Arch Virol (2006) 151: 2475–2494
DOI 10.1007/s00705-006-0794-z

Archives of Virology
Printed in Austria

Molecular epidemiology of African swine fever virus studied by analysis of four variable genome regions

R. J. Nix[1,*], **C. Gallardo**[2,*], **G. Hutchings**[1], **E. Blanco**[2], and **L. K. Dixon**[1]

[1]Institute for Animal Health Pirbright Laboratory, Pirbright, U.K.
[2]Centro de Investigacion en Sanidad Animal (CISA), Madrid, Spain

Received March 14, 2006; accepted April 26, 2006
Published online July 3, 2006

Summary. Variable regions of the African swine fever virus genome, which contain arrays of tandem repeats, were compared in the genomes of isolates obtained over a 40-year period. Comparison of the size of products generated by polymerase chain reaction (PCR) from four different genome regions, within the B602L and KP86R genes and intergenic regions J286L and BtSj, placed 43 closely related isolated from Europe, the Caribbean, West and Central Africa into 17 different virus sub-groups. Sequence analysis of the most variable fragment, within the B602L gene, from 81 different isolates distinguished 31 sub-groups of virus isolates which varied in sequence and number of a tandem repeat encoding 4 amino acids. Thus, each of these analysis methods enabled isolates, which were previously grouped together by sequencing of a more conserved genome region, to be separated into multiple sub-groups. This provided additional information about strains of viruses circulating in different countries. The methods could be used in future to study the epidemiology and evolution of virus isolates and to trace the sources of disease outbreaks.

Introduction

The natural hosts for African swine fever virus are warthogs, bushpigs and soft ticks (*Ornithodoros moubata* and *erraticus*). The virus is very well established in these hosts and causes inapparent, persistent infections. In contrast, ASFV causes an acute haemorrhagic fever in domestic pigs. Since its original descriptions in the 1920s in Kenya, African swine fever has been reported from most African countries south of the Sahara. The disease spread outside Africa to Lisbon, Portugal, in 1957 and was endemic in Spain and Portugal from 1960 [20] until the

*RJN and CG contributed equally to this work.

mid 1990s. Outbreaks occurred in other European countries and in the Caribbean and Brazil, but the virus has been eradicated from all of these, apart from Sardinia, where disease remains endemic. In recent years, ASF has caused severe economic losses in several West African countries and spread to Madagascar for the first time in 1998, where it has caused the loss of about half of the pig population.

ASF is caused by a large, cytoplasmic virus with a double-stranded DNA genome which varies in length between 170 and 192 kbp depending on the virus isolate. ASFV is the only member of the family *Asfarviridae* [9, 10]. The virus genome contains between 160 and 175 open reading frames, depending on the isolate, and these encode proteins and enzymes required for virus replication as well as other proteins that are non-essential for virus replication but have a role in virus survival in and transmission between its hosts.

Analysis of ASFV genomes by restriction enzyme site mapping and by partial sequencing of the gene encoding the major capsid protein p72 has shown that isolates from Europe, the Caribbean, S. America, and W. and C. Africa are closely related to each other. In contrast, isolates from S. and E. Africa are more diverse. The data suggest that isolates from the long-established sylvatic cycle in E. and S. Africa are genetically diverse and that several introductions have occurred from the sylvatic cycle into domestic pig populations in these regions [4, 21, 23]. Once introduced into pig populations, virus can be transmitted between domestic pigs by bites from infected ticks [18, 22], by ingestion of infected meat and by direct contact between pigs [11]. Restriction enzyme site mapping and sequence analysis of virus genomes has established that the central region of the ASFV genome is relatively conserved in length but that large length variations occurred, particularly within 40 kbp from the left end of the genome, but also within 15 kbp from the right end of the genome. Many of the length variations are associated with gain or loss of members of multigene families. In addition, smaller length variations are associated with variation in the number of tandem repeats which are located at a number of genome positions both within coding regions and in intergenic regions between genes [1, 2, 5, 8, 25].

In the current study we have identified variable genome regions, from a diverse range of ASFV isolates, by analysis of the size of PCR fragments amplified from several genome regions containing tandem repeat arrays. We have also sequenced one of the most variable genome regions, from the B602L gene, from a wide range of isolates. The results provide additional information about the ASFV genome variablilty and a method which can be used to distinguish between closely related isolates. This information has provided novel epidemiological information about circulating ASFV strains.

Materials and methods

Virus isolates

Virus isolates selected for study were available in the ASFV collection held at the IAH Pirbright and CISA Madrid. These isolates are described in Tables 1 (European, Caribbean and South American isolates) and 2 (African isolates). Isolates include the Nu 86, Nu 90/1, Ori 90,

Table 1. Summary of European, Caribbean and South American isolates used in this study

Isolate	Country	Year	Origin	Genotype*	p72	B602L	Reference
European							
Bel 85	Belgium	1985	Pig	I	Y	Y*	IAH-P
Fr 64	France	1964		I		Y*	CISA
Hol 85	Holland	1985	Pig	I	Y	Y*	IAH Pirbright
Malta 78	Malta	1978	Pig	I	Y	Y*	Wilkinson et al., 1980
Lis 57	Portugal	1957	Pig	I	Y*	Y*	Dr. J. Vigario LNIV
Lis 60	Portugal	1960	Pig	I	Y	Y*	Dr. J. Vigario LNIV
Por 63	Portugal	1963		I	Y	Y*	Dr. J. Vigario LNIV
Mon 84	Portugal	1984	Pig	I	Y*	Y*	Dr. J. Vigario LNIV
Cas 86	Portugal	1986	Pig	I	Y*	ND	Dr. J. Vigario LNIV
Coi 86	Portugal	1986		I	Y*	Y*	Dr. J. Vigario LNIV
Por 86	Portugal	1986	Pig	I	Y*	Y*	Dr. J. Vigario LNIV
San 86	Portugal	1986	Pig	I	ND	Y*	Dr. J. Vigario LNIV
Tom 86	Portugal	1986	Pig	I	Y*	Y*	Dr. J. Vigario LNIV
Vis 86	Portugal	1986	Pig	I	Y*	Y*	Dr. J. Vigario LNIV
Tom 87	Portugal	1987	Tick (pig)	I	ND	ND	Dr. J. Vigario LNIV
Our T88/1	Portugal	1988	Tick (pig)	I	Y	ND	Boinas et al. [6]
Our T88/2	Portugal	1988	Tick (pig)	I	Y*	ND	Boinas et al. [6]
Our T88/3	Portugal	1988	Tick (pig)	I	Y*	ND	Boinas et al. [6]
Our T91/1	Portugal	1991	Tick (pig)	I	Y*	Y*	Boinas et al. [6]
Port 99	Portugal	1999	Tick (pig)	I	Y*	Y*	Dr. Fernando Portugal LNIV
Ca 78	Sardinia	1978	Pig	I	ND	Y*	CISA
Nu 79	Sardinia	1979	Pig	I	Y*	Y*	Dr. A. Lattumado
Ss 81	Sardinia	1981	Pig	I		Y*	CISA
Nu 81/1	Sardinia	1981	Pig	I	Y*	Y*	Dr. A. Lattumado
Nu 84	Sardinia	1984	Wild Boar	I	Y*	ND	Dr. A. Lattumado
Ori 84	Sardinia	1984	Pig	I	ND	Y*	CISA
Ori 85	Sardinia	1985	Pig	I	Y*	Y*	CISA
Ca 85	Sardinia	1985	Pig	I	ND	Y*	CISA
Nu 86	Sardinia	1986	Pig	I	ND	ND	IAH-P
Ss 88	Sardinia	1988	Pig	I	ND	Y*	IAH-P
Nu 88/3	Sardinia	1988		I	ND	Y*	IAH-P
Nu 90/1	Sardinia	1990	Pig	I	Y	Y*	IAH-P
Nu 90/2	Sardinia	1990	Pig	I	ND	Y*	IAH-P
Ori 90	Sardinia	1990	Pig	I	Y*	Y*	IAH-P
Nu 91	Sardinia	1991	Pig	I	ND	Y*	CISA
Nu 91/3	Sardinia	1991	Pig	I	ND	Y*	CISA
Nu 91/5	Sardinia	1991		I	ND	Y*	CISA
Nu 93	Sardinia	1993	Pig	I	ND	Y*	Mannelli et al., 1998
Ori 93	Sardinia	1993	Pig	I	ND	Y*	CISA
Nu 95/1	Sardinia	1995	Pig	I	ND	Y*	CISA
Nu 95/4	Sardinia	1995	Pig	I	Y*	Y*	CISA
Nu 96	Sardinia	1996	Pig	I	ND	Y*	CISA
Ca 97	Sardinia	1997	Pig	I	ND	Y*	CISA

(*continued*)

Table 1 (*continued*)

Isolate	Country	Year	Origin	Genotype*	p72	B602L	Reference
Nu 97	Sardinia	1997	Pig	I	ND	Y*	CISA
Nu 98	Sardinia	1998	Pig	I	ND	Y*	CISA
Ali 61	Spain	1961	Pig	I	ND	Y*	CISA
M 61	Spain	1961	Pig	I	ND	Y*	CISA
Co 61	Spain	1961	Pig	I	ND	Y*	CISA
Co 62	Spain	1962	Pig	I	ND	Y*	CISA
Mad 62	Spain	1962	Pig	I	Y	Y*	CISA
Ba 68	Spain	1968	Pig	I	ND	Y*	CISA
Co 68	Spain	1968	Pig	I	ND	Y*	CISA
E 70	Spain	1970	Pig	I	ND	Y*	CISA
Ba71V	Spain	1971	T/C	I	Y	Y*	
Av 71	Spain	1971	Pig	I	ND	Y*	CISA
B74	Spain	1974	Pig	I	ND	Y*	J. Plana, PC
E75	Spain	1975	Pig	I	ND	Y*	Sanchez-Vizcaino et al.,1981
Val 76	Spain	1976	Pig	I	Y	Y*	IAH-P
Mu 82	Spain	1982	Pig	I	ND	Y*	IAH-P
Zar 85	Spain	1985	Pig	I	Y	Y*	IAH-P
Sa 88	Spain	1988	Pig	I	ND	Y*	Perez-Sanchez et al., 1994
Se 88	Spain	1988	Pig	I	ND	Y*	CISA
Hu 90	Spain	1990	Pig	I	ND	Y*	CISA
Hu 94	Spain	1994	Pig	I	ND	Y*	CISA
646	Spain	1969	Pig	I	ND	Y*	CISA
South America and the Caribbean							
Brazil 78	Brazil	1978	Pig	I	Y	Y*	Mebus et al., 1978
Dom Rep	Dominican Republic	1978	Pig	I	ND	Y*	Mebus et al., 1978
Haiti	Haiti	1981	Pig	I	Y*	Y*	CISA

Y denotes sequence available. *Denotes genotype determined by partial p72 gene sequencing performed in this study. *ND* not done; *CISA* Centro de Investigacion en Sanidad Animal, Madrid; *IAH-P* Institute for Animal Health, Pirbright; *T/C* tissue-culture-adapted isolate

Nu 84, Nu 86, Nu 95/4, Ori 85 and Nu 79 isolates from Sardinia provided by Drs. Domenico Rutili and Alberto Laddomada. The Vis 86, Tom 86, Por 86, Coi 86, Mon 84, San 86, Lis 60 isolates from Portugal were provided by Dr. J. Vigario, Laboratorio Nacional de Investigaξao Veterinaria (LNIV), Lisbon. The Our T88/1, Our T91/1, Our T88/2, Our T88/3 isolates were provided by Dr. Fernando Boinas, Faculdade de Medicina Veterinaria (FMV), Universidade Tecnica de Lisboa, Lisbon. The Port 99, Cape Verde 93 and Guinea 92 isolates were obtained from Dr. Fernando Portugal LNIV, Lisbon. Additional isolates analysed were available from virus collections held at IAH and CISA, and included ones from Europe; Belgium (Bel 85), France (Fr 64), Holland (Hol 86), Malta (Mal 78), Portugal (Lis 57, Por 63, Cas 86, Tom 87, Port 99), Sardinia (Ca 78, Nu 79, Ss 81, Nu 81/1, Ori 84, Ca 85, Ss 88, Nu 88/3, Nu 90/2, Nu 91, Nu 91/3, Nu 91/5, Nu 93, Ori 93, Nu 95/1, Nu 93, Ori 93, Nu 95/1, Nu 96, Ca 97, Nu 97, Nu 98), Spain (Ali 61, M 61, Co 61, Co 62, Mad 62, Ba 68, E70, Ba71V, Av71, B74,

Table 2. Summary of African isolates used in this study

Isolate	Country	Year	Origin	Genotype	p72	B602L	Reference
African							
Ang 70	Angola	1970	Pig	I	Y*	Y*	IAH-P
Ang 72	Angola	1972	Pig	I	Y	Y*	Vigario et al., 1970
Bur 84/1	Burundi	1984	Pig	X	Y	Y*	IAH-P
Bur 84/2	Burundi	1984	Pig	X	Y	Y*	IAH-P
Bur 90/1	Burundi	1990	Pig	X	Y	Y*	IAH-P
Ben 97/1	Benin	1997	Pig	I	Y	ND	IAH-P
Ben 97/2	Benin	1997	Pig	I	Y*	ND	IAH-P
Ben 97/3	Benin	1997	Pig	I	Y*	Y*	IAH-P
Ben 97/4	Benin	1997	Pig	I	Y*	ND	IAH-P
Ben 97/5	Benin	1997	Pig	I	Y*	ND	IAH-P
Ben 97/6	Benin	1997	Pig	I	Y*	Y*	IAH-P
Bots 1/99	Botswana	1999	Pig	III	Y	Y*	IAH-P
Cam 82	Cameroon	1982	Pig	I	Y	Y*	Wesley and Tuthill [23]
Cam 85/4	Cameroon	1985	Pig	I	Y	ND	IAH-P
CV 93	Cape Verde	1993			ND		IAH-P
CV 97	Cape Verde	1997	Pig		ND	Y*	CISA
CV 98	Cape Verde	1998	Pig		ND	Y*	CISA
Gui 92	Guinea	1992			ND	ND	IAH-P
Cm 96	Ivory Coast	1996	Pig		ND	Y*	CISA
Hinde II	Kenya	1959	Warthog	X	Y	Y*	IAH-P
Ten 60	Malawi	1960	Warthog	V	Y	Y*	IAH-P
NDA 90/1	Malawi	1990	Pig	VIII	Y	Y	Sumption et al. [21]
Zom 84/2	Malawi	1984	Pig	VIII	Y	Y*	Sumption et al. [21]
MwLil20/1	Malawi	1983	Tick (Pig)	VIII	Y	Y*	Haresnape et al., 1988
Mal 78	Malawi	1978		VIII	Y	Y	Sumption et al. [21]
Moz 64	Mozambique	1964	Pig		ND	Y*	Vigario et al., 1970
Moz 94/1	Mozambique	1994	Pig	VI	Y	Y*	IAH-P
Nig 01	Nigeria	2001	Pig		ND	Y*	CISA
Dakar 59	Senegal	1959	Pig	I	Y	ND	IAH-P
KWH12	Tanzania		Warthog	X	Y	ND	IAH-P
Uga 95/1	Uganda	1995	Pig	IX	Y	Y*	IAH-P
Uga 95/3	Uganda	1995	Pig	X	Y	Y*	IAH-P
Kat 63	Zaire	1963	Pig	I	Y	ND	IAH-P
Kat 67	Zaire	1967	Pig		ND	Y*	Vigario et al., 1970
Jon 89/13	Zambia	1989	Pig	VIII	Y	ND	Sumption et al. [21]
Kal 88/1	Zambia	1988	Pig	VIII	Y	Y*	Sumption et al. [21]
Kav 89/1	Zambia	1989	Pig	VIII	Y	ND	Sumption et al. [21]
Vict 90/1	Zambia	1990	Tick (Warthog)	I	Y	Y*	Sumption et al. [21]

Y denotes sequence available. *Denotes sequencing performed in this study. *CISA* Centro de Investigacion en Sanidad Animal, Madrid; *IAH-P* Institute for Animal Health, Pirbright; *ND* not done; *PC* personal communication

E75, Val 76, Mu 82, Zar 85, Sa 88, Se 88, Hu 90, Hu 94, 646), the Caribbean and South America; Dominican Republic (Dom Rep), Haiti (Haiti) and Brazil (Brazil 78); and Africa; Angola (Ang 70, Ang 72) Benin (Ben 97/1, 97/2, 97/3, 97/4, 97/5, 97/6), Botswana (Bots

1/99) Burundi (Bur 84/1, 84/2, 90/1), Cameroon (Cam 82, Cam 85/4), Cape Verde (CV97, CV 98), Ivory Coast (Cm 96), Kenya (Hinde II), Malawi (NDA 90/1, Zom 84/2, MwLil 20/1, Mal 78), Mozambique (Moz 64, Moz 94/1), Nigeria (Nig 01), Senegal (Dakar 59), Tanzania (KWH12), Uganda (Uga 95/3, Uga 95/1), Zaire (Kat 63, Kat 67), Zambia (Kal 88/1, Jon 89/13, Kav 89/1) and Zimbabwe (Vict 90/1).

Preparation of primary porcine alveolar macrophages

Porcine alveolar macrophages were prepared by lung lavage and plated at a concentration of approximately 10^7 cells ml^{-1}. Macrophages were cultured in RPMI medium plus HEPES supplemented with 5% fetal calf serum and incubated at 37 °C in the presence of 5% CO_2.

Serial passage of virus stocks

The growth medium was removed from flasks of alveolar macrophages and the cells were washed once with PBS before inoculation of the flask with an ASFV isolate. Serum-free RPMI plus HEPES medium was added, and the cells were incubated at 37 °C in the presence of 5% CO_2. After 1 h, the flask was supplemented with fresh RPMI + HEPES supplemented with 5% fetal calf serum and incubated at 37 °C in the presence of 5% CO_2. After 5 days, the medium was transferred to a Falcon tube and centrifuged at 3500 × *g* for 15 min to remove the cell debris. Supernatants were stored at 4 °C for short-term storage or −70 °C for long-term storage.

Isolation of virus DNA

Viral DNA was extracted directly from cell culture isolates or from suspensions of clinical samples using the GFX Genomic Blood DNA Purification Kit (Amersham Biosciences) following the protocol for blood.

PCR amplification

Polymerase chain reactions (PCRs) were performed using the Triplemaster PCR system (Eppendorf) as recommended by the manufacturer. Reactions contained high-fidelity buffer, 0.2 mM dNTPs, triplemaster enzyme mix, 100 ng of DNA and 500 nM of each primer in a final reaction volume of 20 µl. Reactions included a 5-min denaturation step at 95 °C followed by 40 cycles including 1 min at 95 °C, 1 min at 50 °C and 1 min at 68 °C followed by a 10-min elongation step at 68 °C. Part of the gene encoding the p72 proteins was isolated using primers p72-D and p72-U [4], which amplify a product of 478 bp. The primer pairs ORF9L-F (5′AATGCGCTCAGGATCTGTTAAATCGG3′, binding site 84635–84659) and ORF9L-R (5′TCTTCATGCTCAAAGTGCGTATACCT3′, binding site 85027–85004) were used to amplify part of the B602L gene [14], the primers Mal-F1 (5′CCAACACAATGGTTATAGACA ACACA3′, binding site 148754–148773) and Mal-R1 (5′TATTGTGCCTGTGTAAACTCC GGCT3′, binding site 122879–122894) [8] were used to amplify the Bt/Sj region, J268L-F (5′GGTTCACTGGTGTCCATGATCAAAA3′, binding site 10535–10559) and J268L-R (5′CCTAAATGATAAAACCGATTACATC3′, binding site 11599–11575) were used to amplify part of the J268L region [2] and KP86R-F (5′TTTCGCTTGATCAAGAAATATACA AAA3′, binding site 1699–1725) and KP86R-R (5′TCTTATACATATCTGTTGTCATACG3′, binding site 2045–2021) were used to amplify the KP86R gene. Primer binding sites and product size estimations are given based on the Ba71V genome (Accession No. U18466).

PCR product purification

Amplified products were gel-purified using GFX PCR DNA and the Gel Band Purification Kit (Amersham Biosciences).

Fragment analysis

PCR reactions were performed as described, substituting one primer of each pair with a Well Red Dye labelled primer (Proligo). Typically, 5 μl of PCR reaction was added to 34.5 μl sample loading solution (Beckman Coulter) containing 0.5 μl CEQ DNA Size Standard-600 (Beckman Coulter). All fragment analysis was conducted on a Beckman Coulter CEQ8000 Genetic Analysis system using the Frag-4 program. Data were processed using the dye mobility ver 1 and size standard 600 analysis parameters and analysed using Beckman Coulter CEQ8000 software.

Nucleotide sequence analysis

Primers used for the sequencing of p72 were p72-D (5′GGCACAAGTTCGGACATGT3′) and p72-U (5′GTACTGTAACGCAGCACAG3′) [4] and for B602L were ORF9LF2 (5′CAT CCGGGCCGGTTTCTTGTATAT3′) and ORF9L-R3 (5′GGAGTTTGGGTGATTGCATCA ATATCG3′). Reactions were prepared using the Dye Terminator Cycle Sequencing (DTCS) QuickStart Mix (Beckman Coulter). Thermal cycling consisted of 30 cycles of 96 °C for 20 sec, 50 °C for 20 sec and 60 °C for 3 min. Completed reactions were processed following the manufacturer's instructions. All sequencing was conducted on a Beckman Coulter CEQ8000 Genetic Analysis System using the LFR-1 program. Data were processed using the default sequence analysis parameters and analysed using GCG and Beckman Coulter CEQ8000 software. The accession numbers for sequences of B602L fragments described are AM259388 to AM259466.

Results

Selection of variable ASFV genome fragments for analysis by PCR amplification

Sequence analysis of ASFV isolate genomes has identified regions which contain tandem repeat arrays either within the coding regions of genes or in intergenic regions. We designed primers from regions flanking eight repeat arrays and tested their ability to amplify fragments from ASFV isolates from different geographical locations. Genome regions including the ORFs DP93R (located at the left-hand end of the genome), B646L (encodes the major viral capsid protein p72), E183L (encodes the structural protein p54) and H171R (located at the right end of the genome) did not show significant length variation when compared between different isolates and were not included in further analysis. The other four genome regions, which showed greatest size variation between isolates, included the B602L (or 9R-L) gene in the centre of the genome, the KP86R gene and a region adjacent to the J268L gene from close to the left genome end and a region between the ORFs E146L and E199L, close to the right end of the genome. This genome region, which we have named BtSj, was previously shown to contain between 8 and 38 copies of a 17-nucleotide repeat [8]. The repeat region next to J268L contains a set of internal repeated sequences composed of 5 types of 200-bp-long tandemly repeated units. These units contain a G-rich core of 10–14 nucleotides surrounded by regions with a high A and T content [2]. The ORF B602L encodes a central region containing twelve-base-pair repeats [14]. The ORF KP86R is cysteine-rich and contains

tandem repeats identical to those in the ORF DP86L at the right-hand end of the genome [24].

1. Comparisons of the lengths of PCR fragments generated from variable genome regions from African swine fever virus isolates

1.1. Comparison of viral isolates from group I (W. Africa and Europe)

Forty-one virus isolates were selected from W. Africa, Europe and the Caribbean. These included isolates from Spain, Portugal, Malta, Belgium, Holland, Sardinia, Dominican Republic, Haiti, Cameroon, Angola, Benin, Democratic Republic of

Table 3. Table of groups I, VIII and X isolates' J268L, 9R-L, Bt/Sj and KP86R tandem repeat fragment sizes by PCR agarose gel estimation. PCRs were performed on DNA extracted from viral suspensions using the primer pairs J268L-F & J268L-R; ORF9L-F & ORF9L-R, Mal-F1 & Mal-R1 and KP86R-F & KP86R-R. Product sizes were estimated by their relative mobility against a molecular weight marker run on an agarose gel

Isolate	Country	PCR product sizes (Kb)				Isolate	Country	PCR product sizes (Kb)			
		J268L	B602L	Bt/Sj	KP86R			J268L	B602L	Bt/Sj	KP86R
Group I						Group I					
Mad 62	Spain	1.00	0.38	0.60	0.32	Tom 86	Portugal	0.80	0.38	0.60	0.32
Val 76	Spain	1.00	0.38	0.60	0.32	Por 86	Portugal	0.80	0.38	0.60	0.32
Zar 85	Spain	1.00	0.38	0.60	0.32	Coi 86	Portugal	0.80	0.38	0.60	0.32
Dom Rep	Dominican Republic	1.00	0.38	0.60	0.32	San 86	Portugal	0.80	0.38	0.60	0.32
						Mon 84	Portugal	0.80	0.38	0.60	0.32
Haiti	Haiti	1.00	0.38	0.60	0.32	Vis 86	Portugal	0.80	0.38	0.60	0.32
Nu 81/1	Sardinia	1.00	0.38	0.60	0.32	Port 99	Portugal	0.80	0.38	0.60	0.32
Ori 85	Sardinia	1.00	0.38	0.60	0.32	Ben 97/3	Benin	1.60	0.15	0.60	0.32
Nu 79	Sardinia	1.00	0.38	0.60	0.32	Ben 97/5	Benin	1.00	0.45	0.60	0.32
Malta 78	Malta	1.00	0.43	0.60	0.32	Ben 97/6	Benin	1.00	0.45	0.60	0.32
Nu 86	Sardinia	1.00	0.20	0.60	0.32	Ben 97/2	Benin	1.00	0.45	0.60	0.32
Ori 90	Sardinia	1.00	0.20	0.60	0.32	Ben 97/1	Benin	1.00	0.45	0.60	0.32
Nu 84	Sardinia	1.00	0.20	0.60	0.32	Cam 85/4	Cameroon	1.00	0.45	0.60	0.32
Nu 90/1	Sardinia	1.00	0.20	0.60	0.32						
Nu 95/4	Sardinia	1.00	0.20	0.60	0.32	Group VIII					
Bel 85	Belgium	1.00	0.30	0.60	0.32	NDA 90/1	Malawi	1.00	0.30	0.60	0.25
Hol 86	Holland	1.00	0.30	0.60	0.32	Zom 84/2	Malawi	1.00	0.30	0.80	0.25
Cam 82	Cameroon	1.00	0.40	0.60	0.32	Kal 88/1	Zambia	1.00	0.35	0.60	0.25
Ang 72	Angola	1.00	0.28	0.60	0.32	Jon 89/13	Zambia	1.00	0.30	0.70	0.25
Ang s70	Angola	1.00	0.28	0.60	0.32	Kav 89/1	Zambia	1.00	0.40	0.60	0.25
Dakar 59	Senegal	1.40	0.28	0.60	0.25						
Katanga 63	Dem. Rep. of Congo	1.40	0.34	0.80	0.25	Group X					
						KWH12	Tanzania	1.00	0.25	0.60	0.25
Lis 60	Portugal	1.00	0.38	0.80	0.25	Uga 95/3	Uganda	1.00	0.25	0.60	0.25
Lis 57	Portugal	1.00	0.28	0.80	0.25	Hinde II	Kenya	1.00	0.20	0.60	0.25
Vict 90/1	Zimbabwe	1.00	0.28	0.60	0.25	Bur 84/1	Burundi	1.00	0.20	0.60	0.27
Our T88/2	Portugal	0.80	0.60	0.60	0.32	Bur 84/2	Burundi	1.00	0.20	0.60	0.27
Our T88/3	Portugal	0.80	0.60	0.60	0.32	Bur 90/1	Burundi	1.00	0.20	0.60	0.27
Our T91/1	Portugal	1.00	0.38	0.80	0.32						
Our T88/1	Portugal	1.00	0.38	0.80	0.32						

Congo and Senegal. Nineteen of these isolates had been analysed by partial sequencing of the p72 gene and defined [4] as belonging to the same group (group I). Partial sequencing of the p72 genes of 23 virus isolates, which had not previously been determined, showed no differences in amino acid sequence compared to the other isolates that had been placed in group I (data not shown) and we therefore assigned these isolates to group I (see Table 1). Primers were used to amplify the four different genome regions described above and the fragments generated were compared by agarose gel electrophoresis (Table 3). Amplification of the J268L genome region from the isolates in group I produced fragments ranging in size from 0.8 to 1.6 kbp. Based on this analysis the viruses could be sub-grouped into 4.

Amplification of the B602L variable region produced fragments ranging in size from 0.15 to 0.6 kbp and enabled ten different groups of viruses to be distinguished (see Fig. 1). Amplification of the BtSj genome region produced fragments of either 0.6 or 0.8 kbp. Using primers flanking the KP86R gene, PCR generated products of 0.25, 0.27 or 0.32 kbp and enabled 3 groups of isolates to be distinguished. Combining the data from the analysis of all four of these variable genome regions, 16 subgroups of group I viruses could be distinguished from the 41 isolates analysed (see Table 3). In several cases, isolates from the same country fell into one or two groups. Thus, 5 isolates from Sardinia isolated between 1984 and

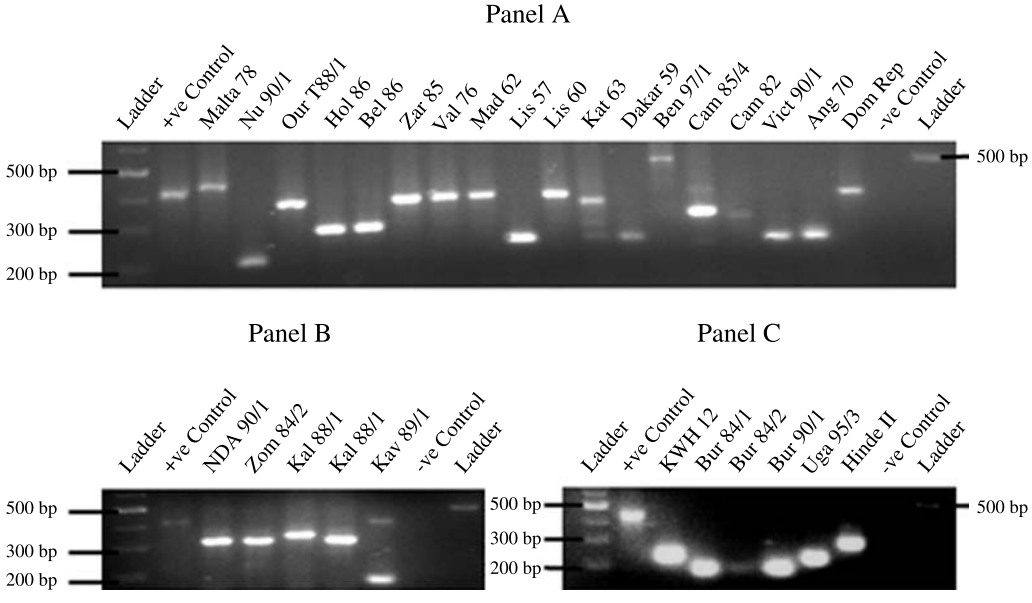

Fig. 1. Agarose gel images of B602L fragments amplified by PCR from different ASFV isolates. PCR was performed using the primers ORF9L-F and ORF9L-R using DNA extracted from virus suspensions as a template for the reaction. **A**: PCR products from European, Caribbean and West African ASFV isolates. **B**: PCR products from ASFV isolates from Malawi and Zambia. **C**: PCR products from ASFV isolates from Burundi, Kenya, Uganda and Tanzania. Negative controls were performed in the absence of DNA template. PCR products were run on a 1.6% agarose gel, and sized against a 100-bp or 1-Kb ladder (Biogene)

1990 were in one group and a further 3 Sardinian isolates were placed in a second group, based on a difference in size of the ORF B602L fragment. The Lisbon 1957 (Lis 57) isolate was placed in a separate group, and so was the Lisbon 1960 isolate (Lis 60). Six other more recent isolates obtained from Portugal in the 1980s were placed in the same group (Vis 86, Tom 86, Por 86, Coi 86, Mon 84, San 86). These isolates are all from a region north of the River Tagus and have previously been grouped together based on restriction enzyme site mapping of the complete genome [6]. Two other virus groups from Portugal were distinguished, one group containing the Our T88/1 and Our T91/1 isolate and a second group containing the Our T88/2 and Our T88/3 isolates. These isolates were obtained from the southern Alentejo region of Portugal from *Ornithodoros erraticus* ticks inhabiting pig houses. It is likely that these had persisted in tick populations for long periods since they were isolated months or years after ASFV had occurred in pigs on the farms concerned [6]. Isolates from the Dominican Republic and Haiti were placed in the same group, and the three Spanish isolates analysed were placed in the same group. Five isolates obtained from Benin from ASF outbreaks in pigs in 1997 were tested. Interestingly, these were placed into 2 different groups based on variation in size of the B602L and J268L fragments. The Katanga 1963 isolate from Democratic Republic of Congo was individually identifiable.

1.2. Comparisons of PCR fragments generated from variable genome regions from isolates from group VIII

Five isolates from Malawi and Zambia in E. Africa, which had been placed into the same group by partial sequencing of the p72 gene (group VIII), were compared by length of the four variable regions described above. The size of the J268L and KP86R genes were the same for all group VIII isolates, although analysis of the ORF B602L and Bt/Sj fragment each placed the isolates into 3 groups. Combining the analysis of these regions allowed each of these isolates to be individually distinguished.

1.3. Comparisons of PCR fragments generated from variable genome regions from isolates from group X (Burundi, Kenya, Uganda and Tanzania)

Six isolates from Uganda, Burundi and Tanzania were compared. These had previously been compared by partial sequencing of the gene encoding the p72 protein and placed in the same group (group X [4, 17]). By comparison of lengths of PCR fragments generated from the four variable genome regions, these isolates could be placed into three different groups. Isolates obtained from Burundi in 1984 and 1990 could not be distinguished from each other, whilst isolates from Tanzania and Uganda grouped together. The Hinde II isolate (Kenya) was individually identifiable. Analysis of the J268L and BtSj regions did not distinguish between the isolates. Analysis of the ORF B602L enabled two sub-groups of viruses to be distinguished, and analysis of KP86R distinguished two sub-groups of viruses.

2. Comparison of PCR product size estimation by agarose gel electrophoresis to fragment analysis using a capillary sequencer

Fragment analysis using the capillary sequencer is accurate to a single nucleotide, so offers some advantages in accuracy. However, this method is applicable only to relatively small fragments, between 60 and 640 bp, and hence could not be used for the larger fragments greater than 650 bp, generated from tandem repeat arrays.

A panel of PCR products obtained from group I isolates were screened by fragment analysis using a capillary sequencing machine to determine the accuracy of fragment size estimation by agarose gel electrophoresis (Table 4). The results show in general a good agreement in the size of fragments estimated by both methods. For example, differences in the size estimate by agarose gel electrophoresis compared to sequencing were between 1 and 22 bp for the B602L fragment, between 2 and 20 bp for the BtSj fragment and between 2 and 4 bp for the KP86R fragment.

Table 4. Comparison of B602L, Bt/Sj and KP86R PCR fragments as estimated by agarose gel electrophoresis to sizes measured by fragment analysis using a capillary sequencer. PCRs were performed on DNA extracted from viral suspensions using the primer pairs ORF9L-F and ORF9L-R, Mal F1 and Mal R1, and KP86R-F and KP86R-R. Product sizes were estimated by their relative mobility against a molecular weight marker run on an agarose gel. Fragment analysis was performed in a PCR reaction using the appropriately Well Red Dye-labeled primer and its corresponding unlabeled primer. Agarose gel-purified PCR products were used as a template for the fragment analysis reaction using a 600-bp size standard (Beckman Coulter). Fragment analysis measurements were taken using a Beckman Coulter CEQ8000 capillary sequencer

Isolate	p72 genotype	B602L (kbp)		Bt/Sj (kbp)		KP86R (kbp)	
		Fragment analysis	Agarose gel	Fragment analysis	Agarose gel	Fragment analysis	Agarose gel
Nu 95/4	I	0.202	0.20	0.619	0.60	0.317	0.32
Nu 86	I	0.202	0.20	0.619	0.60	0.317	0.32
Ori 90	I	0.201	0.20	0.619	0.60	0.317	0.32
Nu 84	I	0.202	0.20	0.619	0.60	0.317	0.32
Kat 63	I	0.331	0.34	ND	ND	ND	ND
Lis 60	I	0.357	0.38	ND	ND	0.318	0.32
Tom 86	I	0.357	0.38	0.598	0.60	ND	ND
Por 86	I	0.357	0.38	ND	ND	0.317	0.32
San 86	I	0.357	0.38	ND	ND	0.318	0.32
Vis 86	I	0.357	0.38	0.620	0.60	0.317	0.32
Ori 85	I	0.358	0.38	0.619	0.60	0.316	0.32
Nu 79	I	0.358	0.38	ND	ND	ND	ND
Uga 95/3	VIII	0.251	0.25	ND	ND	ND	ND
NDA 90/1	VIII	0.314	0.30	ND	ND	ND	ND
Kal 88/1	VIII	0.339	0.35	ND	ND	ND	ND
Bur 84/1	X	0.165	0.20	ND	ND	ND	ND
Hinde II	X	0.259	0.28	ND	ND	ND	ND

Table 5. B602L Amino acid tetramer sequence of ASFV isolates. Key: A, CAST; a, CVST, CTST, CASI; B, CADT, CTDT; C GAST, GANT; D, CASM; F, CANT; N, NVDT, NVGT; T, NVNT; H, RAST; S, SAST; O, NANI, NADI, NASI; and V, NAST, NAVT, NANT, NADT. Dashes indicate gaps introduced to enable similarities between sequences to be more easily visualised

Sub-group	Country	Isolates		P72 genotype	CVR amino acid sequence	No. repeats
I	Spain	M 61	Co 62	I	-A—-BNAAAAAA-BABNABNABNAB-----TD---FA	25
II	Spain	Co 61		I	-AAABNAAAAAAAACBNABNABNAB-----TD---FA	28
III	France	Fr 64		I	-A--BNAAAA----CBNABNABNAB-----TDBNAF	24
	Portugal	Mon 84 Por 86 San 86 OurT91/1	Vis 86 Coi 86 Tom 86 Port 99			
	Sardinia	Ca 78 Nu 81 Ori 84 Ss 88	Nur 79 Ss 81 Ori 85 Nu 98/8B			
	Spain	Ali 61 Av 71 B74 Mu 82 Sa 88 Hu 90 646	Mad 62 Val 76 E75 Zar 85 Se 88 Hu 94			
	Haiti	Haiti				
IV	Portugal	Our T88/1		I	-A--BNAAAA---CBNABTABNAB------TDB	21
V	Spain	Ba71v		I	-A—-BNAAAAAAACBNABNABNAB------TDBNAFA	28
	Portugal	Por 63				
VI	Malta	Malta 78		I	-A—-BNAAAA---CBNABNABNABNABNABTDBNAFA	31
VII	Belgium	Bel 85		I	-A—-BNA-------BNABNAB---------TDBNAFA	18
	Holland	Hol 86				
VIII	Portugal	Lis 60		I	-A--BNAAAA---CBNABNABNA------DTDBNAFA	25
IX	Spain	Co 68		I	-A—-BNAAAAA--CBNABNABNA------BTDBNAFA	26
X	Sardinia	Nu 90/1 Ori 90 Nu 91/5 Nu 95/4 Ca 97 Nu 98/3	Nu 90/2 Nu 91/3 Nu 95/1 Nu 96 Nu 97 Nur 95/4	I	-A--BNAAA----------DBNAFA	12
XI	Dominican Republic	Dom Rep		I	-A—BNAAAA-----CBNABNABN--------DBNAFA	22
XII	Brazil	Brazil 78		I	-A--BNAAAA----CBNA----------BTDBNAFA	19
XIII	Angola	Ang 70	Ang 72	I	-AA-BNA----------BTDBNAAAA	14
	Portugal	Lis 57				
XIV	Zaire	Kat 67		I	-AAAAAAAABNA----------BTDBNAAAAAAA	23
XV	Cape Verde	CV 97	Cv 98	I	-AABNABNA----------BTDBNAAAA	17
XVI	Cameroon	Cam 82		I	-A—BNAAAA-----CBNA----------BTDBNAAAAANA	23
XVII	Zambia	Vict 90/1		I	---BNAAFN----------BTDBTAFF	14
XVIII	Benin	Ben 97/3		I	AAABNABA	8
XIX	Benin	Ben 97/6		I	-A-BNAAAA-----CBNAAAAA-CBNAAAAA-CBNAAAACBNAFA	36
XX	Nigeria	Nig 01			-AAA----------CBNAAAAAACBNAAAAAACBNAAAACBNA	32
XXI	S. Africa	Ten 60		V	-ABNBBVa	7
XXII	Mozambique	Moz 94/1		VI	-AAAA----------BABNABABNBTBA	17
XXIII	Mozambique	Moz 64			-A-----BNAAAaBNBNBABN-BABTAa	21
XXIV	Uganda	Uga 95/1		IX	-AAA---BNABBNABBNABBaaBBNABNBA	26
XXV	Uganda	Uga 95/3		X	-AAAAAABNaABA	12
XXVI	Kenya Burundi	Hinde II Bur 84/1 Bur 90/1	Bur 84/2	X	-AAA---BNAAAAAAAAABA	17
XXVII	Botswana	Bots 1/99		III	-BVSVSV---VNAAABAB	14
XXVIII	Zambia	Kal 88/1		VIII	-AVSVSVSVSVNAAAAAAACBNABNABNABT-	30
XXIX	Malawi	Malawi Lil 20/1		VIII	-AVSVS---OVNA---ONOVVNVOVNAVNOVVN--OVOOV	31
XXX	Malawi	Malawi 78		VIII	-AVSVS-----BVVVAVNOVVNVNVVNAANOVVDBOVOOV	34
XXXI	Malawi	Zom 84/2		VIII	-AVSVS------VVVAVNOVVNV----------OVOOV	21

3. Comparison of ORF B602L sequences from African swine fever virus isolates

3.1. Comparison of viral isolates from West Africa, Europe and the Caribbean

To determine if sequencing of the B602L variable genome region could provide more information about relationships between isolates, we determined the sequence of this genome region from 41 isolates that had been compared by analysing variation in PCR fragment size plus an additional 48 isolates. As previously described, the B602L variable genome region contains twelve-base-pair repeats which encode 4 amino acids that vary in number and sequence when genomes of different isolates are compared. In all, 22 different sequences of amino acid tetramers were identified from the isolates we sequenced. These were given code numbers depending on their sequence, as previously described [14] and shown in Table 5. Fifty isolates from Europe, the Caribbean and Brazil, which were placed into the same group I, were divided into 13 sub-groups based on B602L sequences. For all of these sequences, a process of sequence divergence of individual tetramer sequences combined with unequal crossing over during replication could explain how isolates were derived from each other. The most common tetramer encoded by these isolates was CAST (coded as A in Table 5), and variable numbers of this repeat were encoded by different isolates. The triplet of tetramers coded BNA, where B is CADT or CTDT, N is NVDT or NVGT, was repeated several times in different isolates. Repeat arrays in most European, Caribbean and Brazilian isolates ended at the C-terminus with the sequence of tetramers DBNAF(A), where D is CASM, N is NVDT or NVGT and F is CANT. Exceptions to this were the Portuguese isolate OurT88/1, from which the NAF(A) was missing, and Spanish isolates M61, Co62, Co61, from which BNA was missing, and the Lisbon 57 isolate, which lacked the FA sequence. The largest sub-group of European isolates (sub-group III), contained 30 isolates from Sardinia, France, Spain, Haiti and Portugal. Sardinian isolates could be divided into 2 sub-groups. All of the isolates from prior to 1990 and one isolate from 1998 (Nu 98/8B) were placed in sub-group III, and the remaining 12 Sardinian isolates from 1990 to 1998 were grouped together, and separate from any other isolates, into sub-group X. These Sardinian isolates contained repeat arrays from which 12 tetramer repeats were deleted from the centre of the array compared to the sub-group 2 viruses. The isolate from the Dominican Republic (sub-group XI) differed from group III isolates by deletion of the tetramer sequences ABT from the centre of the array. The Spanish isolates were separated into 4 different sub-groups. The largest sub-group contained 13 isolates obtained between 1962 until the 1990s and were in sub-group III. Another group of 2 Spanish isolates (M 61, Co 62) were placed together and were distinguished from other isolates. The Portuguese isolates grouped into 4 sub-groups. The Lisbon 57 isolate grouped with two isolates from Angola, (Angola 70 and Angola 72), although by analysis of the size of the four variable fragments, the Lisbon 57 isolate was individually distinguishable. For some isolates, B602L sequence analysis also did not distinguish between isolates

that could be distinguished by fragment size analysis. For example, within sub-group III, defined by sequence analysis of B602L, isolates from Spain (Ali 61, Mad 62, Av 71, Val 76, B74, E75, Mu 82, Zar 85, Sa 88, Se 88, Hu 90, Hu 95, 646), Haiti (Haiti), and Sardinia (Ca 78, Nu 79, Nu 81, Ss 81, Ori 84, Ori 85, Ss 88, Nu 98/8B) were indistinguishable from some of the isolates from Portugal (Tom 86, Por 86, Coi 86, San 86, Mon 86, Vis 86). However, these Portuguese isolates could be distinguished from the other isolates based on the size of the J268L fragment (0.8 kbp compared to 1 kbp). In contrast, two Portuguese isolates (Our T91/1 and Our T88/1) that were indistinguishable by tandem repeat fragment size could be distinguished by their B602L sequence. Sequence analysis showed that the two isolates differed by 3 amino acid tetramers (24 compared to 21 repeats), and the small size difference (12 nucleotides) would have been difficult to distinguish by agarose gel electrophoresis. Isolates from Holland and Belgium (sub-group VII) were indistinguishable from each other but differed from other isolates. Epidemiological data also linked these outbreaks together.

Isolates from the Caribbean and Brazil are closely related to those from Europe and W. Africa based on p72 sequencing. The isolate from Haiti (1981) was placed in sub-group III based on B602L sequencing, whereas the isolate from Dominican Republic was placed in a separate group, which contained 3 fewer tetramer repeats.

The 9 West African isolates that were placed into group I by p72 sequencing, could be sub-divided into 7 sub-groups based on the number and sequence of amino acid tetramer repeats. The number of tetramer repeats varied from 9 to 36, and 10 different tetramer sequences were encoded. Two isolates from Benin (Benin 97/3 and Benin 97/6) could be distinguished by sequence, as was expected from analysis of the size of the B602L fragment (Genotypes XVIII and XIX, respectively). The isolates from Benin 97/6 and Nigeria 01 contain a characteristic pattern of tetramer repeats CBNAAAAA(A) which was not present in other isolates and suggests that these isolates are more closely related to each other than to the others.

3.2. Comparisons of B602L sequence data of isolates from South and Eastern Africa

The variable B602L region was sequenced from the genomes of 15 isolates from E. and S. Africa that were placed into groups other than group I by partial sequencing of the p72 gene. The number of amino acid tetramer repeats varied between 14 and 34, and the 15 isolates sequenced could be separated into 10 sub-groups. These isolates encoded amino acid tetramers that were not identified in the genomes of the isolates from group I, including sequences NAST, NAVT, NANT, NADT (labelled V in Table 5) and NANI, NADI and NASI (labelled O on Table 5) as was expected given their genetic diversity from the group I isolates. The V tetramers were encoded by one isolate from Botswana, Bots 1/99, which was placed in group III by partial p72 sequence, 4 isolates from Malawi and Zambia which were placed in group VIII. The latter isolates also encoded the O tetramer. The remaining isolates from Kenya, Mozambique, Uganda and Burundi, which

were placed in groups IX and X by partial p72 sequence, encoded tetramers that more closely resembled those of the group I isolates.

4. Stability of the B602L variable region following passage in pig macrophages or in ticks

We compared the stability of the B602L variable region following passage in either pig macrophages or *O. erraticus* ticks. To do this, pig macrophage cultures were infected with the ASFV isolates CaV 93 or Gui 92, and after 5 days, virus recovered from the supernatant of cell cultures was harvested and used to re-infect fresh macrophage cultures. This was repeated for 15 passages, and the B602L variable fragment was amplified by PCR from virus collected at each passage. This analysis showed that no variation in size of the B602L fragment was detected during these passages (Fig. 2, Panels A and B). This is in agreement with a previous study in which the size of the B602L variable region was shown

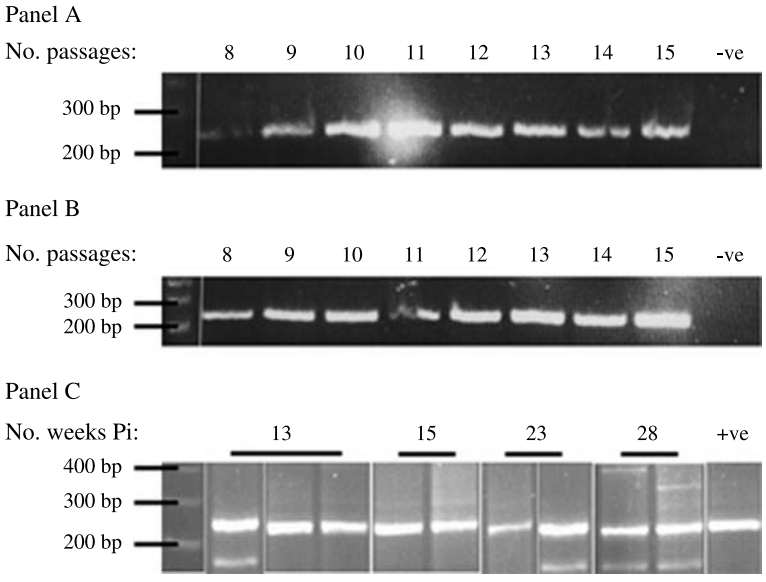

Fig. 2. Agarose gel image showing B602L fragments amplified by PCR following serial passage of virus in porcine alveolar macrophages or following infection of *Ornithodoros erraticus* ticks with African swine fever virus. PCR was performed using the primers ORF9L-F and ORF9L-R using DNA extracted from virus suspensions or tick homogenate as templates for the reaction. **A** and **B** show PCR products of the B602L region following serial passage of virus through porcine alveolar macrophages. Products from up to 15 passages of isolate Gui 92 are shown in **A** and from CaV93 in **B**. Lane headings indicate the number of passages of virus from which the PCR was performed. **C** shows PCR products of the B602L region from DNA extracted from whole-tick homogenates of *Ornithodoros erraticus* ticks which had been membrane fed the recombinant isolate Recombinant 34. Lane headings indicate the number of weeks post ingestion (weeks pi) at which ticks were homogenised, DNA extracted and PCR performed. PCR products were run on a 1.6% agarose gel, and sized against a 100-bp ladder (Biogene). Negative controls were performed in the absence of DNA template. Positive controls were performed by mixing viral DNA with DNA extracted from an uninfected tick

not to change during passage in pig macrophages. In contrast, the size of this fragment increased from 350 to 550 bp during 81 passages in MS cells [14].

To examine the stability of the B602L fragment during passage in *O. erraticus* ticks, ticks were infected by membrane feeding with a recombinant ASFV isolate, Rec 34. At various times post-feeding, ticks were homogenised and DNA was purified from whole-tick homogenates. The variable B602L fragment was amplified by PCR and analysed by agarose gel electrophoresis (Fig. 2, Panel C). This analysis showed that in the virus stock used to feed ticks, only one predominant B602L band was detected. In contrast, in 4 out of 15 infected ticks analysed at 13, 15, 23 or 28 weeks post-feeding, two or more predominant bands were detected (Fig. 2, Panel C). In four of the ticks, a band about 50 bp smaller than the predominant band present in the virus used to feed ticks was amplified, whereas in two other ticks, bands of around 50 and 80 bp larger were amplified. These fragments were sequenced to confirm that they contained sequences from the B602L genome region (data not shown). These fragments may have been generated during passage in ticks or have been present as a minor sub-population in the original virus stock and been amplified during passage in ticks. Virus obtained from homogenized ticks was titrated and found to be 5–6 log 10 HAD_{50}/ml [5, 6] tick homogenate. During virus passage in pig macrophage cultures titres of 6–7 log 10 HAD_{50}/ml [6, 7] were obtained at each passage. Thus, much greater virus replication occurred during passage in macrophage cultures compared to ticks, and the appearance of sub-populations of the B602L fragment in infected ticks was therefore not related to the amount of virus replication.

Discussion

In this study, we have investigated variable regions of the ASFV genome and examined their use in distinguishing between closely related virus isolates. Two approaches were taken to study genome variability. First, four genome regions containing arrays of tandem repeats located either within coding regions or in intergenic regions were amplified by PCR and the size of the fragments produced compared. Secondly, the nucleotide sequence of the most variable fragment, a region of the B602L gene, was determined.

In previous studies [4, 14], a conserved region of the ASFV genome, part of the gene encoding the p72 protein, was sequenced from a range of different isolates. This analysis was useful to distinguish between genetically diverse isolates but did not distinguish between closely related isolates. For example, the largest group of isolates grouped together by this analysis included a wide group of viruses from Europe, the Caribbean and West and Central Africa, obtained over a wide time period [4].

In our study, the comparison of four variable genome regions by analysis of PCR fragment sizes enabled 41 isolates from Europe, the Caribbean and West Africa, which were placed in the same group by partial sequencing of the p72 gene, to be divided into 16 sub-groups. In general, the viruses were sub-grouped according to the country from which they were isolated, and in some cases more

than one virus sub-type was identified from a single country. For example, isolates from Sardinia were placed in two groups. One group included all except one of the isolates obtained since 1990. This evidence supports the epidemiological data that virus is circulating in Sardinia and will provide a method for tracing isolates from Sardinia if they are introduced into another country. Interestingly, the isolates from Benin were placed into 2 different sub-groups based on variation in size of the B602L and J268L fragments. The variation in two separate fragments between these isolates suggests that two different isolates were present in Benin in 1997 rather than one isolate having been derived from the other. In addition, isolates from Portugal were placed in 4 different sub-groups, which in part reflects the relatively large number (13) of isolates from Portugal we analysed. Interestingly, amongst the 5 isolates obtained from ticks inhabiting pig premises in the southern part of Portugal, 3 different sub-groups could be distinguished. Possibly, the persistent infection of ticks over a long time period results in a greater genome variation. Our analysis of virus obtained from ticks at various times after feeding did suggest that virus subpopulations are readily detected in persistently infected ticks. Despite the fewer Southern and Eastern African isolates studied, isolates placed into a single genotype by partial sequencing of the p72 gene could also be sub-divided into groups by this method. For example, 5 isolates from Malawi and Zambia placed in the same group by p72 sequencing were each individually distinguishable. As previously reported, the BtSj fragment in these isolates is more variable than in other isolates [8]. This may reflect an expansion of the tandem repeats in this genome region in these isolates. Although the isolates from Malawi and Zambia were from pigs, *O. moubata* ticks are known to inhabit pig houses in this region and are thought to play a role in virus transmission [13]. Possibly, long-term persistent infection of ticks may also be important for generating genome diversity in this region.

Fragment size analysis identified the B602L as the most variable genome region. The sequencing of this genome region from 81 different isolates confirmed the fragment size data and enabled additional virus genome sub-groups to be identified. Nineteen subgroups within 66 isolates from group I p72 genotype were identified by sequencing the variable B602L fragment. The four isolates from group VIII from Malawi and Zambia were also distinguished from each other. Four group X isolates from Burundi and Kenya grouped together based on B602L sequence, whilst the Ugandan group X isolate was distinguishable from these.

The variable region of ORF B602L consists of repeated amino acid tetramers that vary in number and type We identified 23 different amino acid tetramers, although the tetramers CA(D/N)T and NV(D/N)T were most frequently encoded. Thirty-six different B602L amino acid sequences were identified, with the number of tetramers encoded per virus genome ranging from 8 to 34. The reasons for the variability of the B602L protein are not clear. The protein has been reported to act as a chaperone involved in assembly of the p72 capsid protein into virions, although B602L protein itself was reported not to be incorporated into virions [7]. Analysis of the principal serological determinants detected following infection of pigs with ASFV showed that B602L protein was one of the 14 proteins against

which antibodies were generated [16]. Variation in the sequence of the amino acid tetramers could provide a mechanism to generate antigenic variation and help the virus to evade an antibody response. Although the B602L protein has been reported not to be incorporated into extracellular virions, it may be released from cells that are lysed following infection and thus stimulate an antibody response. Tandem repeat arrays are encoded within several different ASFV proteins [24] including the E183L (also named j13L and p54) protein and EP402R (also named CD2v protein), and variation in the number and sequence of these repeats between different isolates has also been observed. The E183L (or p54 protein) is a virus structural protein which has an important role in virus entry and morphogenesis and is one of the proteins against which antibodies are detected during virus infection of pigs [3, 12, 19]. Proline-rich tandem repeats are located in the cytoplasmic tail of the CD2v protein and act as a binding site for the actin-binding adaptor protein SH3P7/mabp1 [15].

As previously described, partial sequencing of the p72 gene is useful to place ASFV isolates in broad genotypes [4, 17]. The approaches we describe have both helped to distinguish between closely related ASFV isolates. Each approach could distinguish between some isolates that weren't distinguished by the other method and thus both could be used in parallel. The ASFV genome contains a number of other regions that contain tandem repeat arrays, and hence the approach we describe could be extended to the analysis of size variation in additional repeat arrays. In this way it may be possible to distinguish between closely related isolates that haven't been distinguished using either of the methods we describe and in this way extend knowledge of virus evolution and epidemiology.

Acknowledgements

We would like to thank Drs. Carlos Martins, Fernando Boinas, Alexandre Leitao, Chris Oura, Dave Chapman, Charles Abrams, Fuquan Zhang for helpful discussions. This work was funded by EU project. QLK2-CT-2001-02216. We thank E. Martin for technical assistance. Work at CISA has been supported by the EU.

References

1. Aguero M, Blasco R, Wilkinson P, Vinuela E (1990) Analysis of naturally-occurring deletion variants of African swine fever virus – multigene family-110 is not essential for infectivity or virulence in pigs. Virology 176: 195–204
2. Almazan F, Murguia JR, Rodriguez JM, Delavega I, Vinuela E (1995) A set of African swine fever virus tandem repeats shares similarities with Sar-like sequences. J Gen Virol 76: 729–740
3. Alonso C, Miskin J, Hernaez B, Fernandez-Zapatero P, Soto L, Canto C, Rodriguez-Crespo I, Dixon L, Escribano JM (2001) African swine fever virus protein p54 interacts with the microtubular motor complex through direct binding to light-chain dynein. J Virol 75: 9819–9827
4. Bastos ADS, Penrith ML, Cruciere C, Edrich JL, Hutchings G, Roger F, Couacy-Hymann E, Thomson GR (2003) Genotyping field strains of African swine fever virus by partial p72 gene characterisation. Arch Virol 148: 693–706

5. Blasco R, Delavega I, Almazan F, Aguero M, Vinuela E (1989) Genetic-variation of African swine fever virus – variable regions near the ends of the viral-DNA. Virology 173: 251–257
6. Boinas FS, Hutchings GH, Dixon LK, Wilkinson PJ (2004) Characterization of pathogenic and non-pathogenic African swine fever virus isolates from Ornithodoros erraticus inhabiting pig premises in Portugal. J Gen Virol 85: 2177–2187
7. Cobbold C, Windsor M, Wileman T (2001) A virally encoded chaperone specialized for folding of the major capsid protein of African swine fever virus. J Virol 75: 7221–7229
8. Dixon LK, Bristow C, Wilkinson PJ, Sumption KJ (1990) Identification of a variable region of the African swine fever virus genome that has undergone separate DNA rearrangements leading to expansion of minisatellite-like sequences. J Mol Biol 216: 677–688
9. Dixon LK, Costa JV, Escribano JM, Rock DL, Vinuela E, Wilkinson PJ (2000) Family *Asfarviridae*. In: Van Regenmortel MHV, Fauquel CM, Bishop DHL (eds) Virus taxonomy, 7th Report of the ICTV, Academic Press, San Diego, pp 159–165
10. Dixon LK, Escribano JM, Martins C, Rock DL, Salas ML, Wilkinson PJ (2005) Asfarviridae. In: Fauquet CM, Mayo MA, Maniloff J, Desselberger U, Ball LA (eds), Virus taxonomy, VIIIth Report of the ICTV, Elsevier/Academic Press, London, pp 135–143
11. Genovesi EV, Knudsen RC, Whyard TC, Mebus CA (1988) Moderately virulent ASFV infection: blood cell changes and infective virus distribution among blood components. Am J Vet Res 49: 338–344
12. Gomez-Puertas P, Rodriguez F, Oviedo JM, Brun A, Alonso C, Escribano JM (1998) The African swine fever virus proteins p54 and p30 are involved in two distinct steps of virus attachment and both contribute to the antibody-mediated protective immune response. Virology 243: 461–471
13. Haresnape JM, Wilkinson PJ (1989) A study of African swine fever virus-infected ticks (Ornithodoros-Moubata) collected from 3 villages in the Asf enzootic area of Malawi following an outbreak of the disease in domestic pigs. Epidemiol Infect 102: 507–522
14. Irusta PM, Borca MV, Kutish GF, Lu Z, Caler E, Carrillo C, Rock DL (1996) Amino acid tandem repeats within a late viral gene define the central variable region of African swine fever virus. Virology 220: 20–27
15. Kay-Jackson PC, Goatley LC, Cox L, Miskin JE, Parkhouse RM, Wienands J, Dixon LK (2004) The CD2v protein of African swine fever virus interacts with the actin-binding adaptor protein SH3P7. J Gen Virol 85: 119–130
16. Kollnberger SD, Gutierrez-Castaneda B, Foster-Cuevas M, Corteyn A, Parkhouse RME (2002) Identification of the principal serological immunodeterminants of African swine fever virus by screening a virus cDNA library with antibody. J Gen Virol 83: 1331–1342
17. Lubisi BA, Bastos ADS, Dwarka RM, Vosloo W (2005) Molecular epidemiology of African swine fever in East Africa. Arch Virol 150: 2439–2452
18. Plowright WPJ, Pierce MA (1969) African swine fever in ticks (Ornithodoros moubata, Murray) collected from animal burrows in Tanzania. Nature 221: 1071–1073
19. Rodriguez F, Alcaraz C, Eiras A, Yanez RJ, Rodriguez JM, Alonso C, Rodriguez JF, Escribano JM (1994) Characterization and molecular-basis of heterogeneity of the African swine fever virus envelope protein P54. J Virol 68: 7244–7252
20. Sanchez Botija C (1963) Reservatoirios del virus de la peste porcina africana. Investigacion del virus de la PPA en los atropodos mediante prova de la hemadsorcion. Bull Off Int Epiz 60: 895–899
21. Sumption KJ, Hutchings GH, Wilkinson PJ, Dixon LK (1990) Variable regions on the genome of Malawi isolates of African swine fever virus. J Gen Virol 71: 2331–2340

22. Thomson GR (1985) The epidemiology of ASF: the role of free-living hosts in Africa. Onderstepoort J Vet Res 52: 201–209
23. Wesley RD, Tuthill AE (1984) Genome relatedness among African swine fever virus field isolates by restriction endonuclease analysis. Prev Vet Med 2: 53–62
24. Yanez RJ, Rodriguez JM, Nogal ML, Yuste L, Enriquez C, Rodriguez JF, Vinuela E (1995) Analysis of the complete nucleotide-sequence of African swine fever virus. Virology 208: 249–278
25. Yozawa T, Kutish GF, Afonso CL, Lu Z, Rock DL (1994) Two novel multigene families, 530 and 300, in the terminal variable regions of African swine fever virus genome. Virology 202: 997–1002

Author's address: Dr. Linda K. Dixon, Institute for Animal Health Pirbright Laboratory, Ash Road, Pirbright, Woking, Surrey GU24 0NF, U.K.; e-mail: Linda.dixon@bbsrc.ac.uk