generated using heuristic search option and a bootstrap consensus tree was computed. Representative members of Bunyamwera, California, Group C and Simbu serogroups were included in this analysis to root the trees. The results indicated that the overall phylogenetic relationships determined using NJ and MP methods were similar except with one difference: the NJ plot based on complete N sequence and MP plot based on partial sequence predict TCMV in the same clade as ANAV, while the NJ plot based on partial sequence, separated these viruses into different clades. Previous serological classification grouped BERV and GMAV in the Guama serogroup and CAPV in Capim serogroup. However, phylogenetic analysis based on partial amino acid sequence of CAPV, BERV and GMAV N proteins grouped BERV in the same clade as CAPV but separate from GMAV, indicating that BERV was closer to CAPV than GMAV. With the exception of BERV and TCMV viruses, all viruses from the same serogroup were placed in the same clade in NJ and MP derived trees.

4.2.8 Partial M and L sequences

During the attempts to get the reverse-transcription PCR product for S segments, primers S13+ and S13- were also used, which were designed based on conserved sequences of 13 terminal nucleotides of the S, M and L segments of the published sequences of Bunyamwera, California, Group C and Simbu serogroups viruses. Partial sequences of the L and M segments of PATV, and of the L segment of CAPV were obtained (Figure 4.9). For PATV L segment, both primers were found to hybridise at the internal sequences. Only the forward primer hybridised at the 3' end of virus-sense M RNA segment of PATV and L segment of CAPV, while the reverse primer hybridised at internal sequences (Figure 4.10). The lengths of M and L segments obtained for PATV were 505 nucleotides (representing 151 amino acids) and 922 nucleotides (representing 307 amino acids) respectively. For CAPV L segment, a sequence of 742 nucleotides (representing 237 amino acids) was obtained. The sequences of representative virus of each serogroup; BUNV for Bunyamwera, LACV for California, ITQV for Group C and OROV for Simbu. However the comparisons were very limited due to the limited sequences of M and L

A. PATV M sequence

1	AGTA	GTG	TGC	TCC.	ACA	ATA.	AAG	ATT	TCA	AAG	GAA	ATA	TCA	AAA	AGA	TAC	ACC	ATG	TTA.	AC
																		М	L	т
61	TAAA	ATG	CTG	CTG	TTT	GCA	CTG	CTT	ACA	GTA	TGT	GCA	TCA	GTA	CCC	TTG	GAT	AAG	TGC	ТΤ
	K	М	L	L	F	A	L	L	т	v	С	A	S	v	Ρ	L	D	K	С	F
121	CAGT	GGA	GGC	ATA	GTT	ATA.	AAA	AAG	CTT	CAA	ACT	GAT	TAT	GGC	CTC	CCC	CAT	ATG	TGT	СТ
	S	G	G	I	v	I	ĸ	ĸ	L	Q	т	D	Y	G	L	P	н	М	С	L
181	AAGA	GAC	GAT	ATA.	AGC	ATG.	ATA	AAA	ACT	GAT	TCG	GTT	GCA	GTT	CAA	GGT	ACT	TCA	AAT	GΑ
	R	D	D	I	S	М	I	K	т	D	S	v	A	v	Q	G	т	S	N	D
241	TGAG	AAG.	ACA	ATA	TTC	ACA	TCA	AGT	ATC	ACA	AGG	AAA	TTG	CTT	GTG	ACG	GAT	TGG	AAA.	AA
	E	ĸ	т	I	F	т	S	S	I	т	R	ĸ	L	L	v	т	D	W	K	N
301	CTGT	AGG	CCT	GAA.	AAA	ATG.	ATA	GGT	GGT	CCT	ATC	ATG	TTG	CTA	TCT	GTA	GAT	GAC	AGA	GG
	С	R	Ρ	Е	ĸ	М	I	G	G	P	I	М	L	L	S	v	D	D	R	G
361	TCAC	TTG.	AAA'	TCA	GAA	GAG	TAT	GTG	TGT	CAA	AAT	GAT	TGT	GAG	ATA	AAA	TTG	GAG	AAA	GΑ
	H	L	ĸ	S	Е	Е	Y	v	С	Q	N	D	С	E	I	ĸ	L	E	K	Е
421	ATCA	GGA	TTG	ATA.	ATA	TTT	GAA	ACA	GCA	TCA	CTA	AAT	TAC	TAT	CAA	GTG	AGT	GGT	ACC.	AC
	S	G	L	I	I	F	Е	т	A	S	L	N	Y	Y	Q	v	S	G	т	т
481	CATC	AGT	TCT	GGC	TGG	TTT.	AAA	TCA												
	I	S	S	G	W	F	к	S												

B. PATV L sequence

1	AGTA	GTG	TGC	CCC	TCA	GCA	AGT	ATT	CAA	TCT	AAG	AAA	TCA	ACT	GTG	GTG	TTT	TCA	ACA.	AT
	V	V	C	P	<mark>S</mark>	A	<mark>S</mark>	I	Q	S	K	K	S	T	V	V	F	<mark>S</mark>	T	I
61	TTGT	CTA	CAT.	AAG	GAT	TCA	AAG	GAT	gta	TTG	AAA	TGC	GGG	GCA	TTG	TAT	AAA	ACA	TAC:	AA
	C	L	H	<mark>K</mark>	D	S	<mark>K</mark>	D	V	L	<mark>K</mark>	C	<mark>G</mark>	A	L	Y	K	T	Y	<mark>K</mark>
121	AATA	GGA	.GAC	CAA	TAC	ATA	TCT	ATA	TCT	AAA	GCT	ATT	CGT	TTA	.GAT	AAA	.GAG	AGA	TGC(CA
	I	<mark>G</mark>	D	Q	Y	I	<mark>S</mark>	I	<mark>S</mark>	<mark>K</mark>	A	I	R	L	D	K	E	<mark>R</mark>	C	Q
181	GAGG R	TTA L	GTC. V	ACA T	GCT A	CCT P	gga <mark>G</mark>	ATT I	TTC F	ATG M	TTA L	ACC T	ACA T	CTA.	TTG. L	TTG L	AAA K	GGG <mark>G</mark>	GAT. D	AC T
241	AGAA E	ATA I	.GAC D	TTA L	GAT D	GAG E	ATC I	ATG M	GCT A	TTC F	TCA <mark>S</mark>	TTT F	TTC F	ACC T	TCA S	CTA.	TCT S	ATA <mark>I</mark>	ACA. T	AA <mark>K</mark>
301	AAGT	ATG	CTC	TCA	CTA	ACA	GAG	CCA	TCT	AGG	TAT	ATG	ATA	ATG	AAC	TCG	TTG	GCA	GTG	AG
	<mark>S</mark>	M	L	<mark>S</mark>	L	T	E	P	<mark>S</mark>	R	Y	M	I	M	N	S	L	A	V	<mark>S</mark>
361	CAGC	CAT	GTC	CGA	GAA	TAT	ATA	GCC	GAA	AAG	TTT	TCT	CCA	TAT	ACA	AAA	ACC	TTA	TTC'	TC
	<mark>S</mark>	H	V	<mark>R</mark>	E	Y	I	A	E	<mark>K</mark>	F	S	P	Y	T	K	T	L	F	S
421	AGTT	TAT	ATG	ACA	GAA	TTG	ATC	CAC	AAA	GGC	TGC	ATG	TCA	.gca	.AAC	AAC	CAA	CGG	TCT:	AA
	V	Y	M	T	E	L	I	H	<mark>K</mark>	<mark>G</mark>	C	M	S	A	N	N	Q	R	S	<mark>K</mark>
481	GATA	TCT	GTG	AAA	AAT	GTT	TTC	CTC	AAT	GAA	TAT	GAA	ATA	ACT	CAA	AAA	.GGG	GTT	TCA	GA
	I	S	V	<mark>K</mark>	N	V	F	L	N	E	Y	E	I	T	Q	<mark>K</mark>	<mark>G</mark>	V	S	E
541	AGAT	CGC	GAT	CTA	GAA	TCA	ATC	TGG	TTT	CCA	gga	.CAT	GTA	AAT	CTC	AAG	GAA	TAT	ATC:	AA
	D	R	D	L	E	<mark>S</mark>	I	W	F	P	<mark>G</mark>	<mark>H</mark>	V	<mark>N</mark>	L	<mark>K</mark>	E	Y	I	N
601	tcaa	GTG	TAT	CTC	CCC	TTT	TAC	TTC	AAT	GCT	AAA	.GGG	TTA	.CAT	'AAT	AAG	CAT	CAT	GTT:	AT
	<mark>Q</mark>	V	Y	L	P	F	Y	F	<mark>N</mark>	A	<mark>K</mark>	<mark>G</mark>	L	<mark>H</mark>	<mark>N</mark>	<mark>K</mark>	H	<mark>H</mark>	V	M
661	GATA	GAC	TTA	GCT.	AAG	ACA	GTT	TTA	GAA	ATT	GAA	.CTA	GAA	.CAA	AGA	.GAG	CAA	CTG	CCC:	AA
	I	D	L	A	<mark>K</mark>	T	V	L	E	I	E	L	E	Q	<mark>R</mark>	E	Q	L	P	N
721	CCCA	TGG	GGC	AAA	GAT	TTC	CAG	AAA	CAA	TCT	GTG	AAC	CTA	.GAA	ATA.	TTG	ATC	TAT	TCG.	AT
	P	W	<mark>G</mark>	<mark>K</mark>	D	F	Q	<mark>K</mark>	Q	<mark>S</mark>	V	N	L	E	I	L	I	Y	<mark>S</mark>	I
781	AGCA	AAG	ATG	CTG.	AAC	AAT	GAT	ACA	TCT	AAG	CAT	AAT	CAT	CTA	AGA	AGC	AGG	ATT	GAG.	AA
	A	<mark>K</mark>	M	L	N	N	D	T	S	<mark>K</mark>	<mark>H</mark>	N	H	L	R	<mark>S</mark>	R	I	E	N
841	TAGG	AAC	AAC	TTC.	AAG	AGG	TCA	CTG	GCA	AGC	ATA	.TCG	ACA	TTC	ACT	AGC	TCG	AAG	TCA'	TG
	<mark>R</mark>	N	N	F	<mark>K</mark>	R	<mark>S</mark>	L	A	<mark>S</mark>	I	<mark>S</mark>	T	F	T	<mark>S</mark>	S	<mark>K</mark>	<mark>S</mark>	C
901	CATA I	AAA <mark>K</mark>	.GTG V	GGG <mark>G</mark>	GAT D	TTC F	TCG <mark>S</mark>													

C. CAPV L sequence

1	AGTA	GTGI	ACCO	CTGA	.CAT	AGT	ATT	TAG	СТА	TTT	AGA	AGG	GTA	CAC	TCA	AAA	ATG M	GCA. A	ACG T
61	CAAA Q	TTGO I O	GAGAZ	ATTA L	GTT V	AGG R	CAG Q	TTT F	GCT A	GCA A	AGA <mark>R</mark>	GTT. V	AGG R	GCA A	TTA L	AAC N	CCA P	AAT) N	CAA Q
121	CCAG P	AGTI <mark>e i</mark>	AGGC	GAGA <mark>R</mark>	GAT D	CTT L	CTG L	TCT <mark>S</mark>	GAA E	ATA. I	ACT T	gta V	GCT A	CGA R	CAT <mark>H</mark>	AAT N	TAT Y	TTT(F	GCC A
181	CAAG Q	AGTI E F	TTGC	CGAG E	AGC <mark>S</mark>	ATC I	gga <mark>G</mark>	TTA L	GAA E	TAT. Y	AGA <mark>R</mark>	AAC N	GAT D	GTA V	CCA P	GCT A	CTG L	AAT. N	ATA <mark>I</mark>
241	GTAA V	TGGA M E	GAT(M	GAGA R	.CCG P	GAC D	TTC F	GAT D	CCT P	ATG M	ACA T	ATT. I	AAG <mark>K</mark>	GTG V	CCG P	GAA E	ATC I	ACA T	CCA P
301	GACA D	ACTA N Y	ATTA(Y	CAGA R	.GAC D	GGG <mark>G</mark>	GAC D	AAG <mark>K</mark>	ATC I	TAT Y	ATT. I	ATA I	GAT D	TTT F	AAA <mark>K</mark>	GTA V	TCG <mark>S</mark>	GTT: V	AGT <mark>S</mark>
361	GACG. D	AATC E S	CTGCT	FAAT <mark>N</mark>	CAT <mark>H</mark>	ACA T	TTC F	AAG <mark>K</mark>	AAA <mark>K</mark>	TAC Y	AAT <mark>N</mark>	ACT T	TTG L	TTC F	GGC <mark>G</mark>	GAT D	GTC V	TTT(F	GAT D
421	AAGT <mark>K</mark>	TAGO L C	GAT: J	ГGAA <mark>E</mark>	TTC F	GAG E	GTA V	GTC V	ATT I	ATT I	AGG R	ATG M	GAC D	CCC P	ACA T	AGA <mark>R</mark>	gga <mark>G</mark>	GAG. E	ATA <mark>I</mark>
481	TTGA L	AGGO K G	GTTGO W	GGAC D	AAA K	ATG M	GTC V	GAA E	AGA <mark>R</mark>	.GTC V	TCA <mark>S</mark>	GAA E	ACT. T	AGA <mark>R</mark>	GAG E	ATA I	ACC T	AGA R	GAT D
541	GTGA V	CCAA T F	GCAC	gaag <mark>k</mark>	CCA P	AGT <mark>S</mark>	ATA I	CAT <mark>H</mark>	TTC F	ATA I	TGG W	TCT <mark>S</mark>	GGG <mark>G</mark>	CAC H	AAT <mark>N</mark>	TCA <mark>S</mark>	AAA <mark>K</mark>	GAA E	GTT V
601	ACAG T	GCAA G N	CAC	fcag <mark>Q</mark>	AAA K	ATA <mark>I</mark>	TTG L	AGG R	CTC L	TCA S	AAA <mark>K</mark>	TGT C	CTA L	CAA Q	AGA R	ATA I	AAG <mark>K</mark>	GAC. D	ACC T
661	GACC. D	AATI <mark>q F</mark>	CAGO	CACA T	.GCT A	TTT F	AGG R	AAA <mark>K</mark>	ATA I	.GGT <mark>G</mark>	GCA A	CTA L	ATG M	GAT D	TTC F	AGT <mark>S</mark>	GAA E	AAC. N	ATA <mark>I</mark>
721	GCAG A	GCTA G Y	ACGAC	GAAT <mark>N</mark>	TTT F	TGC C	CTG L	AAG <mark>K</mark>	CTA L	.AAA <mark>K</mark>	.gca A	GAC D	GCG A						

Fig. 4.9 Nucleotide and amino acid sequences of positive-sense of partial M (A) and L segments (B) of PATV and partial L of CAPV (C). The nucleotides were translated to amino acid using DNAMAN version 4.0 program (Lynnon Biosoft)



Figure 4.10 Position of L sequence obtained for CAPV and PATV (A), and M sequence obtained for PATV (B). Red letters indicate position relative to BUN sequences

available. For instance, there are no sequences of L and only partial sequences of M segment available for Group C virus and this M sequence was found to be outside the comparison area of PATV. In comparison with L sequences of BUNV, LACV and OROV, the identities of PATV are 65%, 62% and 64% respectively for nucleotides, and 64%, 60% and 61% respectively for amino acids, while for CAPV, the identity values are 43%, 46% and 47% respectively for nucleotides, and 28%, 26% and 30% respectively for amino acids. The nucleotide and amino acid sequence identity of PATV M segment compared to BUNV, LACV and OROV is 30%, 43% and 24% respectively for nucleotides, and 37%, 36% and 33% respectively for amino acids. The L protein sequences alignment of PATV revealed 20 deletion of amino acid at the position of amino acid 664 relative to BUNV.

4.3 Discussion

Even though Dunn et al. (1994), who designed BUN S+/S- successfully obtained the complete sequences of the S segment of seven viruses in Bunyamwera serogroup and Nunes et al. (2005) managed to get complete sequences of 13 S segments of viruses in the Group C serogroup, other researchers like Bowen et al. (1995) failed to use these primers to amplify six of the viruses in California serogroup [CEV, JCV, Jerry Slough (JSV), Melao (MELV), Keystone (KEYV) and TVTV]. An attempt to get complete S segment sequences for Garissa virus with these primers during the hemorrhagic fever outbreak in Kenya and Somalia (Bowen et al., 2001) also failed although partial S segment sequences obtained using BUNCAL1 and BUNCAL2 primers revealed its closeness to BUNV (>98% identity). From the sequences obtained from G-tailed cDNAs of California serogroup viruses, variations in the terminal sequences of 3' and 5'end were detected. In this study, repeated attempts to amplify the complete S segment of some viruses using primers BUN S+/S- were also unsuccessful except with the viruses in Nyando serogroup, although different sources of RNA such as virus infected cell, virion RNA and nucleocapsid RNA were used as a template for cDNA synthesis. Variable annealing temperatures were also applied but yet failed to get the expected products. Perhaps some viruses do not share the same consensus 3'/5' termini sequences of the published orthobunyaviruses. This is further suggested by the data of Ushijima et al. (1981), who showed by direct RNA sequencing that the terminal RNA sequences of Pahayokee virus were different to those of published orthobunyaviruses (UCAUCAAAUGAAGAU). Sequence analysis of partial S sequences

of PATV, which were obtained using designed internal primer S400F and terminal primer S13-, revealed that nucleotides 14 and 15 of the 3' end of the positive-sense RNA were different to the published S sequences (CA instead of AG- refer to Figure 4.1). This could explain why I could not get any product with some of the studied viruses using primers BUN S+/S-.

Attempts to get the product using 3'/5' RACE and RCA-RACE, modified 3'/5' RACE and reverse-transcription PCR on ligated RNA using the internal primers that were designed based on the partial sequences obtained also failed. One possible explanation would be that ribonucleases present in some of the preparation could have degraded the RNA. Although ribonuclease inhibitor was added to the reaction during the reverse-transcription process, degradation could occur during the RNA extraction process. Polidoros et al. (2006) suggested that the failure to get the product with 3'/5' RACE could be due to the use of a universal primer corresponding to the anchor sequence was present in all cDNAs and the presence of many truncated RNAs. Combination of internal primers that were designed based on some conserved motifs of the published sequences also did not yield any product with some of the viruses, suggesting that not all the viruses contain these motifs.

The other major task in this study was to design the universal primers that are conserved for all the viruses and could be used to amplify the S segment of viruses in all 18 serogroups of *Orthobunyavirus* genus. Besides that, cross-contamination of the preparation with nucleic acid of other viruses especially BUNV was also one of the major problems in this study since all the experiments such as RNA extraction, reagent preparation and gel running were conducted in the same laboratory. Therefore, beside separating the place for RNA extraction, reagent preparation and running gel, filtered tip and pipettors treated with RNase AWAY (Molecular BioProducts, Inc) were also used during some of the preparations to prevent this contamination.

This study represents the first report on the complete S segment sequences of viruses in Anopheles A, Anopheles B, Nyando, Bwamba, Turlock, Tete and Gamboa serogroups, partial S sequences of viruses in Patois, Guama, Capim and Minatitlan serogroups, partial L sequences of PATV and CAPV, and partial M sequences of PATV. These partial sequences will be useful as a template to obtain a complete sequence in the future. Analysis for AUG-initiated ORFs (FRAMES analysis) and alignment of N with NSs initiation codon revealed that some viruses lacked on NSs protein initiation codon. This is the first report on the presence of naturally occurring orthobunyaviruses that do not have NSs ORF, i.e. those in Anopheles A, Anopheles B, Tete and Capim serogroups. Whether the NSs function has been replaced by the N protein or by proteins encoded on the other segments is not known and therefore needs to be further investigated. It is interesting that most of these viruses were found to have longer N proteins compared to viruses with an NSs ORF. Similar to GROV NSs protein, a shortening of NSs protein of MPOV was also observed at both ends. It is possible that this is a homoplastic event in the evolution of orthobunyavirus that by having short or without NSs protein, to free the S segment from the constraints involved in encoding N and NSs proteins in the same sequence.

Analysis of the second ORF potentially to encode NSs protein using FRAMES programme of the DNAMAN software has detected the presence of different initiation codons at the same relative spacing as BUNV and CEV in BMA and CAPV, and at the same relative spacing as OROV and ITQV in ANAV, TCMV, ANBV, BORV and TETEV, but they were considered too short to encode for NSs protein. In TCMV, an ORF starts with TTG (Figure 4.3) with the length of 61 amino acids was observed. Whether they are the ORFs for NSs protein need to be further investigated. It is possible that ATG initiation codon of these viruses had been lost or mutated over time due to certain selective pressures.

Sequence alignments of the nucleotide and amino acid sequences of N protein obtained here with the published sequences of viruses in Bunyamwera, California, Group C and Simbu serogroups have identified some conserved regions. These regions might contain possible functional domains which might play a role in N protein interaction with L protein and virus RNA or epitopes for CF antibodies to react.

Previous serological classification has placed TCMV in the same serogroup as ANAV from the Anopheles A serogroup and phylogenetic analysis using NJ plot based on complete S sequence and MP plot based on partial N sequence supported this classification. However based on published data, their nucleotide and amino acid sequence identities were found outside the range to be placed in the same serogroup, and phylogenetic analysis using NJ method based on partial N sequence have placed this virus in different clade from ANAV. However a similar result has been reported before with the viruses in the Simbu serogroup, where the N amino acid sequence identity between

Mermet virus (MERV) and Douglas virus (DOUV), and between Buttonwillow virus (BUTV) and AKAV were as low as 59%, suggesting that greater sequence divergence was observed in Simbu serogroup viruses (Saeed et al., 2001). Serological classification has placed BERV in the same serogroup as GMAV in Guama serogroup. Based on the comparison of partial amino acid sequences of the N protein and phylogenetic analysis using NJ and MP methods, BERV was found to be closer to CAPV compared to GMAV. However sequence identities between these three viruses were found to be more than 70%, suggesting that these viruses originated from the same ancestor; therefore, it is recommended that they be placed into one group. However, comparison of complete N sequences of BERV, CAPV and GMAV are needed to confirm this. It is also possible that TCMV and BERV are reassortant viruses that have different S segment but similar M segment as ANAV and GMAV respectively, and were placed in their current serogroups based on PRNT or HI assays which recognise epitopes encoded by M RNA gene products. Therefore, sequence analyses of M segment of ANAV, TCMV, BERV, CAPV and GMAV are necessary to confirm this speculation. The sequence comparison result of BWAV was in agreement with previous serology classification, which also identified BWAV as 'bridging' virus between California and Bwamba serogroups (Calisher, 1988). Previously, GAMV has been identified as a bridging virus between Gamboa and Capim serogroup (Calisher, 1988). However, sequence comparison result has showed that the sequence identity for GAMV and CAPV was only 35%. Based on the published data, this value was too low to place this virus as a bridging virus between these two serogroups.

The last ten amino acid residues including the LPS motif at the carboxy terminal of NSs protein, which were conserved in Bunyamwera and California serogroup viruses was only detected in BWAV. However this motif was also absent in GROV, Group C and Simbu serogroup viruses (Fig. 4.5). This domain is reported to be involved in the inhibition of cellular RNA polymerase II-driven transcription through the interaction with MED8 component of mediator, a protein complex that is necessary for the host mRNA production (Leonard et al., 2006). Absence of the domain was shown to reduce the ability of the virus to inhibit host protein expression and the interferon response. This suggests that this motif is not conserved in all orthobunyaviruses and it is possible that these LPS minus viruses use a different binding motif or interact with a different cellular partner to inhibit host protein expression and interferon response.

5 Interferon response and Shutoff of Host Protein Synthesis caused by NSs Minus Orthobunyaviruses

5.1 Introduction

Sequencing data in Chapter 4 revealed that viruses in 3 serogroups (Anopheles A, Anoheles B and Tete) as well as CAPV do not show any evidence of having an NSs ORF. As mentioned in Chapter 1, the NSs proteins of BUNV in Bunyamwera serogroup and LACV in California serogroup are necessary to cause the shutoff of host protein synthesis in mammalian cells and also act as interferon antagonists (Bridgen et al., 2001; Weber et al., 2002; Blakqori and Weber, 2005). In this chapter, I analysed the proteins expressed by the S segments of these NSs-minus viruses, and studied the shutoff of host protein metabolism and induction of the interferon response in infected cells. CAPV was omitted in this study because its S sequence was only obtained after these experiments had been conducted.

5.2 **Results**

5.2.1 *In vitro* transcription and translation (TnT)

This TnT system allows coupled *in vitro* transcription and translation of protein coding regions cloned downstream of the T7, SP6 or T3 RNA polymerase promoter in a eukaryotic environment. The S segment cDNAs of the viruses were cloned into pBluescript2 KS plus plasmid (Strategene) under the control of a T7 RNA polymerase promoter. The constructs were used to program protein synthesis in TnT systems (Promega) using wheat or rabbit reticulocyte lysates in the presence of [³⁵S] methionine. The products were then separated by SDS-PAGE. Plasmid without an insert was used as a control.

In vitro TnT reactions using both rabbit reticulocyte and wheat germ lysate systems showed that only N protein was expressed from the S segment of these viruses; in the control BUNV S segment, both N and NSs were translated (Figure 5.1). The N protein bands of these NSs minus viruses were shown to be larger than BUNV and BUNdelNSs N



Fig. 5.1 Proteins expressed from orthobunyavirus S segments in the *in vitro* TnT system using wheat germ lysate and rabbit reticulocyte lysate system. The products were fractionated by polyacrylamide gel electrophoresis [NUPAGE 6-12 % gradient gel (Invitrogen)].

protein bands. In the rabbit retyculocyte lysate protein gel, smaller products less abundant than BUNV NSs band were detected. However, they were not obviously seen in the other gel using wheat germ lysate system. Globin bands of 15–16 kDa were also observed in some of the lanes using this system but not in the wheat germ system. In the wheat germ system, a band of 16 kDa was observed in the BUNdelNSs lane but not in the other system. A non-specific band of more than 26 kDa was seen in the control plasmid of both gels, which was clearly observed in wheat germ system. How this band appeared is unknown.

5.2.2 Shutoff of host protein synthesis

Two assay systems, direct protein radio-labelling of infected cells and a luciferase expressing plasmid based assay were used in this study to examine the shutoff caused by the different viruses.

5.2.2.1 Shutoff in A549 infected cells

In these cells, a drastic shutoff was only observed in cells infected with TCMV and BUNV, but the level of shutoff caused by TCMV was slightly slower compared to BUNV, where the shutoff was almost complete at 24 h pi. In comparison to shutoff in mock infected cells, a slight shutoff was observed in cell infected with ANAV, ANBV, BORV, TETEV, BMAV and BUNdelNSs virus (Figure 5.2).

5.2.2.2 Shutoff in Hep2 infected cells

No shutoff was observed in Hep2 cells infected with any of the virus including BUNV (Fig. 5.3). The N protein bands were only observed in TCMV, BORV, BUNV and BUNdelNSs-infected cells with the highest amount in BUNV-infected cells. No obvious viral protein bands were observed in ANAV, ANBV, TETEV and BMAV-infected cells.

5.2.2.3 Shutoff in Hep2/V cells infected with the viruses

In these infected cells, similar result was observed as in infected Hep2 cells where no shutoff was observed. Unlike in Hep2-infected cells, the N proteins expressed by TCMV,



Fig. 5.2 Shutoff of host protein synthesis in infected A549 cells at 18 h pi. The cells were infected with 1 PFU/cell virus and labeled for 2 h with 30 μ Ci [³⁵S]methionine. Infected cell lysates were harvested and analysed by 15% SDS-PAGE gel.



Fig. 5.3 Shutoff of host protein synthesis in infected Hep2 cells at 18 h pi. The cells were infected with 1 PFU/cell virus and labeled for 2 h with 30 μ Ci [³⁵S]methionine. Infected cell lysates were harvested and analysed by 15% SDS-PAGE gel.

BORV and BUNdelNSs in Hep2/V-infected cells were almost similar in amount to that in BUNV-infected cells (Fig. 5.4). Similar to Hep2-infected cells, no viral protein bands were observed in cells infected with ANAV, ANBV, TETEV and BMAV. Thus, the growth of these viruses was either inhibited or delayed in these cells.

5.2.3 Shutoff of protein synthesis using a luciferase reporter based assay

5.2.3.1 Shutoff in infected BHK-21 cells

Transfection of cells with plasmids expressing luciferase gene has been used to investigate the role of influenza A NS1 and rotavirus NSP3 proteins in shutoff of host protein synthesis (Salvatore et al., 2002; Vende et al., 2000). A similar assay to examine the shutoff caused by BUNV using a luciferase reporter plasmid (pRL-SV40 or phRL-CMV) to investigate the effect of NSs on host protein synthesis in infected BHK-21 cells was first established by Hart (2004). These two plasmids contain the luciferase gene isolated from Renilla reniformis and were obtained from Promega. pRL-SV40 expresses Renilla luciferase under the control of the SV40 early enhancer/promoter while phRL-CMV expresses Renilla luciferase under the control of the CMV immediate early enhancer/promoter. pRL-SV40 was recommended for use in mammalian cells while phRL-CMV was recommended for used in insect cells due to a report of its efficient expression in insect cell lines (Saraiva et al., 2000). BHK-21 cells were transfected with pRL-SV 40. After 5 h incubation at 37°C, the cells were infected with 1 PFU/cell of virus and the luciferase readings were taken after 12 and 24 h pi. There was a decrease in luciferase activity observed in most of the virus infected cells after 12 and 24 h pi compared to mock infected cells. The most reduction of luciferase activity was observed in BUNV-infected cells, followed by ANAV, ANBV and TCMV-infected cells, which were lower than in BUNdelNSs infected cells (Figure 5.5A). The luciferase activities in cells infected with TETEV, BMAV and BORV were found to be higher than in BUNdelNSs-infected cells, indicating that less shutoff of host protein synthesis occurred. For a better comparison between different experiments, Hart (2004) recommended to calculate the luciferase activities by dividing the luciferase activities in the virus-infected cell with those corresponding mock infected cells. The highest percentage reduction was observed in BUNV-infected cells after 12 and 24 h pi (almost 50% reduction in comparison to



Fig. 5.4. Shutoff of host protein synthesis in infected Hep2/V cells at 18 h pi. The cells were infected with 1 PFU/cell virus and labeled for 2 h with 30 μ Ci [³⁵S]methionine. Infected cell lysates were harvested and analysed by 15% SDS-PAGE gel.







Fig. 5.5 Shutoff of protein synthesis in infected BHK-21 cells at 12 and 24 h pi from expression of pRL-SV40. The cells were transfected with 0.1 μ g/dish pRL-SV40 using Lipofectamine 2000 diluted in OptiMEM. After 5 h incubation at 37°C, the cells were infected with 1 PFU/cell virus and the cells were lysed with passive lysis buffer after 12 and 24 h pi. The samples were assayed using manual luminometer and Dual Luciferase assay system (Promega). Results are shown as the luciferase readings from infected cells without manipulation (A) and percentage luciferase activity of infected cells compared to the corresponding mock infected cells (B). Results are the average readings of 3 samples.

BUNdelNSs-infected cells), followed by cells infected with ANAV, ANBV and TCMV, BUNdelNSs, BMAV, TETEV and BORV (Figure 5.5B).

5.2.3.2 Shutoff in C6/36 infected cells

In this study, the cells were transfected with phRL-CMV and the luciferase reading was taken at 18 h pi. Previous studies using radiolabeled protein of infected cells and luciferase based assays have showed no shutoff observed in cells infected with BUNV and BUNdelNSs virus (Elliott & Wilkie, 1986; Scallan & Elliott, 1992; Kohl et al., 2004, Hart, 2004). However, luciferase based assay used in this study revealed that in comparison to mock infected cells, no shutoff was only observed in C6/36 cells infected with BUNdelNSs virus, whereas shutoff less than 50% compared to the mock-infected cells was still observed in the cells infected with BUNV, ANAV, TCMV, ANBV, BORV, TETEV, BMAV (Figure 5.6).

5.2.4 Induction of IFN- α/β in infected cells

It has been shown previously that the NSs protein of BUNV and LACV acts as an IFN antagonist that blocks the transcriptional activation of IFN- α/β . Therefore, the ability of the NSs minus viruses to induce IFN- α/β was investigated by using two assay systems; IFN promoter induction assay and RT-PCR to detect the presence of IFN- α/β mRNA in infected cells.

5.2.4.1 Reporter gene assay

This assay was conducted to determine the ability of the NSs minus viruses to transactivate the IFN- β promoter using a bioassay based on transfected reporter plasmids. The reporter plasmid expressing the firefly luciferase gene under the control of the IFN- β promoter (p125Luc) was used in this study. Confluent A549 cells were transfected with p125Luc plasmid and infected with 1 PFU/cell of virus after 5 h incubation. Luciferase readings in infected cell lysates were taken after 18 h pi. Wt BUNV, BUNdelNSs and mock- infected cells were used as a control. The most activation of IFN- β promoter was observed in TETEV-infected cells, followed by BORV, BUNdelNSs and ANBV-infected cells with more than 10-fold increase of luciferase activity compared to those of mock-infected cells



Fig. 5.6 Inhibition of protein synthesis in infected C6/36 cells from the expression of phRL-CMV at 18 h pi. The cells were transfected with 0.1 μ g/dish phRL-CMV using Lipofectamine 2000 diluted in OptiMEM. After 5 h incubation at 37°C, the cells were infected with 1 PFU/cell virus and the cells were lysed with passive lysis buffer at 18 h pi. The samples were assayed using manual luminometer and Dual Luciferase assay system (Promega). Results are shown as the percentage luciferase activity of infected cells compared to the corresponding mock infected cells and are the average readings of 2 samples.

(Figure 5.7). A slight activation, less than 5-fold increase in luciferase activity compared to mock-infected cells, was observed in ANAV, BMAV and TCMV-infected cells, while BUNV was almost unable to activate IFN- β promoter.

5.2.4.2 IFN-specific RT-PCR

It has been shown previously that NSs protein of BUNV interferes with the production of IFN by blocking transcription of the IFN gene (Weber at al., 2002). To investigate the presence of IFN-B mRNA in infected cells, RT-PCR using primers IFN-BF and IFN-BR (Table 2.1) on total RNA extracted from infected 293 cells was carried out. Controls included RNA from mock infected cells, PCR without preceding RT and a reaction without template, as well as RT-PCR using actin specific primers. A clear positive band was observed using RNA extracted from cells infected with ANBV, BORV, TETEV and BUNdelNSs virus, while a very faint band was observed with RNA from ANAV and BMAV-infected cells (Figure 5.8A). Only TCMV gave a negative IFN- β mRNA signal similar to BUNV and mock-infected cells. These positive bands were not observed in all PCR without preceding RT reactions (Fig.5.8C), suggesting that there was no contamination with residual genomic DNA in the RNA preparations. No band was observed in RT-PCR reaction using water as template (Figure 5.8A), indicating that no cross-contamination occurred in these preparations. RT-PCR to detect the presence of cellular y-actin mRNA was used to examined whether all the preparations contained similar amounts of the RNA. The band for cellular γ -actin mRNA was almost the same intensity in all the samples, indicating that they contained the same amount of RNA (Fig.5.8B). These results demonstrated that with the exception of TCMV, IFN- β mRNA was synthesized in cells infected by all the NSs minus viruses.

5.3 Discussion

Sequencing data for viruses in four serogroups of orthobunyaviruses (Bunyamwera, California, Simbu and Group C) revealed the presence of an NSs protein in their S segments encoded in an overlapping ORF with the N protein (Elliott et al., 1989; Dunn et al., 1994; Bowen et al., 1995; Saeed et al., 2001; Nunes et al., 2005,). The roles of the NSs protein in causing shutoff of host protein synthesis in mammalian cells and as an interferon antagonist were studied previously using mutant viruses of BUN (BUNdelINSs) and LAC



Fig. 5.7. Activation of IFN- β promoter. Human A549 cells were transfected with the reporter plasmid p-125Luc containing the firefly luciferase gene under the control of IFN- β promoter. After 5 h incubation, the cells were infected with 1 PFU/cell viruses or left uninfected (mock) and luciferase activities were taken 18 h later. The samples were assayed using manual luminometer and single Luciferase assay system (Promega). Results are the average reading of 3 samples.

A IFN β -mRNA]





(rLACVdelNSs) that do not express NSs (Bridgen et al., 2001; Weber et al., 2002; Blakqori and Weber., 2005). Here, I report the presence of naturally occurring viruses lacking the NSs protein, and have analysed the proteins expressed by their S segment using *in vitro* transcription-translation kit and their ability to cause shutoff of host protein metabolism and to induce the interferon response in mammalian cells.

5.3.1 In vitro Transcription and Translation

Sequencing and FRAMES analysis data in Chapter 4 revealed that viruses in three serogroups (Anopheles A, Anopheles B, Tete) were without NSs protein as no potential second ORF to encode NSs protein was present in their S segments. *In vitro* TnT of these viruses supported this finding, where no NSs protein was obviously expressed in both TnT protein gels using wheat germ and rabbit reticulocyte lysate expression systems. Less abundant and smaller bands than the NSs product of BUNV were observed in the protein gel using rabbit reticulocyte lysate system, however, these could be artifacts of this *in vitro* system since these bands were absent in the wheat germ lysate system. Globin bands of 15–16 kDa were obviously observed in some of the lanes in rabbit reticulocyte lysate system but these bands were not detected in the other gel system, suggesting that these could be also an artifact of this rabbit reticulocyte system. However the reason why these bands were only obvious in certain lanes is not known.

5.3.2 Shutoff of host Protein synthesis

The shutoff of host protein synthesis induced by bunyaviruses appears to be caused by multiple factors. It has been shown previously that NSs protein of orthobunyaviruses is responsible in contributing to host cell shutoff in mammalian cells (Bridgen et al., 2001; Blakqori and Weber, 2005). In the nucleus, NSs inhibits the transcriptional activity of RNAP II by interfering with CTD phosphorylation (Thomas et al., 2004) and in the cytoplasm, degradation of cellular capped mRNAs by cap-snatching of L protein further contributes to the depletion of mRNAs (Raju and Kolakofsky, 1988).

In agreement with the shutoff observed by direct radio-labelling of infected Vero cells, shutoff was also observed in BHK-21 cells infected with ANAV, TCMV, ANBV and BORV using the luciferase based assay. However, in contrast to the direct radio-labelling

B γ-actin mRNA





500 bp

Fig.5.8 Detection of IFN- β and cellular γ -actin mRNA. Total RNA was extracted from 293 infected cells at 18 h pi, treated with DNase, reversed-transcribed using hexanucleotide primers and amplified with specific primers for IFN- β mRNA (A), cellular γ -actin mRNA (B), PCR without preceding RT step (C). Products were analysed by agarose gel electrophoresis.

experiment where no shutoff was observed in cells infected with TETEV and BMAV, a slight shutoff was seen with luciferase based assay, suggesting that this reporter assay is more sensitive. Furthermore, this assay was found to be more reliable and could be measured quantitatively (Hart, 2004). Compared to the mock-infected cells, reduction of luciferase activities was still observed in cells infected with all of these NSs minus viruses, but not to the same extent as in BUNV-infected cells. These reductions could be because the cells' growth, division and metabolic pathway were impaired following virus infection but were not affected in mock-infected cells (Hart, 2004).

No shutoff was observed in Hep2 or Hep2/V infected with any virus including BUNV. N protein bands were clearly seen in both cell types infected with TCMV, BORV, BUNV and BUNdelNSs, with the same intensity in infected Hep2 cells but less abundant in infected Hep2/V cells compared to BUNV-infected cells. No viral proteins were observed in either Hep2 or Hep2/V cells infected with ANAV, ANBV, TETEV and BMAV, suggesting that the growth of these viruses was limited in these cells as previously observed with the inability of Vero-E6 cells to support the replication of RSV (Young et al., 2003). This indicates that some other factors such as host constraints may be involved in limiting or delaying the replication of these viruses.

In infected A549 cells, a rapid shutoff almost similar to BUNV-infected cells was also observed in cells infected with TCMV which was found does not have NSs protein in its S segment (Figure 5.1). The mechanism of how it caused this shutoff is unknown. This suggests that TCMV uses different strategies other than NSs protein to induce this shutoff. A slight shutoff similar to BUNdelNSs was observed in cells infected with ANAV, ANBV, BORV, TETEV and BMAV.

Although BUNdelNSs and these viruses are without NSs protein, slight shutoff was still detected in their infected cells. This could be due to the degradation of cellular mRNA during the cap-snatching process (Raju and Kolakofsky, 1988). However the effect of cap-snatching on shutoff of host protein synthesis was found to be low compared to the shutoff caused by NSs protein as observed in wt BUNV and LACV which encode NSs protein in their S segment.

Previous studies showed that orthobunyaviruses do not cause shutoff of host protein synthesis in *Aedes albopictus* C6/36 cells where infection is non-cytolytic and leads to long term viral persistence (Elliott and Wilkie, 1986; Hacker et al., 1989; Scallan and Elliott, 1992; Kohl et.al., 2004; Hart, 2004). In this study using the luciferase based assay, no apparent shutoff was observed with any of the viruses, including BUNV, supporting the previous result that the NSs protein is only functional in mammalian cells but not in C6/36 cells. This may contribute to the persistent infection in insect cells (Thomas et al., 2004).

5.3.3 Production of Interferon in infected cells

The NSs protein of BUNV has been shown to block IFN gene expression by targeting components of mammalian RNAP II (Thomas et al., 2004) and is linked to the interaction with MED8, a protein complex necessary for mRNA production (Leonard et al., 2006). Therefore, the presence of NSs protein in the virus was found to be of advantage for virus multiplication in cells and animals with a functional IFN- β system (Weber et al., 2002).

The IFN specific RT-PCR and IFN reporter gene assay results revealed that TCMV which does not encodes NSs protein in its S segment is still capable of counteracting IFN production similar to wt BUNV. In contrast to the previous result where all the bunyaviruses used NSs protein to counteract the IFN production, TCMV which does not have NSs protein might used different strategies to block the induction of IFN- β mRNA in 293 and activation of IFN- β promoter in A549 infected cells. In addition, TCMV was isolated from human with mild symptoms, indicating that this virus has the ability to counteract the IFN production of IFN- β mRNA and little activation of IFN- β promoter were observed in ANAV and BMAV-infected cells. These signals were not observed in PCR without preceding RT, indicating that they did not arise from contaminating residual genomic DNA.

These results led to the conclusion that beside the host constraints, capability to counteract interferon production and shutoff mechanism are required for the virus to grow efficiently in these mammalian cells.

6 General Discussion and Future Works

Previous serological classification has divided members of the *Orthobunyavirus* genus into 18 serogroups. Viruses in only four of these serogroups (Bunyamwera, California, Group C and Simbu) have been studied intensively. Therefore this study was conducted with the aim of acquiring new information about viruses in these unstudied serogroups. To achieve this, I first had to grow and adapt the viruses in different cell culture systems, then observe their plaque sizes, analyze their protein profiles and compare them with other orthobunyaviruses, and investigate the ability of these viruses to cause shutoff of host protein synthesis. This study also aimed to sequence the S segment of these viruses and compare them with the other orthobunyavirus sequences. Finally, I aimed to investigate the ability of the viruses that lacked an NSs protein to induce IFN and shutoff host protein synthesis. Significant achievements and conclusions that can be drawn from this study are discussed in the sections below.

6.1 Virus growth, protein synthesis and ability of the viruses to cause shutoff of host protein synthesis

All the viruses studied except BOTV were successfully grown in Vero-E6, BHK-21 and LLCMK2 cells. The failure to grow BOTV might be due to it already dead or it was unable to replicate in the mammalian cells used in this study.

In general, viruses in the same serogroup were found to produce similar plaque sizes. Whether plaque size is related to virulence could be verified by inoculating the virus into an animal host. However a previous study showed that large and small plaque variants of GROV were similar with respect to pathogenicity in suckling mice (Tauraso, 1969). This led to the conclusion that plaque size of orthobunyaviruses is not related to their virulence.

Although it has been shown that BUNdelNSs has a smaller plaque size compared to wt BUNV (Bridgen et al., 2001), in general, the presence of NSs protein did not correlate with plaque size. For instance, NDV and SJ2441V, which were shown by sequencing and protein profiles of infected cells to encode an NSs protein, produced a smaller plaque

compared to BUNdelNSs. Therefore, besides the presence of NSs protein, other factors such as host restriction and adaptation of the virus to particular cells might possibly be involved in determining these plaque-size differences.

Analysis of virus infected cells by metabolic labeling revealed that all of the viruses have the typical orthobunyavirus protein profile and migration pattern, with L, Gc and N clearly identified. This result confirmed the previous serological classification that grouped these viruses into *Orthobunyavirus* genus.

Previously, the NSs proteins of BUNV and LACV were shown to play a role in causing shutoff of host protein synthesis in infected mammalian cells. In this study, no shutoff was observed in cells infected with TETEV, BMAV, PLSV, MNTV, PGAV and KETV and shutoff similar to BUNdelNSs infected cells was observed in cells infected with ACAV, ANAV, TCMV, ANBV, BORV, E147V, NDV, OLIV, KOOV, GAMV, SJ2441V, TURV, MPOV, SJ2441V, BERV and GMAV. Shutoff at a level almost similar to BUNV-infected cells was only observed in ACAV, BAKV, BWAV, CAPV, PAHV, PATV and WONVinfected cells. Radio-labelled proteins in infected Vero cells, sequencing of the S segment and in vitro TnT results of ANAV, ANBV, BORV, BMAV, TCMV and TETEV further confirmed that these viruses lack an NSs protein. However in CAPV-infected cells, a drastic shutoff was observed but no NSs protein was detected in the radio-labelled protein gel and S segment sequences. The mechanism of shutoff involved by this virus is unclear and needs to be further studied. Furthermore, sequencing of S segment of E147V, NDV, GAMV, SJ2441V and MPOV revealed that these viruses encode NSs protein in their S segment but the shutoff in their infected cells was at the same level as BUNdelNSs. Therefore, it is very difficult to relate the shutoff caused by the virus with the presence of NSs protein.

6.2 Sequence analysis of the S RNA segments of orthobunyaviruses

Complete S RNA sequences of the Anopheles A (ANAV and TCMV), Anopheles B (ANBV and BORV), Nyando (NDV and EREV), Bwamba (BWAV), Turlock (MPOV), Tete (TETEV and BMAV) and Gamboa (SJ2441V) serogroups, and partial S RNA sequences of Gamboa (GAMV), Patois (PATV), Guama (GMAV and BERV, Capim (CAPV) and Minatitlan (PLSV) serogroups were obtained in this study. No S sequence
was obtained for ACAV, BAKV, KETV, MNTV, OLIV, KOOV, PAHV, PGAV, TURV and WONV.

Failure to get complete sequences for some of the viruses using primers BUN S+/S- (Dunn et al., 1994), that were thought to be conserved to all orthobunyaviruses, could be due to that not all orthobunyaviruses share these apparent consensus sequences.

An attempt to get partial sequence using internal primers that were designed based on conserved sequences found by aligning available S sequences also failed on some viruses, indicating that not all viruses contained these conserved sequences. Therefore, designing degenerate primers that could be used to amplify orthobunyaviruses in all the serogroups could be of great advantage.

Attempts to obtain the the 3' and 5' ends using 3'/5' RACE, modified 3'/5' RACE and RT-PCR on ligated RNA were also unsuccessful. This could be due to the presence of ribonucleases in some preparation that resulted in the truncation of viral RNA (Polidoros et al., 2006). Furthermore, the use of a universal primer corresponding to the anchor sequence present in all cDNAs resulted in a high background of nonspecific products. Therefore, strategies to eliminate these problems are required in the future work.

Sequence analyses of complete S segments of these viruses have identified ANAV, TCMV, ANBV, BORV, TETEV, BMAV and CAPV as naturally occurring orthobunyaviruses not having an NSs protein. Most of these viruses except CAPV were found to have longer N proteins compared to those reported of having NSs protein with the biggest N protein detected in TETEV and BMAV. It is possible that these viruses do not encode an NSs protein in their S segment, that their NSs initiation codon disappeared during evolution, their NSs protein is very short or that they used different initiation codons to encode their NSs protein.

Similar to OROV and GROV, the LPS motif that is thought to be important for interaction of NSs with MED8 (Leonard et al., 2006), is absent in most of studied viruses except in BWAV. Therefore, studies on the mechanism or the motif used by these viruses to inhibit host protein synthesis and the interferon response would be of great interest.

The conserved CA rich motif in the 5'NCR of genomic-sense RNA of Bunyamwera, California and Group C serogroups viruses, which was thought to play an important role in transcriptional termination, was also present in all the studied viruses except in BWAV and MPOV. This suggests that BWAV and MPOV might use different motif to terminate their transcription process. However these viruses were found to have an extensive GAU/AU/GU-rich motif in this region. ANBV and BORV were found to have two copies of the CA motif. Repeated sequences in 5' NCR of genomic-sense RNA, as previously observed in LUMV (Dunn et al., 1994) were also detected in BWAV and TCMV.

Sequence comparisons and phylogenetic analyses used in this study were in agreement with previous serological classification of these viruses, except for TCMV and BERV. TCMV together with ANAV were serologically placed in the Anopheles A serogroup, but their nucleotide and amino acid sequences identity was only 59% and 53% respectively, Phylogenetic analyses using NJ plot based on complete N sequence and MP plot based on partial N sequence supported the previous serological classification which placed this virus in the same clade as ANAV, but the NJ plot based on partial N sequence has placed TCMV as an intermediate virus between ANAV and ANBV. However similar results were also obtained with viruses in Simbu serogroup (Saeed et al., 2001) and Group C serogroup (Nunes et al., 2005), suggesting that viruses in these serogroups have greater sequence divergence compared to the other serogroups. BERV was previously placed in Guama serogroup, however partial S sequence and phylogenetic analysis of this virus revealed that it was closer to CAPV than GMAV. Due to high sequence identity obtained between BERV, CAPV and GMAV (>75%), these viruses are suggested to originate from the same ancestor and therefore it is recommended to classify them into one group. It is also possible that TCMV and BERV are bridging viruses between the two serogroups or reassortant viruses that have different S segment but similar M segment as ANAV and GMAV respectively. They were placed in their current serogroups based on PRNT or HI assays which recognise epitopes encoded by M RNA determination rather than those encoded by S RNA. However, complete sequences of their M segments are necessary to confirm this speculation.

High sequence identity was observed between BWAV and CEV (68% identity), therefore supporting serological classification that identified BWAV as a bridging virus between California and Bwamba serogroups (Calisher, 1988). In contrast with GAMV which was

previously identified as a bridging virus between Gamboa and Capim serogroups (Calisher, 1988), the sequence identity of these viruses was only 35%, which is thought too low to classify this virus as a bridging virus between these two serogroups.

Complete S segment sequence and phylogenetic analysis of NDV and MPOV by Yandoko et al. (2007) have placed these viruses into Bunyamwera serogroup, with more than 85% N sequence identity, which is in contrast to the previous serological classification. However, the S segment sequence results obtained in this study for NDV and MPOV were in agreement with previous serological classification which placed NDV in Nyando serogroup and MPOV in Turlock serogroup. Furthermore, NDV was found to have N sequence identity of more than 80% with EREV, which is also in Nyando serogroup. Unfortunately, no S segment sequence for TURV was obtained. It is possible that the viruses analysed by Yandoko et al have been misdiagnosed or are reassortant viruses, which have the S segment close to BUNV but M segment similar to Nyando and Turlock serogroups respectively. Furthermore, all the 10 viruses used in the Yandoko et al study, which were classified previously by serological methods into the Bunyamwera, Nyando and Simbu serogroups, all fell into Bunyamwera serogroup, strongly suggesting that cross-contamination had occurred in their samples.

6.3 Ability of NSs minus viruses in inhibiting the production of interferon and in causing the shutoff of host protein synthesis.

Shutoff in BHK-21 infected cells based on luciferase based assay was in agreement with shutoff observed by direct radio-labelling of infected Vero cells, where shutoff caused by ANAV, TCMV and ANBV was almost similar to the shutoff caused by BUNdelNSs and a slower shutoff was seen with cells infected with BORV, TETEV and BMAV. However, radio-labelling of infected cells showed no shutoff in TETEV and BMAV-infected cells, therefore suggesting that luciferase based assay is more sensitive than radiolabelled protein assay. No shutoff was caused by any of the viruses in Hep2 and Hep2/V infected cells. The growth of ANAV, ANBV, TETEV and BMAV seemed to be very slow or almost restricted in Hep2 and Hep2/V cells because no viral proteins were observed in either cell type.

In infected A549 cells, rapid shutoff was observed following infection with TCMV which does not have an NSs in its S segment, suggesting that TCMV might use a different mechanism to induce shutoff.

Luciferase reporter plasmid based assays in infected C6/36 cells revealed no significant shutoff caused by any of the viruses, suggesting that NSs protein might not function in this cell line.

The specific RT-PCR for IFN- β mRNA in infected 293 cells and IFN reporter gene assay to detect the activation of IFN- β promoter in infected A549 cells revealed that TCMV was still capable of counteracting IFN production similar to wt BUNV. TCMV was isolated from a human with mild symptoms, suggesting that it is capable of circumventing the human IFN response. The mechanism used by this virus to cause shutoff and subvert the IFN response requires further investigation.

The other NSs minus viruses (ANAV, ANBV, TETEV and BMAV) were found to be capable in inducing IFN in infected cells. However, very faint bands of IFN- β mRNA and little activation of IFN- β promoter were detected in ANAV and BMAV-infected cells. The ability of these viruses to cause infection in the host need to be further investigated.

References

- Aigner, T. (2002). Apoptosis, necrosis, or whatever: how to find out what really happens? *The Journal of pathology* **198**, 1-4.
- Akashi, H. & Bishop, D. H. (1983). Comparison of the sequences and coding of La Crosse and snowshoe hare bunyavirus S RNA species. *Journal of virology* **45**, 1155-1158.
- Akashi, H., Gay, M., Ihara, T. & Bishop, D. H. (1984). Localized conserved regions of the S RNA gene products of bunyaviruses are revealed by sequence analyses of the Simbu serogroup Aino virus. *Virus research* 1, 51-63.
- Akashi, H., Kaku, Y., Kong, X. & Pang, H. (1997). Antigenic and genetic comparisons of Japanese and Australian Simbu serogroup viruses: evidence for the recovery of natural virus reassortants. *Virus research* 50, 205-213.
- Andrejeva, J., Young, D. F., Goodbourn, S. & Randall, R. E. (2002). Degradation of STAT1 and STAT2 by the V proteins of simian virus 5 and human parainfluenza virus type 2, respectively: consequences for virus replication in the presence of alpha/beta and gamma interferons. *Journal of virology* 76, 2159-2167.
- Barr, J. N., Rodgers, J. W. & Wertz, G. W. (2006). Identification of the Bunyamwera bunyavirus transcription termination signal. *The Journal of general virology* 87, 189-198.
- Barr, J. N. & Wertz, G. W. (2004). Bunyamwera bunyavirus RNA synthesis requires cooperation of 3'- and 5'-terminal sequences. *Journal of virology* **78**, 1129-1138.
- Barr, J. N. & Wertz, G. W. (2005). Role of the conserved nucleotide mismatch within 3'and 5'-terminal regions of Bunyamwera virus in signaling transcription. *Journal of virology* 79, 3586-3594.
- Beaty, B. J. & Bishop, D. H. (1988). Bunyavirus-vector interactions. Virus research 10, 289-301.
- Beaty, B. J., Sundin, D. R., Chandler, L. J. & Bishop, D. H. (1985). Evolution of bunyaviruses by genome reassortment in dually infected mosquitoes (Aedes triseriatus). *Science* 230, 548-550.
- Beaty, B. J. C., C.H. (1991). *Bunyaviridae*-Natural history. *Curr Top Microbiol Immunol* **169**, 27-78.
- Bellocq, C. & Kolakofsky, D. (1987). Translational requirement for La Crosse virus SmRNA synthesis: a possible mechanism. *Journal of virology* **61**, 3960-3967.
- Bishop, D. H., Calisher, C. H., Casals, J., Chumakov, M. P., Gaidamovich, S. Y., Hannoun, C., Lvov, D. K., Marshall, I. D., Oker-Blom, N., Pettersson, R. F., Porterfield, J. S., Russell, P. K., Shope, R. E. & Westaway, E. G. (1980). Bunyaviridae. *Intervirology* 14, 125-143.

- Bishop, D. H., Gay, M. E. & Matsuoko, Y. (1983). Nonviral heterogeneous sequences are present at the 5' ends of one species of snowshoe hare bunyavirus S complementary RNA. *Nucleic Acids Res* 11, 6409-6418.
- Bishop, D. H. L. (1996). Biology and Molecular Biology of Bunyaviruses: *In*: The *Bunyaviridae* (R.M. Elliott, ed.), pp.19-53. Plenum Press, New York.
- Bishop, D. H. L. & Shope, R.E. (1979). Bunyaviridae. Compr Virology 14, 1-156.
- Blakqori, G., Kochs, G., Haller, O. & Weber, F. (2003). Functional L polymerase of La Crosse virus allows in vivo reconstitution of recombinant nucleocapsids. *The Journal of general virology* 84, 1207-1214.
- Blakqori, G. & Weber, F. (2005). Efficient cDNA-based rescue of La Crosse bunyaviruses expressing or lacking the nonstructural protein NSs. *Journal of virology* **79**, 10420-10428.
- Borucki, M. K., Kempf, B. J., Blitvich, B. J., Blair, C. D. & Beaty, B. J. (2002). La Crosse virus: replication in vertebrate and invertebrate hosts. *Microbes and infection / Institut Pasteur* **4**, 341-350.
- Bouley, M. (1991). Bunyaviridae: Genome Organisation and Replication Strategies. *Advances in Virus Research* **40**, 235-275.
- Bouloy, M. & Hannoun, C. (1976). Studies on lumbo virus replication. I. RNA-dependent RNA polymerase associated with virions. *Virology* **69**, 258-264.
- Bouloy, M., Pardigon, N., Vialat, P., Gerbaud, S. & Girard, M. (1990). Characterization of the 5' and 3' ends of viral messenger RNAs isolated from BHK21 cells infected with Germiston virus (Bunyavirus). *Virology* 175, 50-58.
- Bowen, M. D., Jackson, A. O., Bruns, T. D., Hacker, D. L. & Hardy, J. L. (1995). Determination and comparative analysis of the small RNA genomic sequences of California encephalitis, Jamestown Canyon, Jerry Slough, Melao, Keystone and Trivittatus viruses (Bunyaviridae, genus Bunyavirus, California serogroup). *The Journal of general virology* **76** (Pt 3), 559-572.
- Bowen, M. D., Trappier, S. G., Sanchez, A. J., Meyer, R. F., Goldsmith, C. S., Zaki, S. R., Dunster, L. M., Peters, C. J., Ksiazek, T. G. & Nichol, S. T. (2001). A reassortant bunyavirus isolated from acute hemorrhagic fever cases in Kenya and Somalia. *Virology* 291, 185-190.
- Bridgen, A., Weber, F., Fazakerley, J. K. & Elliott, R. M. (2001). Bunyamwera bunyavirus nonstructural protein NSs is a nonessential gene product that contributes to viral pathogenesis. *Proceedings of the National Academy of Sciences of the United States of America* 98, 664-669.
- Briese, T., Rambaut, A. & Lipkin, W. I. (2004). Analysis of the medium (M) segment sequence of Guaroa virus and its comparison to other orthobunyaviruses. *The Journal of general virology* 85, 3071-3077.
- Brockus, C. L. & Grimstad, P. R. (1999). Sequence analysis of the medium (M) segment of Cache Valley virus, with comparison to other Bunyaviridae. *Virus genes* **19**, 73-83.

- Bupp, K., Stillmock, K. & Gonzalez-Scarano, F. (1996). Analysis of the intracellular transport properties of recombinant La Crosse virus glycoproteins. *Virology* 220, 485-490.
- Calisher, C. H. (1983). Taxonomy, classification, and geographic distribution of California serogroup bunyaviruses. *Progress in clinical and biological research* **123**, 1-16.
- Calisher, C. H. (1988). Evolutionary significance of the taxonomic data regarding bunyaviruses of the family Bunyaviridae. *Intervirology* **29**, 268-276.
- Calisher, C. H. (1996). History, Classification, and Taxonomy of viruses in the Family *Bunyaviridae*: *In:* The *Bunyaviridae* (R.M. Elliott, ed.), pp.1-18. Plenum Press, New York.
- Campbell, G. L., Mataczynski, J. D., Reisdorf, E. S., Powell, J. W., Martin, D. A., Lambert, A. J., Haupt, T. E., Davis, J. P. & Lanciotti, R. S. (2006). Second human case of Cache Valley virus disease. *Emerging infectious diseases* 12, 854-856.
- Campbell, W. P. & Huang, C. (1999). Sequence comparisons of medium RNA segment among 15 California serogroup viruses. *Virus research* **61**, 137-144.
- Casal, J. & Whitman, L. (1961). Group C. A new serological group of hitherto undescribed arthropod-borne viruses. Immunological studies. *The American journal of tropical medicine and hygiene* 10, 250-258.
- Casals, J. (1963). Relationships among Arthropod-Borne Animal Viruses Determined by Cross-Challenge Tests. *The American journal of tropical medicine and hygiene* **12**, 587-596.
- Casals, J. & Whitman, L. (1960). A new antigenic group of arthropod-borne viruses: the Bunyamwera group. *The American journal of tropical medicine and hygiene* **9**, 73-77.
- Cash, P. (1982). Inhibition of La Crosse virus replication by monensin, monovalent ionophore. *The Journal of general virology* **59**, 193-196.
- Chandler, L. J., Hogge, G., Endres, M., Jacoby, D. R., Nathanson, N. & Beaty, B. J. (1991). Reassortment of La Crosse and Tahyna bunyaviruses in *Aedes triseriatus* mosquitoes. *Virus research* 20, 181-191.
- Cheng, L. L., Schultz, K. T., Yuill, T. M. & Israel, B. A. (2000). Identification and localization of conserved antigenic epitopes on the G2 proteins of California serogroup Bunyaviruses. *Viral immunology* **13**, 201-213.
- Childs, J., Shope, R. E., Fish, D., Meslin, F. X., Peters, C. J., Johnson, K., Debess, E., Dennis, D. & Jenkins, S. (1998). Emerging zoonoses. *Emerging infectious diseases* 4, 453-454.
- Colon-Ramos, D. A., Irusta, P. M., Gan, E. C., Olson, M. R., Song, J., Morimoto, R. I., Elliott, R. M., Lombard, M., Hollingsworth, R., Hardwick, J. M., Smith, G. K. & Kornbluth, S. (2003). Inhibition of translation and induction of apoptosis by Bunyaviral nonstructural proteins bearing sequence similarity to reaper. *Molecular biology of the cell* 14, 4162-4172.

- Cunningham, C. & Szilagyi, J. F. (1987). Viral RNAs synthesized in cells infected with Germiston Bunyavirus. *Virology* **157**, 431-439.
- Doherty, R. L., Carley, J. G., Mackerras, M. J. & Marks, E. N. (1963). Studies of arthropod-borne virus infections in Queensland. III. Isolation and characterization of virus strains from wild-caught mosquitoes in North Queensland. *The Australian journal of experimental biology and medical science* **41**, 17-39.
- Douglas, R. G., Jr., Couch, R. B., Baxter, B. D. & Gough, M. (1974). Attenuation of rhinovirus type 15: relation of illness to plaque size. *Infection and immunity* 9, 519-523.
- Dunn, E. F., Pritlove, D. C. & Elliott, R. M. (1994). The S RNA genome segments of Batai, Cache Valley, Guaroa, Kairi, Lumbo, Main Drain and Northway bunyaviruses: sequence determination and analysis. *The Journal of general virology* **75** (**Pt 3**), 597-608.
- Dunn, E. F., Pritlove, D. C., Jin, H. & Elliott, R. M. (1995). Transcription of a recombinant bunyavirus RNA template by transiently expressed bunyavirus proteins. *Virology* 211, 133-143.
- Edwards, J. F., Livingston, C. W., Chung, S. I. & Collisson, E. C. (1989). Ovine arthrogryposis and central nervous system malformations associated with in utero Cache Valley virus infection: spontaneous disease. *Vet Pathol* **26**, 33-39.
- El Said, L. H., Vorndam, V., Gentsch, J. R., Clewley, J. P., Calisher, C. H., Klimas, R. A., Thompson, W. H., Grayson, M., Trent, D. W. & Bishop, D. H. (1979). A comparison of La Crosse virus isolated obtained from different ecological niches and an analysis of the structural components of California encephalitis serogroup viruses and other bunyaviruses. *The American journal of tropical medicine and hygiene* **28**, 364-386.
- Elliott, L. H., Kiley, M. P. & McCormick, J. B. (1984). Hantaan virus: identification of virion proteins. *The Journal of general virology* **65** (**Pt 8**), 1285-1293.
- Elliott, R. M. (1985). Identification of nonstructural proteins encoded by viruses of the Bunyamwera serogroup (family Bunyaviridae). *Virology* **143**, 119-126.
- Elliott, R. M. (1989). Nucleotide sequence analysis of the large (L) genomic RNA segment of Bunyamwera virus, the prototype of the family Bunyaviridae. *Virology* **173**, 426-436.
- Elliott, R. M. (1990). Molecular biology of the Bunyaviridae. *The Journal of general virology* **71** (Pt 3), 501-522.
- Elliott, R. M., Schmaljohn, C.S. and Collett, M.S. (1991). *Bunyaviridae* genome structure and gene expression. *Curr Top Microbiol Immunol* **169**, 91-142.
- Elliott, R. M. (1996). The *Bunyaviridae:* Concluding Remarks and Future Prospects. *In:* The *Bunyaviridae* (R.M. Elliott, ed.), pp.295-332. Plenum Press, New York.
- Elliott, R.M. (1997). Emerging viruses: The *Bunyaviridae*. Molecular medicine **3** (Number 9), 572-577.

- Elliott, R. M., Bouloy, M., Calisher, C.H., Goldbach, R., Moyer, J.T., Nichol, S.T.,Pettersson, R., Plyusnin, A and Schmaljohn, C.S. (2000). *Bunyaviridae. In* Virus Taxonomy. Seventh report of the International Committee on Taxonomy of Viruses. San Diego: Academic Press.
- Elliott, R. M. (2001). The continuing threat of bunyaviruses and hantaviruses. *In* Sixtieth Symposium of the Society for General Microbiology, pp. 53-65. Edited by W. L. I. G.L.Smith, J.W. McCauley & D.J. Rowlands. Heriot-Watt University: Cambridge University Press.
- Elliott, R. M. & Wilkie, M. L. (1986). Persistent infection of Aedes albopictus C6/36 cells by Bunyamwera virus. *Virology* **150**, 21-32.
- Endres, M. J., Griot, C., Gonzalez-Scarano, F. & Nathanson, N. (1991). Neuroattenuation of an avirulent bunyavirus variant maps to the L RNA segment. *Journal of virology* 65, 5465-5470.
- Eshita, Y. & Bishop, D. H. (1984). The complete sequence of the M RNA of snowshoe hare bunyavirus reveals the presence of internal hydrophobic domains in the viral glycoprotein. *Virology* **137**, 227-240.
- Eshita, Y., Ericson, B., Romanowski, V. & Bishop, D. H. (1985). Analyses of the mRNA transcription processes of snowshoe hare bunyavirus S and M RNA species. *Journal of virology* **55**, 681-689.
- Fazakerley, J. K., Gonzalez-Scarano, F., Strickler, J., Dietzschold, B., Karush, F. & Nathanson, N. (1988). Organization of the middle RNA segment of snowshoe hare Bunyavirus. *Virology* 167, 422-432.
- Felsenstein, J. (1993). PHYLIP (Phylogeny Inference Package) Version 3.5C. *In* Department of Genetic: University of Washington.
- Fenner, F. (1975). International Committee on Taxonomy of Viruses: official names for viral families. *Acta Virol* **19**, 93.
- Frugulhetti, I. C. & Rebello, M. A. (1989). Na+ and K+ concentration and regulation of protein synthesis in L-A9 and Aedes albopictus cells infected with Marituba virus (Bunyaviridae). *The Journal of general virology* **70** (Pt 12), 3493-3499.
- Fuller, F. & Bishop, D. H. (1982). Identification of virus-coded nonstructural polypeptides in bunyavirus-infected cells. *Journal of virology* 41, 643-648.
- Gavrilovskaya, I. N., Brown, E. J., Ginsberg, M. H. & Mackow, E. R. (1999). Cellular entry of hantaviruses which cause hemorrhagic fever with renal syndrome is mediated by beta3 integrins. *Journal of virology* 73, 3951-3959.
- Gentsch, J., Bishop, D. H. & Obijeski, J. F. (1977). The virus particle nucleic acids and proteins of four bunyaviruses. *The Journal of general virology* **34**, 257-268.
- Giachetti, C. & Holland, J. J. (1989). Vesicular stomatitis virus and its defective interfering particles exhibit in vitro transcriptional and replicative competition for purified L-NS polymerase molecules. *Virology* 170, 264-267.

- Gonzalez-Scarano, F., Janssen, R. S., Najjar, J. A., Pobjecky, N. & Nathanson, N. (1985). An avirulent G1 glycoprotein variant of La Crosse bunyavirus with defective fusion function. *Journal of virology* 54, 757-763.
- Grady, L. J., Sanders, M. L. & Campbell, W. P. (1987). The sequence of the M RNA of an isolate of La Crosse virus. *The Journal of general virology* **68** (Pt 12), 3057-3071.
- Graham, F. L., Smiley, J., Russell, W. C. & Nairn, R. (1977). Characteristics of a human cell line transformed by DNA from human adenovirus type 5. *The Journal of* general virology 36, 59-74.
- Haaster, C. C. & Bishop, D. H. (1980). Analyses of the 3' terminal sequences of snowshoe hare and La Crosse Bunyaviruses. *Virology* **105**, 564-574.
- Hacker, D., Raju, R. & Kolakofsky, D. (1989). La Crosse virus nucleocapsid protein controls its own synthesis in mosquito cells by encapsidating its mRNA. *Journal of virology* 63, 5166-5174.
- Hacker, D., Rochat, S. & Kolakofsky, D. (1990). Anti-mRNAs in La Crosse bunyavirusinfected cells. *Journal of virology* **64**, 5051-5057.
- Hacker, J. K. & Hardy, J. L. (1997). Adsorptive endocytosis of California encephalitis virus into mosquito and mammalian cells: a role for G1. *Virology* **235**, 40-47.
- Haller, O., Kochs, G. & Weber, F. (2006). The interferon response circuit: induction and suppression by pathogenic viruses. *Virology* 344, 119-130.
- Hammon, W. M., Reeves, W. C. & Sather, G. (1952). California encephalitis virus, a newly described agent. II. Isolations and attempts to identify and characterize the agent. J Immunol 69, 493-510.
- Hart, T. (2004). The role of Bunyamwera nonstructural protein (NSs) in shutoff of host cell protein synthesis. PhD thesis. *In:* Division of Virology, IBLS. University of Glasgow, Glasgow
- Holland, J. & Domingo, E. (1998). Origin and evolution of viruses. Virus genes 16, 13-21.
- Holland, J., Spindler, K., Horodyski, F., Grabau, E., Nichol, S. & VandePol, S. (1982). Rapid evolution of RNA genomes. *Science* **215**, 1577-1585.
- Hunt, A. R. & Calisher, C. H. (1979). Relationships of bunyamwera group viruses by neutralization. *The American journal of tropical medicine and hygiene* **28**, 740-749.
- Igarashi, A. (1978). Isolation of a Singh's Aedes albopictus cell clone sensitive to Dengue and Chikungunya viruses. *The Journal of general virology* **40**, 531-544.
- Ikegami, T., Won, S., Peters, C. J. & Makino, S. (2006). Rescue of infectious rift valley fever virus entirely from cDNA, analysis of virus lacking the NSs gene, and expression of a foreign gene. *Journal of virology* **80**, 2933-2940.

- Iroegbu, C. U. & Pringle, C. R. (1981). Genetic interactions among viruses of the Bunyamwera complex. *Journal of virology* **37**, 383-394.
- Jin, H. & Elliott, R. M. (1991). Expression of functional Bunyamwera virus L protein by recombinant vaccinia viruses. *Journal of virology* **65**, 4182-4189.
- Jin, H. & Elliott, R. M. (1993). Characterization of Bunyamwera virus S RNA that is transcribed and replicated by the L protein expressed from recombinant vaccinia virus. *Journal of virology* 67, 1396-1404.
- Karabatsos, N., (ed) (1985). International catalogue of arbovirus including certain other viruses of vertebrates, 3rd edition. *American Society of Tropical Medicine and Hygiene*, San Antonio, Texas, USA.
- Kappus, K. D., Monath, T. P., Kaminski, R. M. & Calisher, C. H. (1983). Reported encephalitis associated with California serogroup virus infections in the United States, 1963-1981. *Progress in clinical and biological research* 123, 31-41.
- Kascsak, R. J. & Lyons, M. J. (1978). Bunyamwera virus. II. The generation and nature of defective interfering particles. *Virology* 89, 539-546.
- Kimura, M. (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of molecular evolution* **16**, 111-120.
- King, P. & Goodbourn, S. (1994). The beta-interferon promoter responds to priming through multiple independent regulatory elements. *The Journal of biological chemistry* 269, 30609-30615.
- Kingsford, L. (1991). Antigenic variance. Curr Top Microbiol Immunol 169, 183-215.
- Kingsford, L. & Hill, D. W. (1983). The effect of proteolytic cleavage of La Crosse virus G1 glycoprotein on antibody neutralization. *The Journal of general virology* 64 (Pt 10), 2147-2156.
- Klimas, R. A., Thompson, W. H., Calisher, C. H., Clark, G. G., Grimstad, P. R. & Bishop, D. H. (1981). Genotypic varieties of La Crosse virus isolated from different geographic regions of the continental United States and evidence for a naturally occurring intertypic recombinant La Crosse virus. *Am J Epidemiol* **114**, 112-131.
- Klimas, R. A., Ushijima, H., Clerx-Van Haaster, C. M. & Bishop, D. H. (1981). Radioimmune assays and molecular studies that place Anopheles B and Turlock serogroup viruses in the Bunyavirus genus (*Bunyaviridae*). *The American journal* of tropical medicine and hygiene **30**, 876-887.
- Kohl, A., Clayton, R. F., Weber, F., Bridgen, A., Randall, R. E. & Elliott, R. M. (2003). Bunyamwera virus nonstructural protein NSs counteracts interferon regulatory factor 3-mediated induction of early cell death. *Journal of virology* 77, 7999-8008.

- Kohl, A., Dunn, E. F., Lowen, A. C. & Elliott, R. M. (2004). Complementarity, sequence and structural elements within the 3' and 5' non-coding regions of the Bunyamwera orthobunyavirus S segment determine promoter strength. *The Journal of general virology* 85, 3269-3278.
- Kohl, A., Lowen, A. C., Leonard, V. H. & Elliott, R. M. (2006). Genetic elements regulating packaging of the Bunyamwera orthobunyavirus genome. *The Journal of* general virology 87, 177-187.
- Kolakofsky, D. & Hacker, D. (1991). Bunyavirus RNA synthesis: genome transcription and replication. *Curr Top Microbiol Immunol* **169**, 143-159.

Kumar, S., Tamura, K., Jakobsen, I.B., and Nei, M. (2001) MEGA2: Molecular Evolutionary Genetics Analysis software, Version 2.1. Arizona State University, Arizona, USA.

- Lappin, D. F., Nakitare, G. W., Palfreyman, J. W. & Elliott, R. M. (1994). Localization of Bunyamwera bunyavirus G1 glycoprotein to the Golgi requires association with G2 but not with NSm. *The Journal of general virology* **75** (Pt 12), 3441-3451.
- Lazdins, I. & Holmes, I. H. (1979). Protein synthesis in Bunyamwera virus-infected cells. *The Journal of general virology* **44**, 123-133.
- Lees, J. F., Pringle, C. R. & Elliott, R. M. (1986). Nucleotide sequence of the Bunyamwera virus M RNA segment: conservation of structural features in the Bunyavirus glycoprotein gene product. *Virology* 148, 1-14.
- Leonard, V. H., Kohl, A., Hart, T. J. & Elliott, R. M. (2006). Interaction of Bunyamwera Orthobunyavirus NSs protein with mediator protein MED8: a mechanism for inhibiting the interferon response. *Journal of virology* **80**, 9667-9675.
- Leonard, V. H., Kohl, A., Osborne, J. C., McLees, A. & Elliott, R. M. (2005). Homotypic interaction of Bunyamwera virus nucleocapsid protein. *Journal of virology* 79, 13166-13172.
- Leppert, M., Rittenhouse, L., Perrault, J., Summers, D. F. & Kolakofsky, D. (1979). Plus and minus strand leader RNAs in negative strand virus-infected cells. *Cell* **18**, 735-747.
- Li, Z., Yu, M., Zhang, H., Wang, H. Y. & Wang, L. F. (2005). Improved rapid amplification of cDNA ends (RACE) for mapping both the 5' and 3' terminal sequences of paramyxovirus genomes. *Journal of virological methods* **130**, 154-156.
- Liu, W. J., Wang, X. J., Mokhonov, V. V., Shi, P. Y., Randall, R. & Khromykh, A. A. (2005). Inhibition of interferon signaling by the New York 99 strain and Kunjin subtype of West Nile virus involves blockage of STAT1 and STAT2 activation by nonstructural proteins. *Journal of virology* **79**, 1934-1942.
- Lowen, A. C., Boyd, A., Fazakerley, J. K. & Elliott, R. M. (2005). Attenuation of bunyavirus replication by rearrangement of viral coding and noncoding sequences. *Journal of virology* 79, 6940-6946.

- Lowen, A. C. & Elliott, R. M. (2005). Mutational analyses of the nonconserved sequences in the Bunyamwera Orthobunyavirus S segment untranslated regions. *Journal of virology* 79, 12861-12870.
- Ludwig, G. V., Israel, B. A., Christensen, B. M., Yuill, T. M. & Schultz, K. T. (1991). Role of La Crosse virus glycoproteins in attachment of virus to host cells. *Virology* 181, 564-571.
- Lyles, D. S. (2000). Cytopathogenesis and inhibition of host gene expression by RNA viruses. *Microbiol Mol Biol Rev* 64, 709-724.
- Lyons, M. J. & Heyduk, J. (1973). Aspects of the developmental morphology of California encephalitis virus in cultured vertebrae and arthropod cells and in mouse brain. *Virology* **54**, 37-52.
- Macpherson, I. & Stoker, M. (1962). Polyoma transformation of hamster cell clones--an investigation of genetic factors affecting cell competence. *Virology* **16**, 147-151.
- Marriott, A. C., el-Ghorr, A. A. & Nuttall, P. A. (1992). Dugbe Nairovirus M RNA: nucleotide sequence and coding strategy. *Virology* **190**, 606-615.
- Matsuoka, Y., Chen, S. Y. & Compans, R. W. (1991). Bunyavirus protein transport and assembly. *Curr Top Microbiol Immunol* **169**, 161-179.
- Matumoto, M. (1966). Multiplication of measles virus in cell cultures. *Bacteriological reviews* **30**, 152-176.
- McConnell, S., Livingston, C., Jr., Calisher, C. H. & Crandell, R. A. (1987). Isolations of Cache Valley virus in Texas, 1981. *Veterinary microbiology* **13**, 11-18.
- McPhee, D. A. & Della-Porta, A. J. (1988). Biochemical and serological comparisons of Australian bunyaviruses belonging to the Simbu serogroup. *The Journal of general virology* 69 (Pt 5), 1007-1017.
- McPhee, D. A. & Westaway, E. G. (1981). Proteins and glycoproteins specified by bunyamwera virus and by Belmont virus, a possible bunyavirus, in mammalian cells. *The Journal of general virology* 54, 149-159.
- Munoz-Jordan, J. L., Sanchez-Burgos, G. G., Laurent-Rolle, M. & Garcia-Sastre, A. (2003). Inhibition of interferon signaling by dengue virus. *Proceedings of the National Academy of Sciences of the United States of America* 100, 14333-14338.
- Murphy, F. A., Harrison, A. K. & Whitfield, S. G. (1973). Bunyaviridae: morphologic and morphogenetic similarities of Bunyamwera serologic supergroup viruses and several other arthropod-borne viruses. *Intervirology* 1, 297-316.
- Najjar, J. A., Gentsch, J. R., Nathanson, N. & Gonzalez-Scarano, F. (1985). Epitopes of the G1 glycoprotein of La Crosse virus form overlapping clusters within a single antigenic site. *Virology* 144, 426-432.
- Nakitare, G. W. & Elliott, R. M. (1993). Expression of the Bunyamwera virus M genome segment and intracellular localization of NSm. *Virology* 195, 511-520.

- Newton, S. E., Short, N. J. & Dalgarno, L. (1981). Bunyamwera virus replication in cultured Aedes albopictus (mosquito) cells: establishment of a persistent viral infection. *Journal of virology* **38**, 1015-1024.
- Nichol, S. T. (2001). Bunyaviruses. *In*: Field Virology (D.M. Knipe and P.M. Howley, eds.), Vol.2, pp.1603-1668. Lippincott, Williams & Wilkins.
- Nunes, M. R., Travassos da Rosa, A. P., Weaver, S. C., Tesh, R. B. & Vasconcelos, P. F. (2005). Molecular epidemiology of group C viruses (Bunyaviridae, Orthobunyavirus) isolated in the Americas. *Journal of virology* **79**, 10561-10570.
- Obijeski, J. F., Bishop, D. H., Palmer, E. L. & Murphy, F. A. (1976). Segmented genome and nucleocapsid of La Crosse virus. *Journal of virology* **20**, 664-675.
- Obijeski, J. F., McCauley, J. & Skehel, J. J. (1980). Nucleotide sequences at the terminal of La Crosse virus RNAs. *Nucleic Acids Res* **8**, 2431-2438.
- Obijeski, J. F. & Murphy, F. A. (1977). Bunyaviridae: recent biochemical developments. *The Journal of general virology* **37**, 1-14.
- Osborne, J. C. & Elliott, R. M. (2000). RNA binding properties of bunyamwera virus nucleocapsid protein and selective binding to an element in the 5' terminus of the negative-sense S segment. *Journal of virology* **74**, 9946-9952.
- Pardigon, N., Vialat, P., Gerbaud, S., Girard, M. & Bouloy, M. (1988). Nucleotide sequence of the M segment of Germiston virus: comparison of the M gene product of several bunyaviruses. *Virus research* 11, 73-85.
- Patel, A. H. & Elliott, R. M. (1992). Characterization of Bunyamwera virus defective interfering particles. *The Journal of general virology* **73** (**Pt 2**), 389-396.
- Patterson, J. L., Holloway, B. & Kolakofsky, D. (1984). La Crosse virions contain a primer-stimulated RNA polymerase and a methylated cap-dependent endonuclease. *Journal of virology* **52**, 215-222.
- Pekosz, A., Griot, C., Nathanson, N. & Gonzalez-Scarano, F. (1995). Tropism of bunyaviruses: evidence for a G1 glycoprotein-mediated entry pathway common to the California serogroup. *Virology* 214, 339-348.
- Pennington, T. H., Pringle, C. R. & McCrae, M. A. (1977). Bunyamwera virus-induced polypeptide synthesis. *Journal of virology* **24**, 397-400.
- Pettersson, R. F. (1974). Effect of Uukuniemi virus infection on host cell macromolecule synthesis. *Medical biology* **52**, 90-97.
- Pinheiro, F. P., Travassos da Rosa, A. P., Travassos da Rosa, J. F., Ishak, R., Freitas, R. B., Gomes, M. L., LeDuc, J. W. & Oliva, O. F. (1981). Oropouche virus. I. A review of clinical, epidemiological, and ecological findings. *The American journal of tropical medicine and hygiene* **30**, 149-160.
- Plassmeyer, M. L., Soldan, S. S., Stachelek, K. M., Martin-Garcia, J. & Gonzalez-Scarano, F. (2005). California serogroup Gc (G1) glycoprotein is the principal determinant of pH-dependent cell fusion and entry. *Virology* **338**, 121-132.

- Polidoros, A. N., Pasentsis, K. & Tsaftaris, A. S. (2006). Rolling circle amplification-RACE: a method for simultaneous isolation of 5' and 3' cDNA ends from amplified cDNA templates. *BioTechniques* 41, 35-36, 38, 40 passim.
- Pollitt, E., Zhao, J., Muscat, P. & Elliott, R. M. (2006). Characterization of Maguari orthobunyavirus mutants suggests the nonstructural protein NSm is not essential for growth in tissue culture. *Virology* 348, 224-232.
- Pringle, C. R. (1991). The Bunyaviridae and their genetic-An Overview. Curr Top Microbiol Immunol 169, 3-25.
- Pringle, C. R. (1996). Genetic and genome segment reassortment. In The Bunyaviridae (R.M. Elliott, ed.), pp.189-226. Plenum Press, New York.

Raju, R. & Kolakofsky, D. (1988). La Crosse virus infection of mammalian cells induces mRNA instability. *Journal of virology* **62**, 27-32.

- Roberts, A., Rossier, C., Kolakofsky, D., Nathanson, N. & Gonzalez-Scarano, F. (1995). Completion of the La Crosse virus genome sequence and genetic comparisons of the L proteins of the Bunyaviridae. *Virology* 206, 742-745.
- Rossier, C., Patterson, J. & Kolakofsky, D. (1986). La Crosse virus small genome mRNA is made in the cytoplasm. *Journal of virology* **58**, 647-650.
- Rossier, C., Raju, R. & Kolakofsky, D. (1988). LaCrosse virus gene expression in mammalian and mosquito cells. *Virology* 165, 539-548.
- Saeed, M. F., Li, L., Wang, H., Weaver, S. C. & Barrett, A. D. (2001). Phylogeny of the Simbu serogroup of the genus Bunyavirus. *The Journal of general virology* 82, 2173-2181.
- Saeed, M. F., Wang, H., Nunes, M., Vasconcelos, P. F., Weaver, S. C., Shope, R. E., Watts, D. M., Tesh, R. B. & Barrett, A. D. (2000). Nucleotide sequences and phylogeny of the nucleocapsid gene of Oropouche virus. *The Journal of general virology* 81, 743-748.
- Saitou, N. & Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular biology and evolution* **4**, 406-425.
- Salvatore, M., Basler, C. F., Parisien, J. P., Horvath, C. M., Bourmakina, S., Zheng, H., Muster, T., Palese, P. & Garcia-Sastre, A. (2002). Effects of influenza A virus NS1 protein on protein expression: the NS1 protein enhances translation and is not required for shutoff of host protein synthesis. *Journal of virology* 76, 1206-1212.
- Saraiva, E., Fampa, P., Cedeno, V., Bergoin, M., Mialhe, E. & Miller, L. H. (2000). Expression of heterologous promoters in Lutzomyia longipalpis and Phlebotomus papatasi (Diptera: Psychodidae) cell lines. *Journal of medical entomology* 37, 802-806.

- Scallan, M. F. & Elliott, R. M. (1992). Defective RNAs in mosquito cells persistently infected with Bunyamwera virus. *The Journal of general virology* 73 (Pt 1), 53-60.
- Schmaljohn, C. S. & Dalrymple, J. M. (1983). Analysis of Hantaan virus RNA: evidence for a new genus of bunyaviridae. *Virology* 131, 482-491.
- Schmaljohn, C. S., and Hooper, J. W. (2001). Bunyaviridae: The viruses and their replication. In: Field Virology (D.M. Knipe and P.M. Howley, eds.), Vol.2, pp.1581-1602. Lippincott, Williams & Wilkins.
- Sexton, D. J., Rollin, P. E., Breitschwerdt, E. B., Corey, G. R., Myers, S. A., Dumais, M. R., Bowen, M. D., Goldsmith, C. S., Zaki, S. R., Nichol, S. T., Peters, C. J. & Ksiazek, T. G. (1997). Life-threatening Cache Valley virus infection. *The New England journal of medicine* 336, 547-549.
- Shi, X., Kohl, A., Leonard, V. H., Li, P., McLees, A. & Elliott, R. M. (2006). Requirement of the N-terminal region of orthobunyavirus nonstructural protein NSm for virus assembly and morphogenesis. *Journal of virology* 80, 8089-8099.
- Shi, X., Lappin, D. F. & Elliott, R. M. (2004). Mapping the Golgi targeting and retention signal of Bunyamwera virus glycoproteins. *Journal of virology* **78**, 10793-10802.
- Short, N. J., Meek, A. D. & Dalgarno, L. (1982). Seven infection-specific polypeptides in BHK cells infected with Bunyamwera virus. *Journal of virology* **43**, 840-843.
- Singh, K. V., Saad, N. & el-Zein, A. (1970). Sensitivity of the plaque technique for the study of selected vaccine strains and a virulent strain of Newcastle disease virus. *Applied microbiology* 20, 638-640.
- Smith, J. F. & Pifat, D. Y. (1982). Morphogenesis of sandfly viruses (Bunyaviridae family). Virology 121, 61-81.
- Soldan, S. S., Plassmeyer, M. L., Matukonis, M. K. & Gonzalez-Scarano, F. (2005). La Crosse virus nonstructural protein NSs counteracts the effects of short interfering RNA. *Journal of virology* 79, 234-244.
- Spiegel, M., Pichlmair, A., Martinez-Sobrido, L., Cros, J., Garcia-Sastre, A., Haller, O. & Weber, F. (2005). Inhibition of Beta interferon induction by severe acute respiratory syndrome coronavirus suggests a two-step model for activation of interferon regulatory factor 3. *Journal of virology* 79, 2079-2086.
- Streitenfeld, H., Boyd, A., Fazakerley, J. K., Bridgen, A., Elliott, R. M. & Weber, F. (2003). Activation of PKR by Bunyamwera virus is independent of the viral interferon antagonist NSs. *Journal of virology* 77, 5507-5511.
- Suzich, J. A. & Collett, M. S. (1988). Rift Valley fever virus M segment: cell-free transcription and translation of virus-complementary RNA. *Virology* **164**, 478-486.
- Tauraso, N. M. (1969). Identification of two plaque variants of Guaroa virus. Archiv fur die gesamte Virusforschung 28, 212-218.

- Thomas, D., Blakqori, G., Wagner, V., Banholzer, M., Kessler, N., Elliott, R. M., Haller, O. & Weber, F. (2004). Inhibition of RNA polymerase II phosphorylation by a viral interferon antagonist. *The Journal of biological chemistry* 279, 31471-31477.
- Thompson, W. H. & Beaty, B. J. (1977). Venereal transmission of La Crosse (California encephalitis) arbovirus in Aedes triseriatus mosquitoes. *Science* **196**, 530-531.
- Ushijima, H., Clerx-van Haaster, M. & Bishop, D. H. (1981). Analyses of Patois group Bunyaviruses: evidence for naturally occurring recombinant Bunyaviruses and existence of immune precipitable and nonprecipitable nonvirion proteins induced in Burnyavirus-infected cells. *Virology* **110**, 318-332.
- Ushijima, H., Klimas, R., Kim, S., Cash, P. & Bishop, D. H. (1980). Characterization of the viral ribonucleic acids and structural polypeptides of Anopheles A, Bunyamwera, Group C, California, Capim, Guama, Patois, and Simbu bunyaviruses. *The American journal of tropical medicine and hygiene* **29**, 1441-1452.
- Vasconcelos, P. F. C., Travassos da Rosa, A. P. A., Rodrigues, S. G., Degallier, N. & Travassos da Rosa, J. F. S. (2001). Inadequate management of natural ecosystem in the Brazilian Amazon region results in the emergence and reemergence of arbovirus. *Cad Saude Publica* 17.
- Vende, P., Piron, M., Castagne, N. & Poncet, D. (2000). Efficient translation of rotavirus mRNA requires simultaneous interaction of NSP3 with the eukaryotic translation initiation factor eIF4G and the mRNA 3' end. *Journal of virology* **74**, 7064-7071.
- Vezza, A. C., Repik, P. M., Cash, P. & Bishop, D. H. (1979). In vivo transcription and protein synthesis capabilities of bunyaviruses: wild-type snowshoe hare virus and its temperature-sensitive group I, group II, and group I/II mutants. *Journal of virology* 31, 426-436.
- Wang, H., Beasley, D. W., Li, L., Holbrook, M. R. & Barrett, A. D. (2001). Nucleotide sequence and deduced amino acid sequence of the medium RNA segment of Oropouche, a Simbu serogroup virus: comparison with the middle RNA of Bunyamwera and California serogroup viruses. *Virus research* 73, 153-162.
- Wang, M., Pennock, D. G., Spik, K. W. & Schmaljohn, C. S. (1993). Epitope mapping studies with neutralizing and non-neutralizing monoclonal antibodies to the G1 and G2 envelope glycoproteins of Hantaan virus. *Virology* **197**, 757-766.
- Watret, G. E., Pringle, C. R. & Elliott, R. M. (1985). Synthesis of bunyavirus-specific proteins in a continuous cell line (XTC-2) derived from Xenopus laevis. *The Journal of general virology* 66 (Pt 3), 473-482.
- Weber, F., Bridgen, A., Fazakerley, J. K., Streitenfeld, H., Kessler, N., Randall, R. E. & Elliott, R. M. (2002). Bunyamwera bunyavirus nonstructural protein NSs counteracts the induction of alpha/beta interferon. *Journal of virology* 76, 7949-7955.
- Weber, F., Dunn, E. F., Bridgen, A. & Elliott, R. M. (2001). The Bunyamwera virus nonstructural protein NSs inhibits viral RNA synthesis in a minireplicon system. *Virology* 281, 67-74.

- White, A. B. (1975). Structural polypeptides of California encephalitis virus: BFS-283. *Archives of virology* **49**, 281-290.
- Whitman, L. & Shope, R. E. (1962). The California complex of arthropod-borne viruses and its relationship to the Bunyamwera group through Guaroa virus. *The American journal of tropical medicine and hygiene* **11**, 691-696.
- Yanase, T., Yoshida, K., Ohashi, S., Kato, T. & Tsuda, T. (2003). Sequence analysis of the medium RNA segment of three Simbu serogroup viruses, Akabane, Aino, and Peaton viruses. *Virus research* 93, 63-69.
- Yandoko, E. N., Gribaldo, S., Finance, C., Le Faou, A. & Rihn, B. H. (2007). Molecular characterization of African orthobunyaviruses. *The Journal of general virology* 88, 1761-1766.
- Young, D. F., Andrejeva, L., Livingstone, A., Goodbourn, S., Lamb, R. A., Collins, P. L., Elliott, R. M. & Randall, R. E. (2003). Virus replication in engineered human cells that do not respond to interferons. *Journal of virology* 77, 2174-2181.
- Zeller, H. G., Karabatsos, N., Calisher, C. H., Digoutte, J. P., Cropp, C. B., Murphy, F. A. & Shope, R. E. (1989). Electron microscopic and antigenic studies of uncharacterized viruses. II. Evidence suggesting the placement of viruses in the family Bunyaviridae. *Archives of virology* **108**, 211-227.

Appendix

Analyses of the S Segments of Divergent Orthobunyaviruses XIII International Conference on Negative Strand Viruses, June 17th-22nd,2006, Salamanca, Spain