Genetic and Phenotypic Analysis of Novel South African Avian Poxviruses

OLIVIA CARULEI

Supervisors: Prof. Anna-Lise Williamson and Dr Nicola Douglass

Thesis presented for the degree of Doctor of Philosophy

in the Division of Medical Virology, Department of Pathology,

Faculty of Health Sciences, University of Cape Town.

July 2018

TABLE OF CONTENTS

i. DECLARATION 5
ii. ACKNOWLEDGEMENTS 6
iii. ABSTRACT 7
iiii. LIST OF FIGURES AND TABLES 8
iv. ABBREVIATIONS 10

CHAPTER 1: LITERATURE REVIEW 14
1.1 POXVIRUSES 15
1.1.1 Classification and Phylogenetics 15
1.1.2 Genome 25
1.1.3 Virion Structure and Life Cycle 26
1.2 AVIAN POXVIRUSES 29
1.2.1 Classification and Phylogenetics 29
1.2.2 Transmission/Pathology 34
1.2.3 Prevention and Control 35
1.2.4 Growth Characteristics 36
1.2.5 Avian Poxvirus Genome Organisation and Statistics 37
1.2.5.1 FP9 vs FWPV US 38
1.2.5.2 FWPV vs CNPV 41
1.2.5.3 FeP2 Genome 43
1.2.5.4 TKPV Genome 43
1.2.5.5 SWPV-1 and SWPV-2 Genomes 44
1.2.6 South African Avian Poxviruses 44
1.2.6.1 Penguinpox Virus (PEPV) 45
1.2.6.2 Flamingopox Virus (FGPV) 46
1.2.6.3 Pigeonpox Virus (PGPV) 47
1.3 PROJECT RATIONALE 47

CHAPTER 2: HISTOLOGIC AND PHYLOGENETIC ANALYSIS OF SOUTH AFRICAN AVIAN POXVIRUSES 49
2.1 INTRODUCTION 50
2.2 MATERIALS and METHODS 51
2.2.1 Virus Acquisition and Isolation 51
2.2.1.2 PEPV 51
2.2.1.3 FGPV 51
2.2.1.4 Pi1 52

2.2.1.5 Pi2 and Pi3 52
2.2.1.6 Pi4 52
2.2.1.7 Virus Isolation from Scab Material 53
2.2.1.8 Growth of Viruses on Cams 53
2.2.2 Histology 54
2.2.3 DNA Extraction 54
2.2.4 Polymerase Chain Reaction 55
2.2.5 Phylogenetic Analysis 55
2.3 RESULTS 56
2.3.1 Pock Morphology and Membrane Histology 56
2.3.2 Phylogenetic Analysis 60
2.4 DISCUSSION 66

CHAPTER 3: THE GENOME SEQUENCES OF PENGUINPOX VIRUS AND FLAMINGOPOX VIRUS 69
3.1 INTRODUCTION 70
3.2 MATERIALS and METHODS 70
3.2.1 DNA extraction – see chapter 2.2 70
3.2.2 Ion torrent sequencing and bioinformatics 70
3.3 RESULTS 72
3.3.1 Sequencing Statistics 72
3.3.2 Basic Genome Structure and Statistics 72
3.3.3 Core/Conserved Genes 109
3.3.4 Multigene Families 112
3.3.5 Gene Translocations/Duplications 112
3.3.6 Gene Disruptions/Deletions 113
3.3.7 Gene Insertions 114
3.3.8 NOVEL ORFs and Genes of Interest 114

CHAPTER 4: COMPARATIVE GENOMICS 121
4.1 INTRODUCTION 122
4.2 METHODS 122
4.3 RESULTS 123
4.3.1 Dotplot Analysis 123
4.3.2 Regions of Difference 123
4.4 DISCUSSION 139
4.5 CONCLUSIONS 142

REAGENTS 143
REFERENCES 145

DECLARATION

I, **Olivia Carulei,** hereby declare that the work on which this dissertation/thesis is based is my original work (except where acknowledgements indicate otherwise) and that neither the whole nor any part of it has been, is being, or is to be submitted for another degree in this or any other university.

I empower the university to reproduce for the purpose of research either the whole or any portion of the contents in any manner whatsoever.

Signature: Signed by candidate Date: 23/07/2018

ACKNOWLEDGEMENTS

I would like to thank my family and friends for their on-going support and encouragement over the years. I would also like to thank my colleagues in the Department of Pathology and Division of Virology for providing a warm and stimulating environment in which to learn and grow. Lastly, I would like to thank my supervisors, Prof. Anna-Lise Williamson and Dr. Niki Douglass – I really couldn't have done this without you.

ABSTRACT

Avian poxviruses are important pathogens of both wild and domestic birds and exhibit a large degree of intragenus diversity at a genomic level. These viruses are known to differ in growth characteristics (*in vitro* and *in vivo)*, virulence, and cross-protection, with little known about the genomic contributions to these differences. Only six isolates from subclades A and B and one from proposed subclade E have had their genomes completely sequenced. These genomes have been shown to exhibit typical poxvirus genome characteristics with conserved central regions and more variable terminal regions, however all isolates exhibit major differences in defined central regions. This study aimed to analyze and characterize novel isolates from South Africa in terms of growth characteristics and phylogenetic relationships. It also added to the pool of genome sequences available for comparative genomic analyses to further investigate genome architecture.

Poxvirus isolates from lesser flamingo (*Phoenicopterus minor*) and African penguin (*Spheniscus demersus*) were chosen for analysis from a larger pool of donated isolates by comparison of macroscopic growth characteristics on chorioallantoic membranes, membrane histology and phylogenetic analyses based on nucleotide alignment of partial P4b sequences. Flamingopox virus was shown to group in subclade A3, induce membrane thickening and mesodermal hyperplasia while Penguinpox virus grouped in subclade A2, and did not induce membrane thickening or hyperplasia. The genomes of the above isolates were sequenced and compared to other available avipoxvirus genomes. Dotplot comparisons revealed major differences in central regions that have traditionally been thought to be conserved. Further analysis revealed five regions of difference, of varying lengths, spread across the central regions of the various genomes. Although individual gene identities at the nucleotide level did not vary greatly, gene content and synteny between isolates/species at these identified regions were far more divergent than expected. The reasons for these large genomic rearrangements are yet to be elucidated and will need to be considered in future phylogenetic studies and vaccine vector design. Sequencing and analysis of further avian poxvirus genomes will help characterize this complex genus of poxviruses.

LIST OF FIGURES AND TABLES

FIGURES

Figure 1.1 Schematic of poxvirus morphogenesis

Figure 1.2 Schematic of the clades and subclades of the Avipoxvirus genus

Figure 1.3 Schematic depicting the positions, functions and major gene families of FWPV US open reading frames.

Figure 1.4 Schematic showing alignment of the FWPV and CNPV genomes

Figure 1.5 Poxvirus lesion near the eye of a juvenile African penguin

Figure 2.1 Poxvirus lesions near the tibiotarsal joint of a juvenile Lesser Flamingo

Figure 2.2 Poxvirus lesions on the heads of two juvenile racing pigeons

Figure 2.3 Phylogenetic tree based on the nucleotide alignment of partial P4b sequences of the five novel avian poxviruses and representative chordopoxviruses

Figure 2.4 Phylogenetic tree based on the nucleotide alignment of partial P4b sequences of clade A2 and clade A3 avian poxviruses

Figure 3.1 ORF schematic of the PEPV genome

Figure 3.2 ORF schematic of the FGPV genome

Figure 3.3 Gene synteny analysis of the location of orthologues of IL-10 in eight avian poxvirus genomes

Figure 4.1 Dotplots of the PEPV genome (x axis) vs other sequenced avian genomes and VACV (y axis)

Figure 4.2 Dotplots of the FGPV genome (x axis) vs other sequenced avian genomes and VACV (y axis)

Figure 4.3 Schematic of Table 4.1 in the horizontal orientation showing conserved regions and regions of difference

Figure 4.4 Schematic representing the ORFs present in the rearranged regions a) fpv114-126 and b) fpv146-165 in nine avian poxvirus genomes

Figure 4.5 Schematic representing the ORFs present in the rearranged regions a) fpv031-047, b) fpv058-077 and c) fpv193-211 in nine avian poxvirus genomes

Figure 4.6 Genome schematic of VACV and FWPV genomes showing four regions conserved in gene content and synteny and five regions of difference

TABLES

Table 1.1 Genera, species and hosts of the family *Chordopoxvirinae*

Table 1.2 Summary of phylogenetic analyses conducted on poxviruses

Table 1.3 Avian species identified with poxvirus infection

Table 1.4 Basic genome statistics of the seven fully sequenced avian poxvirus genomes

Table 2.1 Macroscopic and histologic comparison of CAMs infected with eight avian poxviruses

Table 2.2 List of 73 avian poxvirus isolates used in comparison of partial P4b nucleotide sequences

Table 2.3 Pairwise comparison of the % identity and number of differences between nucleotide alignments of partial P4b sequences of avian poxviruses

Table 3.1 Genome statistics of PEPV and FGPV compared with those of six avian poxvirus genomes

Table 3.2 PEPV open reading frames

Table 3.3 FGPV open reading frames

Table 3.4 ORFs of interest identified in the PEPV and FGPV genomes as well as their location in six other avian poxvirus genomes, if present.

Table 3.5 47 ORFs found to be conserved in each of the fully sequenced avian poxvirus genomes

Table 3.6 Pairwise comparison of the % identity and number of differences between nucleotide alignments of 130 conserved genes in eight avian poxvirus genomes

Table 3.7 Copy number of genes in each of the 14 multi-gene families identified in eight avian poxvirus genomes

Table 4.1 ORF synteny comparison of 8 avipoxviruses and VACV

Table 4.2 Size comparison of five regions of difference in eight avian poxvirus genomes

ABBREVIATIONS

AT	adenine + thymine
aa	amino acid
ALVAC	canarypox virus vaccine strain
APV	avipoxvirus
ATCC	American Type Culture Collection
BHK-21	Syrian baby hamster kidney cell
bp	base pairs
BPSV	bovine papular stomatitis virus
BSR	bootstrap replicates
CAM	chorioallantoic membrane
CED	chick embryo dermal cell
CEF	chick embryo fibroblast
CEV	cell associated enveloped virus
ChPV	chordopoxvirus
CNPV	canarypox virus
CPE	cytopathic effect
CP	cytopathogenic
CPV	cowpox virus
CRV	crocodylidpox virus
CV-1	simian kidney cell
DEC	duck embryo cell
DMEM	Dulbecco's modified Eagle's medium
DNA	deoxyribonucleic acid
EBTr	bovine embryonic tracheal cells
ECTV	ectromelia virus
EEV	extracellular enveloped virus
EFC	entry-fusion complex
EnPV	entomopoxvirus
EtOH	ethanol
EV	enveloped virus
EEV	extracellular enveloped virus
FCS	foetal calf serum

FeP2	feral pigeon isolate 2
Fp9	attenuated European strain of FWPV derived from HP1
FPVUS	pathogenic American strain of FWPV
FPV/FWPV	fowlpox virus
GPCR	G protein-coupled receptor
HA	haemagglutination
HCL	hydrogen chloride
HEF	human embryonic fibroblast
HeLa	human cervix adenocarcinoma cell
HGP	Hawaiian goosepox virus
HP1	pathogenic European strain of FWPV
IC	intracellular compartment
ICTV	International Committee on Taxonomy of Viruses
IEV	intracellular enveloped virus
IFN	interferon
Ig	immunoglobulin
IL	interleukin
IMV	intracellular mature virion
ISG	interferon stimulated gene
ITR	inverted terminal repeat
IV	intracellular virion
kb	kilobase
kbp	kilobase pairs
kDa	kilodalton
LMH	chicken hepatocellular carcinoma cell line
LSDV	lumpy skin disease virus
MDBK	Madin Darby bovine kidney cells
MHC	major histocompatibility complex
ML	maximum likelihood
MOCV/MCV	molluscum contagiosum virus
m.o.i.	multiplicity of infection
MRC-5	medical research council cell strain 5 (human diploid)
mRNA	messenger ribonucleic acid
MV	mature virion

MVA	modified vaccinia Ankara
MPXV	monkeypox virus
MYXV	myxoma virus
NCLDV	nucleocytoplasmic large DNA virus
NCP	non-cytopathogenic
NJ	neighbour-joining
nt	nucleotide
NYVAC	attenuated strain of VACV Copenhagen
NZ	New Zealand
OPV	orthopoxvirus
ORF	open reading frame
ORFV	Orf virus
PASC	pair-wise sequence comparison
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PEPV	penguinpox virus
PGPV	pigeonpox virus
PKR	protein kinase RNA activated
PRPV	parrotpox virus
PS	penicillin, streptomycin
PSF	penicillin, streptomycin, fungin
QPV	quailpox virus
QT-35	Japanese quail fibrosarcoma cell line
REV	reticuloendotheliosis virus
RK-13	rabbit kidney cell
RNA	ribonucleic acid
ROD	region of difference
RR	ribonucleotide reductase
SA	South Africa
SANCCOB	Southern African Foundation for the Conservation of Coastal Birds
SEM	scanning electron microscopy
SFV	Shope fibroma virus
SLPV	starlingpox virus

SPF	specific pathogen free
SPPV/SPV	sheeppox virus
SWPV	swinepox virus
SWPV1/SWPV2	shearwater poxvirus
TANV	tanapox virus
TEM	transmission electron microscopy
TF	transcription factor
TK	thymidine kinase
TKPV	turkeypox virus
TNF	tumour necrosis factor
VACV/VV	vaccinia virus
VARV/VAR	variola virus
Vero	African green monkey kidney cell line
WV	wrapped virion
YLDV	Yaba-like disease virus

CHAPTER 1: LITERATURE REVIEW

1.1 POXVIRUSES

- 1.1.1 Classification and Phylogenetics
- 1.1.2 Genome
- 1.1.3 Virion Structure and Life Cycle
- 1.1.4 Host Range and Immunomodulation

1.2 AVIAN POXVIRUSES

- 1.2.1 Classification
- 1.2.2 Pathology/Clinical signs/symptoms
- 1.2.3 Prevention/Control
- 1.2.4 Growth Characteristics
- 1.2.5 Genomes
  - 1.2.5.1 FP9 vs FWPV
  - 1.2.5.2 FWPV vs CNPV
  - 1.2.5.3 FEP2 Genome
  - 1.2.5.4 TKPV Genome
  - 1.2.5.5 SWPV-1 AND SWPV-2 Genomes
- 1.2.6 South African Avian Poxviruses
  - 1.2.6.1 Penguinpox virus
  - 1.2.6.2 Flamingopox virus
  - 1.2.6.3 Pigeonpox virus

1.3 PROJECT RATIONALE

POXVIRUSES

1.1.1 Classification and Phylogenetics

Bioinformatic analysis of conserved protein sequences shows that poxviruses are one of seven confirmed (*Ascoviridae, Asfarviridae*, *Iridoviridae*, *Marseilleviridae*, *Mimiviridae* and *Phycodnaviridae)* (Iyer *et al*., 2006, 2001) and four unconfirmed families (Pandoraviridae (Philippe *et al*., 2013), Pithoviridae (Legendre *et al*., 2014), Faustoviridae (Reteno *et al*., 2015), and Molliviridae (Legendre *et al*., 2015) of the monophyletic nucleocytoplasmic large DNA viruses (NCLDV), in the proposed order "Megavirales" (Colson *et al*., 2013)

The *Poxviridae* are a family of large, complex, DNA viruses divided into two subfamilies based on host-range: the *Entomopoxvirinae* (EnPV), which infect insects and the *Chordopoxvirinae* (ChPV), which infect chordates. The EnPVs are further divided into three genera: *Alphaentomopoxvirus, Betaentomopoxvirus and Gammaentomopoxvirus* (ICTV), while the ChPV, the focus of this review, are composed of 11 genera: *Avipoxvirus, Capripoxvirus, Centapoxvirus, Cervidpoxvirus, Crocodylidpoxvirus, Leporipoxvirus, Molluscipoxvirus, Orthopoxvirus, Parapoxvirus, Suipoxvirus,* and *Yatapoxvirus* (Table 1.1). Several other viruses have been described (Cotia poxvirus (Afonso *et al*., 2012; Ueda *et al*., 1978), Salmon gill poxvirus (Gjessing *et al*., 2015), Squirrelpox virus (Thomas *et al*., 2003), Pteropox virus (O'Dea *et al*., 2016), Cetacean poxvirus (Barnett *et al*., 2015; Bracht *et al*., 2006), Embu virus (Ibrahim *et al*., 2014), Grey kangaroo poxvirus (McKenzie *et al*., 1979), Red kangaroo poxvirus (Bagnall and Wilson, 1974), Quokka poxvirus (Papadimitriou and Ashman, 1972), Marmoset poxvirus (Gough *et al*., 1982; Mätz-Rensing *et al*., 2006), Spectacled caiman poxvirus (Ramos *et al*., 2002), Sea Otter poxvirus (Tuomi *et al*., 2014), Eptesipox virus (Tu *et al*., 2017), but remain unclassified (ICTV 9th report, 2011). All species belonging to a particular genus are genetically and antigenically related, have similar morphology, and in most cases host-range, with the exception of the orthopoxviruses which naturally infect several distantly related mammals (Pastoret and Vanderplasschen, 2003)

Table 1.1: Genera, species and hosts of the family *Chordopoxvirinae*.

GENUS	SPECIES	PRIMARY HOST
Avipoxvirus	*Canarypox virus*	Canary
	Crowpox virus	Crow
	Fowlpox virus	Chicken
	Juncopox virus	Junco
	Mynahpox virus	Mynah
	Parrotpox virus	Parrot
	Peacockpox virus	Peafowl
	Penguinpox virus	Penguin
	Pigeonpox virus	Pigeon
	Quailpox virus	Japanese quail
	Shearwaterpox virus	Pacific Shearwater
	Sparrowpox virus	Sparrow
	Starlingpox virus	Starling
	Turkeypox virus	Turkey
Capripoxvirus	*Goatpox virus*	Goat
	Lumpy skin disease virus	Cattle
	Sheeppox virus	Sheep
Centapoxvirus	***Yokapox virus***	Unknown
Cervidpoxvirus	***Mule deerpox virus***	Mule deer
Crocodylid-poxvirus	***Nile crocodile poxvirus***	Nile Crocodile
Leporipoxvirus	*Hare fibroma virus*	Hare
	Myxoma virus	Rabbit
	Rabbit fibroma virus	Rabbit
	Squirrel fibroma virus	Squirrel

GENUS	SPECIES	PRIMARY HOST
Molluscipox-virus	**Molluscum contagiousum virus**	Human
Orthopoxvirus	*Camelpox virus*	Camel
	Cowpox virus	Rodent
	Ectromelia virus	Mice
	Monkeypox virus	Squirrel, Rodent, Primate
	Raccoonpox virus	Raccoon
	Skunkpox virus	Skunk
	Taterapox virus	Rodent
	Uasin gishu disease virus	Horse
	Vaccinia virus	Unknown
	Variola virus	Human
	Volepox virus	Vole
Parapoxvirus	Auzduk disease virus	Unknown
	Bovine popular stomatitis virus	Cattle
	Camel contagious ecthyma virus	Camel
	Chamois contagious ecthyma virus	Sheep
	Orf virus	Sheep and Goat
	Parapoxvirus of red deer in N.Z.	Red Deer
	Pseudocowpox virus	Cattle
	Sealpox virus	Seal
Suipoxvirus	***Swinepox virus***	Swine
Yatapoxvirus	*Tanapox virus*	Unknown
	Yaba monkey tumor virus	Primate

Bold and italic = type species; *italic* = official species.

Several phylogenetic studies have been conducted including different species from different genera, using different loci for analysis as well as different alignment and tree building methods (Table 1.2). (Upton *et al*., 2003) analysed all completely sequenced poxvirus genomes available at the time (n=21) and found that 49 genes located in the central 100kb region of the genome are completely conserved in poxviruses. They are thus proposed to be essential in the poxvirus life cycle and to comprise the minimum essential genome. A further 41 gene families were found to be conserved in ChPVs bringing the total number of conserved genes in ChPVs to 90. These findings were later updated with the reports of the genome sequences of orf virus, strain IA82 (Delhon *et al*., 2004) and deerpox virus, strain W83 (Afonso *et al*., 2005). The D9R (fpv054; cnpv076)) and F15L (fpv105; cnpv132) loci were removed from the list and the A30.5L (fpv194.5; cnpv268.5) locus added (Lefkowitz *et al*., 2006). F16L was added to the list in error as it is not present in avipoxvirus genomes (Senkevich *et al*., 2011). Of the 89 conserved gene families, 33 were found to be responsible for DNA replication and transcription, 29 were associated with virion assembly, morphogenesis and egress, three were involved in both transcription and morphogenesis and 24 had no functional characterisation. In addition, the promoters and 5´-untranslated regions of the above genes were also found to be highly conserved (Lefkowitz *et al*., 2006). No gene in the terminal regions of any virus studied was conserved amongst ChPVs as these genes are mostly concerned with virulence and host range and are therefore species or host-specific (Upton *et al*., 2003).

The proteins encoded by the conserved genes have been the focus of several phylogenetic studies as alignment and analysis of multiple, conserved, concatenated ORFs have been shown to be more accurate for tree building over alignment of single loci because sampling of multiple loci more accurately represents the entire genome and because of the increase in number of phylogenetically informative sites (Nichols, 2001). It must be noted that a viral isolate from farmed Atlantic salmon, which was shown to form the deepest branch of the *Chordopoxvirinae* subfamily by phylogenomic analyses, was shown to be missing several of the genes found to be conserved above (Gjessing *et al*., 2015). The missing genes were mostly involved in viral membrane biogenesis and virus-host interactions.

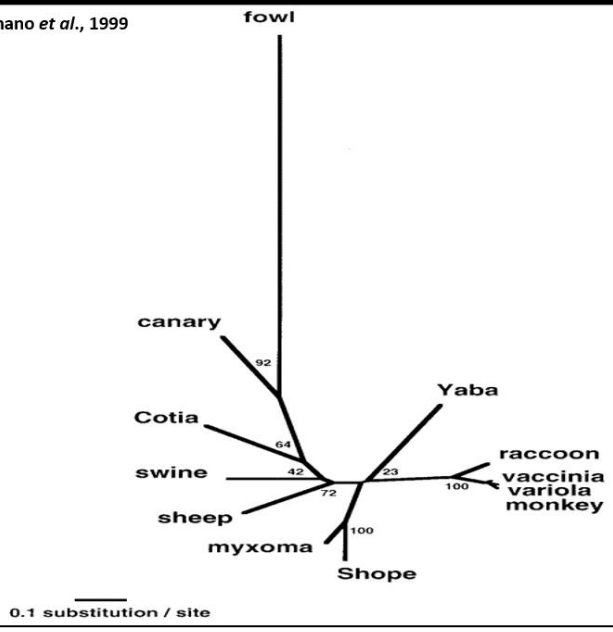
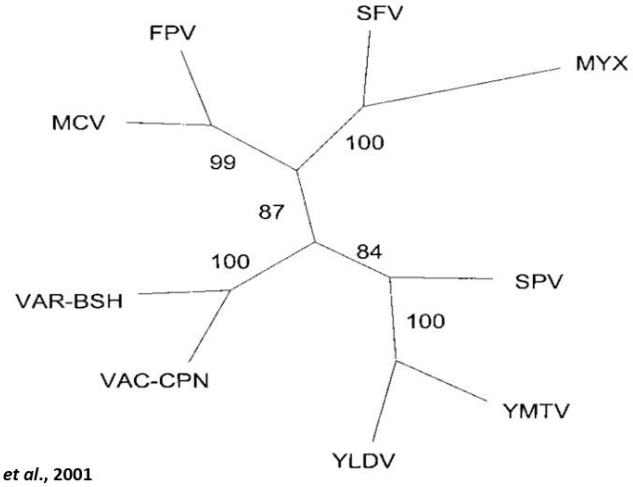
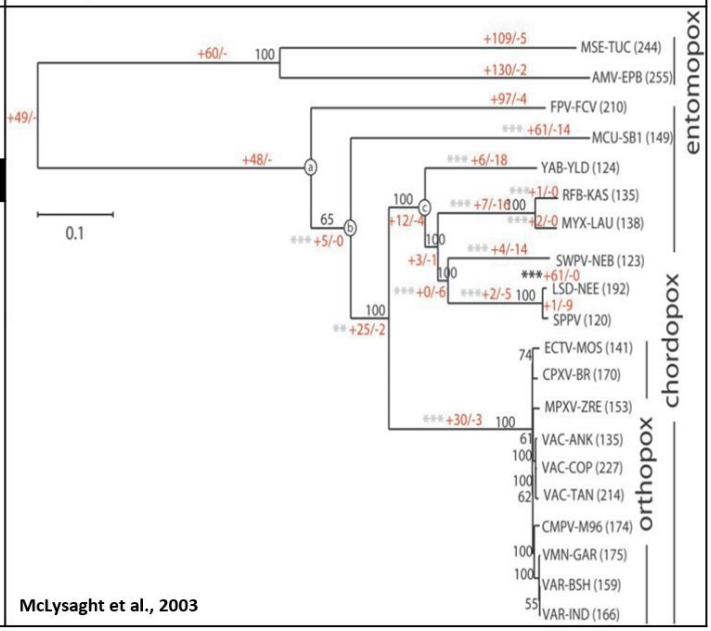
It has been shown however, that concatenation and alignment of a specific subset of the 89 conserved proteins provide a more robust starting point for phylogenetic analysis. This is especially true for analysis of distantly related species containing highly divergent orthologues that result in long-branch attraction (LBA) (Felsenstein, 1978). LBA is a systematic error in the clustering of organisms and does not reflect true evolutionary relationships. This type of error can occur when two or more outgroup organisms mistakenly group together due to convergent evolution rather than common ancestry. LBA can be addressed by specific selection of highly conserved orthologues for analysis and/or inclusion of additional taxa and/or usage of multiple substitution models (Hillis, 1996). Lefkowitz *et al*., 2006 chose a subset of 35 of the 90 conserved proteins in all ChPVs on the basis that they were relatively highly conserved among all species and had similar nucleotide substitution rates. The proteins used are listed in Table 1.2 and include 28 proteins involved in transcription, four in DNA replication and three in the S-S bond formation pathway. Although this method does not result in a significantly different tree topology to those studies using the full complement of conserved proteins, bootstrap values are improved over single gene analyses and with the addition of further divergent species, e.g., crocodilepox, to the family, it will become more important to address long-branch attraction to ensure accurate tree building (Anderson and Swofford, 2004)

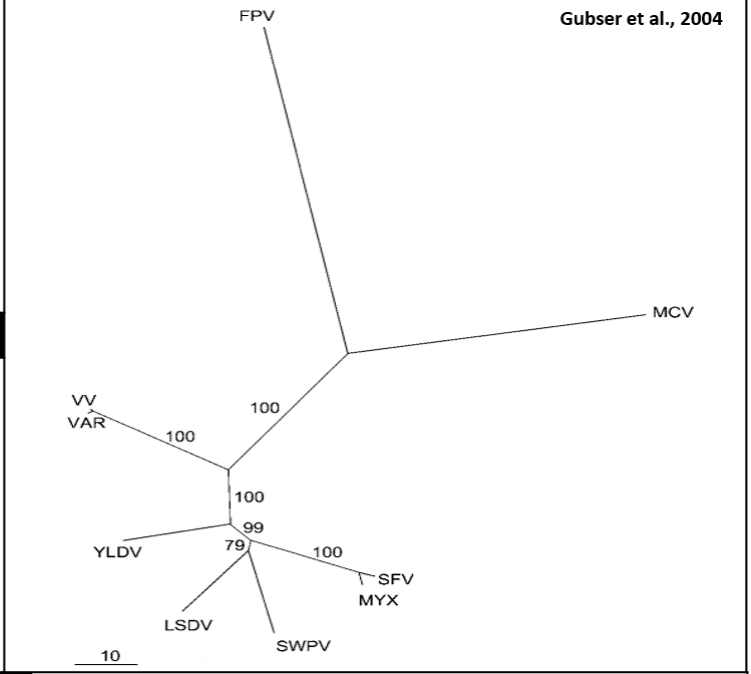
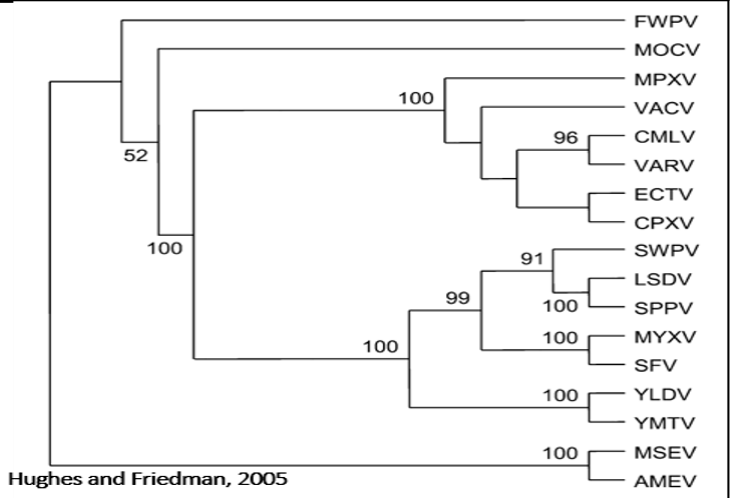
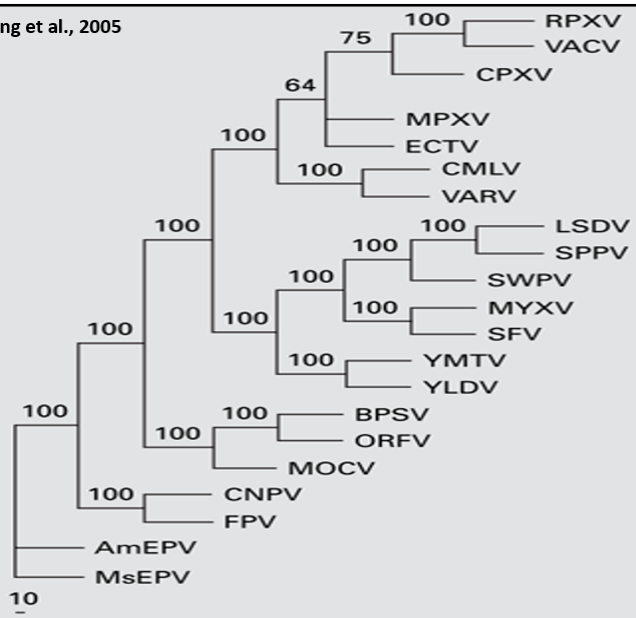
In general, within the *Chordopoxvirinae* subfamily, the avipoxviruses (APVs), crocodilepoxviruses (CRVs), molluscipoxvirus (MCVs), and parapoxviruses (PPVs) have been found to be most divergent. Pair-wise sequence comparison (PASC) analysis of the DNA polymerase gene (VACV E9) in an intergenus comparison, showed that the APVs and PPVs exhibit the lowest levels of amino acid identity (40-45%) at this locus (Lefkowitz *et al*., 2006). In contrast, the orthopoxviurses (OPVs), yatapoxviruses (YPVs), leporipoxviruses (LPVs), suipoxviruses (SPVs), capripoxviruses (CaPVs) and cervidpoxviruses (CePV) showed the highest levels of intergenus identity (65-75%). Comparison of the aforementioned genera with the APVs, PPVs and MOCV reveals intermediate levels of identity (45-55%). In one analysis of molecular dating of ChPVs, members of the GC rich genera (>60% GC; parapoxviruses, molluscipoxviruses and crocodilepox virus) and members of the AT rich chordopoxvirus genera (>60% AT; all remaining genera) were shown to diverge

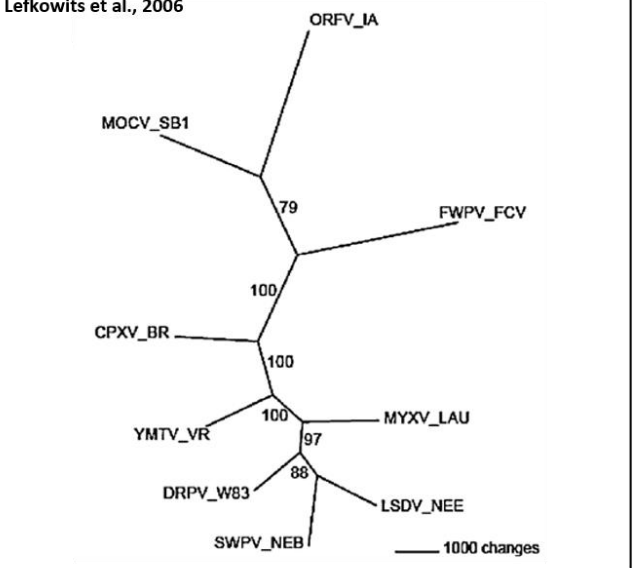
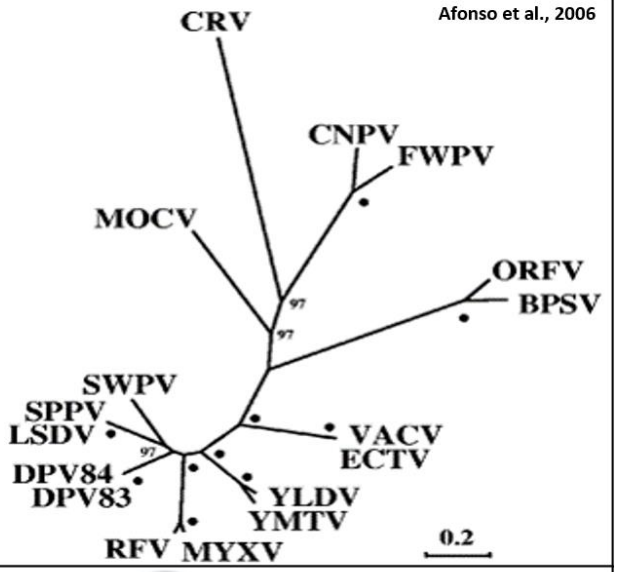
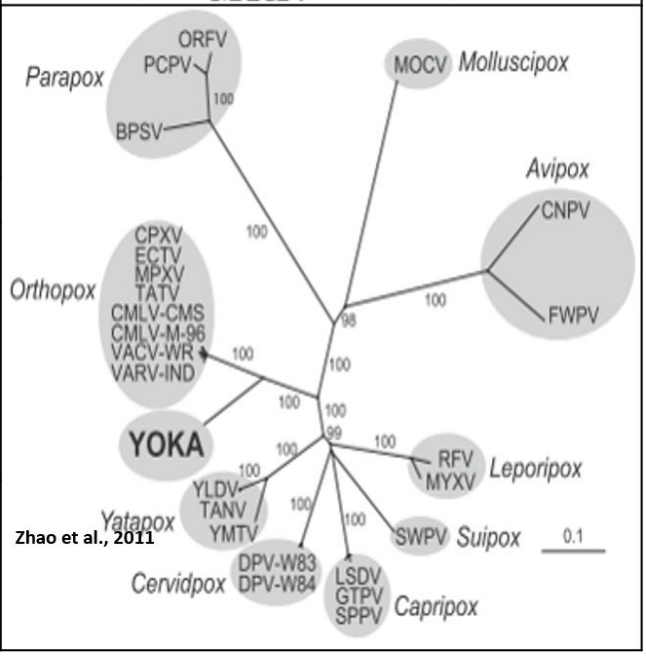
approximately 500 thousand years ago. The avipoxviruses (FWPV and CNPV) then diverged from mammalian viruses approximately 420 thousand years ago with the OPVs diverging around 300 thousand years ago (Babkin and Shchelkunov, 2006). This study was updated in 2011 with APVs diverging ~249 thousand years ago and the OPVs, ~166 thousand years ago (Babkin and Babkina, 2011). Phylogenetic studies have further revealed that the phylogeny of the chordopoxviruses does not follow the phylogeny of the host, indicating that these viruses have changed hosts several times in their history (McLysaght *et al*., 2003).

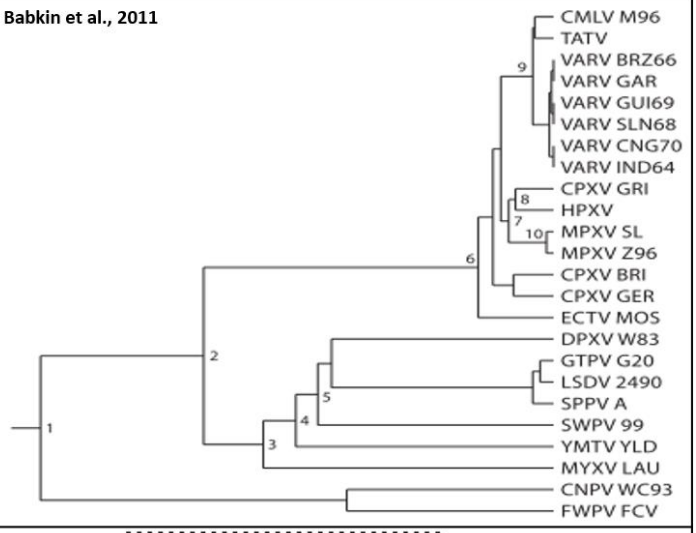
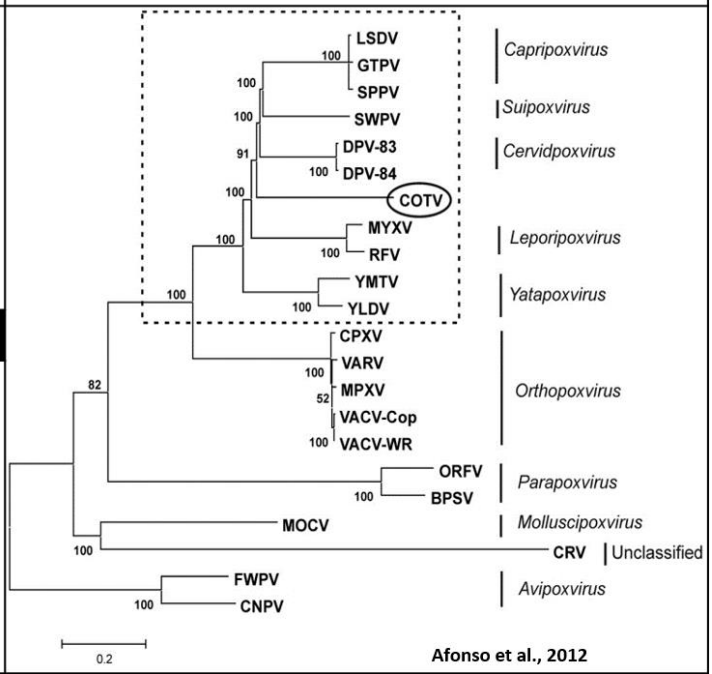
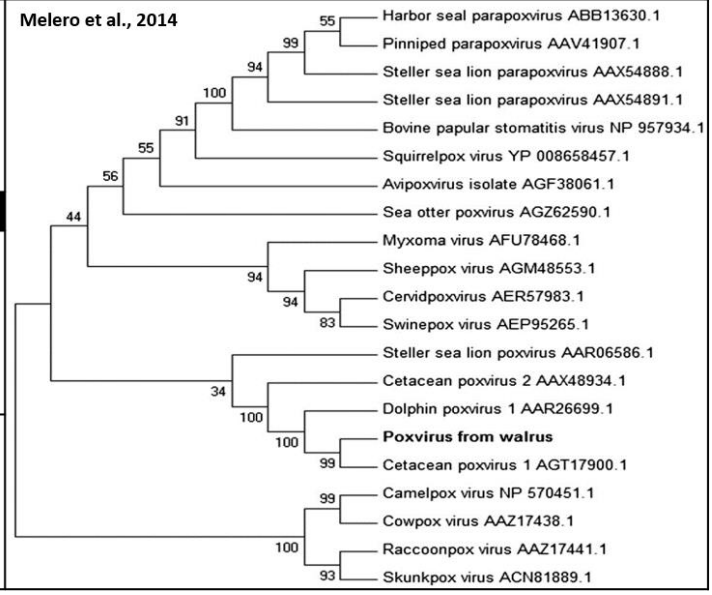
Gene content and synteny have also been used to help elucidate evolutionary relationships between poxviruses. Alignment and comparison of the central genomic regions of viruses from eight ChPV genera (using the VV genome as the reference) showed that MCV and FPV were most divergent, encoding 40 and 33 unique genes respectively in their central genomic regions. In contrast, MYX YLDV, LSDV and SWPV contained three or less unique genes in this region. In terms of gene order, FPV also showed major differences with large blocks of genes being translocated and/or inverted compared to VV and the other ChPVs (Gubser *et al*., 2004a). These findings are in accordance with what has been shown through phylogenetic analysis.

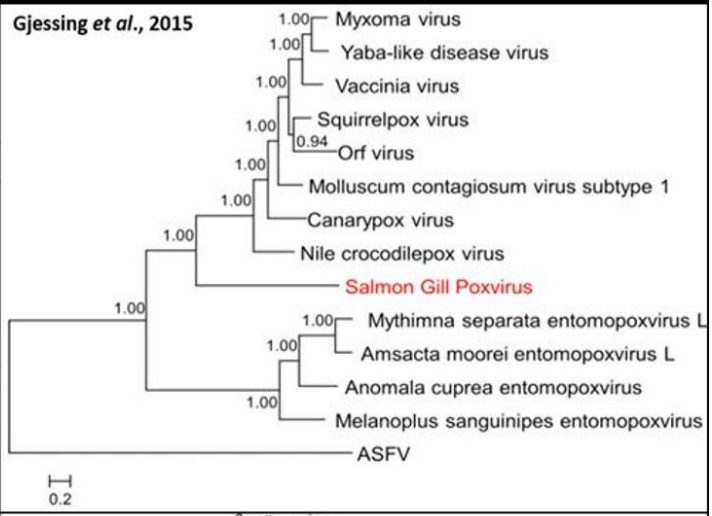
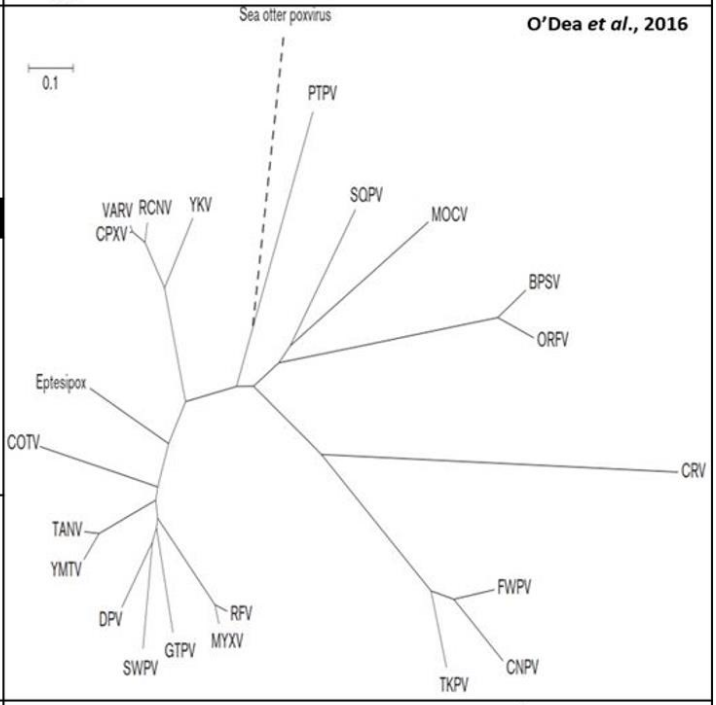
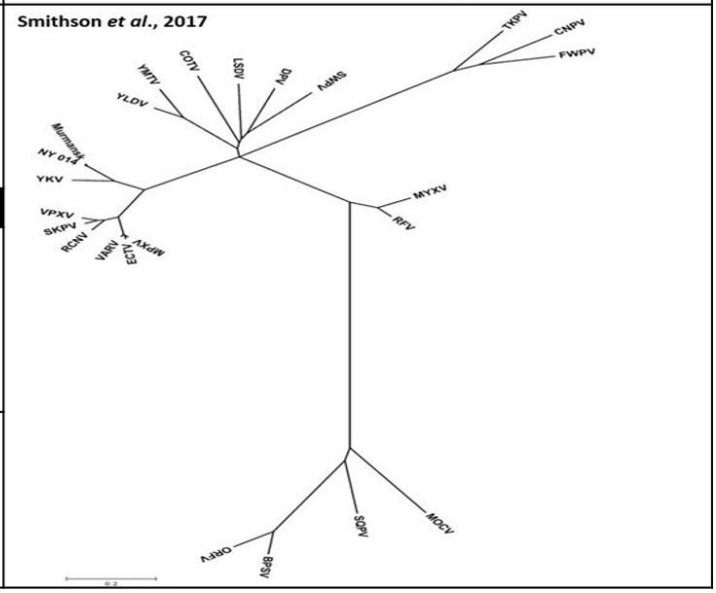
Table 1.2: Summary of phylogenetic analyses conducted on poxviruses

LOCI	METHOD	RESULT
Thymidine Kinase	NJ tree (Kimura 2) 1000 B.S.R. **GENERA** *Avipoxvirus* *Capripoxvirus* *Cotiapoxvirus* *Leporipoxvirus* *Orthopoxvirus* *Suipoxvirus* *Yatapoxvirus*	**Amano *et al.*, 1999** 
D12L	ClustalW Unrooted tree **GENERA** *Avipoxvirus* *Capripoxvirus* *Leporipoxvirus* *Molluscipoxvirus* *Orthopoxvirus* *Yatapoxvirus*	 **Lee *et al.*, 2001**
34 aa	T-COFFEE NJ tree (Poisson correction) **GENERA** *Avipoxvirus* *Capripoxvirus* *Entomopoxvirus* *Leporipoxvirus* *Molluscipoxvirus* *Orthopoxvirus* *Suipoxvirus* *Yatapoxvirus*	 **McLysaght et al., 2003**

LOCI	METHOD	RESULT
E9L I7L I8R G9R J3R J6R H2R H6R D1R D5R D6R D11L D13L A7L A16L A24R	ClustalW NJ trees as starting trees for ML trees (GTR model with gamma distribution) 1000 B.S.R. **GENERA** *Avipoxvirus* *Capripoxvirus* *Leporipoxvirus* *Molluscipoxvirus* *Orthopoxvirus* *Suipoxvirus* *Yatapoxvirus*	**Gubser et al., 2004** 
VARV 50 VARV 62 VARV 122 VARV 96 VARV 101 VARV 120 VARV 89 VARV 139 VARV 126 VARV 85 VARV 135 VARV 82 VARV 51	ClustalW MP tree 1000 B.S.R. **GENERA** *Avipoxvirus* *Capripoxvirus* *Entomopoxvirus* *Leporipoxvirus* *Molluscipoxvirus* *Orthopoxvirus* *Suipoxvirus* *Yatapoxvirus*	 Hughes and Friedman, 2005
49aa	ClustalW NJ tree 1000 B.S.R. **GENERA** *Avipoxvirus* *Capripoxvirus* *Entomopoxvirus* *Leporipoxvirus* *Molluscipoxvirus* *Orthopoxvirus* *Parapoxvirus* *Suipoxvirus* *Yatapoxvirus*	**Xing et al., 2005** 

LOCI		METHOD	RESULT
E1 E4 E9 E10 I8 G2 G4 G5.5 G8 H1 H4-6 D1 D4-7 D11 D12	A1 A2 A2.5 A5 A7 A8 A18 A20 A23 A24 A29 F10 J3 J4 J6	MP tree **GENERA** *Avipoxvirus* *Capripoxvirus* *Cervidpoxvirus* *Leporipoxvirus* *Molluscipoxvirus* *Orthopoxvirus* *Parapoxvirus* *Suipoxvirus* *Yatapoxvirus*	**Lefkowits et al., 2006** 
83 aa		MUSCLE ML tree (WAG correction) 1000 B.S.R. **GENERA** *Avipoxvirus* *Capripoxvirus* *Cervidpoxvirus* *Crocodylidpoxvirus* *Leporipoxvirus* *Molluscipoxvirus* *Orthopoxvirus* *Parapoxvirus* *Suipoxvirus* *Yatapoxvirus*	**Afonso et al., 2006** 
E1 E4 E9 E10 I8 G2 G4 G5.5 G8 H1 H4-6 D1 D4-7 D11 D12	A1 A2 A2.5 A5 A7 A8 A18 A20 A23 A24 A29 F10 J3 J4 J6	ClustalW MP tree 100 B.S.R. **GENERA** *Avipoxvirus* *Capripoxvirus* *Cervidpoxvirus* *Leporipoxvirus* *Molluscipoxvirus* *Orthopoxvirus* *Parapoxvirus* *Suipoxvirus* *Yatapoxvirus* Yokapoxvirus	 **Zhao et al., 2011**

LOCI		METHOD	RESULT
E1 E4 E9 E10 I8 G2 G4 G5.5 G8 H1 H4-6 D1 D4-7 D11 D12	A1 A2 A2.5 A5 A7 A8 A18 A20 A23 A24 A29 F10 J3 J4 J6	ClustalX ML tree 100 B.S.R. **GENERA** *Avipoxvirus* *Capripoxvirus* *Cervidpoxvirus* *Leporipoxvirus* *Molluscipoxvirus* *Orthopoxvirus* *Parapoxvirus* *Suipoxvirus* *Yatapoxvirus*	**Babkin et al., 2011** 
90 aa		ClustalX and MUSCLE NJ tree (JTT substitution model) 2500 B.S.R. **GENERA** *Avipoxvirus* *Capripoxvirus* *Cervidpoxvirus* *Leporipoxvirus* *Molluscipoxvirus* *Orthopoxvirus* *Parapoxvirus* *Suipoxvirus* *Yatapoxvirus*	 **Afonso et al., 2012**
DNA Pol		MP tree (SPR algorithm) 500 B.S.R. **GENERA** *Avipoxvirus* *Capripoxvirus* *Cervidpoxvirus* *Leporipoxvirus* *Orthopoxvirus* *Parapoxvirus* *Suipoxvirus* Squirrelpoxvirus	**Melero et al., 2014** 

LOCI	METHOD	RESULT
A1 A2 A7 A18 A24 A32 D5 D6 D13 E9 I8 J6	MUSCLE aa ML tree **GENERA** *Avipoxvirus* *Crocodylidpoxvirus* *Leporipoxvirus* *Molluscipoxvirus* *Orthopoxvirus* *Parapoxvirus* *Yatapoxvirus* Salmon gill poxvirus Squirrelpoxvirus	**Gjessing *et al.*, 2015** 
A7 A10 A24 D1 D5 J6 H4	MAFFT aa ML tree (LG model with rate parameter as Gamma distributed with invariant sites (G+I) **GENERA** *Avipoxvirus* *Capripoxvirus* *Cervidpoxvirus* *Crocodylidpoxvirus* *Leporipoxvirus* *Molluscipoxvirus* *Orthopoxvirus* *Parapoxvirus* *Suipoxvirus* *Yatapoxvirus* Eptesipoxvirus Cotiapoxvirus Pteropoxvirus Raccoonpoxvirus Sea otter poxvirus Squirrelpoxvirus Yokapoxvirus	**O'Dea *et al.*, 2016** 
VACV-Cop-A2L to VACV-Cop-A24R (25kbp)	ClustalO and MAFFT ML tree 500 B.S.R. **GENERA** *Avipoxvirus* *Capripoxvirus* *Cervidpoxvirus* *Leporipoxvirus* *Orthopoxvirus* *Parapoxvirus* *Suipoxvirus* Squirrelpoxvirus	**Smithson *et al.*, 2017** 

1.1.2 Genome

Poxviruses have linear, double stranded, DNA genomes, which vary in size from ~130kbp in parapoxviruses (Delhon *et al*., 2004) to well over 300kbp in some avipoxviruses (Tulman *et al*., 2004). These viruses replicate solely in the cytoplasm of host cells by encoding the genes needed for replication, in contrast with most other DNA viruses, which replicate in the nucleus using the host replication machinery. The adenine-thymine (AT) content varies both between genera (from 36% in parapoxviruses to 75% in capripoxviruses) and between genes of a particular genome (from 18-61% for genes in the orf virus strain NZ2) (Lefkowitz *et al*., 2006). Poxvirus genomes have inverted terminal repeats (ITRs) at each end consisting of identical but oppositely oriented sequences ranging in size from 0.5kb in some strains of VARV (Massung *et al*., 1995) to almost 19kb in raccoonpox virus (Fleischauer *et al*., 2015) and the two DNA strands are covalently linked resulting in hairpin termini (Baroudy *et al*., 1982) All sequenced ChPVs, except VARV, have one or more genes in the ITR resulting in diploidy at these loci. Also present in the ITRs of most poxvirus genomes are tandemly repeated sequences which vary both in length and copy number (Campbell *et al*., 1989; Wittek and Moss, 1980) Given that the average size of proteins encoded by poxviruses is 30kDa and that the mRNAs remain unspliced, the number of genes in a poxvirus genome can be approximated to 1 gene per 1kb of genome with an average of ±200 genes per genome (Lefkowitz *et al*., 2006).

Genes are spatially organised into groups, which are transcribed in the same direction. Genes near the centre of the genome tend to be transcribed towards the centre, whereas most of the genes toward the ends of the genome are transcribed towards their respective ends. This genetic organisation most likely functions to prevent transcriptional complexes colliding when transcription of multiple, proximal genes is occurring simultaneously (Lefkowitz *et al*., 2006).

The core region forms a continuous block in all ChPV genomes except parapoxviruses and avipoxviruses due to various genome rearrangements. The core region of FWPV for example has broken into four segments, which are now interspersed amongst the variable regions (Gubser *et al*., 2004a), which is in

accordance with the external placement of FWPV in phylogenetic tree topologies. Although basic genome arrangement has been conserved amongst poxviruses, there has been substantial evolution in gene content through extensive gene duplication, gain (retention) and loss (A. McLysaght *et al*., 2003). Many of the gene gain events are the result of horizontal transfer events from other viruses (Dall *et al*., 2001) and from their eukaryotic hosts (Hughes and Friedman, 2005; Bratke and McLysaght, 2008; and Odom *et al*., 2009) with the rate of gene acquisition and/or retention increased in the *Orthopoxvirus* genus relative to other genera. Genes acquired by horizontal transfer are commonly involved in immunomodulation and nucleic acid metabolism and confer a replication advantage. Gene loss varies among gene families with no obvious relationship between the extent of gene loss of any particular family and the dispensability of that gene, although the types of genes that are gained and lost in the process of poxvirus evolution are likely to have host-specific effects (McLysaght *et al*., 2003).

1.1.3 Virion Structure and Life Cycle – review literature

Poxvirus virions are large, and barrel shaped with dimensions of ± 360 X 270 X 250 nm making them just visible by light microscopy with the principal components being protein, lipid and DNA (Moss, 2007). There are several different forms of mature virus, each of which contain a dumbbell shaped, nucleoprotein core containing the DNA genome, and various virally encoded factors and enzymes essential for early transcription, as well as two proteinacious lateral bodies, external to the core, containing at least three proteins implicated in immunomodulation (Schmidt *et al*., 2013).

The morphogenesis and life cycle of poxviruses are summarised in Figure 1 below. There are two distinct forms of infectious poxvirus virions: the mature virion (MV) with one membrane and the extracellular enveloped virion (EEV) with an additional membrane. EEV include enveloped virions (EV) and cell associated enveloped virions (CEV). The method of entry of MVs and EEVs into target cells has not been fully determined, although endosomal entry (Dales, 1963; Townsley *et al*., 2006a; Townsley and Moss, 2007) and plasma membrane fusion (Armstrong et al., 1973; Carter, 2005; Chang and Metz, 1976) have both been reported using four proteins

for attachment and 11 for fusion (these 11 proteins make up the entry fusion complex (EFC) (reviewed in (Moss, 2006). The end result of either of these processes is delivery of the naked virion core into the cytoplasm of the infected cell. It is likely that the MV and EEV can use either entry mechanism but that entry via a low-pH-dependent endosomal pathway, more specifically, macropinocytosis is predominant (Huang et al., 2008; Mercer and Helenius, 2008; Schmidt et al., 2011; Townsley et al., 2006b).

Virion assembly occurs in distinct areas of the cytoplasm called viral factories, where initial crescent structures develop into immature virions (IV) into which the genome is packaged. Proteolytic cleavage of core proteins produces the MV (Byrd and Hruby, 2006) which have a single lipoprotein membrane (HOLLINSHEAD *et al*., n.d.) and are the most abundant form of mature virus, the majority of which remain inside the cell until lysis. This form of the virus is very stable and thought to mediate host to host transmission. A fraction of the MV (5 - 40% depending on the viral strain and the host cell (Payne, 1979; Schmelz *et al*., 1994)) become wrapped virions (WV) after acquiring two additional membranes from the *trans*-Golgi network (Schmelz *et al*., 1994) or early endosomal cisternae (Tooze *et al*., 1993). Alternatively, MV acquire a single membrane by budding through the cellular plasma membrane (Boulanger *et al*., 2001). Virions are transported to the cell surface on microtubules where in the case of WV, the outer membrane fuses with the plasma membrane resulting in exocytosis of extracellular virions (EV) if released from the cell surface, and/or cell associated extracellular virions (CEV) if retained on the cell surface (collectively known as extracellular enveloped virions (EEV)). EEV have a fragile outer membrane and are thought to mediate virus spread within the infected host, either via actin tail formation that pushes them away from the host cell to nearby uninfected cells or via release into the body fluids and subsequent infection of different tissues (Smith *et al*., 2003).

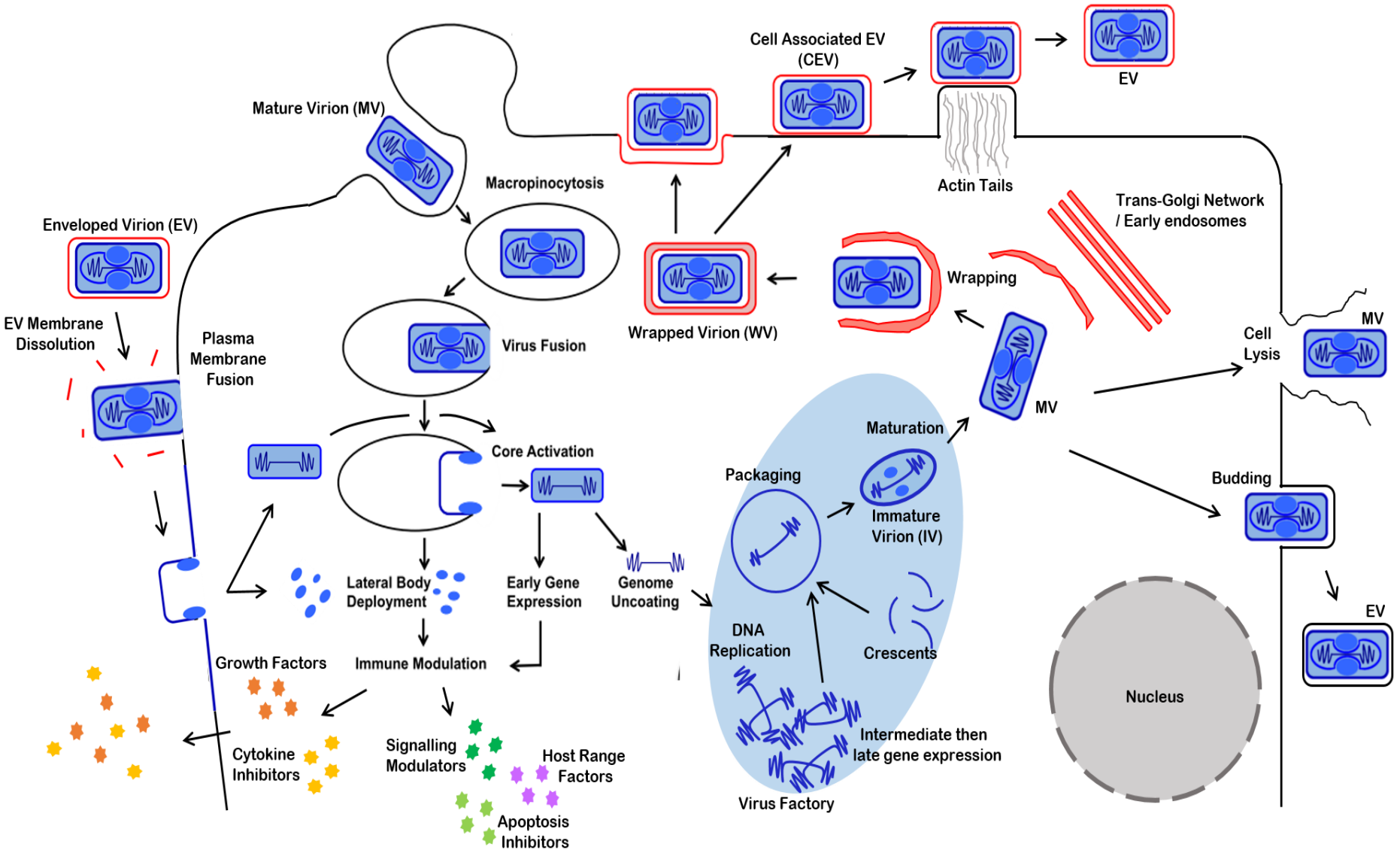


Figure 1.1: Schematic summary of poxvirus morphogenesis (adapted from (Bidgood and Mercer, 2015).

1.2 AVIAN POXVIRUSES

1.2.1 Classification and Phylogenetics

The *Avipoxvirus* genus includes the species *Fowlpox virus* (FWPV) (Afonso *et al*., 2000), the prototype member, *Canarypox virus* (CNPV) (Tulman *et al*., 2004), *Juncopox virus*, *Mynahpox virus*, *Pigeonpox virus* (PGPV), *Psittacinepox virus*, *Quailpox virus* (QPV), *Sparrowpox virus* (SPV), *Starlingpox virus* (SLPV), *Turkeypox virus* (TKPV) and the tentative species: crowpox virus, peacockpox virus and penguinpox virus (PEPV) (ICTV 9th report). Avian poxvirus species have traditionally been distinguished by pock morphology, pathogenicity in different avian hosts, cross protection/neutralisation, complement fixation and agar-gel precipitation tests. These criteria then began to include antigenic and genetic characterisation (Kim *et al*., 2003; Tadese and Reed, 2003) which is now the standard. As of 2007, poxvirus infections had been reported in 278 species of wild and domestic birds from 70 families and 20 orders (Bolte *et al*., 1999; Riper and Forrester, 2008). A literature search reveals the total number of avian species reported to be infected with a poxvirus has risen to at least 337 across 77 families with the addition of those listed in Table 1.3, and it is likely that many more species are susceptible.

As mentioned in chapter 1.1.1, the *Avipoxvirus* genus is predicted to have diverged from other ChPV genera long ago based on phylogenetic analysis as well as analysis of gene content and synteny (Gubser *et al*., 2004b) A monophyletic origin of the avian clade is supported by phylogenetic analysis (Amano *et al*., 1999; Gubser, 2004). Comparison of the relative divergence of partial amino acid P4b sequences from avipoxviruses and those from mammalian poxviruses showed that interclade divergence between avipoxviruses is equivalent to the intergenus divergence seen between ChPV genera (Jarmin, 2006).

Table 1.3: Avian species identified with poxvirus infection

ORDER	FAMILY	COMMON FAMILY NAME	SPECIES	COMMON SPECIES NAME	COUNTRY REPORTED	REFERENCE
Struthioniformes	**Apterygidae**	Kiwis	*Apteryx mantelli*	North Island Brown Kiwi	New Zealand	Ha *et al.*, 2012
Sphenisciformes	Spheniscidae	Penguins	*Spheniscus magellanicus*	Magellanic Penguin	Argentina Brazil	Kane *et al.*, 2012; Niemeyer *et al.*, 2013
			Pygoscelis paua	Gentoo Penguin	Falkland Islands	Unpublished
Procellariiformes	Diomedeidae	Albatrosses	*Thalassarche melanophris*	Black-browed Albatross	Falkland Islands	Jarmin *et al.*, 2006
	Procellariidae	Shearwaters and Petrels	*Macronectes giganteus*	Southern Giant Petrel	Antarctica	Gyuranecz *et al.*, 2013
Pelecaniformes / Suliformes	Phalacrocoracidae	Cormorants	*Phalacrocorax pelagicus*	Pelagic Cormorant	Alaska, USA	Gyuranecz *et al.*, 2013
Phoenicopteriformes	Phoenicopteridae	Flamingos	*Phoenicopterus roseus*	Greater Flamingo	Japan	Terasaki *et al.*, 2010
			Phoenicopterus ruber	American Flamingo	Spain	Henriques *et al.*, 2016
Anseriformes	Anatidae	Ducks, Geese and Swans	*Cygnus buccinators*	Trumpeter Swan	Wisconsin, USA	Gyuranecz *et al.*, 2013
			Anas discors	Blue-winged Teal	Wisconsin, USA	Gyuranecz *et al.*, 2013
			Anas fulvigula	Mottled Duck	Texas	Gyuranecz *et al.*, 2013
			Tadorna variegata	Paradise Duck	New Zealand	Ha *et al.*, 2011
			Aythya americana	Redhead Duck	Wisconsin	Gyuranecz *et al.*, 2013
Falconiformes	Accipitridae	Hawks, Eagles and Kites	*Parabuteo unicinctus*	Harris Hawk (Order Accipitriformes?)	Peru	Lujan-Vega et al., 2011 (unpublished)
			Haliaeetus albicilla	White-tailed Sea Eagle	Japan	Saito *et al.*, 2009
			Hieraaetus pennatus	Booted Eagle	Spain	Gyuranecz *et al.*, 2013
			Spilornis cheela	Crested Serpent Eagle	Taiwan	Chen *et al.*, 2011
			Haliaeetus leucocephalus	Bald Eagle	Florida, Alaska and Minnesota, USA	Gyuranecz *et al.*, 2013
			Milvus milvus	Red Kite	Spain	Gyuranecz *et al.*, 2013
	Falconidae	Falcons and Caracaras	*Falco vespertinus*	Red-footed Falcon	Hungary	Gyuranecz *et al.*, 2013
			Falco pelegrinoides	Barbary Falcon	Kuwait and Dubai	Tarello, 2008
Galliformes	Odontophoridae	New World Quail	*Colinus virginianus ridgwayi*	Northern Masked Bobwhite	Arizona, USA	Gyuranecz *et al.*, 2013
	Phasianidae	Pheasants and Partridge	*Lophura nycthemera*	Silver Pheasant	India	Pawar *et al.*, 2011
			Coturnix japonica	Japanese Quail	Northern Italy	Manarolla *et al.*, 2010
	Tetraonidae	Grouse	*Tympanuchus cupido attwateri*	Attwater's Prairie Chicken	Texas, USA	Zavala *et al.*, 2006
Gruiformes	Otididae	Bustards	*Chlamydotis macqueenii*	Macqueen's Bustard	U.A.E., Morocco and Uzbekistan	Le Loc'h *et al.*, 2014
Charadriiformes	Haematopodidae	Oystercatchers	*Haematopus unicolor*	Variable Oystercatcher	New Zealand	Ha *et al.*, 2011
	Burhinidae	Thick-knees	*Burhinus oedicnemus*	Eurasian Stone-curlew	Spain	Perez-Tris *et al.*, 2011
	Charadriidae	Plovers and Lapwings	*Thinornis novaeseelandiae*	Shore Plover	New Zealand	Ha *et al.*, 2011
			Anarhynchus frontalis	Wrybill	New Zealand	Ha *et al.*, 2011
Columbiformes	Columbidae	Pigeons and Doves	*Streptopelia orientalis*	Oriental Turtle Dove	South Korea	Gyuranecz *et al.*, 2013
			Streptopelia senegalensis	Laughing dove	Oman	Tageldin *et al.*, 2006
			Hemiphaga novaeseelandiae	New Zealand Pigeon	New Zealand	Ha *et al.*, 2011
			Leptotila rufaxilla	Grey-fronted Dove	Peru	Perez-Tris *et al.*, 2011
Psittaciformes	Cacatuidae	Cockatoos and Allies	*Cacatua galerita galerita*	Greater Sulphur-crested Cockatoo	Saudi Arabie	Tarello, 2004
	Psittacidae	Parrots, Macaws and Allies	*Amazona brasiliensis*	Red-tailed Amazon	Brazil	Esteves *et al.*, 2017
			Polytelis swainsonii	Superb Parrot	Chile	Gyuranecz *et al.*, 2013
			Psittacus erithacus	African Grey Parrot	Saudi Arabia	Tarello, 2004
			Platycercus elegans	Crimson Rosella	Australia	Slocombe *et al.*, 2013
			Ara macao	Scarlet Macaw	Georgia, USA	Nemeth *et al.*, 2016
			Guaruba guarouba	Golden conure	Brazil	Esteves *et al.*, 2017
			Pionies leucogaster	White-bellied caique	Brazil	Esteves *et al.*, 2017
Apodiformes / Trochiliformes	**Trochilidae**	Hummingbirds	*Calypte anna*	Anna's Hummingbird	California, USA	Godoy *et al.*, 2013
Passeriformes	**Polioptidae**	Gnatcatchers	*Polioptila caerulea*	Blue-grey Gnatcatcher	Virginia, USA	Adams et al., 2005
	Paridae	Chickadees and Tits	*Periparus ater*	Coal Tit	Spain	Perez-Tris *et al.*, 2011
			Cyanistes caeruleus	Eurasian Blue Tit	Spain	Perez-Tris *et al.*, 2011
			Poecile palustris	Marsh Tit	Oxford, UK	Lachish *et al.*, 2012
	Pteroicidae	Australasian Robins	*Petroica longipes*	North Island Robin	New Zealand	Ha *et al.*, 2011
			Petroica traversi	Black Robin	New Zealand	Ha *et al.*, 2011
	Turdidae	Thrush	*Erithacus Rubecula*	European Robin	Spain	Williams *et al.*, 2014
	Callaeidae	Wattlebirds	*Philesturnus carunculatus*	Saddleback	New Zealand	Ha *et al.*, 2011
	Corvidae	Crows, Jays and Magpies	*Corvus brachyrhynchos*	American Crow	Virginia, USA Washington DC, USA	Grove *et al.*, 2005 Gyuranecz *et al.*, 2013
			Corvus macrorhynchos	Large-billed Crow	Japan	Fukui et al., 2016
	Fringillidae	Siskins, Crossbills and Allies	*Spinus magellanica*	Black-hooded Siskin	The Netherlands	Gyuranecz *et al.*, 2013
			Loxia curvirostra	Red Crossbill	Spain	Perez-Tris *et al.*, 2011
			Loxioides bailleui	Palila	Hawaii, USA	Jarvi *et al.*, 2008
Passeriformes	Thraupidae	Tanagers and Allies	*Camarhynchus pallidus*	Woodpecker Finch	Galapagos Islands, Ecuador	Gyuranecz *et al.*, 2013
			Geospiza fortis	Medium Ground Finch	Galapagos Islands, Ecuador	Gyuranecz *et al.*, 2013
	Icteridae	Troupials and Allies	*Quiscalus major*	Boat-tailed Grackle	Texas, USA	Gyuranecz *et al.*, 2013

This table includes all species discovered in the literature since the publication by van Riper and Forrester in 2007, as well as any species missing in that publication. Families in which poxvirus infections have been noted for the first time are in bold.

Because of the lack of complete genome sequences of avian poxviruses, construction of phylogenies has to date relied on single gene analyses, namely VACV A3L (fpv167; cnpv240) (Binns *et al*., 1989), one of the 49 genes conserved in all poxviruses (Carulei *et al*., 2009; Gyuranecz *et al*., 2013; Jarmin, 2006; Lüschow *et al*., 2004; Manarolla *et al*., 2010; Offerman *et al*., 2013; S. C. Weli *et al*., 2004). This gene encodes the 75.2 kDa P4b protein, a precursor of the 4b virion core protein which is a major structural protein. Proteolytic maturation of P4b and two other precursor proteins (P4a encoded by A10L (fpv174; cnpv247) and VP8 encoded by L4R (fpv131; cnpv176)) is essential for the formation of mature, infectious progeny and thus for completion of the infectious life cycle (Katz and Moss, 1970; Katz and Moss, 1970; Vanslyke *et al*., 1991). In avian poxviruses, this gene is ~2kb in length and located in the central, conserved region of the genome. Other loci used for phylogenetic analysis include VACV H3 (fpv140; cnpv186) which encodes an IMV envelope protein (discussed below), VACV E9 (fpv094; cnpv121) encoding DNA polymerase, VACV A11 (fpv175; cnpv248) encoding a non-structural protein involved in virus assembly (Maruri-Avidal *et al*., 2013) and A12 (fpv175 and fpv176; cnpv248 and cnpv249), a core protein involved in morphogenesis and mRNA synthesis (Satheshkumar *et al*., 2013; Yang and Hruby, 2007).

Due to small sample sizes, initial phylogenetic studies on avipoxviruses offered little reliable information on subclade resolution (Cary J. Adams *et al*., 2005; Lüschow *et al*., 2004; S. C. Weli *et al*., 2004). Acquisition and analysis of several more isolates suggested that the genus could be divided into at least two major clades (A and B) and one minor clade (C) (Jarmin *et al*., 2006). In general, clade A consists of FWPV-like viruses, clade B of CNPV-like viruses and the minor clade C of Psittacinepox (Parrotpox (PRPV)) viruses. Clades A and B were initially divided into sub-clades A1-A4 and B1 and B2, but further subclade divisions have been suggested in more recent studies with the addition of sequences from novel isolates (Fig.1.2 and discussed further in 2.1) (Bányai *et al*., 2015; Gyuranecz *et al*., 2013; Le Loc'h *et al*., 2015; Manarolla *et al*., 2010; Offerman *et al*., 2013; Sarker *et al*., 2017). Two new clades have also been proposed: clade D – containing a single isolate from a Japanese quail (Manarolla *et al*., 2010) and clade E – containing an isolate from a Turkey in Hungary (Bányai *et al*., 2015), and two isolates from layer chickens in Mozambique (Mapaco *et al*., 2017).

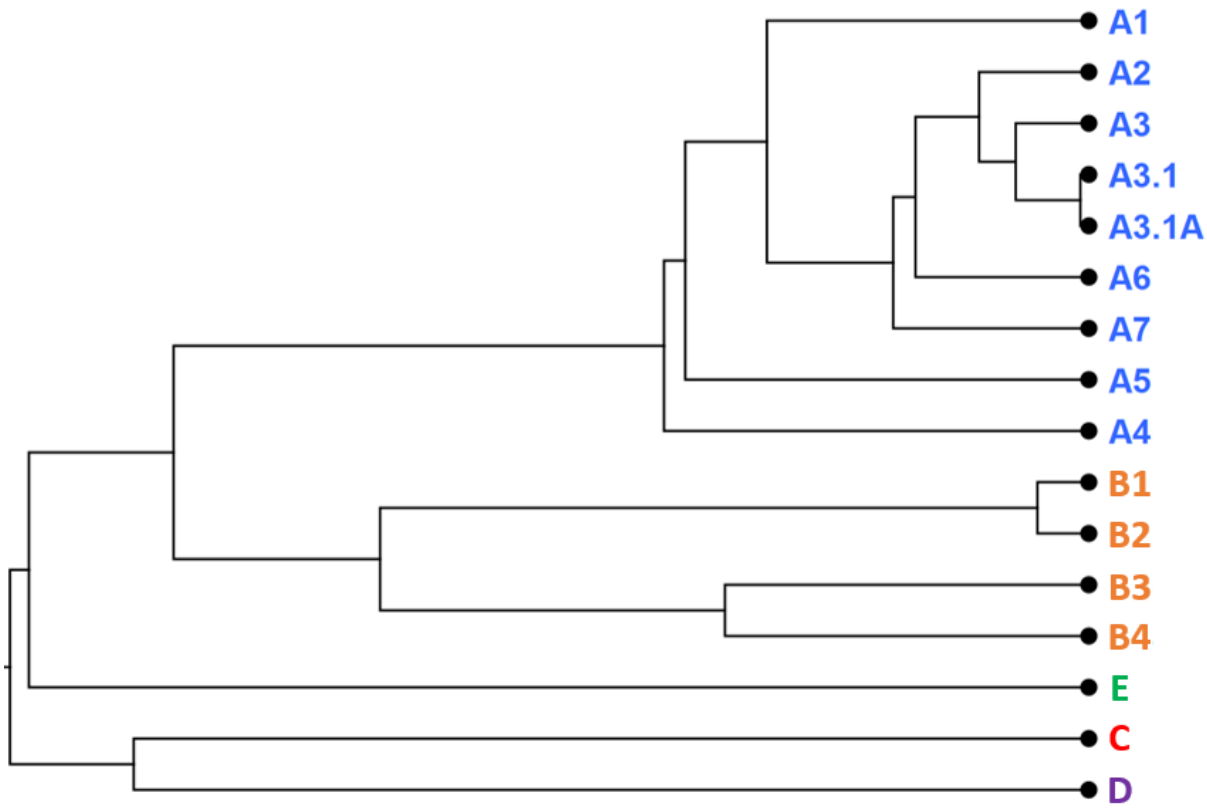


Figure 1.2: Schematic of the clades and subclades of the Avipoxvirus genus proposed to date. Blue - Clade A: Fowlpox-like viruses, Yellow - Clade B: Canarypox-like viruses, Red - Clade C: Parrotpox-like viruses, Purple - Clade D: Quailpox virus, Green - Clade E: Turkeypox virus. (Adapted from Offerman ***et al***., 2013 and Le Loc'h ***et al***., 2014)

Jarmin *et al*., 2006, also sought to find a locus that could be used to easily discriminate between viruses from different clades based on a PCR fragment length polymorphism. Genes encoding the three immunodominant FWPV antigens (fpv140, fpv168 and fpv191) were investigated for this purpose. The region spanning fpv139-fpv141 and containing fpv140 (VACV H3L) was found to provide these diagnostic criteria. fpv140, one of the 49 genes conserved in all ChPV genome encodes the 38.01kDa p35 protein which is an immunodominant envelope protein expressed on the envelope of intracellular mature virions (MV)(Boulanger *et al*., 2002). This protein has been studied in VACV and been shown to bind to cell surface heparin sulphate and to be important for MV morphogenesis and virus infectivity. The length polymorphism was revealed by analysis of PCR products at this locus from 22 avipoxviruses, isolated from 11 different species of birds. It was found that viruses belonging to clade A produced a PCR product of 1800bp compared to the 2400bp product produced by viruses belonging to clade B due to the insertion of cnpv185. Clade C isolates could not be amplified at this locus. This

polymorphism thus allows facile differentiation between clade A and B viruses and also provides better resolution of subclade A2 (Jarmin *et al*., 2006).

An assessment of the concordance of host and virus phylogenies used the cytochrome B locus of clade A vs clade B hosts to determine genetic diversity. This analysis showed that there was greater between-group than within-group diversity overall, showing that clade A and clade B hosts form two distinct groups. However, host specificities are not definitive as several cases were noted in this study where hosts of clade B viruses grouped with hosts of clade A viruses (but not the other way around) (Gyuranecz *et al*., 2013). Individual viral species can infect multiple avian species, and individual avian species can be infected by multiple viral species.

This observation is becoming more evident with increasing numbers of phylogenetic analyses (Adams *et al*., 2005; Jarmin *et al*., 2006; Manarolla *et al*., 2010; Gyuranecz *et al*., 2013; Le Loc'h *et al*., 2014). It was noted by Jarmin *et al*., (2006) that isolates from pigeons grouped in clades A2 and B2 and with the addition of the following study from Italy, B1. Manarolla *et al*., (2010) documented the phylogeny of avipoxviruses isolated over more than four decades (1963-2007) in Northern Italy. It was found that of the three viruses isolated from canaries from 1985-1993, one grouped in clade A1, one in clade A2 and one in clade B1. In a study of viruses isolated from Houbara Bustard (*Chlamydotis undulata* and *Chlamydotis macqueenii*) in captive breeding projects in the UAE, Uzbekistan and Morocco, it was found that isolates from C.*undulata* grouped in clades B1 and B2 while isolates from C.*macqueenii* grouped in five different subclades (A1-3, B1 and B4 (a proposed subclade)) (Le Loc'h *et al*., 2014).

As previously mentioned, phylogenies produced using the alignment of whole genomes or several concatenated genes are the gold standard, but these types of analyses are currently expensive, considering the large size of avian poxvirus genomes. It cannot be assumed that the virus phylogenies produced to date are accurate based on single gene alignments, but it appears that such analyses produce trees that can reliably distinguish between major clades. As more avian poxvirus genomes are sequenced, confirmation and further delineation of these relationships can be initiated.

1.2.2 Transmission/Pathology

Transmission of avian poxviruses between birds occurs through both direct (physical contact with an infected bird or ingestion of infected tissues) and indirect contact (contaminated fomites (feed, water, housing structures/materials, scab particles, dust and dander) and biting arthropods including various species of mosquito, mites, fleas and stable-flies (DaMassa, 1966). Mechanical transmission is considered to be an important method of disease transmission and vectors include mosquitoes which can harbour the virus for over a month after feeding on an infected bird (Ritchie, 1995).

Both susceptibility to infection and symptoms upon infection can vary depending on the age of the bird (juveniles are more frequently affected) route of infection, immune status of the infected bird, presence of secondary infections, virulence of the infecting strain and host susceptibility to that specific strain. Infection of broken skin results in the cutaneous form of disease whereas infection of mucosal surfaces results in the less common, diptheric form. Birds can be infected with both forms of disease simultaneously and systemic infection also occurs (Tripathy and Reed, 2003), although reported cases are rare, possibly due to high mortality rates. Infection can also be asymptomatic in some cases. The cutaneous form of disease is typically self-limiting, non-fatal and characterised by discrete, proliferative, nodular lesions on the un-feathered regions (feet, legs, beak/bill and head) (Hansen, 1999). The diptheric form involves infection of the oral cavity, digestive tract, pharynx and trachea and has a higher mortality rate than the cutaneous form due to obstruction of the oesophagus and resulting starvation as well as obstruction of the trachea and resulting asphyxiation (Bolte *et al*., 1999). In domestic poultry, avian pox infection can cause significant economic losses due to transient decrease in egg production, compromised vision/blindness, reduced growth in young birds and increased mortality (Tripathy and Reed, 2003). In wild bird populations, including endangered and endemic species, poxvirus infection may lead to secondary bacterial or fungal infections and affect vision and/or the ability to feed making them prone to predation (Tripathy *et al*., 2000).

Diagnosis of APV infection can be suggested and confirmed by several methods. Observation of pock like lesions is suggestive of poxvirus infection but is not definitive because several other diseases can cause similar looking lesions. Positive diagnosis of poxvirus infection by growth of virus on CAMs is indicative of infection but infection cannot be ruled out by a negative result as not all avipoxviruses are capable of growth on chicken CAMs (van Riper *et al*., 2002; Krone *et al*., 2004). Histopathology of infected CAMs or tissues and observations of Bollinger bodies as well as serological methods (virus neutralisation in CEFs or other permissive cell lines, agar gel immunodiffusion, passive haemagglutination, immunofluorescence, ELISA and immunoblotting) and restriction-fragment length polymorphism (RFLP) add further confirmation to diagnosis of avian poxvirus infection but electron microscopy and PCR and sequencing are the gold standards for clinical diagnostics.

1.2.3 Prevention and Control

There is currently no treatment available for the treatment of poxvirus infection beyond therapeutic support including the debridement of lesions and prevention/treatment of secondary bacterial and fungal infections with antimicrobials. In commercial settings as well as aviaries, reduction in vector numbers or exposure to vectors can help limit infection as can cleaning and disinfecting of living areas and fomites.

Commercial vaccines are available to protect against both FWPV and PGPV, but they show variable levels of protection depending on the similarity of the vaccine and infecting strain. This has been complicated in the case of FWPV vaccines by the integration of Avian Reticuloendotheliosis Virus (REV) into both field and vaccine strains. REV was first isolated from an adult turkey in 1958 though descriptive reports in the literature only occurred several years later (Robinson and Twiehaus, 1974; Zeigel *et al*., 1966). REVs are oncogenic avian gamma retroviruses belonging to the family *Retroviridae.* They have small, ~9kb genomes that encode gag, pol and env ORFs (Rice *et al*., 1982). REVs infect several different avian species including game birds (Galliformes) and waterfowl (Anseriformes) with infection resulting in a diversity of symptoms including anaemia, immunosuppression, neoplasia (tumor formation), runting and nakanuke (feathering abnormalities) (Dren *et al*., 1988; Witter

and Fadly, 2003). REV integration into a commercial FWPV vaccine strain was first reported in 1996 by Fadly *et al*., and several other discoveries of contaminated vaccines have since been published (Awad *et al*., 2010; Diallo *et al*., 1998; Fadly *et al*., 1996; García *et al*., 2003; Hertig *et al*., 1997; Moore *et al*., 2000; Pratik Singh *et al*., 2005; Singh *et al*., 2003; Wei *et al*., 2012).

1.2.4 Growth Characteristics

Avipoxviruses are generally capable of productive infection on the CAM of embryonated chicken eggs and in primary embryonic fibroblast cells of both fowl (CEFs) and duck origin (Bang *et al*., 1950; Mayr and Kalcher, 1960; Mayr, 1963; Gelenczei and Lasher, 1968; Docherty and Slota, 1988) as well as embryonic dermal cells (CEDs) of fowl origin (El-Zein *et al*., 1974). Growth in permanent cell lines is preferred due to simplified maintenance and relative stability. A permanent duck embryo cell line, DEC99 was shown to be permissive to growth of two vaccine strains FWPV (FK) and PGPV (Dessau) (Ivanov *et al*., 2001). A permanent cell line of quail origin (QT-35) has been shown to be permissive for growth of Juncopox, pigeonpox and both field and vaccine strains of fowlpox virus (Schnitzlin *et al*., 1988; Singh *et al*., 2003). QT-35 and LMH cells (chicken hepatocellular carcinoma) were shown to be permissive to viruses from Hawaiian crows (*Corvus hawaiiensis)* and Apapane *(Himatione sanguinea)* although it must be noted that the authors performed blind passage and only showed evidence of cytopathic effect (CPE) (LMH: cell rounding, enlargement and detachment; QT-35: plaques); they did not titrate the virus on CAMs and no CPE was evident by the fifth passage in CEFs (Tripathy *et al*., 2000).

Abortive infection/replication of avian poxviruses has been shown in various mammalian cell lines including Vero, CV-1, MRC-5, HeLa (Somogyi *et al*., 1993; Stannard *et al*., 1998). Bovine embryonic tracheal cells (EBTr) were shown to support production of IEV which were able to further infect and replicate in DEC99 cells (Sainova *et al*., 2001). Following infection of BHK-21 (Syrian baby hamster kidney) cell monolayers with FWPV-VR250, four distinct morphological forms (viral crescent, IV, MV and IEV/WV) as well as virus budding were identified by transmission electron microscopy (TEM) and scanning electron microscopy (SEM)

(Weli *et al*., 2004). This was further investigated by infection of various mammalian cell lines with isolates from a chicken (FPV-VR250), sparrow (SPV-A468) and a pigeon (PPV-B7). BHK-21 was the only cell line of the eight mammalian cell lines tested, shown to be permissive to growth of the selected avian poxviruses. It must be noted that these cells are unusually permissive to virus replication and are not representative of mammalian cells in general. The authors state "…the fact that modified vaccinia virus Ankara (Drexler et al., 1998) and avipoxvirus (two viruses) both showing an otherwise limited host range can multiply in BHK-21 cells, but not on other mammalian cells suggest that BHK-21 cells may possess some unique proteins, which if investigated may open a new chapter in our understanding of the so-called avipoxvirus abortive replication in mammalian cells" (Weli *et al*., 2005). Furthermore, two of the APVs tested in this study multiplied by less than 50 fold in BHK-21 cells and infection studies of BHK-21 cells in our own laboratory using several different APVs (data not shown) have not been able to replicate this result. It remains questionable to conclude that APVs can replicate in mammalian cells. The growth and replication properties of Penguinpox virus (PEPV; discussed in chapter 1.2.6.1) were studied in CEFs and CV-1 (monkey kidney) cells. All expected stages of virus morphogenesis were observed by TEM in infected CEFs but passage of PEPV in CEFs failed. PEPV was able to enter CV-1 cells and replicate its DNA though infectious progeny virus could not be recovered (Stannard *et al*., 1998).

1.2.5 Avian Poxvirus Genome Organisation and Statistics

FWPV (FPVUS and FP9) (Afonso *et al*., 2000; Laidlaw and Skinner, 2004), CNPV (Wheatley C93) (Tulman *et al*., 2004), FeP2 (Offerman *et al*., 2014), TKPV (Banyai *et al*, 2015), SWPV1 and SWPV2 (Sarker *et al*., 2017) are the only avian poxviruses to have their genomes completely sequenced and analysed. FP9 is an attenuated European Fowl pox virus that has been plaque purified and tissue culture adapted (Mockett *et al.,* 1992). Avian poxvirus genomes ((FP9 (266kbp); FWPV (288kbp); CNPV (365kbp); FeP2 (282kbp); TKPV (188kbp); SWPV1 (327kbp) and SWPV2 (351kbp)) (Table 1.4) are generally larger than those of other chordopoxviruses (130-191kbp) and accordingly, contain more open reading frames (171 – 328 ORFs in avian poxviruses and 133 – 236 ORFs in other chordopoxviruses).

Table 1.4: Comparison of basic genome statistics of the seven fully sequenced avian poxvirus genomes

STATISTIC	FWPV	FP9	FeP2	TKPV	SWPV1	SWPV2	CNPV
Length (kbp)	289	266	282	189	327	351	365
A+T (%)	69.1	69.2	70.5	70.2	72.4	69.8	69.6
# of ORFs	260	244	271	171	310	312	328

As mentioned earlier, only 89 of these genes are considered essential, meaning the remainder of the genetic component of avian poxviruses is largely made up of immunomodulatory and host specific genes that have allowed them to take advantage of their avian hosts. Figure 1.3 is a schematic of the FWPV genome showing individual open reading frames organised into functional groups or gene families. As is seen in other ChPV genomes, genes involved in the poxvirus life cycle are located in the central conserved region with the immunomodulatory and host specific genes located near the termini. The three boxed regions in Figure 1.3 indicate novel coding regions in the junctions between areas of major rearrangement relative to VACV. This is discussed further in chapter 4. The *Avipoxvirus* genus has been shown to be highly divergent to other chordopoxvirus genera (Offerman *et al*., 2013; Gubser *et al*., 2004; McLysaght *et al*., 2003) and genus members also exhibit large intra-genus divergence (Jarmin *et al*., 2006).

1.2.5.1 FP9 vs FWPV US

The genome sequence of the attenuated, European FP9 was compared to FWPV (pathogenic American strain, FPVUS) (Laidlaw and Skinner, 2004). FP9 was found to contain 244 of the 260 ORFs reported in the FWPV genome and to be 22kbp shorter. Of the 244 identified ORFs 9 were disrupted and 189 were identical. The proteins of the remaining 71 genes were found to be affected by 118 differences including in/dels, substitutions, terminations and frame-shifts (19 were shorter in FP9, 5 were longer and 30 had missense mutations). By comparison of each of the 118 differences in FP9, to HP1, the virulent progenitor of FP9, it was determined that more than half of the differences identified between FWPV and FP9 were lineage dependent and of the 71 affected ORFs, 25 had mutations that were lineage dependent. At 50 of the 118 loci examined, the HP1 sequence differed to FP9 indicating that these differences are passage specific. These passage specific mutations affected 46 ORFs including 25 ORFs belonging to gene families (12 ankyrin repeat, 3 C-type lectins, 2 C4L/C10L, 1 GPCR, 1 V-type Ig domain, 1 EFc, 2

B22R, 1 mutT and 2 N1R/p28). This analysis reinforces the idea that attenuation of poxviruses is afforded by the loss of gene family proteins such as the ankyrin repeat family proteins as was seen in this study and in the analysis of sheeppox and goatpox viruses (Tulman *et al*., 2002), and African swine fever virus (Neilan *et al*., 2002).

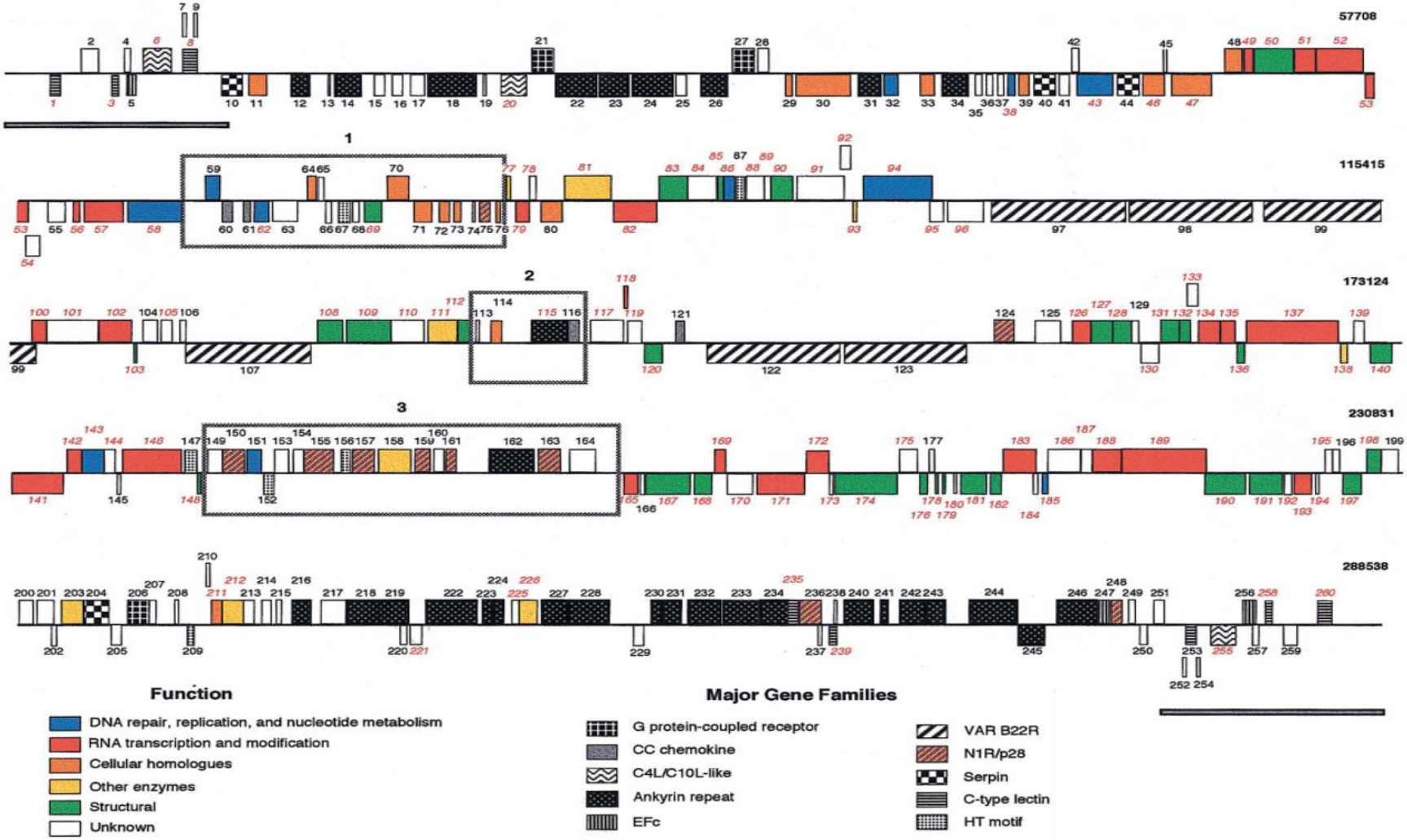


Figure 1.3: Schematic depicting the positions, functions and major gene families of FWPV US open reading frames. Blocks below are transcribed to the left and blocks above transcribed to the right. Larger boxes (1-3) around groups of ORFs depict areas of rearrangement relative to MCV. VACV homologues are indicated with red, italicised numbers. ITRs are depicted as solid grey lines at each end of the genome. (Image taken from Afonso *et al.*, 2000)

1.2.5.2 FWPV vs CNPV

FWPV and CNPV exhibit overall synteny in genome arrangement and have similar genetic complements. They do however, exhibit significant differences in the terminal, variable regions as well as in 3 internal, variable regions (Fig.1.4). These variable, internal regions have been shown to correspond to junctions of major genome rearrangements relative to other ChPVs which show overall gene synteny and conservation (Afonso *et al*., 2000; Gubser, 2004). This variability is significant considering the high degree of conservation of these regions in other ChPVs. It is suggested that this variability could be a result of either increased selective pressures on host range genes found in these regions or unique aspects of avipoxvirus genome replication (Tulman *et al*., 2004). Amino acid sequence identity of the TK gene from each species was found to be 64.2%, which is significantly lower than similarities between species of other genera (Amano *et al*., 1999) further suggesting significant genomic differences amongst avipoxviruses.

The CNPV genome contains 39 ORFs not present in FWPV, 29 of which encode unique, hypothetical proteins. CNPV contains two additional nucleotide metabolism genes (thymidylate kinase and the small subunit of ribonucleotide reductase), a TNFR (CNPV086), an IL-10 like protein (CNPV018), cellular ubiquitin (CNPV096), a protein tyrosine phosphatase (CNPV085), a thioredoxin binding protein (CNPV149), and two Rep like proteins (CNPV153 and CNPV200). FWPV contains 15 ORFs not present in CNPV, 13 of which encode hypothetical proteins. Homologues of fpv217 and fpv250 are notably absent.

Figure 1.4 Schematic showing alignment of the FWPV and CNPV genomes: Coloured ORFs indicate differences between CNPV and FWPV: ORFs used to introduce gaps or lacking discernible orthologous sequence in the other virus are marked in green; nonhomologous ORFs in similar genomic positions are marked in blue; ORFs severely disrupted in the opposite virus are marked in yellow; and, due to extensive variability, ORFs in terminal regions marked in orange are unaligned. CNPV ORFs lacking discernible homology to any FWPV ORF are marked above with an asterisk; FWPV ORFs lacking discernible homology to any CNPV ORF are marked above with a triangle. Thick black bars at genomic termini represent ITRs. Boxed regions indicate novel coding regions at junction sites of major genome rearrangements previously identified in FWPV (Afonso *et al*., 2000), with white indicating gaps between grey sequence.

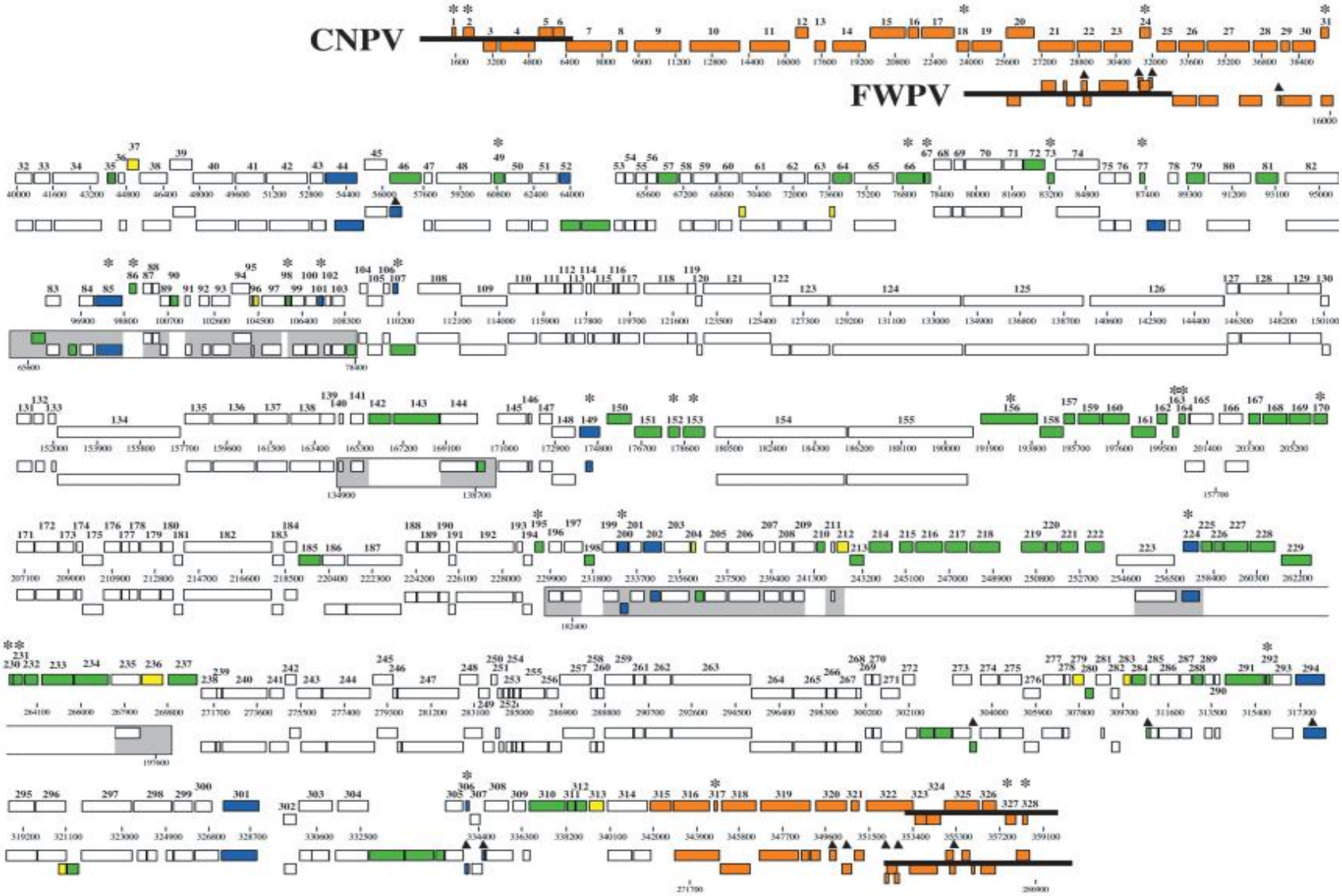


Figure 1.4: Schematic showing alignment of the FWPV and CNPV genomes (Image taken from Tulman *et al*., 2004).

1.2.5.3 FeP2 Genome

FeP2 was isolated from a feral pigeon (*Columba livia*) in Port Elizabeth, South Africa in 2011. Phylogenetically it groups in subclade A2 based on alignment of partial P4b sequences (Offerman *et al*., 2013). The genome was found to be 282kbp in length, encoding 271 ORFs including the 89 ORFs considered conserved amongst ChPV (Offerman *et al*., 2014). Typical poxvirus genome architecture was noted although several differences to the FWPV genome were discovered and a full genome alignment revealed 84% nt identity. The most significant difference between FeP2 and FWPV is a large deletion of 16 kbp in FeP2 relative to FWPV resulting in ORFs FPV121-125 being absent from the FeP2 genome. Relative to FWPV, FeP2 was found to have a total of 36 ORFs deleted as well as 8 ORFs disrupted by truncations and fragmentations. The FeP2 genome was also found to contain several orthologues of genes found in CNPV but not in FWPV. These include thymidylate kinase (CNPV170), interleukin-10 (IL-10) (CNPV018), and a TNFR-like protein (CNPV086). It is proposed that although the FeP2 genome is more similar to the FWPV genome overall, the presence of these CNPV orthologues suggests that FeP2 originates from the common ancestor of FWPV and CNPV. (Offerman *et al*., 2014).

1.2.5.4 TKPV Genome

TKPV-HU1124/2011 (TKPV), was isolated from a previously vaccinated, 38-week-old turkey (*Meleagris gallopavo*) in Hungary, in 2011 (Bányai *et al*., 2015). The TKPV genome was found to be significantly different to other avipoxvirus genomes largely due to its small size (Table 1.4). The genome was found to be 188kbp in length, with ~1.5kb ITRs and encoding 171 potential ORFs. The overall genome structure was as expected and gene synteny was evident but rearranged regions affecting several ORFs were identified. Relative to FWPV, ~ 20 ORFs at each end of the genome are absent. Two ORFs previously considered to be conserved in ChPV (fpv194 and fpv194.1) are not present in the TKPV genome and a third (fpv103; tkpv074) is truncated/fragmented. Two ORFS (tkpv049 and tkpv114) were identified with no orthologues identified in other avipoxvirus genomes. Phylogenetically, TKPV groups on its own in a proposed new subclade, subclade E, based on partial P4b and DNA polymerase nucleotide sequence alignments. Although smaller genomes have been associated with reduced virulence (Laidlaw and Skinner, 2004), it is likely that this

vaccine breakthrough occurred due to the degree of divergence compared to the vaccine strain.

1.2.5.5 SWPV-1 and SWPV-2 Genomes

SWPV-1 was isolated from a Flesh-footed Shearwater (*Ardenna carneipes*) and SWPV-2 from a Wedge-tailed Shearwater (*Ardenna pacificus*) living in breeding colonies on Lord Howe Island off the east coast of Australia in 2015 (Sarker *et al*., 2017). The SWPV-1 and SWPV-2 genomes were found to be 327 and 351 kbp, encoding 310 and 312 ORFs respectively. Overall, both genomes showed relative synteny to the CNPV genome but several rearranged blocks of 1-6 genes were noted in SWPV-1 as well as various fragmentations, truncations, insertions and deletions in both viruses. SWPV-2 lacks orthologues of 18 ORFs found in CNPV and a further 15 have been fragmented. SWPV-2 also contains four hypothetical proteins with no significant similarity to ORFs in Genbank. SWPV-1 showed greater divergence in gene content with orthologues of 43 CNPV ORFs absent and a further six, fragmented. SWPV-1 was found to contain all 89 conserved ChPV genes while SWPV-2 lacks an orthologue of fpv095, previously considered conserved in ChPV. Compared to CNPV, SWPV-2 ORFs were found to be very similar with an average aa identity of >98% while SWPV-1 orthologues have a lower average aa identity of 67% and full genome alignments show nt identities of 99% and 79% respectively. The authors suggest evidence of recombination in both viruses, based on differences in sequence identity within short syntenic regions but insufficient information was provided in the publication to evaluate that claim and no further information was provided by the authors upon request.

1.2.6 South African Avian Poxviruses

Documented cases of poxvirus infections in birds in South Africa date back to the early 1960's when infections were noted in Cape turtle doves (*Streptopelia capicola*) and a Cape thrush (*Turdus olivaceus*) (Middlemiss, 1961). Infections were later noted in an African penguin (penguinpox virus (PEPV) - *Spheniscus demersus*) (Kow, 1992) ostriches (*Struthio camelus australis*) (Allwright *et al*., 1994), several species of pigeon and dove, (Bwala *et al*., 2015; Offerman *et al*., 2013) and Cape

sugarbirds (*Promerops cafer*) (Unpublished; P4b sequences submitted directly to Genbank KP760476 and KP760477).

1.2.6.1 Penguinpox Virus (PEPV)

Poxvirus infection has been documented in five penguin species: African (*Spheniscus demersus*) (Stannard *et al*., 1998), Humboldt (*Spheniscus humboldti*) (Bolte *et al*., 1999), Gentoo (*Pygoscelis papua*) (Munro, 2007), Magellanic (*Spheniscus magellanicus*) (Kane *et al*., 2012; Niemeyer *et al*., 2013), Little Blue (*Eudyptula minor*) (Hunter, 2014). Penguinpox virus (PEPV) described in this study is a novel avipoxvirus isolated from African penguins (*Spheniscus demersus*) that were brought into the Southern African Foundation for the Conservation of Coastal Birds (SANCCOB) following an oil spill off the coast of Cape Town, South Africa (Kow, 1992). Lesions around the eyes, (Figure 1.5) typical of avipoxvirus infection were noted and scrapings were taken.

Figure 1.5: **Poxvirus lesion near the unfeathered eye-region of a juvenile African penguin at SANCCOB.**

Virus was cultured from these scrapings and histological studies and restriction enzyme analysis and profile comparison to other known avipoxviruses namely FWPV, CNPV and quailpox virus, confirmed that it was indeed a novel avipoxvirus. Infectivity studies of heterologous mammalian cell lines (CV-1, Vero, HeLa, MDBK, RK-13 and HEF) showed that early stages of virus replication were supported but no infectious progeny virus could be recovered (Stannard *et al*., 1998). Also reported was the fact that PEPV transcriptases could recognise the Vaccinia Virus derived late promoter P11 linked to the β-galactosidase reporter gene thus resulting in transient gene expression. Other VACV promoters also gave positive results but P11 was found to be most effective.

In 2007, an adult male King Penguin (*Aptenodytes patagonicus*) underwent rehabilitation at SANCCOB. Small, pale reddish papules/nodules on both eyelids were noted to be similar to lesions seen in the African penguin described above (Parsons *et al*., 2018) but infection with a poxvirus was not confirmed. Topical treatment with chloramphenicol resulted in recession and healing of the lesions one month after admission. As the lesions started developing approximately one week after admission, it is suggested that infection occurred at the centre. Mosquito netting was introduced at SANCCOB in 2008 to protect housed penguins from contracting avian malaria. There has been no evidence of pox infections since then, suggesting that the reported infections were transmitted by mosquitoes. Poxvirus infections have not been observed in wild penguins living in coastal waters around Cape Town (Personal communication with Nola Parsons and Tertius Gous, SANCCOB).

1.2.6.2 Flamingopox Virus (FGPV)

Poxvirus infections have been documented in four species of flamingo to date. The first documented case occurred in Chilean flamingos (*Phoenicopterus chilensis*) that were housed at a zoo in Hino City, Tokyo, Japan (Arai *et al*., 1991) but the virus was not isolated or characterised. Two separate cases of infections in American flamingos (*Phoenicopterus ruber*) have been reported with the first occurring in a bird housed at the National Zoological Park in Washington DC (Mondal *et al*., 2008). In this case, a 4.5kb *Hin*dIII fragment ranging from the equivalent of fpv193-203, was reported to show 99.7% nucleotide identity to an isolate from an Andean condor (*Vultur gryphus*) which groups in clade B phylogenetically. A second case was

reported in a young American flamingo housed at the Lisbon zoo. Phylogenetic analysis based on P4b and the CNPV 186-187 fragment showed this isolate to group in clade B2 with the highest identity to isolates from various species of bustard (Henriques *et al*., 2016). Another case occurred at a zoo in Japan, but the infection was noted in two, young Greater flamingos (*Phoenicopterus roseus*) (Terasaki *et al*., 2010). In this case, 132 flamingos, 23 of which were lesser flamingos, were housed in an aviary with no introduction of flamingos from the outside in the preceding year. The breeding area was enclosed by netting which only allowed entry of small, wild birds. Two, four-week-old flamingos were found to have nodules on their beaks, but no other clinical signs of disease. The birds died several weeks later, and the nodules were removed for viral characterisation. Homogenate was inoculated onto CAMs and grown for 5 days showing typical, raised, poxviral lesions. Based on analysis of the P4b gene this isolate was shown to group with two isolates from pigeons (PPV-B7 and CVL950), also in clade B2. All the above cases occurred in captive flamingos.

1.2.6.3 Pigeonpox Virus (PGPV)

Poxvirus infection has been identified in at least five genera and ten species of columbiformes including pigeons and doves (Table 1.3) (van Riper and Forrester, 2007). Several of these isolates from columbiformes have come from South Africa (FeP1 and FeP2 from feral pigeons (*Columba livia*), Pi5 from a racing pigeon (*Columba livia domestica*), RP1, RP2 and SP1 from rock/speckled pigeons (*Columba guinea*)), and LD1 and LD2 from laughing doves (*Spiopelia senegalensis*)) (Offerman *et al*., 2013). As mentioned in 1.2.1, poxviruses isolated from pigeons have been found to group in two different clades (A and B) and four subclades (A2, A3, B1, B2) phylogenetically. FeP2, belonging to subclade A2/A3 (depending on which locus is used for analysis), is the only isolate from a pigeon to have its genome completely sequenced (Offerman *et al*., 2014).

1.3 PROJECT RATIONALE

Avian poxviruses are important pathogens of wild and domestic birds worldwide. Economic losses due to infection of commercial flocks and the threat infection poses to endangered and endemic species is of great concern. Compared to viruses from

other ChPV genera, the *Avipoxvirus* genus is relatively uncharacterised especially in terms of genomic content and structure. As a large and divergent genus these topics are of particular interest. To date, six avian poxvirus genomes have been sequenced and analysis has shown major differences in genome length, gene content and synteny, and the presence of novel ORFs that have yet to be characterised. This thesis presents characterisation of novel avipoxviruses isolated from birds in South Africa. The genomes of two isolates have been sequenced and analysed in context with what has been published to date.

CHAPTER 2: HISTOLOGIC AND PHYLOGENETIC ANALYSIS OF SOUTH AFRICAN AVIAN POXVIRUSES

2.1 INTRODUCTION

2.2 MATERIALS and METHODS

2.2.1 Virus Acquisition and Isolation

2.2.1.1 FWPV and CNPV

2.2.1.2 PEPV

2.2.1.3 FGPV

2.2.1.4 Pi1

2.2.1.5 Pi2 and Pi3

2.2.1.6 Pi4

2.2.1.7 Virus Isolation from Scab Material

2.2.1.8 Growth of Viruses on CAMs

2.2.2 Histology

2.2.3 DNA Extraction

2.2.4 PCR and DNA Sequencing

2.2.5 Phylogenetic Analysis

2.3 RESULTS

2.3.1 Pock Morphology and Membrane Histology

2.3.2 Phylogenetic Analysis

2.4 DISCUSSION

2.1 INTRODUCTION

Analysis of macroscopic growth characteristics and histology of infected CAMs forms part of the basic characterisation of novel avian poxviruses. These viruses are known to produce pocks of different sizes and colours and to cause variable degrees of thickening and inflammation of the CAM. Histologically, APVs generally cause hypertrophy and hyperplasia to varying degrees and intracytoplasmic inclusion bodies are readily seen in the cytoplasm of infected cells (Gilhare *et al*., 2015; Abdallah and Hassanin, 2013; Diallo *et al*., 2010; Manarolla *et al*., 2010; Halıgür *et al*., 2009; Kulich *et al*., 2008; Rampin *et al*., 2007; Tripathy *et al*., 2000). Few studies are available that directly compare growth characteristics of different viruses under the same infection conditions such as egg source, inoculum reagents and titres, and incubation time. An analysis of both macroscopic growth characteristics and histology of infected CAMs was therefore undertaken comparing the novel, South African isolates to FWPV, FeP2, CNPV and uninfected controls.

As shown by the sequencing and comparison of the FWPV, FP9, CNPV, FeP2, TKPV, SWPV1 and SWPV2 genomes, viruses belonging to different clades and subclades show many differences in gene content and synteny. However, phylogenetic analysis of single genes such as P4b has been shown to reliably classify viruses into three major clades and several minor clades which are still being elucidated. Currently, the seven isolates with completely sequenced genomes represent clades A1 (FWPV and FP9), A2 (FeP2), B (CNPV, SWPV1 and SWPV2) and the proposed clade E (TKPV). A phylogenetic analysis was conducted to determine the phyletic relationships amongst and between the novel viruses and those already classified.

Basic characterisation of viral isolates in terms of growth characteristics on CAMs, CAM histology and phylogenetic analyses was conducted to inform selection of isolates for further characterisation by genome sequencing and comparative genomics as only two isolates could be selected due to time and budget constraints.

2.2 MATERIALS and METHODS

2.2.1 Virus Acquisition and Isolation

2.2.1.1 FWPV and CNPV

The FWPV isolate used was the DCEP 25 modified strain (Sanofi Merial, Duluth, GA, USA) sourced from a licensed veterinarian and the CNPV isolate was obtained from Prof. Keith Dumbell's collection at the University of Cape Town. The virus was originally sourced from Prof. A. Mayr (Veterinary Faculty, Universtiy of Munich, Munich, Germany).

2.2.1.2 PEPV

The penguinpox virus isolate was obtained from an African penguin as previously described in chapter 1.2.6 and in Stannard *et al*., (1998).

2.2.1.3 FGPV

The FGPV isolate characterised in this study was obtained during a poxvirus outbreak that occurred in a permanent, breeding population of lesser flamingos (*Phoenicopterus minor*), living at Kamfers Dam, a perennial wetland near Kimberley, South Africa (Zimmerman *et al*., 2011). Approximately 30% of the juvenile population was estimated to have developed lesions. The virus sample was taken from a lesion near the tibiotarsal joint of a juvenile flamingo (Figure 2.1) on post mortem by Dr. David Zimmermann and Dr. Mark Anderson. The flamingo was euthanised and examined further as part of an investigation into the cause of the dermal lesions. The sample was donated by Dr. Emily Lane of National Zoological Gardens as scab tissue and stored at -20ºC until further processing.

Figure 2.1: Poxvirus lesions near the tibiotarsal joint of a juvenile Lesser Flamingo.

2.2.1.4 Pi1

Isolate Pi1 was isolated from a sick pigeon brought into Elsenburg Veterinary Laboratory Services in Stellenbosch, South Africa by a member of the public. No history of the infected bird was available. The sample was donated as homogenised, infected CAM tissue by Magdalene Dreyer and Dr. James P. Kitching in February of 2009 and stored at -20ºC until further processing.

2.2.1.5 Pi2 and Pi3

Isolates Pi2 and Pi3 were isolated from lesions on the heads of two juvenile (25 days old) racing pigeons (Figure 2.2) handed over by the breeder, Eric Brown in Cape Town, South Africa in March 2012. The pigeons were not vaccinated and were housed in a coop with other birds, but no other infections were identified. It was noted that the breeder had experienced poxvirus infections in his birds in the past but had never seen such gross pathology in birds that young (Personal communication – Eric Brown). The pigeons were taken to the Blue Cross Veterinary Hospital in Newlands, Cape Town and euthanised. Scab material was removed and stored at -20ºC until further processing. Pi2 was used for characterisation as the viruses infecting both birds were assumed to be identical and more scab material was available from this individual.

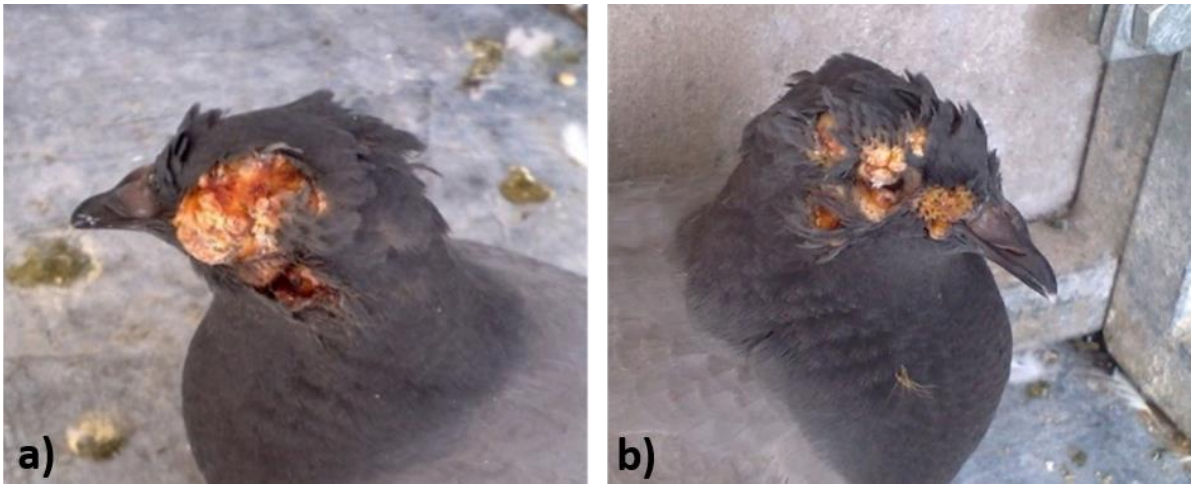


Figure 2.2: Poxvirus lesions on the heads of two juvenile racing pigeons a) Pi2 and b) Pi3.

2.2.1.6 Pi4

Isolate Pi4 was isolated from an oral lesion of an infected racing pigeon owned by Mike Orman, a pigeon fancier in Cape Town, South Africa. The infected pigeon was vaccinated with Medipox at 2 months of age and was 6 months old at the time of diagnosis in August 2010. This bird was in contact with other birds, but their

infection status was not known. The sample was supplied as scab tissue which was stored at -20ºC until further processing.

2.2.1.7 Virus Isolation from Scab Material

For isolates FGPV, Pi2 and Pi4, the scabs were diced with a scalpel and added to a Dounce homogeniser in 1ml of McIlvains buffer containing penicillin (500U/ml), streptomycin (100µg/ml) and fungin (1µg/ml). The homogenate was centrifuged at 800rpm for 5 mins and the supernatant and pellet were stored separately at -20ºC.

2.2.1.8 Growth of Viruses on Cams

To produce high tire virus stocks for use in downstream applications, all viruses were grown on the chorioallantoic membranes (CAMs) of embryonated hen's eggs according to a method first described by (Joklik, 1962) and described in its current form by (Kotwal and Abrahams, 2004). Commercial, farm eggs were sourced from a commercial company in the Western Cape, South Africa and specific pathogen free eggs (SPF) eggs were sourced from Avifarms (Pty) in Lyttelton, South Africa. The following modifications were made to the method described by Kotwal and Abrahams (2004). Eggs were inoculated with 100µl purified virus diluted from 10^{-2} to 10^{-4} (FWPV, CNPV, PEPV), supernatant from CAM homogenates (Pi1) diluted from 10^{-1}-10^{-2} or scab homogenates (FGPV, Pi2 and Pi4), diluted from 10^{-2} to 10^{-4} per egg. All samples were diluted in phosphate buffered saline (PBS) containing penicillin (500U/ml), streptomycin (100µg/ml) and fungin (1µg/ml). Embryos were 10-11 days old at the time of inoculation. Eggs were incubated at 37°C for 4 days and either placed at 4°C for 24hrs to kill the embryo or were harvested immediately and embryos were decapitated using sterile scissors. Membranes were removed using sterile forceps and virus was harvested as previously described (Kotwal and Abrahams, 2004). To determine viral titre, purified stocks were serially diluted in PBS as above and inoculated onto CAMs in triplicate or quadruplicate and incubated as above. The p.f.u./ml was determined by spreading CAMs out on Petri dishes, counting the mean number of pocks per dilution and multiplying by 10 x the dilution factor.

2.2.2 Histology

Ten-day old, non-SPF eggs were inoculated in triplicate with 10^3p.f.u. of each virus and incubated at 37ºC for 5 days. The uninfected control was inoculated with PBS containing penicillin (500U/ml), streptomycin (100µg/ml) and fungin (1µg/ml). Membranes were harvested as above and examined macroscopically. One representative membrane which showed typical virus growth patterns was chosen for analysis, photographed and placed in ~50ml of 10% buffered formalin [formaldehyde (37-40%), $NaH_2PO_4.H_2O$ (35.03M), Na_2HPO_4 (anhydrous, 21.84M), made up to 1L with distilled water; pH 7.4]. Membranes were stored in formalin for 24hrs before removing infected areas of membrane measuring ~ 0.5cm in width x 1-2cm in length and exhibiting relative conformity in pock density between viruses. These sections were rolled up, embedded in paraffin, sectioned (4µm) and stained with haematoxylin and eosin (H&E) by Dr. Ross Millin and Ms. Anna Marie Beukes, Pathcare, Cape Town, South Africa and Morea Peterson, Department of Human Biology, University of Cape Town, Cape Town, South Africa). Slides were examined using a light microscope at various magnifications and photographed. Dr. Tertius Gous, a consultant avian pathologist was consulted on all findings.

2.2.3 DNA Extraction

DNA was extracted by mixing (by inversion) 200-300ul of purified virus stock with an equal volume of vertrel (1,1,1,2,3,4,4,5,5,5-decafluoropentane) (DuPont), followed by centrifugation at 13000 rpm in a benchtop centrifuge (Eppendorf 5417C) for 5 mins. The supernatant was retrieved and freeze/thawed three times to break open any cells of CAM origin. The solution was then treated with DNase (250U/ml) for 30 mins at 37°C to degrade host DNA, followed by inactivation of DNase for 30mins at 80°C. Proteinase K was then added (1µg/ml) and the virus incubated for 30mins at 55°C. An equal volume of lysis buffer (10% N-lauryl sarcosinate, 50mM Tris/HCL (pH7.8) and 200mM β-mercaptoethanol) was added followed by incubation at 55°C overnight. An equal volume of phenol:chloroform (1:1) was then added, the tube inverted several times and centrifuged at 13 000rpm for 5 mins. The aqueous phase was removed followed by addition of RNase (100µg/ml) and incubation at 37°C for 60mins. A further phenol:chloroform (1:1) extraction step was performed followed by

one phenol:chloroform:isoamyl alcohol (1:1:24) extraction step. DNA was then precipitated using ice cold NaAc (3M) and EtOH (70%) and pelleted as per standard protocol.

2.2.4 Polymerase Chain Reaction

P4b sequences of the novel isolates were amplified by PCR using 2x Immomix PCR mix (Bioline, London, U.K.) according to the manufacturer's instructions in a GeneAmp PCR thermocycler (Applied Biosystems, Foster City, California, USA). All reactions used the previously described primers M2925: 5'-CAGCAGGTGCTAAACAACAA-3' and M2926: 5'-CGGTAGCTTAACGCCGAATA-3' (Weli *et al*., 2004) and standard thermocycling conditions. PCR products were purified using the DNA Clean and Concentrator-25 kit (Zymo Research, Irvine, California, USA) and sequenced using the BigDye Terminator sequencing kit v3.1 (Applied Biosystems) and ABI3130XL sequencer (Applied Biosystems) at the Central Analytical Facility at the University of Stellenbosch, Stellenbosch, South Africa (http://academic.sun.ac.za/saf.html).

2.2.5 Phylogenetic Analysis

Additional sequences (n=22) representative of each ChPV genus as well as representative avian poxvirus sequences (n=8) were obtained from Genbank and all sequences were truncated to 402nt to match the shortest amplicon length. An MSA was produced using the CLC Bio proprietary progressive alignment algorithm in the CLC Main Workbench 7.6.1 (www.clcbio.com, Aarhus, Denmark) to compare the five novel sequences to the 30 sequences from Genbank. A neighbour-joining (NJ) phylogenetic tree (Saitou and Nei, 1987) was constructed using MEGA version 7 (Kumar *et al*., 2016), with 100 replicate bootstrap samplings (Felsenstein, 1985). The same analysis was then repeated using all avian poxvirus P4b sequences available from Genbank (n=365). Sequences were truncated to 402nt, aligned, and an NJ phylogenetic tree constructed, as above. Of the 365 sequences, 73 were selected that grouped in subclades A2 and A3 as well as one sequence representing each of the remaining subclades. The 73 selected isolates are listed in Table 2.2. For clarity, the tree was further collapsed to only show sequences from selected

isolates. For both trees, bootstrap values >80% are indicated on the tree at branch points and branch lengths are indicated by substitutions per site on the scale bar, bottom left.

2.3 RESULTS

2.3.1 Pock Morphology and Membrane Histology

Compared to the uninfected membrane, each of the viruses tested including the controls (FWPV and CNPV) exhibited different macroscopic growth characteristics on CAMs. These characteristics were classified as pock morphology (size and colour) and degree of thickening of the membrane. FWPV, CNPV, PEPV and Pi2 did not cause a significant degree of membrane thickening compared to the uninfected control. FGPV infection caused extensive membrane thickening but small, white pocks were still visible whereas Pi1 and Pi4 stood out, as at this titre, no individual pocks were visible due to the extreme inflammation of the membrane. Pocks were visible at lower titres (data not shown) as there was a lesser degree of tissue inflammation present and exhibited the same typical features. FWPV and PEPV infection produced similar looking small, white pocks with those of PEPV appearing slightly less raised. CNPV pocks were small and yellow in colour while Pi2 infection resulted in large, raised, white pocks with slightly pink centres.

Each of the viruses also induced different degrees and types of tissue pathology in the CAM although the features that are hallmarks of poxvirus infection were present in all cases. These common features included intracytoplasmic inclusions/Bollinger bodies which represent sites of viral replication and some level of both mesodermal and epidermal hyperplasia which results from an increase in the rate of cellular reproduction. Histological analysis of infected membranes also revealed varying degrees of angiogenesis (formation of new blood vessels), fibroplasia (formation of fibrous tissue), vacuolisation (increased vacuole formation), ballooning degeneration (cell swelling), sloughing (shedding of dead cells) and necrosis (tissue death) as well as infiltration of heterophils (the avian analogue of mammalian neutrophils) and

lymphocytes. Table 2.1 below outlines the results of this analysis as well as the clade that each virus was found to belong to according to 2.3.2 below.

The viruses that caused significant macroscopic CAM inflammation (Pi1, Pi4 and FGPV) caused greater mesodermal hyperplasia and less epidermal hyperplasia as well as low levels of fibroplasia and few inclusion bodies. FGPV infection also resulted in focal areas of hyperkeratosis. CAM infection with isolates Pi2 and Pi4 resulted in marked immune cell (lymphocytes and heterophils) infiltration compared to the other isolates. Infection with PEPV resulted in generalised hyperplasia of the allantoic epithelium and infection with both PEPV and Pi1 resulted in papilliform projections of allantoic epithelium, outwards away from the mesoderm. Infection with CNPV resulted in several focal regions exhibiting one or more of the following: necrotic or keratinaceous crusts, fibroplasia and chorionic epithelial hyperplasia.

Table 2.1: Comparison of pock morphology and histology of uninfected and infected CAMs of embryonated chicken eggs. Two photographs are shown for each isolate: one macroscopic view of the CAM and one histologic section of the CAM stained with H&E at 100x magnification. Thirteen criteria were used to assess CAM infection and tissue response, and these are classified on a gradient using one to four + marks (+: minimal; ++: moderate; +++: extensive; ++++: extreme) and – if the characteristic was not evident. EH = epidermal hyperplasia. This work was done in conjunction with Dr Kristy Offerman and as such, photographs of the uninfected and CNPV infected CAMs are also present in her PhD thesis “Investigation of local South African avipoxviruses as potential vaccine vectors” (2014).

Table 2.1: Comparison of pock morphology and histology of uninfected and infected CAMs of embryonated chicken eggs.

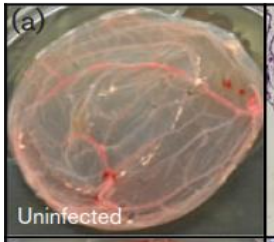
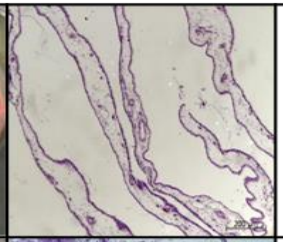
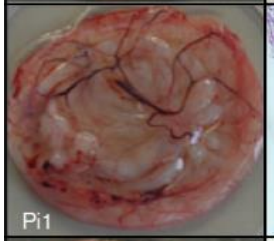
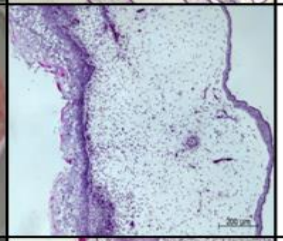
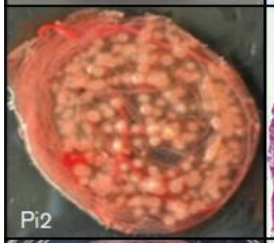
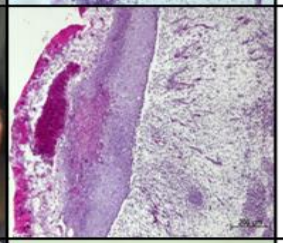
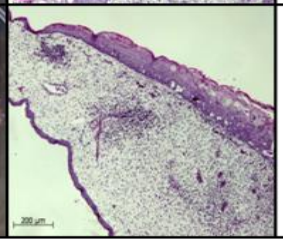
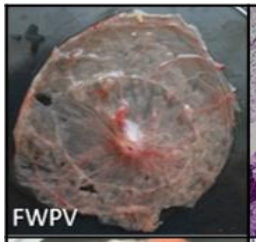
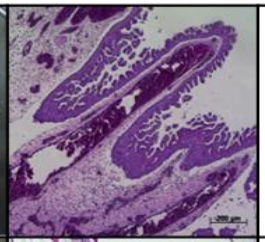
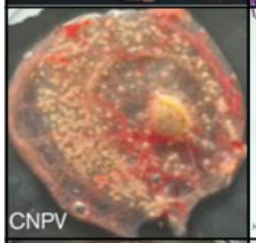
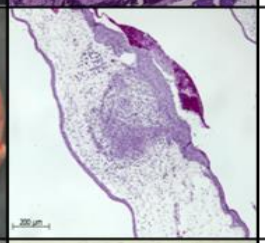
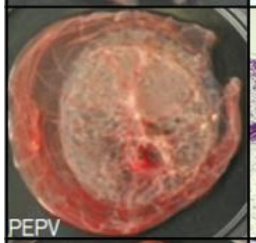
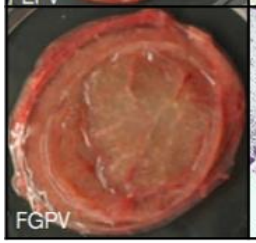
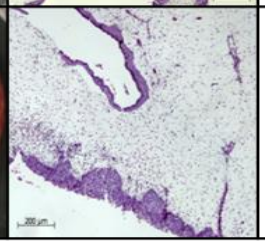
CAM	HISTOLOGY	Macroscopic thickening	Chorionic EH	Allantoic EH	Mesodermal hyperplasia	Angiogenesis	Fibroplasia	Inclusions	Vacuolisation	Ballooning degeneration	Sloughing	Necrosis	Lymphocytes	Heterophils	CLADE
		-	-	-	-	-	-	-	-	-	-	-	-	-	N/A
		+ + + +	+ +	+	+ +	+ + +	+	+	+ +	+ + +	+ +	+ +	+ +	+ +	A3.1
		+	+ + +	+ + + +	+ + +	+ + + +	+ +	+ + +	+ +	+ + +	+ +	+ + +	+ + + +	+ + +	+ + + +	A2
		+ + + +	+ + +	+ +	+ +	+	+	+	+ +	+ +	+	+	+ + +	+ + +	A2

Table 2.1 continued…

CAM	HISTOLOGY	Macroscopic thickening	Chorionic EH	Allantoic EH	Mesodermal hyperplasia	Angiogenesis	Fibroplasia	Inclusions	Vacuolisation	Ballooning degeneration	Sloughing	Necrosis	Lymphocytes	Heterophils	CLADE
 FWPV		+	+ + +	+ +	+	+	+ +	+ +	+ +	+	+ + +	+ + +	+ +	+	A1
 CNPV		+	+ +	+ + +	+ +	+	+ + +	+ +	+ +	+ +	+ +	+ + +	+ +	+	B1
 PEPV		+	+ +	+ +	+	+	-	+ + +	+	+ +	+	+	+	+	A2
 FGPV		+ + +	+	+	+ + +	+ +	-	-	+	+ +	+ + +	+	+	+	A3

2.3.2 Phylogenetic Analysis

Avian poxviruses are known to be host restricted to avian species, but a preliminary analysis was conducted to ensure that all isolated viruses were in fact avian poxviruses. A nucleotide alignment of P4b sequences and phylogenetic tree were constructed using the five, novel, South African isolates, representative sequences from all genera of the *Poxviridae* as well as representative sequences from each of the three confirmed (A, B, and C) and two proposed (D and E) avipoxvirus subclades. This analysis showed all of the novel, South African isolates to group in clade A in the avian poxvirus genus with strong bootstrap support and therefore viruses from other clades were excluded from further phylogenetic analysis except for use of representative sequences as placeholders (Figure 2.3). GC content is known to vary between poxvirus genera and can affect construction of phylogenetic trees. In this case, nucleotide alignments were used regardless, as the tree generated (Fig.2.3) closely resembled trees generated using amino acid alignments.

Although other genes (H3L, DNA polymerase, VLTF-1) have been used for phylogenetic analyses their usefulness is limited due to the small number of isolates that have been sequenced at these loci. To assess their placement within the avian clade, a second tree was constructed (Figure 2.4) including all available P4b sequences from Genbank (isolates with sequences shorter than 402nt were not included in the analysis and isolates from Crimson Rosella were also excluded as the gene was sequenced in a different region). Sequences that grouped outside of clade A were removed and the alignment redone to include only clade A sequences and representative sequences from the other clades. The novel isolates were found to group in three different subclades as discussed below. PEPV, Pi2 and Pi4 grouped in subclade A2 with 100% nt identity to each other and all other A2 isolates except three with which they had two nucleotide differences (Little brown dove|India|2009|HM481408; Indian Peafowl|Hungary|2003|KC017975; Great bustard|Hungary|2003|KC017970). The lesser flamingo isolate grouped in clade A3, with 100% nt identity to six other isolates and between 1-4 differences with the remaining isolates. Isolate Pi1 grouped in A3.1, a minor subclade annotation proposed by Offerman *et al*., 2013. All isolates in A3.1 show 100% nt identity to each other as do the three isolates in A3.1a with only one nucleotide difference

between isolates in A3.1 and A3.1a. 36 of the 39 isolates that grouped in clade A2 were 100% identical at the nucleotide level so clade A2 was collapsed in the phylogenetic tree and comparison table (Table 2.2) for clarity. Only clade A2 isolates from South Africa are shown as well as the three isolates that differ in percentage identity.

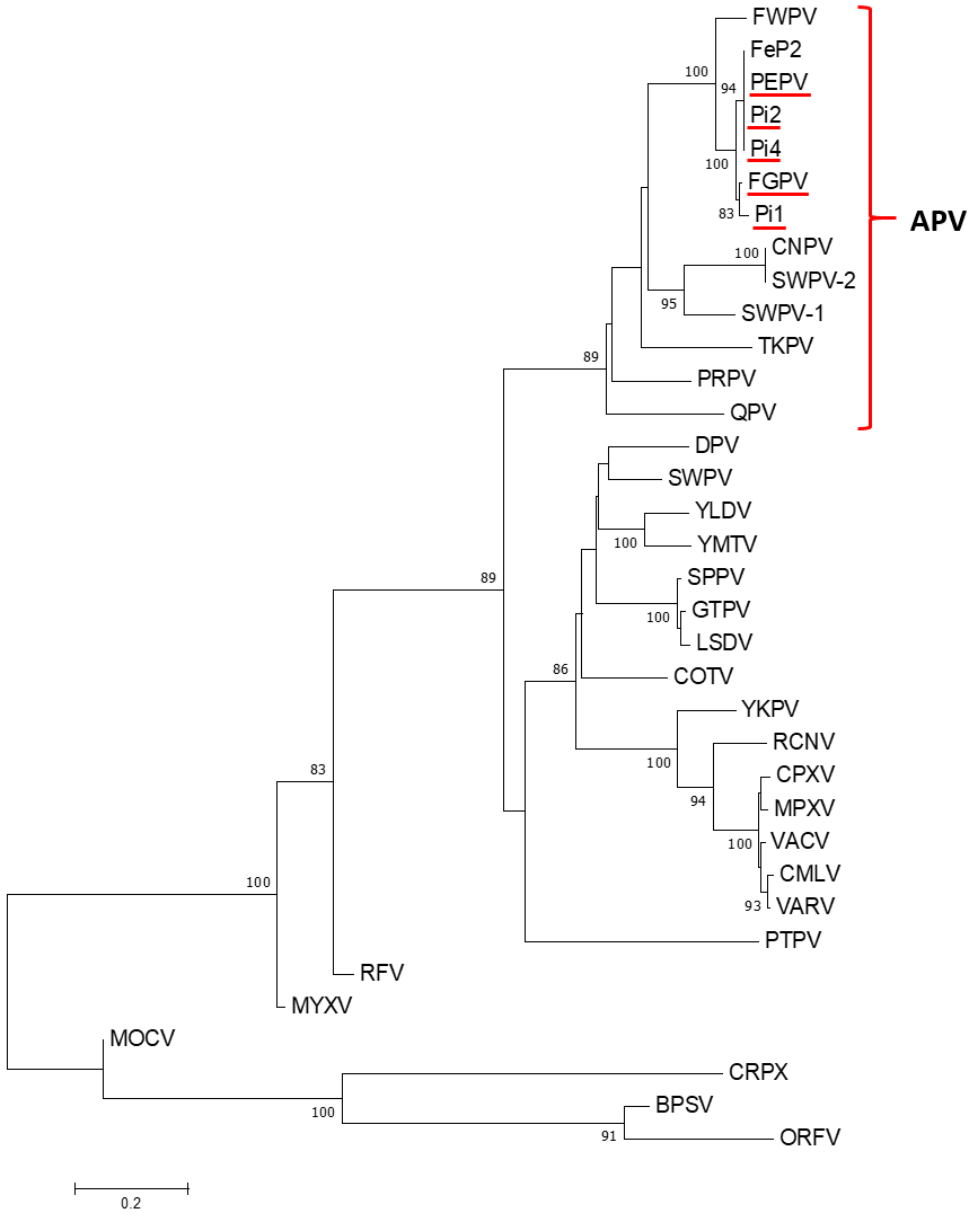


Figure 2.3: Phylogenetic tree based on the nucleotide alignment of partial P4b sequences of the five novel avian poxviruses and representative chordopoxviruses. The avian poxvirus genus is highlighted with a red bracket and isolates Pi1, Pi2, Pi4, PEPV and FGPV are underlined in red. Bootstrap values ≥80% and branch length in number of substitutions per site are shown.

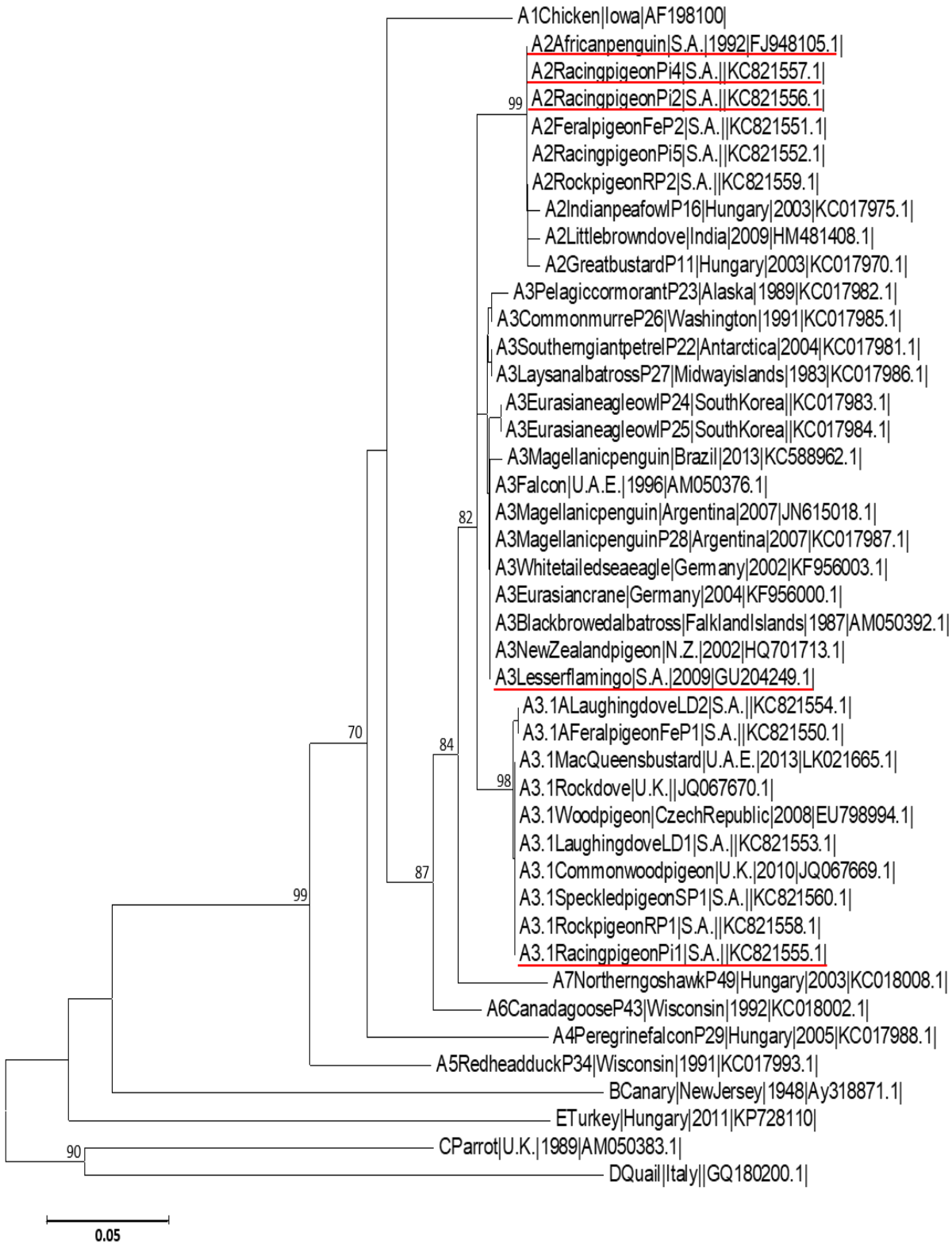


Figure 2.4: Phylogenetic tree based on the nucleotide alignment of partial P4b sequences of clade A2 and clade A3 avian poxviruses. Isolates Pi1, Pi2, Pi4, PEPV and FGPV are underlined in red. The clade that each isolate belongs to precedes its name. Clades A1, A2, A4-A7, B and C are collapsed for clarity and use representative sequences. Bootstrap values ≥80% and branch length in number of substitutions per site are shown.

Table 2.2 shows the full list of 73 sequences used to construct the subclade A2 and A3 tree as well as the representative sequences used for the other clades and subclades before removing isolates that showed 100% nucleotide identities.

A pairwise comparison table (Table 2.3) was constructed based on the nucleotide alignment of representative sequences from each of the clades, subclades and the minor subclade A3.1. Subclades within clade A showed a greater heterogeneity between individual subclades ranging from 85-90% between representative isolates from subclades A4 and A5 compared to each other and the representatives from other subclades and 90-99% between representative isolates from subclades A1-A3, A6 and A7. All clade A subclades showed between 72-79% identity to representative isolates from other clades and subclades. Within clade B, subclades B1 and B2 showed the greatest degree of identity at ~98%, followed by subclade B3 and B4 and lastly subclades B1 and B2, compared to B3 and B4 at 81-82%. Subclade C, and the two proposed subclades D and E all showed between 71-77% identity between each other and representatives from subclades A and B.

Table 2.2: List of 73 avian poxvirus isolates used in comparison of partial P4b nucleotide sequences

A1 Chicken|Iowa|AF198100|	A2 Little brown dove|India|2009|HM481408.1|
A2 Rock dove P7|Georgia|1995|KC017966.1|	A2 Indian peafowl P16|Hungary|2003|KC017975.1|
A2 Eastern imperial eagle P8|Hungary|2000|KC017967.1|	A2 Great bustard P11|Hungary|2003|KC017970.1|
A2 Booted eagle P17|Spain|2000|KC017976.1|	A3 Falcon|U.A.E.|1996|AM050376.1|
A2 Red legged partridge P18|Spain|2000|KC017977.1|	A3 Magellanic penguin|Argentina|2007|JN615018.1|
A2 Red kite P19|Spain|2003|KC017978.1|	A3 Magellanic penguin P28|Argentina|2007|KC017987.1|
A2 Rock dove P9|Hungary|2003|KC017968.1|	A3 White tailed sea eagle|Germany|2002|KF956003.1|
A2 Pigeon|||AY530303.1|	A3 Eurasian crane|Germany|2004|KF956000.1|
A2 Common chaffinch|Spain|1936|HM627225.1|	A3 Black browed albatross|Falkland Islands|1987|AM050392.1|
A2 Eurasian stone curlew|Spain|1980|HM627224.1|	A3 Lesser flamingo|S.A.|2009|GU204249.1|
A2 Ostrich|||AY530305.1|	A3 New Zealand pigeon|N.Z.|2002|HQ701713.1|
A2 Common buzzard|Italy|2005| EF016108.1|	A3 Common murre P26|Washington|1991|KC017985.1|
A2 Gyrfalcon|N.Italy|2006|GQ180210.1|	A3 Laysan albatross P27|Midway islands|1983|KC017986.1|
A2 Canary|N.Italy|1993|GQ180208.1|	A3 Southern giant petrel P22|Antarctica|2004|KC017981.1|
A2 Grey partridge|N.Italy|1986|GQ180204.1|	A3 Magellanic penguin|Brazil|2013|KC588962.1|
A2 Red legged partridge P21|Spain|2002|KC017980.1|	A3 Eurasian eagle owl P25|South Korea||KC017984.1|
A2 Booted eagle P20|Spain|2003|KC017979.1|	A3 Eurasian eagle owl P24|South Korea||KC017983.1|
A2 Fowl|India||DQ873810.1|	A3 Pelagic cormorant P23|Alaska|1989|KC017982.1|
A2 Quail|India||DQ873809.1|	A3.1 MacQueen's bustard|U.A.E.|2013|LK021665.1|
A2 Common wood pigeon|India|2009|HM481409.1|	A3.1 Rock dove|U.K.||JQ067670.1|
A2 Pigeon|India||DQ873811.1|	**A2 Great bustard P15|Spain|2003|KC017974.1|**
A2 Pigeon|Egypt|2011|JQ665840.1|	A3.1 Wood pigeon|Czech Republic|2008|EU798994.1|
A2 African penguin|S.A.|1992|FJ948105.1|	A3.1 Laughing dove LD1|S.A.||KC821553.1|
A2 Pigeon|Tanzania|2013|KJ913659.1|	A3.1 Common wood pigeon|U.K.|2010|JQ067669.1|
A2 Rock dove P6|Hawaii|1994|KC017965.1|	A3.1 Speckled pigeon SP1|S.A.||KC821560.1|
A2 Racing pigeon Pi4|S.A.||KC821557.1|	A3.1 Rock pigeon RP1|S.A.||KC821558.1|
A2 Racing pigeon Pi2|S.A.||KC821556.1|	A3.1 Racing pigeon Pi1|S.A.||KC821555.1|
A2 Feral pigeon FeP2|S.A.||KC821551.1|	A3.1A Laughing dove LD2|S.A.||KC821554.1|
A2 Racing pigeon Pi5|S.A.||KC821552.1|	A3.1A Feral pigeon FeP1|S.A.||KC821550.1|
A2 Turkey|Italy|1998|AM050388.1|	A6 Canada goose P43|Wisconsin|1992|KC018002.1|
A2 Pigeon||1975|AM050385.1|	A7 Northern goshawk P49|Hungary|2003|KC018008.1|
A2 Rock dove P12|Hungary|2005|KC017971.1|	A4 Peregrine falcon P29|Hungary|2005|KC017988.1|
A2 Rock dove P10|Hungary||KC017969.1|	A5 Redhead duck P34|Wisconsin|1991|KC017993.1|
A2 Turkey|U.K.|1966|AM050387.1|	B Canary|New Jersey|1948|AY318871.1|
A2 Houbara bustard|Morocco|2011|LK021666.1|	C Parrot|U.K.|1989|AM050383.1|
A2 MacQueen's bustar|U.A.E.|2012|LK021662.1|	D Quail|Italy||GQ180200.1|
A2 Rock pigeon RP2|S.A.||KC821559.1|	E Turkey|Hungary|2011|KP728110|

Table 2.3: Pairwise comparison of nucleotide alignments of partial P4b sequences of avian poxviruses. Sequences belonging to subclades A2 and A3 as well as representative sequences from subclades A1, A4, A5, A6, A7 and clades B, C, D and E are shown. Percentage identities are shown in the bottom left and number of nucleotide differences in the top right.

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
A1	1		37	33	36	35	33	40	44	53	89	95	89	89	98	103	103
A2	2	90.80		10	14	14	21	25	41	46	95	97	81	85	99	100	109
A3	3	91.79	97.51		8	9	15	19	43	41	96	100	81	86	101	99	107
A3.1	4	91.04	96.52	98.01		1	21	21	47	43	100	102	85	89	101	101	109
A3.1A	5	91.29	96.52	97.76	99.75		22	22	47	42	99	101	84	88	101	101	108
A6	6	91.79	94.78	96.27	94.78	94.53		24	41	41	91	95	80	87	95	101	101
A7	7	90.05	93.78	95.27	94.78	94.53	94.03		52	42	97	103	85	94	104	108	104
A4	8	89.05	89.80	89.30	88.31	88.31	89.80	87.06		54	94	92	85	81	102	99	108
A5	9	86.82	88.56	89.80	89.30	89.55	89.80	89.55	86.57		86	88	84	90	98	99	108
B1	10	77.86	76.37	76.12	75.12	75.37	77.36	75.87	76.62	78.61		6	76	73	98	110	113
B2	11	76.37	75.87	75.12	74.63	74.88	76.37	74.38	77.11	78.11	98.51		73	70	96	108	115
B3	12	77.86	79.85	79.85	78.86	79.10	80.10	78.86	78.86	79.10	81.09	81.84		40	103	101	113
B4	13	77.86	78.86	78.61	77.86	78.11	78.36	76.62	79.85	77.61	81.84	82.59	90.05		92	100	105
C	14	75.62	75.37	74.88	74.88	74.88	76.37	74.13	74.63	75.62	75.62	76.12	74.38	77.11		97	94
E	15	74.38	75.12	75.37	74.88	74.88	74.88	73.13	75.37	75.37	72.64	73.13	74.88	75.12	75.87		116
D	16	74.38	72.89	73.38	72.89	73.13	74.88	74.13	73.13	73.13	71.89	71.39	71.89	73.88	76.62	71.14	

2.4 DISCUSSION

This basic characterisation of the novel isolates was undertaken in order to select two viruses for full genome sequencing and analysis. In terms of macroscopic CAM morphology and pock formation, FWPV, PEPV and FGPV produced similar looking pocks with minimal membrane inflammation for FWPV and PEPV and extensive inflammation for FGPV while infection with Pi1 and Pi4 resulted in extreme membrane inflammation with no individual pocks visible. Both CNPV and Pi2 produced distinct pocks with minimal membrane inflammation.

Comparison of CAM pathology induced by infection with isolates from different clades and subclades, host species, host disease/symptoms and location of isolation revealed no significant trends except that the CAMs infected with the three isolates from pigeons showed increased infiltration of heterophils compared to the other isolates. Avian heterophils are involved in both detecting and destroying pathogens and play an important role in avian immunology. Also noted was that the infections resulting in extensive macroscopic thickening of the membrane showed increased mesodermal hyperplasia compared to epidermal hyperplasia.

Overall, PEPV and FGPV gave minimal cumulative responses of 17+ and 19+, while isolate Pi2 gave the maximum response of 38+. The remaining isolates gave cumulative responses of 24-28+. This result links to the pathology seen in infected birds with the disease in penguins and flamingos being very mild and the disease in the pigeon infected with Pi2 being much more severe. Comparison of the presence/absence of TGF-β, β-NGF, EGF, Serpin and IL-18 BP homologues (reported in Chapter 3) showed no correlation to the presence/absence or degree of CAM hyperplasia. This is not a definitive result due to the small number of isolates tested with complete genome sequence data available. Once more genomes become available, a similar head to head study would be useful.

Phylogenetically, the three novel isolates grouped in clade A with Pi2, Pi4 and PEPV grouping in subclade A2, FGPV grouping in subclade A3, and Pi1 grouping in minor subclade A3.1. Clade A3 consists largely of isolates obtained from seabirds and

wading birds. This was the first reported case of poxvirus infection in free-living flamingos and the first reported case of a poxvirus isolated from a flamingo grouping outside of clade B. No correlation between phylogeny and geographic distribution was noted in this study as was also found in a larger study conducted in South Africa with several more isolates from different geographic locations (Offerman et al., 2013).

The above mentioned study undertaken by Kristy Offerman resulted in the genome sequencing of isolate FeP2 from a feral pigeon which grouped in clade A2 according to P4b analysis and clade A3iv according to analysis of three other loci (VLTF-1, H3L and FPV175-176) (Offerman *et al*., 2013). Infection with this isolate also caused substantial thickening of the CAM.

In this phylogenetic analysis, the isolate from a Quail (QP-241) shares a common ancestor with clade C isolates but formed an entirely new branch in the analysis by Manarolla *et al*. (2010), sharing a common ancestor with clades A, B and C. This is likely to be due to the difference in lengths of analysed amplicons (402nt in this analysis vs 504nt) as there are no gaps in the alignments and the phylogenetic trees were constructed using the same methods. Nevertheless, QP-241 shows a large degree of divergence to clade C isolates (24%) which is equivalent to inter-clade divergences suggesting that QP-241 forms a fourth clade and does not form a subclade within clade C. This would need to be confirmed by sequencing and analysis of more loci. Manarolla *et al*. (2010) note that isolate QP-241 was known to have peculiar characteristics compared to other avipoxviruses (Rinaldi *et al*., 1972 in Manarolla *et al*., 2010). Nucleotide substitutions in isolates from clade C as well as the QP-241 isolate were approximately twice as likely to result in non-synonymous amino acid changes compared to isolates in clade A (data not shown).

In order to select isolates for full genome sequencing, several factors were taken into consideration. In terms of CAM response to viral infection, comparison of genomes of isolates that caused extensive inflammation and isolates that caused minimal inflammation were desired and phylogenetically, intra- and inter-subclade comparisons were desired. Lastly, isolates that had a full history of host pathology and isolation details were preferred. Given the study of FeP2, undertaken by Kristy

Offerman, PEPV and FGPV were chosen for genome sequencing as FGPV and FeP2 both caused extensive membrane thickening and PEPV almost no thickening response. PEPV and FeP2 also group in subclade A2 allowing intra-subclade comparison and FGPV groups in subclade A3, allowing inter-subclade comparison. Full histories of both the PEPV and FGPV isolates are available (Stannard *et al*., 1998; Zimmermann *et al*., 2011).

CHAPTER 3: THE GENOME SEQUENCES OF PENGUINPOX VIRUS AND FLAMINGOPOX VIRUS

3.1 INTRODUCTION

3.2 MATERIALS AND METHODS

3.2.1 DNA extraction

3.2.2 Ion torrent Sequencing and Bioinformatics

3.3 RESULTS

3.3.1 Sequencing Statistics

3.3.2 Basic Genome Structure and Statistics

3.3.3 Core/Conserved Genes

3.3.4 Multigene Families

3.3.5 Gene Translocations/Duplications

3.3.6 Gene Disruptions/Deletions

3.3.7 Gene Insertions

3.3.8 Novel ORFs and Genes of Interest

3.3.9 Reticuloendotheliosis Virus

3.4 DISCUSSION

3.1 INTRODUCTION

Very little is currently known about the intra- and inter-subclade, genetic relationships of avian poxviruses due to the paucity of genomic sequence data available. Several more isolates from different clades and subclades need to be sequenced and compared to begin to elucidate these relationships as has been done within the Orthopoxvirus genus. Comparison of full genome sequences is the gold standard to determine phylogenetic and evolutionary relationships. As such, two isolates were chosen for full genome sequencing, and analysed and compared to published sequences.

3.2 MATERIALS and METHODS

3.2.1 DNA extraction – see chapter 2.2

3.2.2 Ion torrent sequencing and bioinformatics

DNA was sent to the Central Analytical Facility (CAF) at the University of Stellenbosch, in Stellenbosch, South Africa for full genome sequencing. One 316 chip was used for each single genome (PEPV, FGPV) followed by use of half of one 318 chip (FGPV) on the Ion Torrent Personal Genome Machine (PGM) (Life Technologies), according to the manufacturer's instructions. DNA was sheared ultrasonically using the Covaris S2 sample preparation system (Covaris Inc., USA).

Basic quality control was performed by Anelda van der Walt (CAF, University of Stellenbosch, Stellenbosch, South Africa) using Torrent Suite software (version 3.2.1). Reads were trimmed of adaptor sequences and further trimmed if average base quality values (Q value) were < 25 with window size = 11. Reads were discarded if read length was < 50nt and filtered to remove polyclonal reads. All reads that passed the above quality control filters were mapped to the chicken genome (*Gallus gallus* WASHUC2) using Newbler 2.6 (PEPV) or BLASTed to the chicken genome using CLC Genomics Workbench 4.7.1 (http://www.clcbio.com) (FGPV) to filter out reads of host origin.

Unmapped reads were used as input data for de novo assembly in Mira 3.4.0 (PEPV) and CLC Genomics Workbench 4.7.1 (FGPV). Open reading frames (ORFs) longer than 90nt with a methionine start codon (ATG) that were not contained within other ORFs were identified using CLC Genomics Workbench 4.7.1. These ORFs were annotated as potential genes and numbered from left to right if alignment to the NCBI nr database using BLASTn and/or BLASTp and/or BLASTx (Altschul *et al*., 1990) gave BLAST expect values of E ≤ 1e-5. ORFs were annotated as described by (Hendrickson *et al*., 2010). ORFs were annotated as intact (I) if the 5' end is intact and the ORF is ≥ 80% the length of the closest homologue. If the 5' end is intact but the ORF is < 80% the length of the closest homologue, it was annotated as truncated (T). If the 5' end is not intact the ORF was annotated as a fragment (F). If the ORF is > 20% of the closest homologue it was annotated as extended (E) at the 5' or 3' end. Inverted terminal repeats were located and confirmed by PCR and Sanger sequencing for PEPV but not for FGPV. Expression studies and functional analysis would be needed to determine whether fragmented and truncated ORFs are expressed and/or functional. The left most nucleotide was nominated as base 1, as the terminal hairpin loops were not sequenced.

The PEPV genome sequence was submitted to Genbank (accession # KJ859677) and described and discussed in an article published in BMC genomics (Offerman ***et al***., 2014). The PEPV and FeP2 gene annotations in Genbank have been changed relative to the annotations found in the journal article describing the sequences (Offerman *et al*., 2014). When referring to PEPV and FeP2 ORFs in this study, the original annotations from the publication are used which are referred to in Genbank as "old locus tag". The FGPV genome sequence was submitted to Genbank (accession # MF678796) and described and discussed in an article published in BMC genomics (Carulei *et al*., 2017).

3.3 RESULTS

3.3.1 Sequencing Statistics

PEPV: Ion torrent sequencing resulted in 3 197 371 reads with a mean read length of 203bp. 1 119 080 reads (35%) remained after QC and filtering of reads of host origin. Read assembly with MIRA resulted in one contiguous sequence of 306 862bp with an average coverage of 743x.

FGPV: Ion torrent sequencing on the 316-chip resulted in 3 745 381 reads with a mean read length of 180bp. 1 309 385 (35%) remained after QC and filtering of reads of host origin. On the 318 chip, 2 148 517 reads were generated with a mean read length of 251 bp. Only 570 143 (27%) remained after QC and filtering of reads of host origin. Read assembly of all reads from both sequencing runs resulted in one contiguous sequence of 293 130 bp with an average of 1090x coverage using CLC Genomics workbench.

3.3.2 Basic Genome Structure and Statistics

Table 3.1 outlines basic genome statistics of the PEPV and FGPV genomes compared to other published avian poxvirus genomes. All statistics for the PEPV and FGPV genomes were found to be within the ranges of previously analysed genome statistics. Statistics for the inverted terminal repeats (ITRs) of FGPV could not be determined as the ITRs were not resolved.

PEPV: The nucleotide composition across the whole genome was found to be 70.5% AT, while the composition of the potential coding regions was 69.8% AT. PEPV has identical, inverted, terminal repeat sequences of 5 766nt with AT contents of 65.7%. The ITRs contain a repeat region of 165bp, with 3.5 copies of a 47bp tandem repeat. PEPV contains 284 ORFs encoding proteins ranging from 34 to 1 922 amino acids in length and representing an approximate coding density of 85%. Relative to their closest orthologues, 241 of these ORFs are annotated as complete genes, 14 as truncated genes, 28 as fragments of larger genes and one as extended. Relative to

FWPV, 26 ORFs are absent from the PEPV genome and 42 are inserted including 23 intact genes, 17 gene fragments and 2 truncated genes (Table 3.2).

Table 3.1: Genome statistics of PEPV and FGPV compared with each of the fully sequenced avian poxvirus genomes available in Genbank.

STATISTIC	FWPV	FP9	CNPV	TKPV	FeP2	PEPV	FGPV	SWPV1	SWPV2
Genome length (kbp)	288.5	266.1	359.9	188.5	282.3	**306.9**	**293.1**	326.9	351.1
Genome A + T content (%)	69.1	69.2	69.6	70.2	70.5	**70.5**	**70.5**	72.4	69.8
ITR length (bp)	9 520	10 158	6 491	1 540	4 682	**5 766**	**-**	6 106	4 992
# ORFs in the ITR	10	9	6	1	4	**5**	**-**	6	4
ITR A + T content (%)	65.5	66.9	63.0	67.1	64.8	**65.7**	**-**	69.0	64.9
# ORFs (genome)	260	244	328	171	271	**284**	**285**	310	312
Coding density (%)	85	83	91	88/90*	83	**85**	**92**	91	86
Shortest ORF (aa)	60	53	42	54	34	**34**	**37**	47	45
Longest ORF (aa)	1 949	1 949	1 951	1 810	1 937	**1 922**	**1 894**	1 939	1 916

*Coding density for TKPV is equal to 88% if calculated using the ORFs described in the initial paper (Banyai *et al*., 2015), but is equal to 90% when ORFs discovered in this research are included (Chapter 3.3.3)

Figure 3.1 is a schematic of the PEPV genome showing coloured blocks as ORFs numbered from left to right. ORFs depicted by white blocks are intact relative to their closest orthologue. ORFs that have been disrupted resulting in two fragments are named with the ORF number followed by a letter e.g. 066a and 066b. This schematic shows 50% of the truncated/fragmented ORFs to be located within the last 60kb of the genome. A cluster of eight truncated/fragmented ORFs is also present between pepv119-pepv131 which is within a region of difference identified in chapter 4.3.2. The remaining truncated/fragmented ORFs are dispersed throughout the genome.

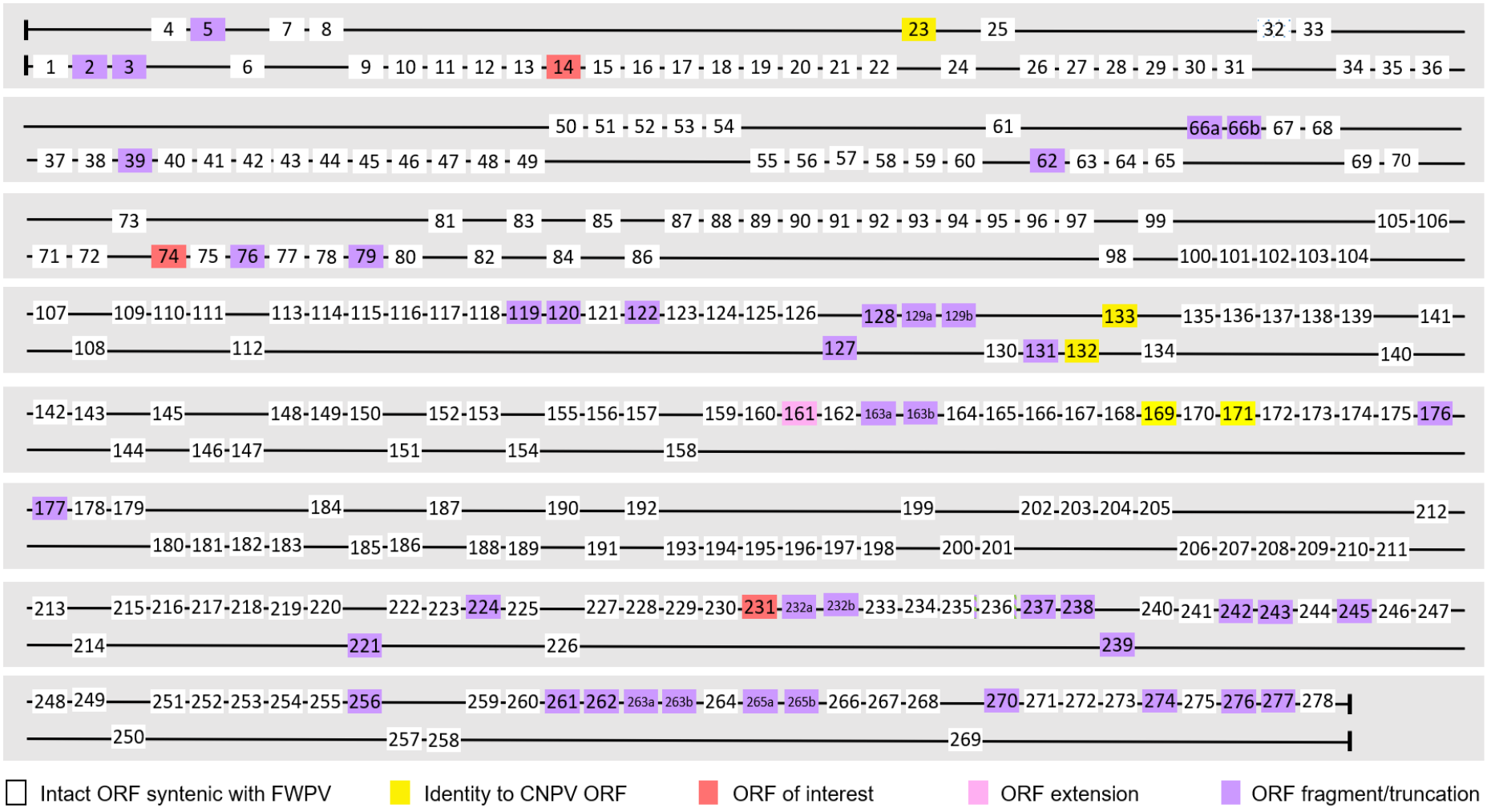


Figure 3.1: ORF schematic of the PEPV genome. The genome is depicted as double stranded, with ORFs shown as coloured blocks (not to scale), numbered from left to right. ORFs transcribed from left to right are depicted above and those transcribed from right to left depicted below.

Table 3.2: PEPV Open Reading Frames

ORF	ORF length (aa)	Status (I, F, T, E)	Closest homologue	Homologue length (aa)	PEPV vs homologue (% aa identity)	VACV homologue	Description / Putative function
1	204	I	fpv001	205	77,7	A40R	C-type lectin family
2	164	F	cnpv319	739	8,3		Ankyrin repeat family
3	115	F	fpv246	537	8,7		Hypothetical protein
4	222	I	fpv002	222	91.0		Hypothetical protein
5	46	F	cnpv006	182	14.8		Hypothetical protein
6	191	I	cnpv309/fpv241	196	33.7	M1L	Ankyrin repeat family
7	139	I	fpv250	140	80.9		Hypothetical protein
8	411	I	fpv006	418	85.0	C10L	C4L/C10L-like gene family protein
9	508	I	cnpv015/fpv162	528	42.7	B4R	Ankyrin repeat family
10	551	I	fpv244	668	24.5		Ankyrin repeat family
11	349	I	fpv010	355	82.8		Serpin family
12	288	I	fpv011	278	74.1		α-SNAP
13	518	I	fpv246	592	29.4		Hypothetical protein
14	174	I	IL10 (Ficedula albicollis$^{\alpha}$)	177	31.1		Interleukin-10
15	329	I	fpv012	331	78.0		Ankyrin repeat family
16	402	I	cnpv028	362	35.7	B4R	Ankyrin repeat family
17	437	I	fpv014	437	86.3		Ankyrin repeat family
18	171	I	fpv015	177	74.6		Hypothetical protein
19	238	I	fpv016	238	80.7		Hypothetical protein
20	245	I	fpv017	245	77.7		V-type Ig domain
21	683	I	fpv018	700	78.0		Ankyrin repeat family
22	113	I	fpv019	104	64.6		Hypothetical protein

ORF	ORF length (aa)	Status (I, F, T, E)	Closest homologue	Homologue length (aa)	PEPV vs homologue (% aa identity)	VACV homologue	Description / Putative function
23	189	I	cnpv037	171	43.0		Hypothetical protein (fragment in FPV)
24	429	I	fpv020	426	85.6	C10L	C4L/C10L-like family
25	334	I	fpv021	320	79.1		G-protein-coupled receptor family
26	579	I	fpv022	578	83.8		Ankyrin repeat family
27	434	I	fpv023	434	91.5		Ankyrin repeat family
28	595	I	fpv024	596	89.6		Ankyrin repeat family
29	203	I	fpv025	203	90.2		Hypothetical protein
30	509	I	cnpv044/fpv024	480	39.1	B4R	Ankyrin repeat family
31	406	I	fpv026	436	78.5		Ankyrin repeat family
32	332	I	fpv027	336	82.4		G-protein-coupled receptor family
33	180	I	fpv028	180	88.3		hypothetical protein
34	126	I	fpv029	124	90.5		Hypothetical protein
35	815	I	fpv030	817	83.4		Alkaline phosphodiesterase
36	341	I	fpv031	341	90.3		Ankyrin repeat family
37	368	I	Fp9.032	375	85.6		DNAse II
38	291	I	fpv033	287	85.9		α-SNAP
39	69	F	fpv034	415	6.5		Ankyrin repeat family
40	135	I	fpv035	135	96.3		Hypothetical protein
41	165	I	fpv037	162	67.9		Hypothetical protein
42	145	I	fpv038	145	91.7	F2L	dUTP pyrophosphatase
43	175	I	fpv039	175	80.0		Bcl-2
44	337	I	fpv040	337	92.3		Serpin family

ORF	ORF length (aa)	Status (I, F, T, E)	Closest homologue	Homologue length (aa)	PEPV vs homologue (% aa identity)	VACV homologue	Description / Putative function
45	220	I	fpv041	206	60.5		Hypothetical protein
46	564	I	fpv043	564	91.1	A50R	DNA ligase
47	358	I	fpv044	358	90.2		Serpin family
48	370	I	fpv046	370	82.8	A44L	Hydroxysteroid dehydrogenase
49	576	I	fpv047	612	74.1	A39R	Semaphorin
50	261	I	fpv048	261	94.6		GNS1/SUR4
51	154	I	fpv049	154	95.5	A1L	Late transcription factor VLTF2
52	552	I	fpv050	552	96.7	D13L	Rifampicin resistance, N3L protein
53	289	I	fpv051	289	94.5	D12L	mRNA capping enzyme
54	577	I	fpv052	637	86.7	D11L	NPH-1 transcription termination factor
55	225	I	fpv053	225	95.1	D10L	mutT motif; gene expression regulation
56	236	I	fpv054	231	92.4	D9R	mutT motif
57	274	I	fpv055	275	63.8		V-type Ig Domain
58	161	I	fpv056	161	96.3	D7R	RNA polymerase subunit RPO18
59	633	I	fpv057	633	98.6	D6R	Early transcription factor VETFs
60	791	I	fpv058	791	97.9	D5R	NTPase; DNA replication
61	219	I	fpv059	219	86.3		Deoxycytidine kinase
62	35	F	fpv060	188	14.9		CC chemokine family
63	112	I	fpv061	129	68.5		CC chemokine family
64	218	I	fpv062	218	97.7	D4R	Uracil DNA glycosylase

ORF	ORF length (aa)	Status (I, F, T, E)	Closest homologue	Homologue length (aa)	PEPV vs homologue (% aa identity)	VACV homologue	Description / Putative function
65	401	I	fpv063	400	79.4		Hypothetical protein
66a **66b**	42 134	T T	fpv064	200	18.0 62.7		Glutathione peroxidase
67	110	I	fpv065	111	78.4		Hypothetical protein
68	137	I	fpv066	122	73.7		Hypothetical protein
69	85	I	fpv067	90	86.8		HT motif family
70	143	I	fpv068	133	52.6		Hypothetical protein
71	269	I	fpv069	270	92.6	D3R	Virion protein
72	273	I	fpv070	273	84.6		T10 gene product
73	43	I	cnpv095/fpv070.5	45	51.1		Hypothetical protein
74	77	T	UBIQUITIN	98	76.5		Ubiquitin
75	287	I	fpv071	289	88.9		Hypothetical protein
76	130	F	cnpv012/fpv229	189	22.6		Hypothetical protein
77	104	I	fpv074	104	78.1		Hypothetical protein
78	211	I	fpv075	199	74.9		N1R/p28 family
79	51	F	fpv076	144	22.22		Beta-Nerve growth factor
80	125	I	fpv077	125	95.2	G4L	Glutaredoxin
81	225	I	fpv079	225	93.3	G2R	Putative elongation factor
82	103	I	fpv078	103	94.2	G3L	Hypothetical protein
83	336	I	fpv080	363	68.9		Transforming Growth Factor β (TGF-B)
84	624	I	fpv081	626	95.1	G1L	Metalloprotease
85	682	I	fpv082	682	94.1	I8R	DNA/RNA helicase/NPH-11
86	421	I	fpv083	421	96.2	I7L	Virion core proteinase

ORF	ORF length (aa)	Status (I, F, T, E)	Closest homologue	Homologue length (aa)	PEPV vs homologue (% aa identity)	VACV homologue	Description / Putative function
87	390	I	fpv084	390	94.4	I6L	DNA-binding protein
88	81	I	fpv085	81	90.1	I5L	IMV membrane protein
89	183	I	fpv086	183	89.1	J2R	Thymidine kinase
90	91	I	fpv087	91	92.3		HT motif family
91	290	I	fpv088	290	95.5	I3L	DNA binding phosphoprotein
92	65	I	fpv089	65	95.4	I2L	Hypothetical protein
93	311	I	fpv090	311	98.4	I1L	Virion protein
94	34	I	fpO3L	34	91.2	O3L	Orthologue of VACV O3L & MC043.1L
95	656	I	fpv091	656	88.3	O1L	Hypothetical protein
96	131	I	fpv092	131	93.9	E11L	Hypothetical protein
97	94	I	fpv093	94	90.4	E10R	Sulfhydryl oxidase ERV1
98	988	I	fpv094	988	92.9	E9L	DNA polymerase
99	280	I	fpv095	272	89.3	E8R	Hypothetical protein
100	571	I	fpv096	571	96.2	E6R	Hypothetical protein
101	1894	I	fpv097	1912	86.2		VARV B22R family
102	1812	I	fpv098	1802	86.7		VARV B22R family
103	1922	I	fpv099	1949	84.9		VARV B22R family
104	182	I	fpv100	182	97.8	E4L	RNA pol subunit RPO30
105	717	I	fpv101	717	93.4	E2L	Hypothetical protein
106	472	I	fpv102	472	97.5	E1L	Poly(A) polymerase large subunit, PAP-L
107	114	I	fpv103	114	100.0	F17R	DNA binding virion core phosphoprotein
108	210	I	fpv104	210	81.4		Hypothetical protein

ORF	ORF length (aa)	Status (I, F, T, E)	Closest homologue	Homologue length (aa)	PEPV vs homologue (% aa identity)	VACV homologue	Description / Putative function
109	149	I	fpv105	148	96.6	F15L	Hypothetical protein
110	99	I	fpv106	99	65.0		Conserved hypothetical protein
111	1779	I	fpv107	1777	85.4		VARV B22R family
112	377	I	fpv108	377	94.2	F13L	Virion envelope protein
113	639	I	fpv109	630	85.5	F12L	Virion release protein
114	452	I	fpv110	451	84.7	F11L	Hypothetical protein
115	443	I	fpv111	444	96.4	F10L	SER/THR protein kinase (virus assembly)
116	213	I	fpv112	213	93.0	F9L	Hypothetical protein
117	66	I	fpv113	66	98.5		Hypothetical protein
118	183	I	fpv114	183	95.6		HAL3 domain
119	84	F	cnpv004/fpv246	514	8.0		Hypothetical protien
120	66	F	Neosartorya fischeri NRRL 181	1174	1.8		Pfs, NACHT and Ankyrin domain protein
121	541	I	fpv115	542	78.0		Ankyrin repeat family
122	82	F	fpv116	120	56.7		CC-chemokine family
123	440	I	fpv117	440	90.9	G5R	Hypothetical protein
124	63	I	fpv118	63	98.4	G5.5R	RNA pol subunit RPO7
125	188	I	fpv119	188	99.5	G6R	Hypothetical protein
126	343	I	fpv120	343	95.9	G7L	Virion core protein
127	83	T	cnpv221/fpv124	281	26.0		N1R/p28 family
128	148	F	cnpv165/fpv124	346	35.8		N1R/p28 family
129a	1442	T	fpv122	1870	51.6		VARV B22R family
129b	501	T			22.9		VARV B22R family

ORF	ORF length (aa)	Status (I, F, T, E)	Closest homologue	Homologue length (aa)	PEPV vs homologue (% aa identity)	VACV homologue	Description / Putative function
130	1771	I	fpv123	1766	77.0		VARV B22R family
131	62	F	cnpv162/fpv080	149	29.5		Transforming Growth Factor (TGF)-beta-like protein
132	116	I	cnpv086	117	46.2		Tumour Necrosis Factor Receptor (TNFR)-like protein
133	217	I	cnpv170	212	80.2		Thymidylate Kinase
134	260	I	fpv126	260	99.2	G8R	VLTF-1
135	322	I	fpv127	336	93.8	G9R	Myristylated protein
136	243	I	fpv128	243	97.1	L1R	Myristylated protein
137	96	I	fpv129	96	87.5	L2R	Hypothetical protein
138	301	I	fpv130	301	93.7	L3L	Hypothetical protein
139	253	I	fpv131	253	94.1	L4R	DNA binding virion core VP8
140	129	I	fpv132	129	93.8	L5R	Putative membrane protein
141	148	I	fpv133	148	96.6	J1R	Hypothetical protein
142	308	I	fpv134	308	95.8	J3R	PolyA polymerase (PAPs)
143	186	I	fpv135	186	95.2	J4R	RNA pol Subunit RPO22
144	137	I	fpv136	137	90.5	J5L	Membrane protein
145	1287	I	fpv137	1287	97.6	J6R	RNA pol Subunit RPO147
146	166	I	fpv138	166	95.2	H1L	Protein tyrosine Phosphatase
147	190	I	fpv139	190	97.4	H2R	Hypothetical protein
148	333	I	fpv140	327	87.4	H3L	Virion env protein(p35)
149	799	I	fpv141	798	96.6	H4L	RNA polymerase associated protein RAP94
150	174	I	fpv142	174	90.2	H5R	Late transcription factor, VLTF-4

ORF	ORF length (aa)	Status (I, F, T, E)	Closest homologue	Homologue length (aa)	PEPV vs homologue (% aa identity)	VACV homologue	Description / Putative function
151	316	I	fpv143	316	97.2	H6R	DNA topoisomerase
152	152	I	fpv144	152	92.8	H7R	Putative 17 kDa protein
153	103	I	fpv145	103	87.38		Hypothetical protein
154	822	I	fpv146	851	94.5	D1R	mRNA capping enzyme, large subunit
155	84	I	fpv147	104	61.5		HT motif family
156	140	I	fpv148	139	90.7	D2L	virion protein
157	189	I	fpv149	186	85.7		Hypothetical protein
158	283	I	fpv150	276	82.9		N1R/p28 gene family protein
159	238	I	fpv151	235	83.7		Deoxycytidine Kinase (dCK)
160	127	I	fpv152	127	81.1		HT motif family
161	270	E	fpv154	150	41.2		Hypothetical protein
162	410	I	fpv155	408	85.9		N1R/p28 family
163a	51	T	fpv156	132	27.3		HT motif family
163b	96	T			60.6		Hypothetical protein
164	329	I	fpv157	311	79.6		N1R/p28 family
165	464	I	fpv158	464	93.3		Photolyase
166	246	I	fpv159	241	84.6		N1R/p28
167	156	I	fpv160	156	94.2		Hypothetical protein
168	149	I	fpv161	157	82.8		N1R/p28 family
169	133	I	cnpv210/fpv075	131	43.0		N1R/p28 family
170	45	I	cnpv211/fpv161.5	54	32.7		Hypothetical protein
171	160	I	cnpv212/fpv124	176	52.8		N1R/p28 family
172	603	I	fpv162	603	81.1		Ankyrin repeat family

ORF	ORF length (aa)	Status (I, F, T, E)	Closest homologue	Homologue length (aa)	PEPV vs homologue (% aa identity)	VACV homologue	Description / Putative function
173	256	I	fpv163	263	77.2		N1R/p28 family
174	741	I	fpv162	603	28.4		Ankyrin repeat family
175	386	I	fpv164	383	65.5		Hypothetical protein
176	150	F	cnpv041/fpv131	430	13.0		Ankyrin repeat family
177	225	I	fpv165	225	94.7	A2L	Late transcription factor VLTF-3
178	72	I	fpv166	72	95.8	A2.5L	Virus redox protein
179	658	I	fpv167	658	98.6	A3L	Virion core protein P4b
180	255	I	fpv168	288	67.7	A4L	Immunodominant virion protein
181	169	I	fpv169	167	97.0	A5R	RNA pol subunit RP019
182	374	I	fpv170	375	94.1	A6L	Hypothetical protein
183	709	I	fpv171	709	97.3	A7L	Early transcription factor large subunit, VETF-L
184	301	I	fpv172	301	98.0	A8R	Intermediate transcription factor VITF-3
185	76	I	fpv173	76	96.1	A9L	Hypothetical protein
186	891	I	fpv174	891	96.0	A10L	Virion core protein P4a
187	272	I	fpv175	274	96.7	A11R	Hypothetical protein
188	174	I	fpv176	171	90.8	A12L	Virion protein
189	70	I	fpv177	68	82.9		Hypothetical protein
190	71	I	fpv178	71	84.5	A13L	Virion protein
191	91	I	fpv179	91	95.6	A14L	Virion envelope protein
192	53	I	fpv179.1	54	98.1	a14.5	Virion envelope protein
193	97	I	fpv180	97	95.9	A15L	Hypothetical protein
194	369	I	fpv181	369	92.4	A16L	Putative myristylated protein

ORF	ORF length (aa)	Status (I, F, T, E)	Closest homologue	Homologue length (aa)	PEPV vs homologue (% aa identity)	VACV homologue	Description / Putative function
195	198	I	fpv182	198	97.5	A17L	Phosphorylated virion membrane protein
196	462	I	fpv183	462	96.5	A18R	DNA helicase (transcription elongation)
197	88	I	fpv184	88	95.5	A19L	Hypothetical protein
198	113	I	fpv186	113	92.0	A21L	Hypothetical protein
199	432	I	fpv185	433	94.5	A20R	Processivity factor
200	161	I	fpv187	156	88.8	A22R	Hypothetical protein
201	383	I	fpv188	383	94.8	A23R	Intermediate transcription factor VITF-3
202	1157	I	fpv189	1161	97.9	A24R	RNA pol subunit RPO132
203	610	I	fpv190	620	87.2	A25L	A-type inclusion protein
204	472	I	fpv191	474	92.2	A26L	A-type inclusion protein
205	140	I	fpv192	141	95.7	A28L	Hypothetical protein
206	302	I	fpv193	302	90.4	A29L	RNA pol subunit RPO35
207	74	I	fpv194	74	96.0	A30L	Hypothetical protein
208	38	I	fp9.194.1		84.2		A30.5L orthologue
209	113	I	fpv195	113	90.3	A31R	Hypothetical protein
210	120	I	fpv196	120	80.8		Hypothetical protein
211	304	I	fpv197	301	93.1	A32L	Virion assembly protein
212	151	I	fpv198	173	91.9	A34R	C-type lectin-like protein
213	220	I	fpv199	219	79.1		V-type Ig Domain
214	263	I	fpv200	265	82.6		V-type Ig domain
215	276	I	fpv201	283	85.9		Hypothetical protein
216	285	I	fpv203	285	83.2		Tyrosine protein kinase

ORF	ORF length (aa)	Status (I, F, T, E)	Closest homologue	Homologue length (aa)	PEPV vs homologue (% aa identity)	VACV homologue	Description / Putative function
217	342	I	fpv204	342	92.4		Serpin family
218	220	I	fpv205	218	82.7		Hypothetical protein
219	308	I	fpv206	308	88.3		G-protein-coupled receptor
220	92	I	fpv207	100	79.0		Hypothetical protein
221	150	F	cnpv279	169	45.8		Beta-NGF-like protein
222	193	I	fpv208	214	41.1		Hypothetical protein
223	137	I	fpv209	130	72.1		HT motif family
224	107	F	cnpv 283	111	48.7		CC chemokine family
225	123	I	fpv211	125	79.4	C11R	Epidermal Growth Factor-like protein
226	304	I	fpv212	303	82.2	B1R	Serine/threonine protein kinase
227	162	I	fpv213	162	90.7		Hypothetical protein
228	124	I	fpv214	124	79.2		IL-18BP
229	74	I	fpv215	74	94.6		Hypothetical protein
230	294	I	fpv216	296	86.2		Ankyrin repeat family
231	143	I	Tanapox 67R	178	27.1		67R Tanapox host range protein
232a	92	T	fpv217	328	24.1		Hypothetical protein
232b	210	T			55.5		Hypothetical protien
233	474	I	fpv218	461	77.0		Ankyrin repeat family
234	441	I	fpv219	434	83.7		Ankyrin repeat family
235	183	I	fpv221	183	88.5	A47L	A47L homolog
236	747	I	fpv222	747	83.4		Ankyrin repeat family
237	69	F	fpv223	141	33.3		Ankyrin repeat family

ORF	ORF length (aa)	Status (I, F, T, E)	Closest homologue	Homologue length (aa)	PEPV vs homologue (% aa identity)	VACV homologue	Description / Putative function
238	59	F	fpv224	146	32.9		Ankyrin repeat family
239	61	F	fpv225	104	45.7		Vaccinia B20R homolog
240	293	I	fpv226	292	88.4	B1R	Serine/threonine protein kinase
241	360	I	fpv227	361	86.2		Ankyrin repeat family
242	48	F	fpv228	525	5.1		Ankyrin repeat family
243	319	F	Orientia tsutsugamu	500	22.2		Hypothetical protein
244	505	I	cnpv301/fpv233	527	36.6		Ankyrin repeat family
245	75	F	cnpv022/fpv027	358	8.5		Serpin family
246	182	I	fpv229	180	77.8		Hypothetical A47L-like protein
247	514	I	cnpv303/fpv230-231	256	47.8		Ankyrin repeat family
248	481	I	fpv232	482	86.3		Ankyrin repeat family
249	414	I	fpv233	512	69.7		Ankyrin repeat family
250	422	I	fpv234	428	80.4		Ankyrin repeat family
251	148	I	fpv235	143	65.5		C-type lectin family
252	274	I	fpv236	280	77.7		N1R/p28 family
253	70	I	fpv237	67	57.1		Hypothetical protein
254	163	I	fpv239	163	79.1	A40R	C-type lectin family
255	411	I	fpv240	410	78.8		Ankyrin repeat family
256	109	F	GTPV_gp138	634	5.1		Hypothetical protein
257	185	I	fpv241	186	48.4		Ankyrin repeat family
258	210	I	cnpv313	218	51.6		Ig domain protein
259	643	I	cnpv314/fpv242-243	584	50.4		Ankyrin repeat family
260	192	I	cnpv309/fpv241	196	29.8		Ankyrin repeat family

ORF	ORF length (aa)	Status (I, F, T, E)	Closest homologue	Homologue length (aa)	PEPV vs homologue (% aa identity)	VACV homologue	Description / Putative function
261	53	F	cnpv014/fpv017	490	5.5		Hypothetical protein
262	58	F	cnpv320/fpv017	469	5.1		Ig domain protein
263a	366	T	fpv244	688	42.8		Hypothetical protein
263b	246	T	fpv244	592	29.8		Ankyrin repeat family
264	440	I	fpv245	463	76.8		Ankyrin repeat family
265a	274	T	cnpv014/fpv017	490	27.9		Ig Domain protein
265b	235	T	cnpv014/fpv017	490	25.1		Ig Domain protein
266	584	I	fpv246	592	85.1		Ankyrin repeat family
267	124	I	fpv247	124	83.0		Efc family
268	149	I	fpv248	151	79.5		N1R/p28 family
269	628	I	fpv162	603	41.3		Ankyrin repeat family
270	52	F	fpv249	105	31.1		Hypothetical protein
271	124	I	fpv258	123	67.2	A40R	C-type lectin family
272	122	I	fpv256	122	78.7		C-type lectin family
273	467	I	cnpv320/fpv017	469	82.7		Hypothetical protein
274	46	F	cnpv006/fpv002.5	182	14.3		Hypothetical protein
275	222	I	fpv259	222	91.4		Hypothetical protein
276	115	F	cnpv310/fpv246	537	8.7		Hypothetical protein
277	164	F	cnpv319/fpv246	739	8.3		Hypothetical protein
278	204	I	fpv260	205	77.7	A40R	C-type lectin family

Table 3.2: I = intact, F = fragmented (shaded red), T = truncated (shaded blue), E = extended (shaded purple). ORFs shaded in yellow are conserved in all ChPV; ORFs shaded in green are conserved in all Avipoxviruses.

FGPV: The AT content across the whole sequenced region was found to be 70.5%, and the composition of the potential coding regions to be 69.8%. The sequenced region of FGPV was found to contain 285 potential ORFs encoding proteins ranging from 37 to 1984 amino acids in length and representing an approximate coding density of 91%. The inverted terminal repeats were not resolved meaning more ORFs will likely be identified on completion of the sequence. Relative to their closest orthologues, 252 ORFs have been annotated as intact, 21 as fragmented, eight as truncated, and 4 as extended (Table 3.3).

Figure 3.2 is a schematic of the FGPV genome showing coloured blocks as ORFs numbered from left to right. As in Figure 3.1, ORFs depicted by white blocks are intact relative to their closest orthologue. In the FGPV genome, multiple fragments of a single intact orthologue have been numbered individually. These fragments are identified in Table 3.3 where they populate the same row in the table. FGPV orthologues of 20 ORFs are truncated/fragmented resulting in 29 fragments. Several ORFs in the FGPV genome show greater nucleotide identity to ORFs found in clade B viruses than FWPV or other clade A ORFs, many of which are truncated or fragmented compared to their closest orthologues (Table 3.3).

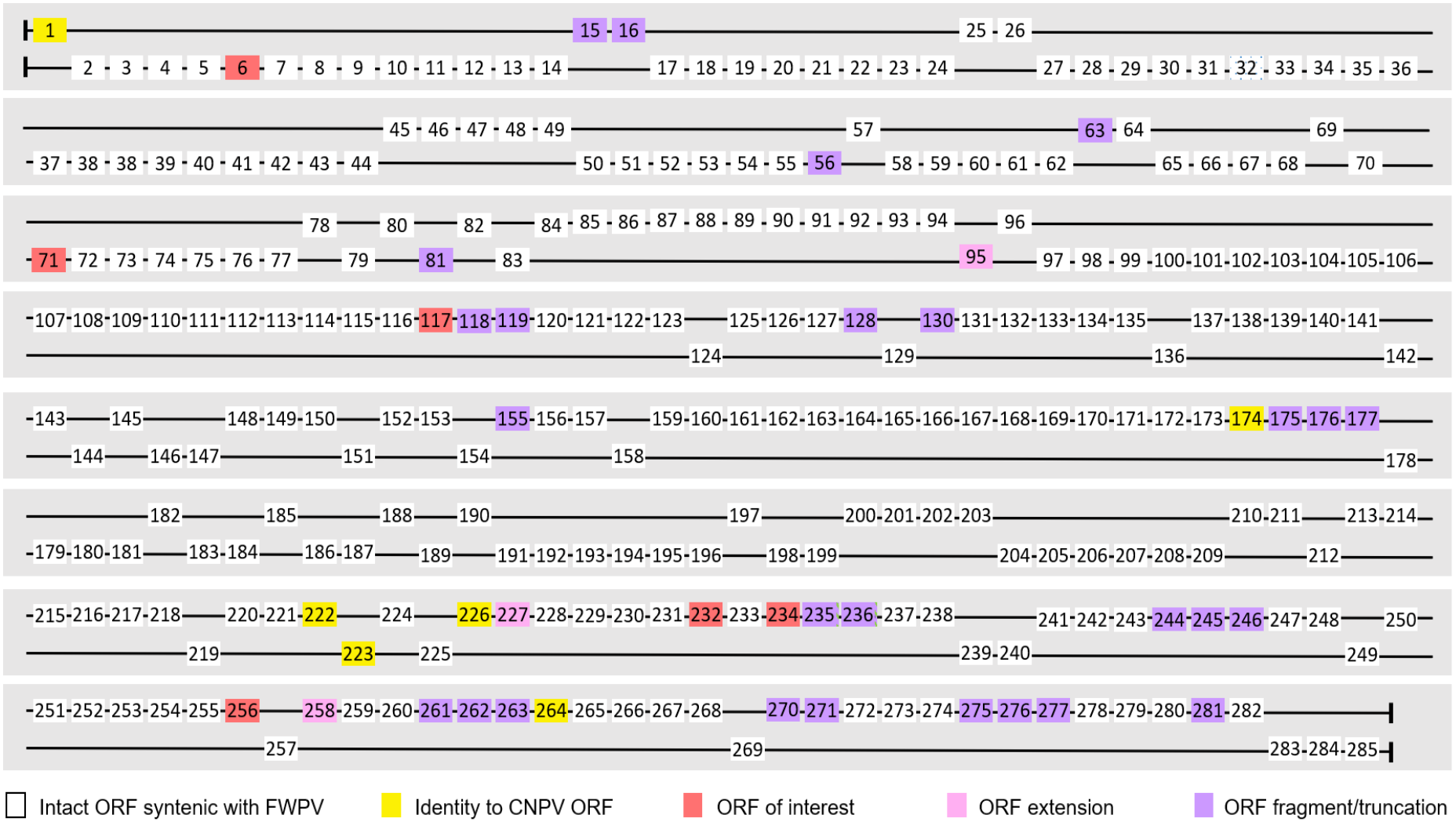


Figure 3.2: ORF schematic of the FGPV genome. The genome is depicted as double stranded, with ORFs shown as coloured blocks (not to scale), numbered from left to right. ORFs transcribed from left to right are depicted above and those transcribed from right to left depicted below.

Table 3.3: FGPV Open Reading Frames

ORF	ORF length (aa)	Status (I, F, T, E)	Closest homologues	Homologue length (aa)	PEPV vs homologue BLASTx identity (%)	VACV homologue	Description / Putative function
1	474	I	cnpv 017/tkpv009/fpv034/fep226	486	52/28/29/29	B4R	Ankyrin repeat family
2	682	I	fep010/pepv010/cnpv009/fpv244/tkpv009	680	95/96/32/32/28	B4R	Ankyrin repeat family
3	355	I	fep011/pepv011/fpv010/cnpv022	355	95/96/86/65	C12L	Serpin family
4	293	I	pepv012/fep012/fpv011/cnpv025	288	95/94/81/58	-	α-SNAP
5	520	I	pepv013/fep013/cnpv026/fpv246	518	89/89/53/31	B4R	Ankyrin repeat family
6	174	I	pepv014/fep014/*Ficedula albicollis*/cnpv018	174	88/86/37/30	-	Interleukin 10
7	330	I	pepv015/fep015/fpv012/cnpv030	329	91/88/80/46	K1L	Ankyrin repeat family
8	402	I	pepv016/fep016/cnpv028/fpv240	402	93/92/40/31	B4R	Ankyrin repeat family
9	437	I	pepv017/fep017/fpv014/cnpv019/tkpv166	437	95/95/87/50/39	B4R	Ankyrin repeat family
10	171	I	pepv018/fep018/fpv015/cnpv013	171	90/88/78/40	-	Hypothetical protein
11	239	I	fep019/pepv019/fpv016/cnpv032	237	92/89/84/32	-	Ig-like domain (putative IFN gamma BP)
12	245	I	fep020/pepv020/fpv017/tkpv002/cnpv033	245	91/90/78/40/41	-	V-type Ig domain
13	684	I	fep021/pepv021/fpv018/cnpv034/tkpv166	683	92/92/85/43/25	C9L	Ankyrin repeat family
14	113	I	pepv022/fep022/fpv019/cnpv036	113	92/91/73/29	-	Hypothetical protein
15	100	T	fep023/pepv023/cnpv037	189	-	-	Hypothetical protein

ORF	ORF length (aa)	Status (I, F, T, E)	Closest homologues	Homologue length (aa)	PEPV vs homologue BLASTx identity (%)	VACV homologue	Description / Putative function
16	47	F					
17	428	I	fep024/pepv024/fpv020/cnpv038/tkpv119	427	95/93/87/43/34	C10L	C4L/C10L protein
18	337	I	pepv025/fep025/fpv021/cnpv039/tkpv006	334	95/92/81/46/37	-	G protein-coupled receptor family
19	580	I	fep026/pepv026/fpv022/cnpv040/tkpv007	581	95/91/88/37/31	B4R	Ankyrin repeat family
20	434	I	fep027/pepv027/fpv023/cnpv041/tkpv008	434	97/96/91/58/49	M1L	Ankyrin repeat family
21	595	I	fep028/pepv028/fpv024/cnpv042/tkpv009	594	97/96/91/50/47	B4R	Ankyrin repeat family
22	203	I	pepv029/fep029/fpv025/cnpv043	203	99/95/90/55	-	Hypothetical protein
23	498	I	fep030/pepv030/cnpv044/fpv024/tkpv009	498	92/93/43/34/34	B4R	Ankyrin repeat family
24	406	I	pepv031/fpv026/tkpv010/cnpv151/fep031	406	94/85/40/30/29	B4R	Ankyrin repeat family
25	333	I	pepv032/fpv027/cnpv045/tkpv011	332	94/86/43/44	-	G-protein-coupled receptor family
26	180	I	pepv033/fpv028	180	95/89	-	Hypothetical protein
27	464	I	fep033/cnpv046/tkpv163/fpv232/pepv252	464	96/51/32/28/29	-	Ankyrin repeat family
28	126	I	pepv034/fep034/fpv029/cnpv047	126	98/96/93/53	-	Peptidyl-tRNA hydrolase
29	815	I	pepv035/fep035/fpv030/cnpv048/tkpv012	815	94/93/85/59/42	-	Alkaline phosphodiesterase
30	345	I	fep036/fpv031/pepv036/cnpv050/tkpv013	341	96/95/95/44/38	B4R	Ankyrin repeat

ORF	ORF length (aa)	Status (I, F, T, E)	Closest homologues	Homologue length (aa)	PEPV vs homologue BLASTx identity (%)	VACV homologue	Description / Putative function
31	382	I	pepv037/fpv032/cnpv051/tkpv014	375	91/88/50/43	-	DNase II
32	291	I	pepv038/fpv033	291	95/89	-	α-SNAP
33	411	I	fpv034/cnpv017 (fep and pepv fragments)	415	84/35	B4R	Ankyrin repeat family
34	135	I	pepv040/fpv035/fep038/cnpv053/tkpv016	135	97/96/96/51/51	-	Hypothetical protein
35	131	I	fpv036/cnpv054	153	79/35	-	Hypothetical protein
36	164	I	fep039/pepv041/fpv037/tkpv017/cnpv055	163	88/87/70/39/42	-	Hypothetical protein
37	145	I	fep040/pepv042/fpv038/cnpv056	145	97/97/93/71	F2L	dUTP pyrophosphatase
38	175	I	fep041/pepv043/fpv039/cnpv058/tkpv020	175	94/94/83/40/40	-	B-cell lymphoma 2 (Bcl-2)
39	337	I	fep042/pepv044/fpv040/tkpv021/cnpv059	337	97/96/92/49/51	C12L	Serpin family
40	212	I	pepv045/fep043/fpv041/cnpv060	220	89/89/67/24	-	Hypothetical protein
41	564	I	fep044/pepv046/fpv043/cnpv061/tkpv022	564	98/98/91/70/64	A50R	DNA ligase
42	358	I	fep045/pepv047/fpv044/cnpv062/tkpv023	358	98/97/92/48/49	C12L	Serpin family
43	368	I	fep046/pepv048/fpv046/cnpv063/tkpv024	370	96/95/83/62/56	A44L	Hydroxysteroid dehydrogenase
44	576	I	fep047/pepv049/fpv047/cnpv065/tkpv025	612	95/95/80/52/46	A39R	Semaphorin
45	261	I	pepv050/fep048/fpv048/cnpv068/tkpv026	261	99/98/95/83/71	-	GNS1/SUR4
46	154	I	pepv051/fpv049/fep049/tkpv027/cnpv069	154	97/95/96/75/71	A1L	Late transcription factor VLTF-2
47	552	I	fep050/pepv052/fpv050/tkpv028/cnpv070	552	99/99/97/82/82	D13L	Rifampicin resistance, N3L protein

ORF	ORF length (aa)	Status (I, F, T, E)	Closest homologues	Homologue length (aa)	PEPV vs homologue BLASTx identity (%)	VACV homologue	Description / Putative function
48	289	I	fep051/pepv053/fpv051/cnpv071/tkpv029	289	99/99/96/76/77	D12L	mRNA capping enzyme, small subunit
49	637	I	fep052/fpv052/pepv054/cnpv074/tkpv030	637	99/96/99/83/84	D11L	NPH-1 transcription termination factor
50	225	I	pepv055/fep053/fpv053/tkpv031/cnpv075	225	98/98/96/70/67	D10R	mutT motif
51	237	I	fep054/pepv056/fpv054/cnpv076/tkpv032	237	97/97/96/77/73	D9R	mutT motif
52	274	I	pepv057/fep055/fpv055/cnpv166	274	93/92/66/41	-	V-type Ig domain
53	161	I	fep056/pepv058/fpv056/tkpv033/cnpv078	161	99/98/97/76/75	D7R	RNA polymerase subunit RPO18
54	633	I	fep057/pepv059/fpv057/cnpv080/tkpv034	633	99/99/99/94/92	D6R	Early transcription factor small subunit, VETFS
55	791	I	pepv060/fep058/fpv058/tkpv035/cnpv082	791	99/99/98/84/80	D5R	NTPase; DNA replication
56	251	F	cnpv072	312	36	-	CC chemokine family
57	219	I	pepv061/fpv059	219	92/90	-	Deoxycytidine kinase
58	196	I	fep059/fpv060	200	90/84	-	CC chemokine family
59	128	I	fep060/fpv061/pepv063/tkpv059	109	93/76/82/61	-	CC chemokine family
60	196	I	fep061	199	90	-	CC chemokine family
61	218	I	fpv062/pepv064/fep062/cnpv084/tkpv037	218	99/98/98/83/82	D4R	Uracil DNA glycosylase

ORF	ORF length (aa)	Status (I, F, T, E)	Closest homologues	Homologue length (aa)	PEPV vs homologue BLASTx identity (%)	VACV homologue	Description / Putative function
63	134	F	fpv064/cnpv087/tkpv039	200	94/78/74	-	Glutathione peroxidase
64	110	I	fep065/fpv065/pepv067/tkpv040/cnpv088	110	97/95/81/38/37	-	Hypothetical protein
65	137	I	pepv068/fep066/fpv066/cnpv089/tkpv041	137	95/95/84/54/51	-	Hypothetical protein
66	85	I	fep067/pepv069/fpv067/cnpv091	93	100/99/96/63	-	HT motif family
67	130	I	fep068/pepv070/fpv068/tkpv042/cnpv092	131	88/84/66/61/59	-	Hypothetical protein
68	269	I	pepv071/fep069/fpv069/cnpv093/tkpv043	269	97/97/95/54/54	D3R	Virion protein
69	273	I	pepv072/fep070/fpv070/cnpv094	273	97/95/87/54	-	T10-like protein
70	43	I	fep071/pepv073/fpv70.5/cnpv095	43	100/93	-	Hypothetical protein
71	77	I	Ubiquitin 40S ribosomal protein	110	100	-	Ubiquitin family
72	287	I	pepv075/fep072/fpv071/tkpv046/cnpv097	287	94/93/89/51/51	-	Hypothetical protein
73	186	I	fpv072/cnpv099/tkpv047	186	88/57/58	-	β-NGF-like family
74	176	I	fpv073/cnpv100/tkpv048	174	75/55/38	-	Hypothetical protein
75	105	I	pepv077/fep076/fpv074	104	94/95/83	-	Hypothetical protein
76	187	I	pepv078/fep077/fpv075/cnpv103/tkpv050	211	97/96/83/61/41	-	N1R/p28 family
77	147	I	fpv076/cnpv279	144	67/33	-	β-NGF-like family
78	125	I	pepv080/fpv077/tkpv051/cnpv104	125	100/95/64/61	G4L	Glutaredoxin
79	225	I	fep079/pepv081/fpv079/cnpv106/tkpv052	225	98/98/96/63/56	G2R	Putative transcriptional elongation factor

ORF	ORF length (aa)	Status (I, F, T, E)	Closest homologues	Homologue length (aa)	PEPV vs homologue BLASTx identity (%)	VACV homologue	Description / Putative function
80	104	I	pepv082/fep080/fpv078/cnpv105/tkpv053	103	99/97/92/65/51	G3L	Hypothetical protein
81	238	F	pepv083/fep081/fpv080	336	91/90/73		Transforming Growth Factor (TGF- β)
82	626	I	pepv084/fep082/fpv081/cnpv108/tkpv054 04578	624	98/97/95/71/67	G1L	Metalloprotease / G1 glycoprotein
83	682	I	pepv085/fep083/fpv082/cnpv109/tkpv055	682	98/98/95/75/73	I8R	DNA/RNA helicase/NPH-11
84	421	I	fep084/pepv086/fpv083/cnpv110/tkpv056	421	99/98/97/74/70	I7L	Virion core peptidase
85	390	I	pepv087/fep085/fpv084/cnpv111/tkpv057	390	99/99/95/69/63	I6L	DNA-binding protein
86	81	I	pepv088/fpv085/cnpv112/tkpv058	81	100/90/74/80	I5L	IMV membrane protein
87	185	I	pepv089/fep087/fpv086/cnpv113/tkpv060	183	95/94/88/65/61	J2R	Thymidine kinase
88	91	I	pepv090/fep088/fpv087/cnpv114/tkpv062	91	96/93/92/56/54	-	HT motif family
89	291	I	fep089/fpv088/pepv091/tkpv063/cnpv115	291	96/95/95/60/61	I3L	ssDNA binding phosphoprotein
90	65	I	fpv089/pepv092/fep090/cnpv116	65	98/97/95/74	I2L	Hypothetical protein
91	311	I	pepv093/fep091/fpv090/cnpv117/tkpv064	311	99/99/98/85/86	I1L	DNA binding virion core protein
92	34	I	fpO3L/pepv094/fep092	34	94/97/97	O3L	MV entry/fusion complex
93	656	I	pepv095/fep093/fpv091/cnpv118/tkpv065	656	97/97/90/55/49	O1L	Hypothetical protein
94	131	I	pepv096/fpv092/fep094/cnpv119/tkpv066	131	98/96/97/64/56	E11L	Hypothetical virion core protein

ORF	ORF length (aa)	Status (I, F, T, E)	Closest homologues	Homologue length (aa)	PEPV vs homologue BLASTx identity (%)	VACV homologue	Description / Putative function
95	130	3' E	pepv097/fep095/fpv093/cnpv120/tkpv067	94	100/96/92/78/82	E10R	Sulfhydryl oxidase
96	989	I	fpv094/pepv098/fep096/cnpv121/tkpv068	988	96/96/96/78/75	E9L	DNA polymerase
97	282	I	fep097/pepv099/fpv095/tkpv069/cnpv122	282	99/97/93/66/63	E8R	Hypothetical protein
98	571	I	pepv100/fep098/fpv096/cnpv123/tkpv070	571	99/99/97/77/74	E6R	Hypothetical protein
99	1893	I	pepv101/fep099/fpv097/cnpv124	1894	94/94/85/78	-	VARV B22R family
100	1802	I	pepv102/fpv098/fep100/cnpv125	1812	96/88/97/75	-	VARV B22R family
101	1935	I	pepv103/fep101/fpv099/cnpv126	1922	94/94/85/59	-	VARV B22R family
102	182	I	pepv104/fep102/fpv100/cnpv127/tkpv071	182	99/99/98/87/77	E4L	RNA pol subunit RPO30
103	717	I	pepv105/fep103/fpv101/cnpv128/tkpv072	717	98/96/94/61/52	E2L	Hypothetical protein
104	472	I	fep104/pepv106/fpv102/cnpv129/tkpv073	472	99/99/98/75/74	E1L	Poly(A) polymerase large subunit, PAP-L
105	114	I	fpv103/cnpv130	114	100/71	F17R	DNA binding virion core phosphoprotein
106	210	I	fep106/pepv108/fpv104/cnpv131/tkpv075	210	95/95/81/39/41	-	Hypothetical protein
107	144	I	pepv109/fep107/fpv105/tkpv076/cnpv132	149	97/97/94/74/73	F15L	Hypothetical protein
108	99	I	fep108/pepv110/cnpv133/fpv106/tkpv077	99	96/93/68/77/46	-	Hypothetical protein

ORF	ORF length (aa)	Status (I, F, T, E)	Closest homologues	Homologue length (aa)	PEPV vs homologue BLASTx identity (%)	VACV homologue	Description / Putative function
109	1780	I	pepv111/fep109/fpv107/cnpv134	1779	97/97/86/58	-	VARV B22R family
110	376	I	fep110/pepv112/fpv108/cnpv135/tkpv078	377	98/96/94/84/75	F13L	Virion envelope protein
111	639	I	pepv113/fep111/fpv109/cnpv136/tkpv079	639	96/96/87/50/45	F12L	Virion release protein
112	453	I	pepv114/fpv110/fep112/cnpv137/tkpv080	452	96/85/92/45/44	F11L	Hypothetical protein
113	444	I	fep113/pepv115/fpv111/cnpv138/tkpv081	444	99/98/97/76/76	F10L	Serine/threonine protein kinase (virus assembly)
114	213	I	pepv116/fep114/fpv112/cnpv139/tkpv082	213	97/97/94/74/70	F9L	Lipid membrane protein of NCLDV
115	66	I	fep115/pepv117/fpv113/cnpv140/tkpv083	66	94/97/95/71/62	F8L	Hypothetical protein
116	174	I	pepv118/fep116/fpv114/cnpv141/tkpv084	183	97/95/95/77	-	HAL3 domain
117	769	I	cnpv223/pepv269/fpv244/tkpv015/fep010	847	28/29/30/27/27	B4R	Ankyrin repeat family
118	299	T	pepv121/fpv115/cnpv144	542	-	-	Ankyrin repeat family
119	171	F					
120	123	I	fep119/fpv116	122	94/73	-	CC-chemokine family
121	440	I	fep120/pepv123/fpv117/cnpv145/tkpv086	440	98/97/93/66/60	G5R	Flap endonuclease (FEN-1)
122	63	I	pepv124/fpv118/fep121/cnpv146	63	100/98/98/74	G5.5R	RNA pol subunit RPO7
123	188	I	fep122/pepv125/fpv119/cnpv147/tkpv087	188	99/97/96/69/65	G6R	Hypothetical protein

ORF	ORF length (aa)	Status (I, F, T, E)	Closest homologues	Homologue length (aa)	PEPV vs homologue BLASTx identity (%)	VACV homologue	Description / Putative function
124	343	I	fep123/pepv126/fpv120/cnpv148/tkpv088	343	99/99/97/80/65	G7L	Virion core protein
125	174	I	fep124/cnpv012	203	77/33	-	Hypothetical protein
126	341	I	cnpv227/fpv124/fep166/pepv174/tkpv115	359	69/70/45/45/44	-	N1R/p28 family
127	357	I	cnpv166/fpv125/fep056/pepv057	345	73/61/40/40	B19R	V-type Ig domain
128	164	F	cnpv166/fpv125/pepv057/fep056	345	48/49/42/41	B19R	V-type Ig domain
129	116	I	pepv132/cnpv086/tkpv038	116	94/50/45	C22L/B28R	TNF receptor like protein
130	279	F	cnpv216/fpv063/pepv065	404	54/42/41	-	Hypothetical protein
131	217	I	pepv133/fep126/cnpv170	217	97/93/82	A48R	Thymidylate kinase
132	260	I	pepv134/fpv126/fep127/cnpv171/tkpv089	260	100/99/99/96/84	G8R	Late transcription factor VLTF-1
133	336	I	pepv135/fep128/fpv127/cnpv172/tkpv090	322	99/96/93/73/62	G9R	Myristylated protein
134	243	I	pepv136/fep129/fpv128/cnpv173/tkpv091	243	99/99/97/87/84	L1R	Myristylated MV membrane protein
135	96	I	pepv137/fep130/fpv129/cnpv174/tkpv092	96	91/91/81/30/	L2R	Hypothetical protein
136	297	I	pepv138/fep131/fpv130/cnpv175/tkpv093	301	97/96/92/82/79	L3L	Hypothetical protein
137	253	I	fep132/pepv139/fpv131/cnpv176/tkpv094	253	98/97/95/74/67	L4R	DNA binding virion core VP8
138	129	I	fep133/pepv140/fpv132/cnpv177/tkpv095	129	98/98/95/77/65	L5R	Putative membrane protein

ORF	ORF length (aa)	Status (I, F, T, E)	Closest homologues	Homologue length (aa)	PEPV vs homologue BLASTx identity (%)	VACV homologue	Description / Putative function
139	148	I	fep134/pepv141/fpv133/cnpv178/tkpv096	148	99/99/97/76/76	J1R	Hypothetical protein
140	308	I	pepv142/fep135/fpv134/cnpv179/tkpv097	308	98/97/96/96/70	J3R	PolyA polymerase (PAPs)
141	186	I	pepv143/fep136/fpv135/tkpv098/cnpv180	186	98/98/96/76/75	J4R	RNA pol Subunit RPO22
142	137	I	fep137/pepv144/fpv136/cnpv181/tkpv099	138	97/96/94/79/75	J5L	Membrane protein
143	1287	I	pepv145/fep138/fpv137/cnpv182/tkpv100	1287	99/99/98/89/87	J6R	RNA pol Subunit RPO147
144	166	I	fep139/pepv146/fpv138/cnpv183/tkpv101	166	99/98/95/81/80	H1L	Protein tyrosine Phosphatase
145	190	I	pepv147/fep140/fpv139/cnpv184/tkpv102	190	100/99/97/87/83	H2R	Hypothetical protein
146	333	I	fpv140/pepv148/fep141/cnpv186/tkpv104	327	99/97/96/57/55	H3L	Virion envelope protein (p35)
147	799	I	pepv149/fep142/fpv141/cnpv187/tkpv105	799	99/98/97/82/79	H4L	RNA polymerase associated protein RAP94
148	172	I	pepv150/fep143/fpv142/cnpv188/tkpv106	174	95/95/91/51/52	H5R	Late transcription factor VLTF-4
149	318	I	pepv151/fep144/fpv143/cnpv189/tkpv107	316	99/99/97/81/75	H6R	DNA topoisomerase
150	152	I	pepv152/fep145/fpv144/cnpv190/tkpv108	152	97/97/93/53/49	H7R	Putative 17 kDa protein
151	103	I	pepv153/fep146/fpv145/cnpv191/tkpv109	103	93/90/90/47/45		Hypothetical protein

ORF	ORF length (aa)	Status (I, F, T, E)	Closest homologues	Homologue length (aa)	PEPV vs homologue BLASTx identity (%)	VACV homologue	Description / Putative function
152	852	I	pepv154/fep147/fpv146/ tkpv110/cnpv192	822	98/98/95/76/73	D1R	mRNA capping enzyme, large subunit
153	109	I	pepv155/fpv147/cnpv193	104	96/89/58	-	HT motif family
154	140	I	pepv156/fep149/fpv148/cnpv194/tkpv111	140	99/98/90/60/56	D2L	Virion protein
155	145	T	pepv157/fep151/fpv149/cnpv196/tkpv112	189	98/98/86/56	-	Hypothetical protein
156	284	I	fep152/pepv158/fpv150/cnpv197/tkpv113	284	93/93/85/44/49	-	N1R/p28 family
157	238	I	pepv159/fep153/fpv151/cnpv199	238	95/94/86/52	-	Deoxycytidine kinase
158	127	I	pepv160/fpv152/cnpv282	127	94/83/36	-	HT motif family
159	208	I	fep154/fpv153/cnpv201	209	93/79/55	-	Hypothetical protein
160	276	I	pepv161/fpv154/cnpv202	270	93/80/31	-	N1R/p28 family
161	411	I	fep155/pepv162/fpv155/cnpv203	408	94/93/88/36	-	N1R/p28 family
162	136	I	fep156/fpv156	132	96/86	-	HT motif family
163	328	I	fep157/pepv164/fpv157/cnpv205	327	94/94/85/48	-	N1R/p28 family
164	464	I	fep158/pepv165/fpv158/cnpv206	464	98/97/95/74	-	Photolyase
165	246	I	fep159/pepv166/fpv159/cnpv207/tkpv121	246	97/96/85/46/40	-	N1R/p28 family
166	156	I	pepv167/fep160/fpv160/cnpv208/tkpv116	156	99/98/95/42/36	-	Hypothetical protein
167	149	I	fep161/pepv168/fpv161/cnpv209/tkpv117	149	98/97/92/48/43	-	N1R/p28 family
168	133	I	pepv169/fep162/cnpv210/fpv124	133	98/95/45/40	-	N1R/p28 family
169	160	I	pepv171/fep164/cnpv212/fpv124	160	90/89/58/44	-	N1R/p28 family
170	611	I	fpv162/pepv172/fep165/cnpv223/tkpv015	603	89/88/86/32/30	-	Ankyrin repeat family

ORF	ORF length (aa)	Status (I, F, T, E)	Closest homologues	Homologue length (aa)	PEPV vs homologue BLASTx identity (%)	VACV homologue	Description / Putative function
171	249	I	pepv173/fep166/fpv163/cnpv159/tkpv050	256	94/90/76/33/33	-	N1R/p28 family
172	741	I	pepv174/fpv162/cnpv223/fep167/tkpv120	741	95/33/35/32/30	-	Ankyrin repeat family
173	379	I	pepv175/fep168/fpv164/cnpv235	386	90/90/67/31	-	Hypothetical protein
174	300	I	cnpv236/*Elephantulus edwardii*	324	83/72	F4L	Ribonucleotide reductase small subunit
175 **176** **177**	134 93 71	F F F	cnpv237	441	-	C9L	Ankyrin repeat family
178	225	I	fpv165/pepv177/fep169/tkpv121/cnpv238	225	98/96/95/91/91	A2L	Late transcription factor VLTF-3
179	72	I	pepv178/fpv166/cnpv239/fep170???	72	100/96/73	A2.5L	Virus redox protein
180	660	I	pepv179/fep171/fpv167/cnpv240/tkpv122	657	99/99/98/84/82	A3L	Virion core protein P4b
181	265	I	pepv180/fep172/fpv168/cnpv241/tkpv123	255	83/80/60/48/43	A4L	Immunodominant virion protein
182	169	I	pepv181/fep173/fpv169/cnpv242/tkpv124	169	100/99/98/75/77	A5R	RNA pol subunit RP019
183	374	I	pepv182/fep174/fpv170/cnpv243/tkpv125	374	98/98/95/68/63	A6L	Hypothetical protein
184	709	I	pepv183/fep175/fpv171/cnpv244/tkpv126	709	99/99/97/88/85	A7L	Early transcription factor large subunit VETF-L
185	301	I	pepv184/fep176/fpv172/cnpv245/tkpv127	301	99/99/99/79/75	A8R	Intermediate transcription factor VITF-3

ORF	ORF length (aa)	Status (I, F, T, E)	Closest homologues	Homologue length (aa)	PEPV vs homologue BLASTx identity (%)	VACV homologue	Description / Putative function
186	76	I	fpv173/fep177/pepv185/cnpv246	76	88/87/87/57	A9L	Hypothetical protein
187	891	I	fep178/pepv186/fpv174/cnpv247/tkpv128	891	98/98/96/71/67	A10L	Virion core protein P4a
188	272	I	pepv187/fep179/fpv175/tkpv129/cnpv248	272	99/99/97/63/66	A11R	Hypothetical protein
189	175	I	pepv188/fep180 /fpv176/cnpv249/tkpv130	175	99/98/91/61/58	A12L	Virion core protein
190	70	I	pepv189/fpv177	70	97/85	-	Hypothetical protein
191	71	I	pepv190/fep182/fpv178/cnpv251	71	97/96/87/56	A13L	Virion membrane protein
192	91	I	pepv191/fep183/fpv179/cnpv252	91	100/99/96/61	A14L	Virion envelope protein
193	53	I	pepv192/fpv179.1/fep184	53	100/98	a14.5	Virion envelope protein
194	97	I	fep185/fpv180/pepv193/cnpv254	97	98/95/95/60	A15L	Hypothetical protein
195	368	I	fep186/pepv194/fpv181/cnpv255/tkpv131	369	94/94/92/77/75	A16L	Putative myristoylated membrane protein
196	198	I	fep187/pepv195/fpv182/cnpv256/tkpv132	198	99/99/98/79/64	A17L	Phosphorylated virion membrane protein
197	462	I	pepv196/fep188/fpv183/cnpv257/tkpv133	462	99/98/97/84/76	A18R	DNA helicase
198	88	I	pepv197/fep189/fpv184/cnpv258/tkpv134	88	100/98/95/68/57	A19L	Zinc finger like protein
199	113	I	fpv186/fep190/pepv198/cnpv260/tkpv135	113	97/95/94/85/80	A21L	Hypothetical protein

ORF	ORF length (aa)	Status (I, F, T, E)	Closest homologues	Homologue length (aa)	PEPV vs homologue BLASTx identity (%)	VACV homologue	Description / Putative function
200	432	I	pepv199/fep191/fpv185/cnpv259/tkpv136	432	99/99/99/98/56	A20R	Processivity factor
201	161	I	pepv200/fep192/fpv187/cnpv261/tkpv137	161	99/97/91/72/72	A22R	Holliday junction resolvase
202	383	I	pepv201/fep193/fpv188/cnpv262/tkpv138	383	98/98/96/73/72	A23R	Intermediate transcription factor VITF-3
203	1157	I	fep194/pepv202/fpv189/cnpv263/tkpv139	1157	99/99/98/91/91	A24R	RNA pol subunit RPO132
204	612	I	fep195/pepv203/fpv190/cnpv264/tkpv140	608	97/97/89/68/62	A25L	A-type inclusion protein
205	474	I	fep196/pepv204/fpv191/cnpv265/tkpv141	472	97/97/92/78/48	A26L	A-type inclusion protein
206	140	I	pepv205/fpv192/cnpv266/hgpv192/tkpv142	140	99/96/84/84/76	A28L	Hypothetical protein
207	302	I	pepv206/fep198/fpv193/hgpv193/cnpv267/tkpv143	302	97/95/91/65/64/61	A29L	RNA pol subunit RPO35
208	74	I	fep199/pepv207/fpv194/tkpv143/cnpv268	74	100/100/98/71/83	A30L	Hypothetical protein
209	38	I	pepv208/fep200/fpv194.1	38	97/95/84	A30.5L	A30.5L orthologue
210	113	I	pepv209/fep201/fpv195/cnpv269	113	100/97/90/64	A31R	Hypothetical protein
211	120	I	fep202/pepv210/fpv196/cnpv270	120	94/94/83/42		Hypothetical protein
212	304	I	pepv211/fep203/fpv197/cp197L (*vultur gryphus* poxvirus)/cnpv271/tkpv145	304	98/97/92/78/75/67	A32L	Virion assembly protein
213	173	I	pepv212/fep204/fpv198/cnpv272/tkpv146	173	97/97/93/66/57	A34R	C-type lectin family
214	220	I	pepv213/fpv199/tkpv147	220	93/81/41	-	V-type Ig Domain

ORF	ORF length (aa)	Status (I, F, T, E)	Closest homologues	Homologue length (aa)	PEPV vs homologue BLASTx identity (%)	VACV homologue	Description / Putative function
215	263	I	pepv214/fpv200/tkpv148	263	95/85	-	V-type Ig domain
216	277	I	fep207/pepv215/fpv201/cnpv273/tkpv149	277	96/95/91/45	-	Hypothetical protein
217	285	I	fep208/pepv216/fpv203/cnpv274/tkpv150	285	98/97/84/51	-	Tyrosine protein kinase
218	342	I	fep209/pepv217/fpv204/cnpv275	342	97/96/92/54	C12L	Serpin family
219	220	I	pepv218/fep210/fpv205/cnpv276/tkpv151	220	95/93/85/38	-	Hypothetical protein
220	308	I	pepv219/fep211/fpv206/cnpv277	308	99/98/89/67	-	G-protein-coupled receptor family
221	92	I	pepv220/fep212/fpv207/cnpv278	92	96/95/90/66	-	Hypothetical protein
222	179	I	cnpv279/fpv072	169	52/34	-	β-NGF-like family
223	137	I	cnvp280/pepv162/fpv152	130	41/28/28	-	HT motif family
224	212	I	pepv222/fep214/cnpv281/fpv208/tkpv152	193	98/96/47/67	-	Hypothetical protein
225	145	I	pepv223/fep215/fpv209/cnpv282	137	99/91/83/45	-	HT motif family
226	108	I	cnpv283	111	50	-	CC chemokine family
227	157	E 5’	pepv225/fep216/fpv211/cnpv285/tkpv153	125	96/93/81/43	C11R	Epidermal Growth Factor-like protein
228	303	I	fep217/fpv212/pepv226/cnpv286/tkpv154	303	95/92/86/55	B1R	Serine/threonine protein kinase
229	162	I	fep218/pepv227/fpv213/cnpv287/tkpv155	162	97/96/91/54	-	Hypothetical protein
230	124	I	fep219/pepv228/fpv214/cnpv289/tkpv156	125	91/93/85/47	-	IL-18BP

ORF	ORF length (aa)	Status (I, F, T, E)	Closest homologues	Homologue length (aa)	PEPV vs homologue BLASTx identity (%)	VACV homologue	Description / Putative function
231	74	I	pepv229/fep220/fpv215/cnpv290	74	100/95/81	-	Hypothetical protein
232	174	F	brevican core protein - *Xenopus Silurana tropicalis* (fep221 fragment)	1146	38/23	-	Hypothetical brevican core protein
233	294	I	pepv230/fep222/fpv216/cnpv293/tkpv157	294	96/95/88/37	C18L	Ankyrin repeat family
234	143	I	fep223/pepv231/Tanapox 67R (178aa)	143	98/96/36	C7L	67R Tanapox host range protein
235	237	T	fpv217	328	-	-	Hypothetical protein
236	59	F					
237	474	I	pepv233/fpv218/cnpv295	474	94/84/37	B4R	Ankyrin repeat family
238	442	I	fep226/fpv219/pepv234/cnpv296/tkpv161	440	95/95/87/46	B4R	Ankyrin repeat family
239	117	I	fpv220	117	100	-	Hypothetical protein
240	183	I	pepv235/fep227/fpv221	183	98/96/90	A47L	Hypothetical protein
241	745	I	fep228/pepv236/fpv222/cnpv297	747	84/83/79/36	B4R	Ankyrin repeat family
242	292	I	pepv240/fpv226/fep231/cnpv299/tkpv154	293	91/89/91/57/43	B1R	Serine/threonine protein kinase
243	361	I	pepv241/fep232/fpv227/cnpv300/tkpv162	361	96/96/88/46	-	Ankyrin repeat family
244	204	T	fpv228/trichomonas vaginalis DUF3447/pepv275/cnpv044	525	-	-	Ankyrin repeat family
245	239	F					
246	297	F					

ORF	ORF length (aa)	Status (I, F, T, E)	Closest homologues	Homologue length (aa)	PEPV vs homologue BLASTx identity (%)	VACV homologue	Description / Putative function
247	502	I	fep248/pepv253/cnpv301/fpv233	503	97/95/39/30	-	Ankyrin repeat family
248	348	I	cnpv022/fpv010/fep011/pepv011	358	37/36/36/35	C12L	Serpin family
249	187	I	fep236/pepv246/fpv229/cnpv302	185	97/93/82/33	-	Hypothetical A47L-like protein
250	502	I	pepv247/fpv231/cnpv303	514	96/87/50	B4R	Ankyrin repeat family
251	489	I	pepv248/fep238/fpv232/cnpv304	481	96/96/86/46	-	Ankyrin repeat family
252	503	I	fep239/pepv249/fpv233	502	96/90/96	-	Ankyrin repeat family
253	430	I	pepv250/fpv234	422	95/86	-	Ankyrin repeat family
254	148	I	pepv251/fpv235/cnpv029	148	91/71/32	A40R	C-type lectin family
255	275	I	pepv252/fep241/fpv236/cnpv305	274	94/93/80/36	-	N1R/p28 family
256	72	I	pepv253/fpv237	70	94/58	-	Hypothetical protein
257	163	I	pepv254/fpv239/cnpv307/tkpv165	163	94/83/46	A40R	C-type lectin family
258	88	E 5’	fpv238	61	57	-	Hypothetical protein
259	411	I	fep244/pepv255/fpv240/cnpv308	411	90/86/83/35	-	Ankyrin repeat family
260	185	I	pepv257/fpv241/cnpv309	185	96/90/46	-	Ankyrin repeat family
261	159	T	cnpv310/fpv246/pepv266/fep252	592	-	-	Ankyrin repeat family
262	317	F					

ORF	ORF length (aa)	Status (I, F, T, E)	Closest homologues	Homologue length (aa)	PEPV vs homologue BLASTx identity (%)	VACV homologue	Description / Putative function
263	88	F	cnpv312	166			Hypothetical protein
264	210	E 5’	cnpv312	166	36	-	Hypothetical protein
265	643	I	pepv259/fep246/cnpv314/fpv242	643	97/97/56/90	-	Ankyrin repeat family
266	192	I	fep258/pepv264/cnpv309/fpv241	192	95/95/42/33	-	Ankyrin repeat family
267	463	I	fep006/pepv273/cnpv320/fpv017	467	33/32/32/30	-	V-type Ig domain
268	667	I	fep260/fpv244/cnpv009/pepv010	667	91/81/79/31	-	Ankyrin repeat family
269	444	I	pepv264/fpv245/cnpv020	440	92/81/45	-	Ankyrin repeat family
270	254	T	cnpv014/pepv262/fep256/fpv017	490	-	-	V-type Ig domain
271	235	F					
272	584	I	fep264/pepv272/fpv246/cnpv011	585	93/92/89/46	-	Ankyrin repeat family
273	124	I	fep253/pepv267/fpv247/cnpv321	122	96/95/85/60	-	EFc family
274	149	I	fep254/pepv268/fpv248/cnpv218	149	87/88/86/43	-	N1R/p28 family
275	98	F	trichomonas vaginalis G3 ankyrin repeat protein	732	-	-	Ankyrin repeat family
276	288	T	fpv162/fep167/pepv175/cnpv150	603	40/38/37/32	-	Ankyrin repeat family
277	544	F					
278	108	F	pepv269	628	88	B20R	Ankyrin repeat family
279	110	I	fep267/fpv255	107	89/71	-	Hypothetical protein
280	124	I	pepv271/fep256/fpv258/cnpv035	124	86/86/71/29	A40R	C-type lectin

ORF	ORF length (aa)	Status (I, F, T, E)	Closest homologues	Homologue length (aa)	PEPV vs homologue BLASTx identity (%)	VACV homologue	Description / Putative function
281	169	F	pepv275/fpv162/cnpv021/fep167	628	-	-	Ankyrin repeat family
282	508	I	pepv009/fep008/cnpv015	508	93/93/45	-	Ankyrin repeat family
283	201	I	fep126 /cnpv012	203	47/37	A47L	Hypothetical protein
284	419	I	fep007/pepv008/fpv006/cnpv214	410	95/95/85/43	C10L	C4L/C10L family
285	140	I	pepv007/fpv250	139	93/80	A40R	US ORF2

Table 3.3: I = intact, F = fragmented (shaded red), T = truncated (shaded blue), E = extended (shaded purple). ORFs shaded in yellow are conserved in all ChPV; ORFs shaded in green are conserved in all Avipoxviruses.

•

3.3.3 Core/Conserved Genes

90 ORFs have been noted to be conserved in all ChPV genomes (Gubser *et al*., 2004; Lefkowitz *et al*., 2006; Upton *et al*., 2003). VACV F16L was erroneously added to the above list, as this ORF is not present in avian poxvirus genomes; and two ORFs (fpv194 and fpv194.1) and their equivalents in the other genomes, previously considered conserved among ChPV, have been excluded from the list as they are not present in the TKPV genome. A fourth ORF, fpv103 was also removed as it is truncated/fragmented in the TKPV genome (tkpv074). ORF fpv095 and the relative equivalents were also removed as no orthologue is present in SWPV2. Lastly, orthologues fpv168 (288aa) and cnpv241 (215aa) differ in length by ≥25% and were excluded. Therefore, in this study, 83 ORFs are considered conserved amongst ChPV.

It was previously noted that a further 89 ORFs were conserved between FWPV, CNPV, FeP2 and PEPV (Offerman *et al*., 2014). This list has been updated with the addition of the four, more recent genomes (TKPV, SWPV-1, SWPV-2 and FGPV) and exclusion of ORFs that differ in length by more than 20% bringing the total to 47 (Table 3.4) and bringing the total number of conserved ORFs in the sequenced avian poxvirus genomes to 130. The TKPV genome is considerably smaller than the other avian poxvirus genomes and as such, is the only genome missing ORFs (n=32) that would otherwise be conserved in avian poxvirus genomes (Table 4.1). Of these 32 ORFs, 28 are either hypothetical proteins or members of gene families suggesting that they are unlikely to be essential to the viral life cycle. Several ORFs were identified that were unreported in the study describing the TKPV genome (Bányai *et al*., 2015). Eight of these (tkpv63.1, tkpv86.1, tkpv121.1, tkpv127.1 and tkpv130.1-130.4) are of the 83 ORFs considered to be conserved in all ChPV genomes. A further two ORFs (tkpv1.1 and tkpv151.1) in common amongst avian poxviruses and ORF tkpv60.1, which is not conserved, were also identified on further inspection. The eleven ORFs mentioned above are named after the ORF directly preceding them with ".1" appended.

Table 3.4: 47 ORFs found to be conserved in each of the fully sequenced avian poxvirus genomes

FWPV	CNPV	PEPV	FeP2	FGPV	TKPV	SWPV1	SWPV2	FUNCTION
016	032	019	019	011	001.1*	024	028	Hypothetical protein
017	033	020	020	012	002	025	029	V-type Ig domain
020	038	024	024	017	005	028	034	C4L/C10L protein
021	039	025	025	018	006	029	035	GPCR
022	040	026	026	019	007	030	036	Ankyrin repeat
023	041	027	027	020	008	031	037	Ankyrin repeat
024	042	028	028	021	009	032	038	Ankyrin repeat
030	048	035	035	029	012	038	044	Alkaline phosphodiesterase
031	050	036	036	030	013	040	046	Ankyrin repeat
035	053	040	038	034	016	044	049	Hypothetical protein
037	055	041	039	036	017	046	051	Hypothetical protein
039	058	043	041	038	020	049	054	B-cell lymphoma 2 (Bcl-2)
040	059	044	042	039	021	050	055	Serpin
043	061	046	044	041	022	052	057	DNA ligase
044	062	047	045	042	023	053	058	Serpin family
046	063	048	046	043	024	054	059	Hydroxysteroid dehydrogenase
047	065	049	047	044	025	056	061	Semaphorin
048	068	050	048	045	026	059	064	GNS1/SUR4
054	076	056	054	051	032	066	072	mutT motif
065	088	067	065	064	040	078	083	Hypothetical protein
068	092	070	068	067	042	082	087	Hypothetical protein
070	094	072	070	069	044	084	089	T10-like protein
071	097	075	072	072	046	087	092	Hypothetical protein
075	103	078	077	076	050	092	098	N1R/p28
086	113	089	087	087	060	102	108	Thymidine kinase
091	118	095	093	093	065	107	113	Hypothetical protein
092	119	096	094	094	066	108	114	Hypothetical virion core protein
104	131	108	106	106	075	120	126	Hypothetical protein
105	132	109	107	107	076	121	127	Hypothetical protein
110	137	114	112	112	080	126	132	Hypothetical protein
113	140	117	115	115	083	129	135	Hypothetical protein
145	191	153	146	151	109	167	179	Hypothetical protein
151	199	159	153	157	113	175	187	Deoxycytidine kinase
190	264	203	195	204	140	237	250	A-type inclusion protein
191	265	204	196	205	141	238	251	A-type inclusion protein

196	270	210	202	211	144	243	256	Hypothetical protein
201	273	215	207	216	149	247	259	Hypothetical protein
203	274	216	208	217	150	248	260	Tyrosine kinase
205	276	218	210	219	151	250	262	Hypothetical protein
207	278	220	212	221	151.1*	252	264	Hypothetical protein
208	281	222	214	224	152	255	267	Hypothetical protein
211	285	225	216	227	153	259	271	Epidermal Growth Factor
212	286	226	217	228	154	260	272	Serine/threonine protein kinase
213	287	227	218	229	155	261	273	Hypothetical protein
214	289	228	219	230	156	263	275	Putative 13.7 kDa protein
219	296	234	226	238	161	272	282	Ankyrin repeat
232	304	248	238	251	164	283	290	Ankyrin repeat

*tkpv 001.1 and tkpv151.1 were not reported in the literature (Bányai *et al*., 2015; Sarker *et al*., 2017), but were identified on inspection of the genome sequences deposited in Genbank.

Concatenation of the 130 ORFs conserved in all sequenced avian poxviruses resulted in sequences ranging in length from 127 092bp in TKPV to 127 756bp in FGPV. A nucleotide alignment of the concatenated sequences from each of the sequenced virus genomes shows PEPV and FGPV to have the greatest degree of identity to each other and FeP2 (~96%), followed by FWPV (~90%), the clade B isolates (~69%) and lastly TKPV (~65%) (Table 3.5). Also noted is the high degree of similarity between CNPV and SWPV-2 ORFs with 99% nt identity compared to ~79% between CNPV and SWPV-1, and SWPV-2 and SWPV-1.

Table 3.5: Pairwise comparison of nucleotide alignments of 130 conserved genes in eight avian poxvirus genomes.

		1	2	3	4	5	6	7	8
FWPV	1		12486	12550	11684	40283	39648	40454	44591
FeP2	2	90.23		4967	4557	39532	39000	39705	44412
PEPV	3	90.19	96.11		4469	39558	39041	39733	44468
FGPV	4	90.88	96.44	96.51		39640	39059	39814	44523
CNPV	5	68.84	69.39	69.40	69.37		27136	1228	45260
SWPV1	6	69.29	69.76	69.76	69.78	78.86		27405	44385
SWPV2	7	68.71	69.25	69.26	69.23	99.04	78.65		45466
TKPV	8	65.64	65.77	65.76	65.75	65.15	65.78	64.99	

% identities are shown in the lower left and the number of nucleotide differences shown in the upper right.

A concatenated amino acid alignment of the same conserved ORFs showed a similar pattern of inter-clade and intra-clade identities (Table 3.6)

Table 3.6: Pairwise comparison of amino acid alignments of 130 conserved proteins in eight avian poxvirus genomes.

		1	2	3	4	5	6	7	8
FWPV	1		3595	3608	3306	14954	15107	15019	17005
FeP2	2	91.56		1654	1467	14815	15025	14878	17055
PEPV	3	91.54	96.11		1479	14855	15035	14918	17077
FGPV	4	92.26	96.56	96.53		14846	15054	14907	17050
CNPV	5	65.20	65.50	65.42	65.48		9074	467	17601
SWPV1	6	64.80	64.98	64.97	64.96	78.76		9185	17625
SWPV2	7	65.05	65.35	65.27	65.34	98.90	78.50		17664
TKPV	8	60.56	60.45	60.43	60.50	59.21	59.13	59.06	

% identities are shown in the lower left and the number of amino acid differences shown in the upper right.

3.3.4 Multigene Families

Avipoxviruses contain several, large multigene families with immune-related functions that can make up close to 50% of the genome. Table 3.7 below outlines the copy numbers of each of the 14 multigene families identified in each of the sequenced avipoxvirus genomes. Overall PEPV and FGPV were found to have a similar complement of multi-gene families but FGPV has considerably more ankyrin repeat family genes than are found in other clade A viruses.

3.3.5 Gene Translocations/Duplications

In many cases, in both the PEPV and FGPV genomes, ORFs that appear to have been translocated belong to multigene families. This makes it difficult to determine if these ORFs have truly translocated or if they are unique members of the gene family showing a high degree of identity to another member of the family in a different genome location.

Table 3.7: Copy number of genes in each of the 14 multi-gene families identified in each of the fully sequenced Avipoxvirus genomes.

GENE FAMILY	FWPV	FP9	CNPV	PEPV	FeP2	TKPV	FGPV	SWPV1	SWPV2
Ankyrin Repeat	31	22	51	33	26	16	45	50	46
B22R	6	5	6	5	4	1	4	6	7
N1R/p28	10	8	26	11	11	3	13	20	20
C4L/C10L	3	3	3	2	2	2	2	2	3
CC chemokine	4	4	5	1	4	2	6	6	5
C-type lectin	9	6	11	7	4	2	4	13	11
GPCR	3	2	4	3	2	2	3	4	4
HT motif	6	6	5	5	4	1	7	4	4
Ig-like domain	5	4	9	6	4	3	9	9	8
Serpin	5	5	5	4	4	3	5	5	5
EFc	3	2	2	1	1	1	1	2	2
TGF-β	1	1	5	1	1	1	1	3	4
B-NGF	2	2	2	0	0	2	3	2	2
IL-18 BP	1	1	3	1	0	2	0	3	3
TOTAL	89	71	137	80	67	41	103	129	124
% of TOTAL ORFs	34	29	42	28	25	24	36	42	40

3.3.6 Gene Disruptions/Deletions

PEPV: Relative to FWPV, PEPV contains 5 truncated and 8 fragmented orthologues as well as several truncated and fragmented ORFs with similarity to CNPV and/or FeP2. PEPV also lacks orthologues of 26 genes found in FWPV.

FGPV: Relative to FWPV, FPGV contains 7 truncated and 8 fragmented orthologues as well as several truncated and fragmented ORFs with similarity to orthologues found uniquely in CNPV, PEPV and/or FeP2. FGPV lacks orthologue of 21 genes found in FWPV, with many of them commonly missing in all three South African isolates.

3.3.7 Gene Insertions

PEPV and FGPV contain several ORFs that have been inserted relative to the FWPV genome. These include orthologues of cnpv086 (pepv132 and fgpv129) which encodes a protein similar to a tumour necrosis factor receptor (TNFR). The avipoxvirus TNFR orthologues are more similar to each other (50-72% aa identity) than to eukaryotic orthologues (28-38% aa identity). Also noted are an orthologue of ubiquitin (pepv074 and fgpv071), an orthologue of thymidylate kinase (pepv133 and fgpv131) which is only found in the CNPV and FeP2 genomes and not FWPV (cnpv170; VACV A48R), and putative orthologues of interleukin-10 (pepv014 and fgpv006). Both viruses also contain orthologues (pepv231 and fgpv234) of orthopoxvirus C7L which shows similarity to tanapox virus (TANV) and Yaba-like disease virus (YLDV) 67R. The FGPV genome was also found to contain an orthologue (fgpv174) of ribonucleotide reductase found only in the clade B viruses (cnpv236, swpv1-209 and swpv2-222).

3.3.8 NOVEL ORFs and Genes of Interest

ORFs of interest have been highlighted (red blocks in Figure 3.2) that have either been identified in poxviruses for the first time (fgpv232 and fgpv256) or were previously identified and noted in other avian poxvirus genomes (pepv014/fgpv006, pepv074/fgpv071, pepv231/fgpv234) (Table 3.8).

An investigation into the potential incorporation of IL-10 orthologues into poxvirus genomes by horizontal gene transfer showed that based on the location of the IL-10 orthologue in CNPV compared to the other poxviruses, it is possible that the transfer of IL-10 into CNPV was an independent event (Bratke and McLysaght, 2008). Because this region is not highly conserved among poxviruses the analysis was inconclusive. With the discovery of IL-10 orthologues in other avian poxviruses, the analysis was repeated using only avian poxvirus genomes. Here, the IL-10 orthologues from CNPV and SWPV-2, and the clade A viruses are shown to be in slightly different locations but in a similar genomic region (Figure 3.3).

Table 3.8: ORFs of interest identified in the PEPV and FGPV genomes as well as their location in seven other avian poxvirus genomes if present.

ORF	FWPV	FGPV	PEPV	FeP2	TKPV	SWPV1	SWPV2	CNPV	NOTES
IL-10	-	006	014	014	-	-	014	018	Found in similar genomic locations
UBQ	Frag.	071	074	Frag.	-	086	091	096	Found in the same genomic location
ANK	-	117	-	-	-	-	-	-	Ankyrin repeat family with limited (30%) identity to avipoxvirus clade B ANKs
C-type lectin	-	232	-	-	-	-	-	-	C-type lectin / Brevican core protein/ NK receptor like
C7L	-	234	231	223	159	-	-	-	Host range gene thought to be restricted to OPVs
P-type ATPase	-	256	-	-	-	-	-	-	No significant identity to known ORFs in Genbank

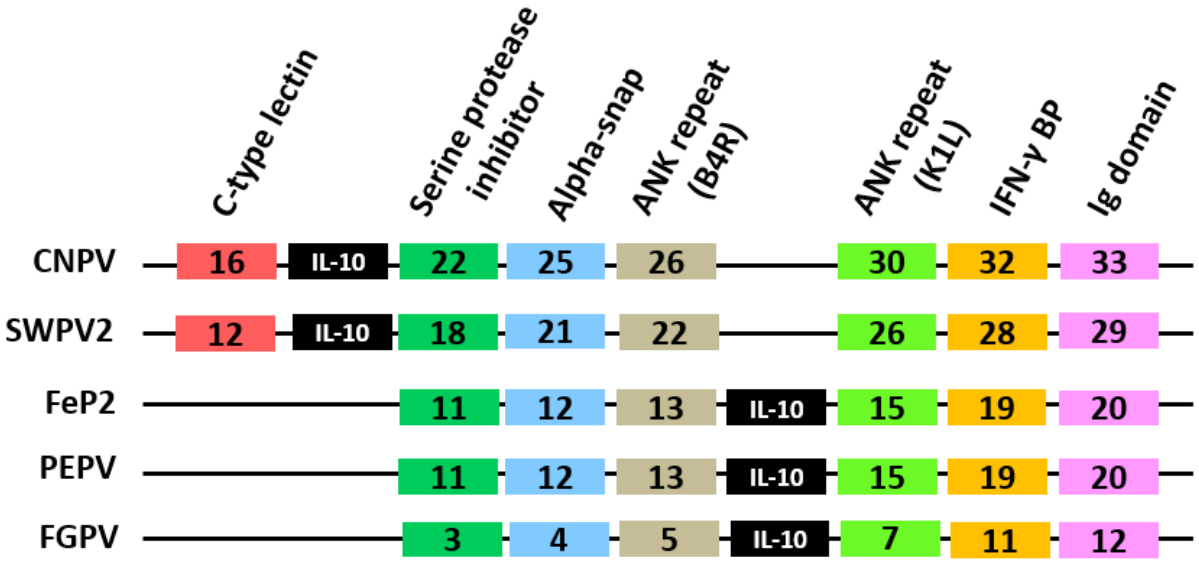


Figure 3.3: Schematic of the relative location of IL-10 orthologues in five avian poxvirus genomes. ORFs are depicted as coloured blocks with syntenic orthologues present in all five genomes in the same colour. Any intervening ORFs that are not present

and syntenic in all five genomes have been removed for clarity, except CNPV016 and SWPV-2 012.

UBQ: The two novel avipoxviruses encode an ubiquitin homologue (pepv074 and fgpv071) most similar to human ubiquitin, S5a ubiquitin interacting motif-1 (UIM1, accession: 1YX5_B) (Wang *et al*., 2005) in the same genomic location as CNPV, SWPV1 and SWPV2. FWPV and FeP2 were shown to have fragmented remains of this homologue in the same location while TKPV showed no evidence of a ubiquitin like protein at this location. The CNPV, PEPV and FGPV homologues show 100% amino acid sequence identity to each other and ubiquitin sequences from eukaryotes. The FGPV sequence is identical while the PEPV sequence has two amino acid changes at positions 20 and 55.

Brevican core protein: FGPV encodes an orthologue (fgpv232) of a C-type lectin with similarity to brevican core proteins encoded by eukaryotes (including reptiles, amphibians and fish). Amino acid identity is relatively low at 30-40% with the region of identity corresponding to the ligand binding surface. C-type lectins of this variety have not been identified in viral species to date.

C7L: Like FeP2 and TKPV, PEPV and FGPV contain orthologues of orthopoxvirus C7L (pepv231 and fgpv234) which are found in equivalent genome positions between orthologues of fpv216 and fpv217. These orthologues are highly conserved with 96-98% aa identity. Although there is no ORF present, the equivalent region in the FWPV genome shows 67-68% nt identity to the above ORFs suggesting that FWPV may have once contained a C7L orthologue.

3.3.9 Reticuloendotheliosis Virus

Genome sequence analysis was used to interrogate the avipoxvirus genomes for the presence of reticuloendotheliosis virus (REV) sequences. REV insertions are typically found between fpv201 and fpv203 although the entire genomes were searched for the presence of REV LTRs as well as the *gag*, *pol* and *env* genes. None of the REV elements were found in either of the genomes.

3.4 DISCUSSION

Overall, the genome sequences from the two, novel, South African avian poxviruses exhibited genome structures typical of poxviruses more generally, and more specifically of avian poxviruses. Importantly, both PEPV and FGPV contained the 89 genes considered to be conserved and essential among chordopoxviruses. Both the PEPV and FGPV genomes were found to have unique complements of gene family proteins as well as several ORFs that were truncated, fragmented or extended relative to their closest orthologues.

The avipoxvirus genomes vary in length from 189kb (TKPV) to over 300kb (CNPV, SWPV-1, SWPV-2 and PEPV), as well as number of ORFs from 171 (TKPV) to over 300 (CNPV, SWPV-1, SWPV-2). It has been shown with the comparison of the FWPV and CNPV genomes that this kind of variation is not unusual within this genus. What was less clear at the beginning of this study was the kind of variation that could be expected within and between subclades.

Phylogenetically, based on both single and multiple gene alignments, PEPV grouped in subclade A2 along with FeP2 allowing intra-clade comparisons and FGPV grouped in A3 allowing further inter-clade comparisons along with FWPV. One obvious point to note is the fact that the FGPV genome was not fully sequenced, with the ITRs remaining undetermined. Because these regions are known to be variable among other poxviruses this was not considered a barrier to determining the degree of relatedness of these viruses. The most notable differences between isolates were present as deletions of large areas of sequence in the central regions.

A nucleotide alignment of 130 conserved ORFs showed PEPV and FGPV to be most closely related to each other and the other South African isolate, FeP2 with ~96% nt identity. This high level of identity between conserved ORFs is expected as these ORFs are important and/or essential to the virus life cycle. Comparison to non-clade A viruses showed nucleotide identities of ~65-70% (Table 3.5).

As has been noted in other avian poxvirus genomes, the central regions of both viruses are more conserved and terminal regions are more variable. Figures 3.1 and

3.2 clearly show these variabilities in the terminal regions with several ORFs having been truncated, fragmented or extended relative to their closest homologues. One FGPV ORF was discovered (fgpv232) that has yet to be identified in other avian poxvirus genomes. Of the 43 PEPV ORFs noted to be truncated, fragmented or extended 15 (34%) encode hypothetical proteins, and 23 (53%) encode proteins belonging to gene families (ankyrin repeat, CC-chemokine, V-type Ig domain, β-NGF, N1R/p28, VARV B22R. HT motif, serpin,). Similarly, of the 25 FGPV ORFs noted to be truncated, fragmented or extended, eight (33%) encode hypothetical proteins, and 12 (46%) encode proteins belonging to gene families (ankyrin repeat, CC-chemokine and V-type Ig domain). Also present in the terminal regions are the majority of the genes of interest that are only present in a subset of avian poxvirus genomes, if at all (summarised below).

IL-10: Interleukin 10 (IL-10) is a cytokine with both immunostimulatory but more significantly, immunosuppressive effects. Viruses are known to modulate host immune responses by both up-regulating expression of host IL-10 and by encoding their own virally produced IL-10. IL-10 orthologues have been identified in several viral species including several species of herpes and alloherpes viruses as well as several parapoxvirses, capripoxviruses and avipoxviruses (Ouyang *et al.,* 2014).

To date, IL-10 had been identified in the genomes of CNPV, SWPV-2 and FeP2 and has now also been identified in PEPV and FGPV. This ORF is found in the same location as the copy found in Fep2, between ORF equivalents pepv013-pepv015 and fgpv005-fgpv007 which is in the terminal regions of these viruses (Fig. 3.3). This region is highly conserved between the three African isolates but the lack of conservation between clades makes it difficult to confirm whether these introductions were from the same or two separate HGT events

UBIQUITIN: Ubiquitin is a highly conserved protein of 76aa that when post-translationally added to target proteins directs a diverse array of cellular processes including but not limited to endocytosis, stress responses, gene transcription and silencing and protein degradation (reviewed in Komander and Rape, 2012). Although interference with and manipulation of the host ubiquitin-proteasome pathway by viruses is common, virally encoded ubiquitin has been identified in species from only

five viral families: bovine viral diarrhoea virus (BVDV), a pestivirus (flavivirus) (Meyers *et al*., 1989), several species of baculovirus including *Autographa californica* nucleopolyhedro virus (AcNPV) (Guarino, 1990) and *Orygia pseudotsugata* multinucleocapsid nuclear polyhedrosis virus (OpMNPV) (Russell and Rohrmann, 1993), the Marseille viruses, Lausanne virus and Port Miou virus, and in the poxviruses CNPV (Tulman *et al*., 2004), as well as MSEV (Afonso *et al*., 1999), and AMEV (Bawden *et al*., 2000), both entomopoxviruses.

In BVDV, acquisition of a cellular UBQ sequence by recombination is implicated in the evolution of non-cytopathogenic (ncp) biotypes to cytopathogenic (cp) biotypes which are associated with mucosal disease (Becher *et al*., 1998). In AcNPV, which has been shown to differ by 18 amino acid residues to animal ubiquitin it was shown that ubiquitin was not essential for productive virus infection but that both infectious and total titres were reduced in the mutant viruses that did not express ubiquitin (Reilly and Guarino, 1996).

Avipoxviruses are known to encode several copies (Table 3.7) of genes encoding RING finger motif proteins (N1R/p28) that are similar to the p28 protein encoded by OPVs. The functions of these proteins are unclear, but it has been suggested that they function similarly to p28 (Huang et al., 2004,) showing the importance of these ubiquitin-proteasome pathway modifying enzymes in the avipoxvirus life cycle. Although it is currently unclear if this protein is expressed and/or functional, it may warrant further investigation given the complex role ubiquitination can play in the viral life cycle.

Brevican core protein: FGPV encodes an orthologue (fgpv232) of a C-type lectin with similarity to brevican core proteins encoded by eukaryotes. Amino acid identity is relatively low at 30-40%. The FGPV ORF is fragmented relative to the closest orthologue but contains a C-type lectin-like domain (CTLD) of the type found in natural killer cell receptors (NKRs). NKRs are associated with both activation and inhibition of NK cells. To our knowledge, C-type lectins of this variety have not been identified in any viral species to date.

C7L: The C7L host-range gene was first identified in VACV and shown to function equivalently to K1L (Perkus *et al*., 1990). OPV C7L orthologues are relatively highly conserved in terms of aa sequence identity and are found in the terminal regions of the OPV genomes. In contrast, orthologues from the *Leporipoxvirus, Yatapoxvirus, Capripoxvirus* and *Suipoxvirus* genomes show a lesser degree of identity between each other and compared to the OPVs and are also found in the central regions of these genomes (Gubser, 2004). In NYVAC, which lacks a copy of C7L, insertion of C7L into the NYVAC genome restores replicative ability in HeLa cells and demonstrates anti-apoptotic effects (Najera *et al*., 2006). The first 10 amino acids at the N-terminus of the protein were shown to be essential for C7L function while the last 10aa at the C-terminus were shown to be dispensable. A further ten amino acids, including tyrosine and isoleucine residues at 135 and 136 were also found to be essential (Terajima *et al*., 2013). In terms of the avian poxviruses, the two sets of 10 amino acid residues are conserved within the APV genus but not between the APV and OPV isolates. Functional studies would need to be conducted to determine any potential function of C7L in avipoxviruses.

Ankyrin repeat family: The FGPV genome was noted to have significantly more ankyrin repeat family proteins than what has been reported in other, clade A avian poxviruses but if truncated and fragmented ORFs are assumed to be non-functional and therefore excluded, the number of ankyrin repeat family proteins decreases to levels previously noted in other, clade A, avian poxvirus genomes (FWPV and PEPV). The large number of disrupted ankyrin repeat proteins found in the FGPV genome could be due to gradual loss of these ORFs that were once the result of genomic accordion gene expansions. In the left hand, terminal region, are four ORFs containing ankyrin repeats (fgpv002, fgpv005, fgpv008 and fgpv0023) that are only found in the genomes of South African isolates and five of the six FGPV ORFs that are homologues of ORFs only found in CNPV, are found in the terminal regions of the FGPV genome.

CHAPTER 4: COMPARATIVE GENOMICS

4.1 INTRODUCTION

4.2 METHODS

4.3 RESULTS

4.3.1 Dotplot analysis

4.3.2 Regions of Difference

4.4 DISCUSSION

4.5 CONCLUSIONS

4.1 INTRODUCTION

The FWPV genome was first shown to exhibit major organisational differences compared to the genome of Vaccinia virus (VACV) using restriction enzyme mapping. It was shown that large segments of the FWPV genome had been reversed and/or translocated relative to VACV although gene content appeared to be largely maintained (Mockett *et al*., 1992). Sequencing of the FWPV genome and other ChPV genomes allowed for more detailed comparisons which showed that the core region forms a continuous block in all ChPVs except parapoxviruses and avian poxviruses due to various genome rearrangements. It was specifically noted that the core region of avian poxviruses has broken into four segments, two of which have been reversed and one of which has been translocated (Afonso *et al*., 2000; Gubser *et al*., 2004a). Sequencing of the CNPV genome allowed comparison of the regions found between these four segments, to the equivalent FWPV regions, and showed major differences in gene content (Tulman *et al*., 2004). At the time, it was unclear if these differences were due to subclade specificities or were a feature of all avian poxviruses. With seven genomes now available, an analysis of genome content and gene synteny was undertaken.

4.2 METHODS

Dotplots were created using Gepard software, with word length = 10 (Krumsiek *et al*., 2007). Pairwise and multiple sequence alignments were created using MAFFT version 7 (Katoh *et al*., 2002) with default settings. Genome sequences of each of the fully sequenced avian poxvirus genomes (FWPV - AF198100; Fp9 - AJ581527; CNPV - AY318871; FeP2 - KJ801920; PEPV - KJ859677; TKPV - KP728110; SWPV1 - KX857216; SWPV2 - KX857215) as well as VACV strain Copenhagen (M35027) were used in dotplot analysis.

To identify regions of difference between genomes, a table of orthologues was constructed for every ORF present in every genome (Table 4.1. This table was manually curated by examining published ORF tables (Afonso *et al*., 2000; Bányai *et*

al., 2015; Offerman *et al*., 2014; Sarker *et al*., 2017; Tulman *et al*., 2004) and the relative genome sequences found in Genbank.

4.3 RESULTS

4.3.1 Dotplot Analysis

Dotplots were created comparing PEPV and FGPV with other sequenced avian poxvirus genomes and VACV (Copenhagen) to compare overall genomic synteny. This analysis showed both the PEPV and FGPV genomes to be highly syntenic with FWPV, FeP2, and each other overall, and to show major differences compared to CNPV, TKPV, and the two SWPV genomes, as expected, due to the large differences in genome size. Also notable are the two large breaks in synteny located in the central regions of the dotplots, indicated by arrows (Figs. 4.1 and 4.2).

Dotplot analysis of PEPV (Fig.4.1H) and FGPV (Fig.4.2H) vs VACV shows the fragmentation of the core region and inversion of the two of the fragments in the avian poxviruses relative to VACV.

4.3.2 Regions of Difference

To further investigate the breaks in synteny identified in the dotplots in 4.3.1, full genome alignments (data not shown) and ORF tables from published genome sequences were used to produce a synteny table (Table 4.1). This table starts at the first ORF conserved in all avipoxviruses (fpv016; cnpv032) and ends at the second last ORF conserved in all avipoxviruses (fpv219; cnpv296). If conserved in all ChPVs, the VACV homologues are included in the last row of the table. If an ORF is conserved and syntenic across all eight avipoxvirus genomes the row is shaded. In cases where only one genome is missing a particular ORF, or that ORF is truncated/ fragmented (T/F), or differs in length by >20%, the entry is in bold and underlined. In cases where multiple ORFs differ in length by >20%, all entries are in bold. Regions of difference are identified by the presence of cell borders. For the purposes of this table, if an ORF in one genome is not syntenic with an ORF in at least one other genome it is recorded as absent. This does not necessarily mean that the ORF is not present in that genome, just that it is not present at the same location.

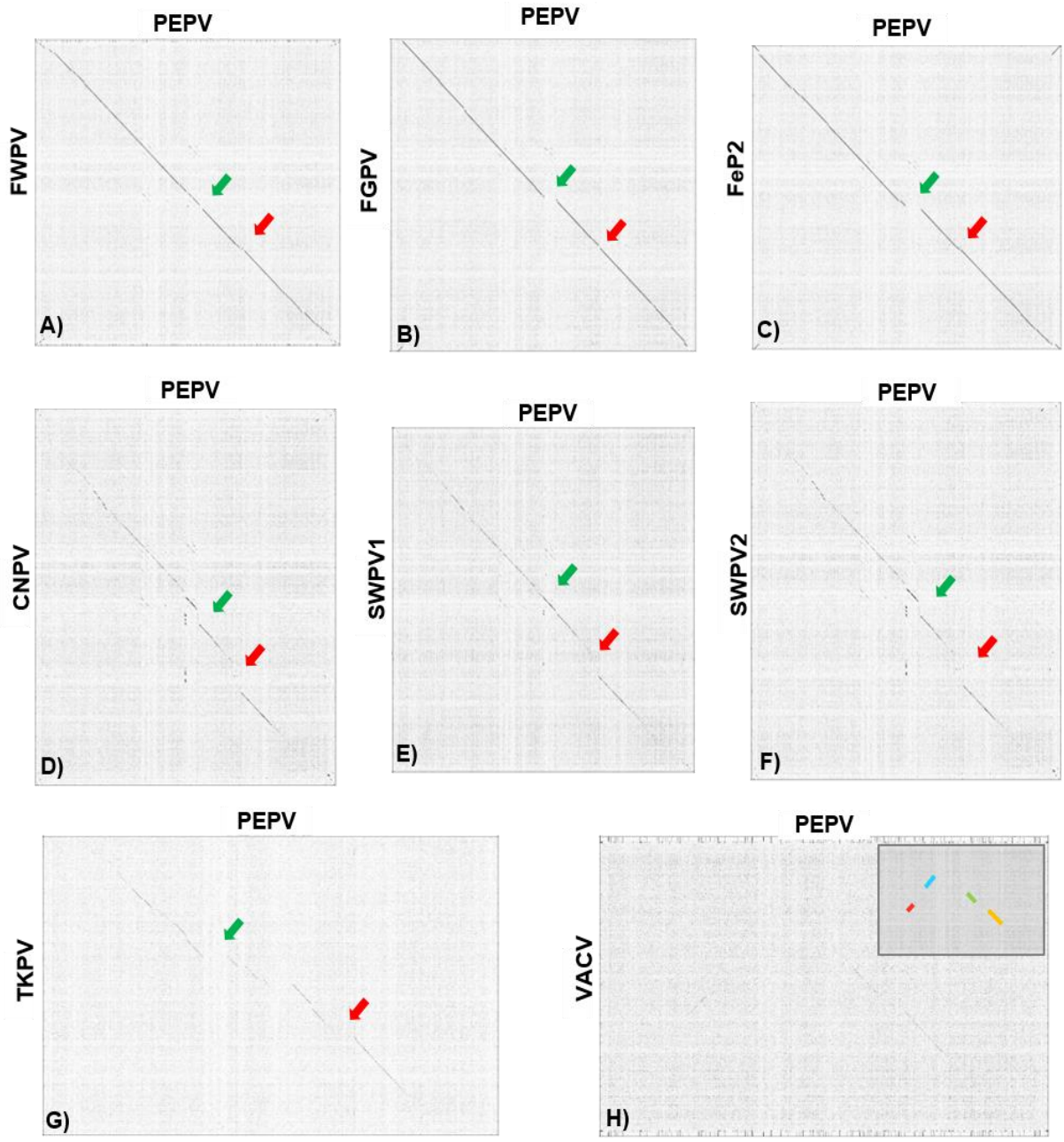


Figure 4.1: Dotplots of the PEPV genome (x axis) vs other sequenced avian poxvirus genomes (y axis). A) PEPV vs FWPV **B)** PEPV vs PEPV **C)** PEPV vs FeP2 **D)** PEPV vs CNPV **E)** PEPV vs SWPV1 **F)** PEPV vs SWPV2 **G)** PEPV vs TKPV **H)** PEPV vs VACV (Copenhagen) **H inset)** Conserved areas of dotplot 4.1H) highlighted in colours corresponding to Fig. 4.6. Green arrows indicate the first region of difference (fpv114-126) and red arrows indicate the second region of difference (fpv146-165). Plots are not to scale. Window size = 10

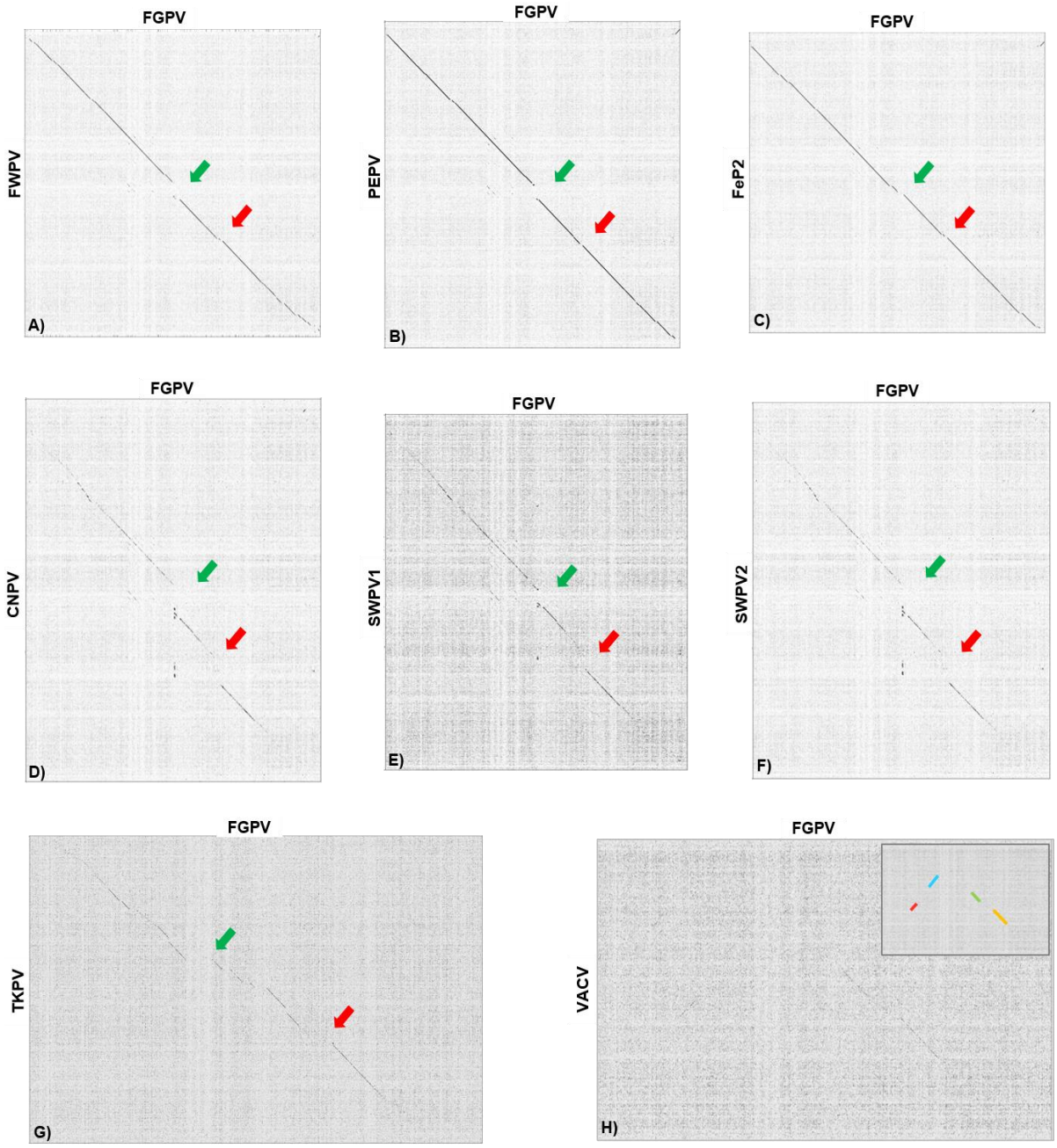


Figure 4.2: Dotplots of the FGPV genome (x axis) vs other sequenced avian poxvirus genomes (y axis). A) FGPV vs FWPV **B)** FGPV vs PEPV **C)** FGPV vs FeP2 **D)** FGPV vs CNPV **E)** FGPV vs SWPV1 **F)** FGPV vs SWPV2 **G)** FGPV vs TKPV **H)** FGPV vs VACV (Copenhagen) **H inset)** Conserved areas of dotplot (4.2H) highlighted in colours corresponding to Fig. 4.6 below. Green arrows indicate the first region of difference (fpv114-126) and red arrows indicate the second region of difference (fpv146-165). Plots are not to scale. Window size = 10.

Figures 4.1H and 4.2H show dotplot comparisons of PEPV and FGPV to VACV (strain Copenhagen) respectively. Regions of identity have been highlighted as coloured lines in the inset image which correspond to the conserved regions depicted in Figure 4.6.

Table 4.1: ORF synteny comparison of eight Avipoxviruses and VACV

FWPV	FeP2	PEPV	FGPV	CNPV	TKPV	SWPV-1	SWPV-2	VACV
16	19	19	11	32	1.1	24	28	
17	20	20	12	33	2	25	29	
absent	absent	absent	absent	absent	3	absent	absent	
absent	absent	absent	absent	absent	4	absent	absent	
18	21	21	13	34	absent	absent	30	
absent	absent	absent	absent	35	absent	absent	31	
19	22	22	14	36	absent	26	32	
absent	23	23	15/16	37	absent	27	33	
20	24	24	17	38	5	28	34	
21	25	25	18	39	6	29	35	
22	26	26	19	40	7	30	36	
23	27	27	20	41	8	31	37	
24	28	28	21	42	9	32	38	
25	29	29	22	43	**absent**	33	39	
absent	30	30	23	44	absent	34	40	
26	31F	31	24	absent	10	absent	absent	
27	**absent**	32	25	45	11	35	41	
28	32T	33	26	absent	absent	absent	absent	
absent	33	absent	27	46	absent	36	42	
29	34	34	28	47	**absent**	37	43	
30	35	35	29	48	12	38	44	
31	36	36	30	50	13	40	46	
32	**absent**	37	31	51	14	41	47	
absent	absent	absent	absent	52	absent	42	48	
33	absent	38	32	absent	absent	absent	absent	
34	37F	39F	33	absent	absent	43T	absent	
absent	absent	absent	absent	absent	15	absent	absent	
35	38	40	34	53	16	44	49	
36	absent	absent	35	54	absent	45	50	
37	39	41	36	55	17	46	51	
38	40	42	37	56	**18**	47	52	
absent	absent	absent	absent	57	19	48	53	
39	41	43	38	58	20	49	54	
40	42	44	39	59	21	50	55	
41	43	45	40	60	**absent**	51	56	
42	absent	absent	absent	absent	absent	absent	absent	
43	44	46	41	61	22	52	57	
44	45	47	42	62	23	53	58	
45	absent	absent	absent	absent	absent	absent	absent	
46	46	48	43	63	24	54	59	
absent	absent	absent	absent	64	absent	55	60	
45i	absent	absent	44F	absent	absent	absent	absent	
47	47	49	45	65	25	56	61	
48	48	50	46	68	26	59	64	
49	49	51	47	69	27	60	65	A1
50	50	52	48	70	28	61	66	D13
51	51	53	49	71	29	62	67	D12
absent	absent	absent	absent	72	absent	absent	68	
absent	absent	absent	absent	73	absent	63	69	
52	52	54	50	74	30	64	70	D11
53	53	55	51	75	31	65	71	D10
54	54	56	52	76	32	66	72	
55	55	57	53	absent	absent	absent	absent	
absent	absent	absent	absent	77	absent	absent	absent	
absent	absent	absent	absent	absent	absent	67	absent	
56	56	58	54	78	33	68	73	D7

FWPV	FeP2	PEPV	FGPV	CNPV	TKPV	SWPV-1	SWPV-2	VACV
57	57	59	55	80	34	69	75	D6
58	58	60	56	82	35	70	77	D5
59	absent	61	57	absent	absent	absent	absent	
60	59	**62F**	58	83	36	71	78	
absent	absent	absent	absent	absent	absent	72	absent	
61	60	63	59	absent	absent	absent	absent	
absent	61	absent	60	absent	absent	absent	absent	
62	62	64	61	84	37	73	79	D4
63	63TF	65	62	absent	absent	absent	absent	
absent	absent	absent	absent	85	absent	74	80	
absent	absent	absent	absent	absent	absent	75	absent	
absent	absent	absent	absent	86	38	76	81	
64	64F	66TF	63F	87	39	absent	82	
absent	absent	absent	absent	absent	absent	77	absent	
65	65	67	64	88	40	78	83	
66	**66**	**68**	**65**	**89**	**41**	**79**	**84**	
absent	absent	absent	absent	90	absent	80	85	
67	67	69	66	91	**absent**	81	86	
68	68	70	67	92	42	82	87	
69	69	71	68	93	43	83	88	D3
70	70	72	69	94	44	84	89	
70.5	**71**	**73**	**70**	**95**	**45**	**85**	**90**	
absent	absent	74	71	96	absent	86	91	
71	72	75	72	97	46	87	92	
absent	73	absent	absent	98	absent	88	93	
72	74F	absent	73	99	47	89	94	
73	75T	absent	74	100	48	90	95	
absent	absent	absent	absent	101	absent	absent	96	
74	76	77	75	102	**absent**	91	97	
75	77	78	76	103	50	92	98	
76	absent	79F	77	absent	85	absent	absent	
77	78?	80	78	104	51	93	99	G4
78	80	82	80	105	53	94	100	G3
79	79	81	79	106	52	95	101	G2
80	81	83	81	**absent**	61	96	102	
81	82	84	82	108	54	97	103	G1
82	83	85	83	109	55	98	104	I8
83	84	86	84	110	56	99	105	I7
84	85	87	85	111	57	100	106	I6
85	86	88	86	112	58	101	107	I5
86	87	89	87	113	60	102	108	
87	**88**	**90**	**88**	**114**	**62**	**103**	**109**	
88	89	91	89	115	63	104	110	I3
89	*90*	*92*	*90*	*116*	*63.1*	*105*	*111*	*I2*
90	91	93	91	117	64	106	112	I1
absent	92	94	92	absent	absent	absent	absent	
91	93	95	93	118	65	107	113	
92	94	96	94	119	66	108	114	
93	95	97	95	120	67	109	115	E10
94	96	98	96	121	68	110	116	E9
95	**97**	**99**	**97**	**122**	**69**	**111**	**absent**	**E8**
96	98	100	98	123	70	112	117	E6
97	99	101	99	124	**absent**	113	118	
98	100	102	100	125	**absent**	114	119	
99	101	103	101	126	**absent**	115	120	
100	102	104	102	127	71	116	121	E4
101	103	105	103	128	72	117	122	E2
102	104	106	104	129	73	118	123	E10

FWPV	FeP2	PEPV	FGPV	CNPV	TKPV	SWPV-1	SWPV-2	VACV
103	**105**	**107**	**105**	**130**	**TF**	**119**	**124**	**F17**
104	106	108	106	131	75	120	125	
105	107	109	107	132	76	121	126	
106	**108**	**110**	**108**	**133**	**77**	**122**	**128**	
107	109	111	109	134	**absent**	123	129	
108	110	112	110	135	78	124	130	F13
109	111	113	111	136	79	125	131	F12
110	112	114	112	137	80	126	132	
111	113	115	113	138	81	127	133	F10
112	114	116	114	139	82	128	134	F9
113	115	117	115	140	83	129	135	
114	**116**	**118**	**116**	**141**	**84**	**130**	**136**	
absent	117	absent	absent	absent	absent	absent	absent	
absent	118	119F	absent	absent	absent	absent	absent	
absent	absent	absent	117	absent	absent	absent	absent	
absent	absent	120F	absent	absent	absent	absent	absent	
absent	absent	absent	absent	absent	absent	131	absent	
absent	absent	absent	absent	142	absent	132	137	
absent	absent	absent	absent	143	absent	133	138	
115	absent	121	118/119	144	absent	134	139	
116	119	122F	120	absent	absent	absent	absent	
absent	absent	absent	absent	absent	85	absent	absent	
117	120	123	121	145	86	135	140	G5
118	*121*	*124*	*122*	*146*	*86,1*	*136*	*141*	*G5.5*
119	122	125	123	147	87	137	142	G6
120	123	126	124	148	88	138	143	G7
absent	124	absent	125	absent	absent	absent	absent	
121	absent	absent	absent	absent	absent	absent	absent	
absent	absent	absent	absent	149	absent	absent	144	
absent	absent	absent	absent	150	absent	absent	absent	
absent	absent	absent	absent	151	absent	absent	145	
absent	absent	absent	absent	152	absent	absent	146	
absent	absent	absent	absent	153	absent	absent	147	
absent	absent	128	absent	absent	absent	absent	absent	
122	absent	129ab	absent	154	absent	141	148	
123	absent	130	absent	155	absent	142	149	
absent	absent	absent	absent	156	absent	absent	150	
absent	absent	absent	absent	157	absent	absent	151	
absent	absent	absent	absent	158	absent	absent	absent	
absent	absent	absent	absent	159	absent	139	absent	
absent	absent	absent	absent	absent	absent	140	absent	
absent	absent	absent	absent	160	absent	absent	absent	
absent	absent	absent	absent	161	absent	absent	152	
absent	absent	131	absent	162	absent	absent	153	
absent	absent	absent	absent	163	absent	absent	absent	
absent	absent	absent	absent	164	absent	absent	absent	
124	absent	absent	126	165	absent	absent	154	
125	absent	absent	127	166	**absent**	143	155	
absent	absent	absent	128F	absent	**absent**	absent	absent	
absent	absent	132	129	absent	absent	absent	absent	
absent	absent	absent	130	absent	absent	absent	absent	
absent	125	absent	absent	absent	absent	absent	absent	
absent	absent	absent	absent	167	absent	144	156	
absent	absent	absent	absent	168	absent	absent	157	
absent	absent	absent	absent	169	absent	145	absent	
absent	126	133	131	170	absent	146	158	
126	127	134	132	171	89	147	159	G8

FWPV	FeP2	PEPV	FGPV	CNPV	TKPV	SWPV-1	SWPV-2	VACV
127	128	135	133	172	90	148	160	G9
128	129	136	134	173	91	149	161	L1
129	130	137	135	174	92	150	162	L2
130	131	138	136	175	93	151	163	L3
131	132	139	137	176	94	152	164	L4
132	133	140	138	177	95	153	165	L5
133	134	141	139	178	96	154	166	J1
134	135	142	140	179	97	155	167	J3
135	136	143	141	180	98	156	168	J4
136	137	144	142	181	99	157	169	J5
137	138	145	143	182	100	158	170	J6
138	139	146	144	183	101	159	171	H1
139	140	147	145	184	102	160	172	H2
absent	absent	absent	absent	absent	103	absent	absent	
absent	absent	absent	absent	185	absent	161	173	
140	141	148	146	186	104	162	174	H3
141	142	149	147	187	105	163	175	H4
142	143	150	148	188	106	164	176	H5
143	144	151	149	189	107	165	177	H6
144	145	152	150	190	108	166	178	H7
145	146	153	151	191	109	167	179	
146	147	154	152	192	110	168	180	D1
147	148F	155	153	193	absent	169	181	
148	149	156	154	194	111	170	182	D2
absent	150F	absent	absent	absent	absent	absent	absent	
absent	absent	absent	absent	195	absent	171	183	
149	151	157	**155T**	196	112	172	184	
150	152	158	156	197	**absent**	173	185	
absent	absent	absent	absent	198	absent	174	186	
151	153	159	157	199	113	175	187	
152	absent	160	158	absent	absent	absent	absent	
absent	absent	absent	absent	200	absent	176	188	
153	154	absent	159	201	absent	177	189	
154	absent	161	160	202	absent	178	190	
155	155	162	161	203	**absent**	179	191	
156	156	1 T; 1F	162	absent	absent	absent	absent	
absent	absent	absent	absent	204F	absent	180	192	
157	157	164	163	205	**absent**	181	193	
158	158	165	164	206	**absent**	182	194	
absent	absent	absent	absent	absent	absent	183	absent	
159	**159**	**166**	**165**	**207**	**115**	**184**	**195**	
160	**160**	**167**	**166**	**208**	**116**	**185**	**196**	
161	**161**	**168**	**167**	**209**	**117**	**186**	**197**	
absent	162	169	168	210	absent	187	198	
161,5	163	170	absent	211	absent	188	199	
2 frags	164	171	169	212	absent	189	200	
absent	absent	absent	absent	absent	absent	190	absent	
absent	absent	absent	absent	213	absent	191	201	
absent	absent	absent	absent	absent	118	absent	absent	
absent	absent	absent	absent	214	119	absent	202	
absent	absent	absent	absent	absent	absent	192	absent	
absent	absent	absent	absent	absent	absent	193	absent	
absent	absent	absent	absent	215	absent	194	203	
absent	absent	absent	absent	216	absent	absent	204	
absent	absent	absent	absent	217	absent	absent	205	
absent	absent	absent	absent	absent	absent	195	absent	
absent	absent	absent	absent	218	absent	196	206	
absent	absent	absent	absent	absent	absent	197	absent	

FWPV	FeP2	PEPV	FGPV	CNPV	TKPV	SWPV-1	SWPV-2	VACV
absent	absent	absent	absent	absent	absent	198	absent	
absent	absent	absent	absent	absent	absent	199	absent	
absent	absent	absent	absent	absent	absent	200	absent	
absent	absent	absent	absent	absent	absent	201	absent	
absent	absent	absent	absent	absent	absent	202	absent	
absent	absent	absent	absent	219	absent	absent	207	
absent	absent	absent	absent	220	absent	absent	208	
absent	absent	absent	absent	221	absent	absent	209	
absent	absent	absent	absent	222	absent	absent	210	
absent	absent	absent	absent	223	absent	absent	211	
162	165	172	170	absent	120	absent	absent	
163	166	173	171	absent	absent	absent	absent	
absent	167	174	172	absent	absent	absent	absent	
absent	absent	absent	absent	224	absent	203	212	
absent	absent	absent	absent	225	absent	absent	213	
absent	absent	absent	absent	226	absent	absent	214	
absent	absent	absent	absent	227	absent	absent	absent	
absent	absent	absent	absent	228	absent	absent	absent	
absent	absent	absent	absent	229	absent	204	215	
absent	absent	absent	absent	230	absent	absent	216	
absent	absent	absent	absent	231	absent	205	217	
absent	absent	absent	absent	232	absent	206	218	
absent	absent	absent	absent	233	absent	207	219	
absent	absent	absent	absent	234	absent	absent	220	
164	168	175	173	235	absent	absent	221	
absent	absent	absent	174	236	absent	209	222	
absent	absent	absent	absent	absent	absent	210	absent	
absent	absent	absent	175-177F	237	absent	absent	223	
165	169	177	178	238	121	211	224	A2
166	*170*	*178*	*179*	*239*	*121.1*	*212*	*225*	*A2.5*
167	171	179	180	240	122	213	226	A3
168	**172**	**180**	**181**	**241**	**123**	**214**	**227**	**A4**
169	173	181	182	242	124	215	228	A5
170	174	182	183	243	125	216	229	A6
171	175	183	184	244	126	217	230	A7
172	176	184	185	245	127	218	231	A8
173	*177*	*185*	*186*	*246*	*127.1*	*219*	*232*	*A9*
174	178	186	187	247	128	220	233	A10
175	179	187	188	248	129	221	234	A11
176	180	188	189	249	130	222	235	A12
177	181T	189	190	250	absent	223	236	
178	*182*	*190*	*191*	*251*	*130.1*	*224*	*237*	*A13*
179	*183*	*191*	*192*	*252*	*130.2*	*225*	*238*	*A14*
179,1	*184*	*192*	*193*	*253*	*130.3*	*226*	*239*	*A14.5*
180	*185*	*193*	*194*	*254*	*130.4*	*227*	*240*	*A15*
181	186	194	195	255	131	228	241	*A16*
182	187	195	196	256	132	229	242	*A17*
183	188	196	197	257	133	230	243	*A18*
184	189	197	198	258	134	231	244	*A19*
185	190	198	199	259	136	232	245	*A20*
186	191	199	200	260	135	233	246	*A21*
187	192	200	201	261	137	234	247	*A22*
188	193	201	202	262	138	235	248	*A23*
189	194	202	203	263	139	236	249	*A24*
190	195	203	204	264	140	237	250	
191	196	204	205	265	141	238	251	
192	197	205	206	266	142	239	252	*A28*

FWPV	FeP2	PEPV	FGPV	CNPV	TKPV	SWPV-1	SWPV-2	VACV
193	198	206	207	267	143	240	253	*A29*
194	199	207	208	268	**absent**	241	254	*A30*
194.1	200	208	209	268.1	absent	absent	absent	*A30.5*
195	201	209	210	269	**absent**	242	255	
196	202	210	211	270	144	243	256	
197	203	211	212	271	145	244	257	*A32*
198	204	212	213	272	146	245	258	*A34*
199	205F	213	214	absent	147	absent	absent	
200	206TF	214	215	absent	148	absent	absent	
absent	absent	absent	absent	absent	absent	246	absent	
201	207	215	216	273	149	247	259	
202	absent	absent	absent	absent	absent	absent	absent	
203	208	216	217	274	150	248	260	
204	209	217	218	275	**absent**	249	261	
205	210	218	219	276	151	250	262	
206	211	219	220	277	**absent**	251	263	
207	*212*	*220*	*221*	*278*	*151.1*	*252*	*264*	
absent	213	221	222	279	absent	253	265	
absent	absent	absent	223	280	absent	254	266	
208	214	222	224	281	152	255	267	
209	215	223	225	282	**absent**	256	268	
210	absent	absent	absent	absent	absent	absent	absent	
absent	absent	absent	226	283	absent	257	269	
absent	absent	absent	absent	284	absent	258	270	
211	216	225	227	285	153	259	271	
212	217	226	228	286	154	260	272	
213	218	227	229	287	155	261	273	
absent	absent	absent	absent	288	absent	262	274	
214	219	228	230	289	156	263	275	
215	220	229	231	290	**absent**	264	276	
absent	absent	absent	absent	291	absent	265	277	
absent	absent	absent	absent	292	absent	266	278	
absent	221F	absent	232	absent	absent	absent	absent	
216	**222**	**230**	**233**	**293**	**157**	**267**	**279**	
absent	absent	absent	absent	294	absent	268	280	
absent	absent	absent	absent	absent	absent	269	absent	
absent	absent	absent	absent	absent	absent	270	absent	
absent	absent	absent	absent	absent	158	absent	absent	
absent	223	231	234	absent	159	absent	absent	
217	224F	1 T; 1F	235/236 F	absent	160	absent	absent	
218	3 frags	233	237	295	absent	271	281	
219	226	234	238	296	161	272	282	

The differences between genomes are best seen as a compressed schematic (Figure 4.3) of the above table (Table 4.1). ORF names were removed from Table 4.1 but the cell shading and bordering were maintained. The table was shifted to a horizontal orientation in order to look at the comparison as a genome alignment reading from left to right. The five regions of difference indicated by bordered cells in Table 4.1 are further identified by curly brackets in Figure 4.3 below.

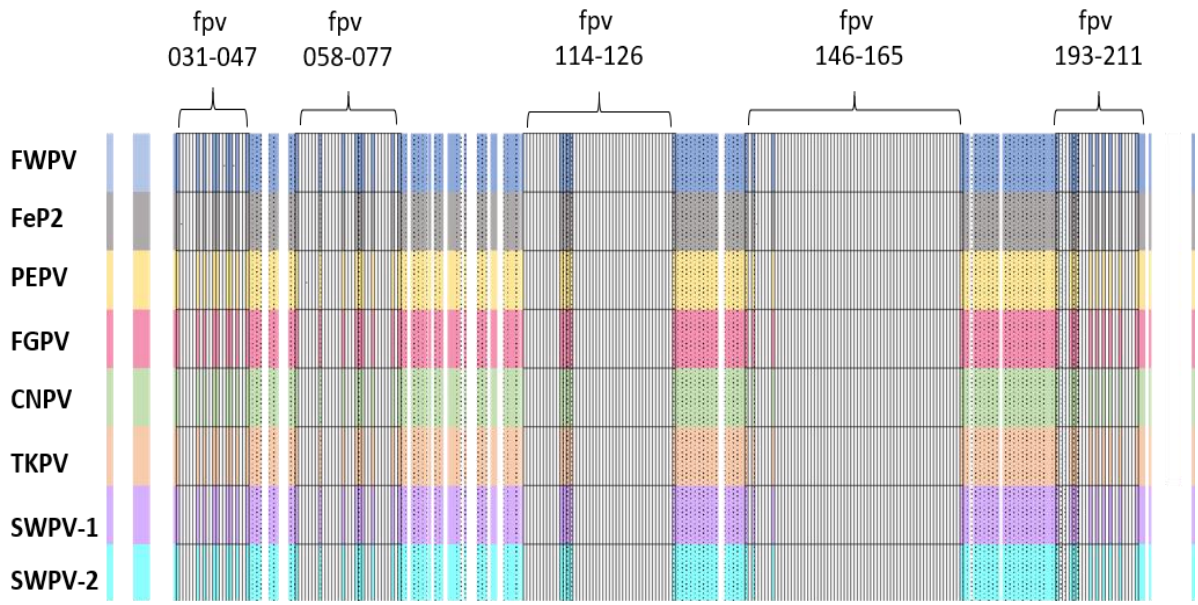
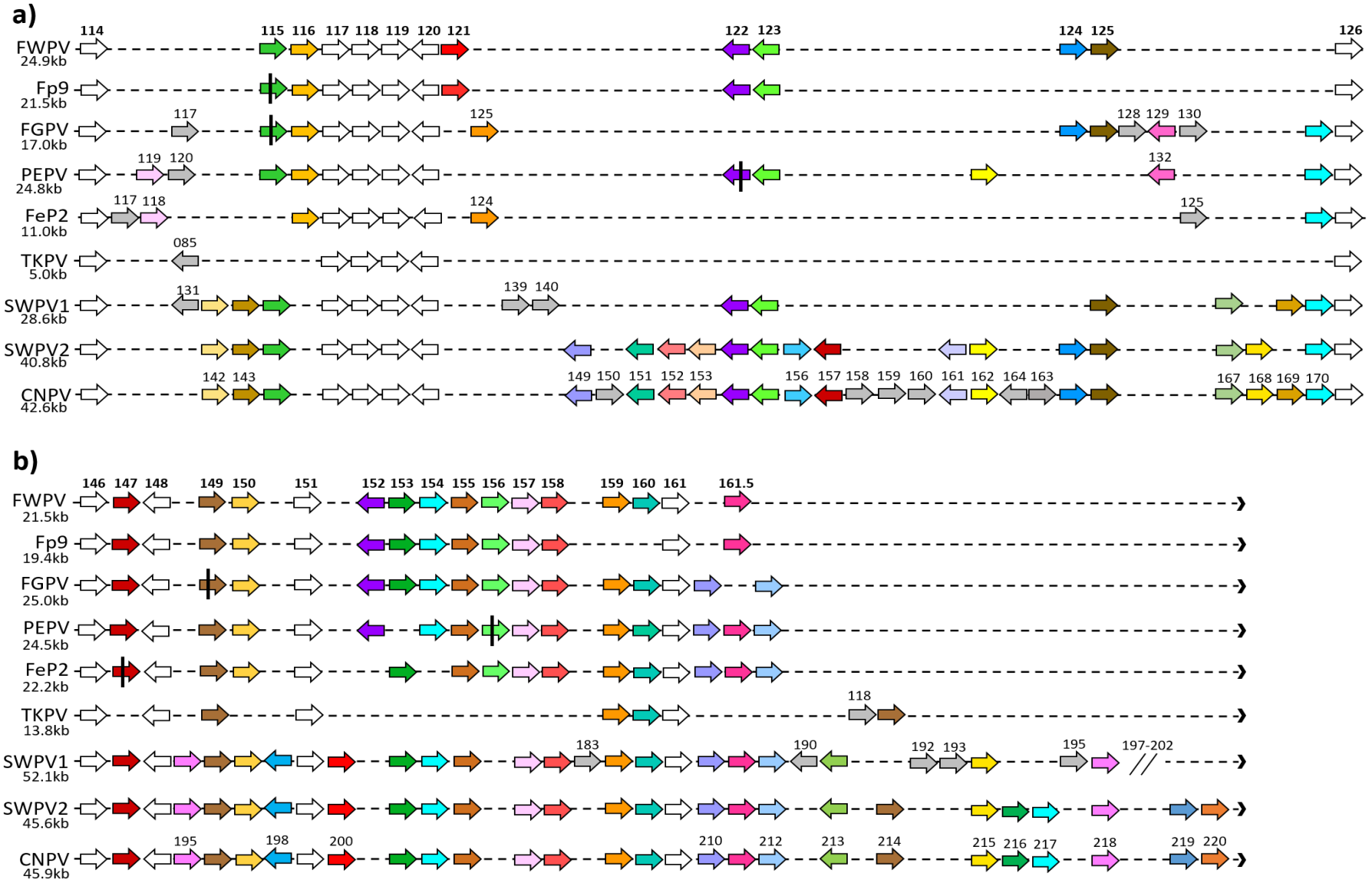


Figure 4.3: **ORF schematic of Table 4.1, compressed and with horizontal orientation.** As in the table, coloured blocks represent conservation of content and synteny; white blocks within the coloured areas represent a lack of conservation of gene content and synteny and outlined, white blocks represent the 5 regions of difference (curly brackets) identified in Table 4.1 above.

As was seen in the FeP2 and TKPV genomes (Offerman *et al*., 2014; Bányai *et al*., 2015), a large, rearranged region is present in the PEPV and FGPV genomes between pepv118-pepv134 and fgpv116-fgpv132 respectively (fpv114-fpv126; cnpv141-cnpv171 – green arrows in Figure 4.1), shown as an alignment schematic in Figure 4.4a. A second area of rearrangement is also noted between pepv154-pepv177 and fgpv152-fgpv178 respectively (fpv146-fpv165; cnpv192-cnpv238 – red arrows in Figure 4.2), shown as an alignment schematic in Figure 4.4b. These are the two large breaks in synteny identified in the dotplots (Figures 4.1 and 4.2). The alignment schematics show ORFs as arrows pointing in the direction of transcription. In regions where multiple genomes have unique ORFs (grey arrows), the ORFs have been placed in line with each other in order to conserve space, not to indicate synteny. Regions of rearrangement are referred to from here onward by the FWPV gene annotations as it is the prototype of the genus.

Three smaller regions of difference closer to the boundaries of the cores and the termini are also present. These regions fpv031-047 (pepv036-049; fgpv030-044; cnpv050-065), fpv058-077 (pepev060-080; fgpv055-078; cnpv082-104) and fpv193-211 (pepe206-225; fgpv207-227; cnpv267-285) were more similar in size across

genomes and therefore not as easily visible in the dotplots. These regions are shown in alignment schematics in Figures 4.5 a-c.


a)
FWPV 24.9kb
Fp9 21.5kb
FGPV 17.0kb
PEPV 24.8kb
FeP2 11.0kb
TKPV 5.0kb
SWPV1 28.6kb
SWPV2 40.8kb
CNPV 42.6kb
b)
FWPV 21.5kb
Fp9 19.4kb
FGPV 25.0kb
PEPV 24.5kb
FeP2 22.2kb
TKPV 13.8kb
SWPV1 52.1kb
SWPV2 45.6kb
CNPV 45.9kb

b) continued…

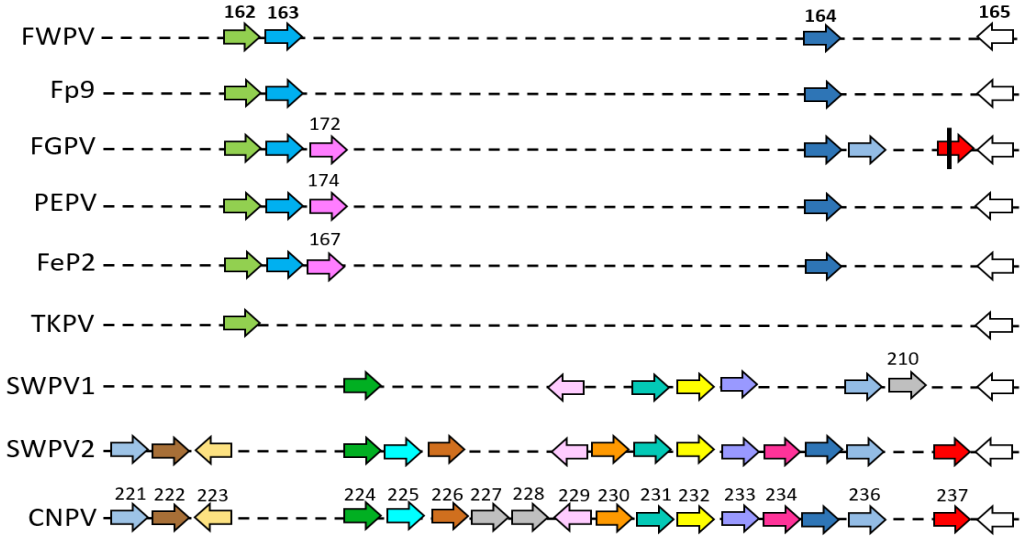
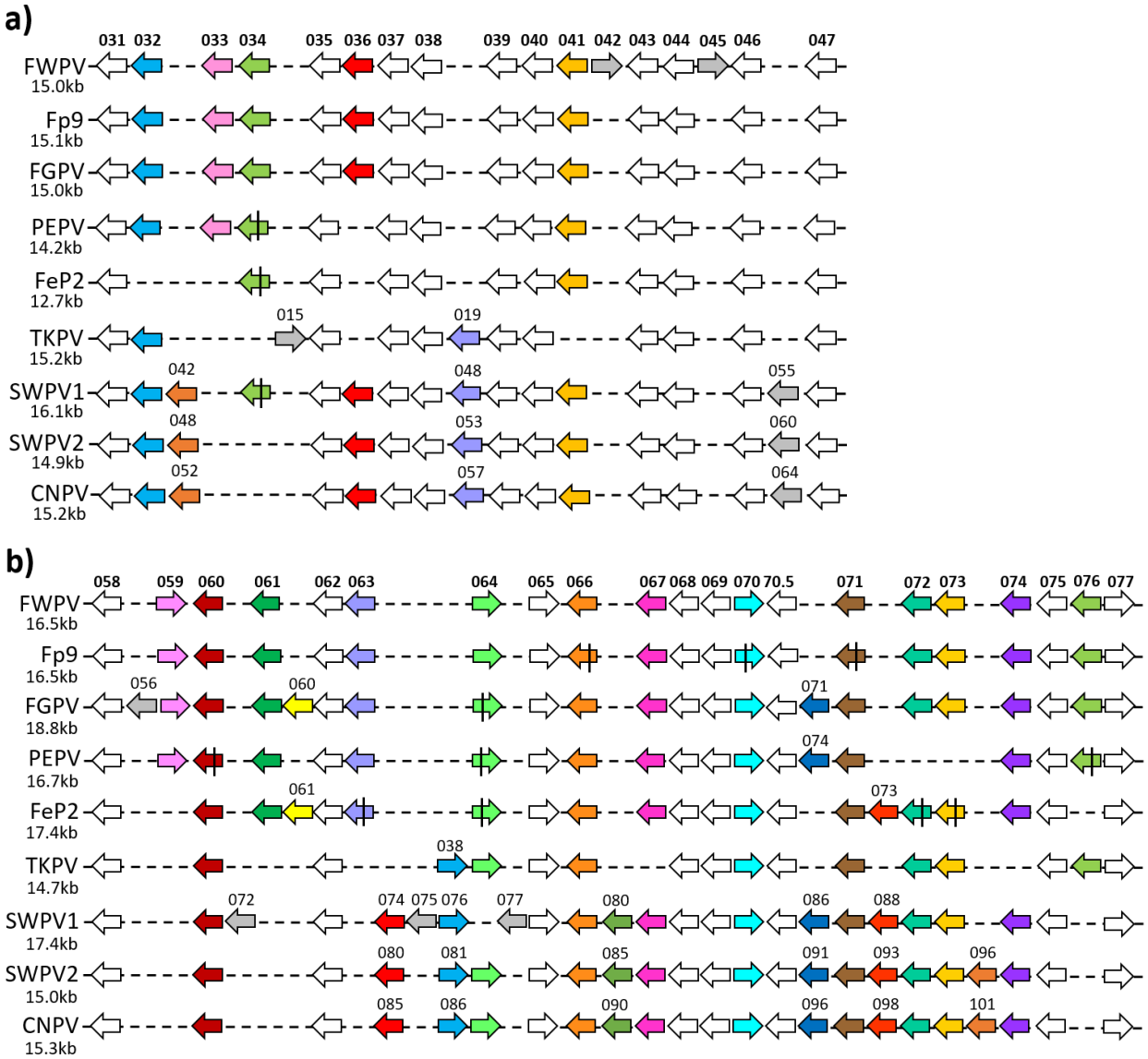


Figure 4.4: Schematic representing the ORFs present in the rearranged regions a) fpv114-126 and b) fpv146-165 in nine avian poxvirus genomes. ORFs are represented as arrows pointing in the direction of transcription. Numbers below the virus labels show the length of each region in kilobase pairs. (White = present in all genomes; Grey = unique to one genome; Coloured = present in 2-6 genomes or present in all genomes but with one or more orthologues not intact. Homologous, syntenic ORFs are shaded in the same colour across genomes (some colours have been repeated across the length of the genomes but do not indicate synteny or homology – only ORFs of the same colour (excluding grey) and directly above or below each other are syntenic homologues); Black vertical bar = fragmented and/or truncated ORF). Alignment is not to scale and ORF colours do not correspond between figures. Figure 4.4b) is continued over two pages.


a)
031 032 033 034 035 036 037 038 039 040 041 042 043 044 045 046 047
FWPV
15.0kb
Fp9
15.1kb
FGPV
15.0kb
PEPV
14.2kb
FeP2
12.7kb
015 019
TKPV
15.2kb
042 048 055
SWPV1
16.1kb
048 053 060
SWPV2
14.9kb
052 057 064
CNPV
15.2kb
b)
058 059 060 061 062 063 064 065 066 067 068 069 070 70.5 071 072 073 074 075 076 077
FWPV
16.5kb
Fp9
16.5kb
056 060 071
FGPV
18.8kb
074
PEPV
16.7kb
061 073
FeP2
17.4kb
038
TKPV
14.7kb
072 074 075 076 077 080 086 088
SWPV1
17.4kb
080 081 085 091 093 096
SWPV2
15.0kb
085 086 090 096 098 101
CNPV
15.3kb

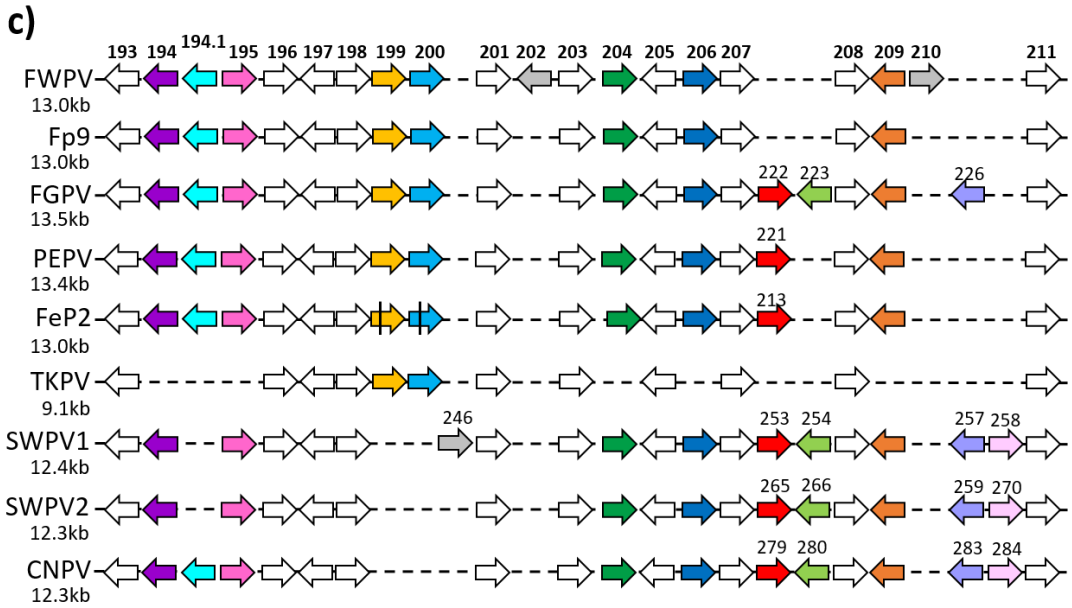


Figure 4.5: Schematic representing the ORFs present in the rearranged regions a) fpv031-047, b) fpv058-077 and c) fpv193-211 in nine avian poxvirus genomes. Annotations are depicted as in Figure 4.3.

To determine how much of an impact these regions of difference have on genome length comparisons, the ROD lengths were totalled and subtracted from both total genome lengths and central region lengths (Table 4.2). CNPV has the largest genome at 359.9kb and TKPV the smallest at 188.5kb equating to a size difference of 171.4kb. If the regions of difference are removed that difference decreases to 101.1kb and if the RODs are removed from the central region the size difference further decreases to just 41kb. For the clade A viruses, the maximum size difference between the central regions of FWPV and FGPV, excluding the RODs is 2.6kb compared to the maximum size difference of 10.8kb between full genomes of PEPV and FeP2.

Table 4.2: Size comparison of five regions of difference in eight avian poxvirus genomes

	FWPV	FeP2	PEPV	FGPV	CNPV	SWPV1	SWPV2	TKPV
Genome Length (kb)	288.5	282.3	306.9	293.1	359.9	326.9	351.1	188.5
Central Region (ORF #)	16-219	19-226	19-234	11-238	32-296	24-272	28-282	1.1-161
Central Region length (kb)	231.1	216.9	236.3	232.1	281.2	259.5	278.0	169.9
ROD Lengths (kb)								
#1	24.9	11.0	24.8	17.0	42.6	28.6	40.8	5.0
#2	21.5	22.2	24.5	25.0	45.9	52.1	45.6	13.8
#3	15.0	12.7	14.2	15.0	15.2	16.1	14.9	15.2
#4	16.5	17.4	16.7	18.8	15.3	17.4	15.0	14.7
#5	13.0	13.0	13.4	13.5	12.3	12.4	12.3	12.3
RODs total length (kb)	**90.9**	**76.3**	**93.6**	**89.3**	**131.3**	**126.6**	**128.6**	**61.0**
Total Genome minus ROD (kb)	197.6	206.0	213.3	203.8	228.6	200.3	222.5	127.5
Central Region minus ROD (kb)	140.2	140.6	142.7	142.8	149.9	132.9	149.4	108.9

Figure 4.6 shows a genome schematic of VACV compared to FWPV with each of the four conserved segments previously identified (Gubser *et al*., 2004) in different colours as in Figures 4.1H and 4.2H inset). The regions of difference identified in this study are shown below these segments, with ORF numbers and size in kilobases for FWPV. The regions of difference largely correspond with the definitions of the conserved regions identified previously although some overlap is present.

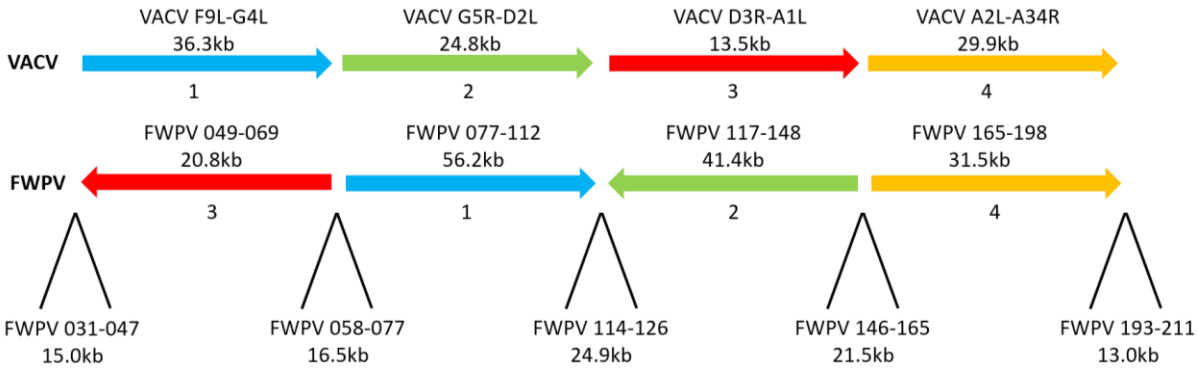


Figure 4.6: Genome schematic of VACV and FWPV genomes showing four regions conserved in gene content and synteny and five regions of difference. The four conserved regions are annotated as in Gubser ***et al***., 2004 and are not drawn to scale.

4.4 DISCUSSION

This analysis confirms the differences between avipoxvirus and orthopoxvirus genomes in terms of gene synteny. Although the overall genome architecture of avian poxviruses is largely conserved, with the expected variability in the termini, a pattern is emerging with all sequenced isolates exhibiting major differences in multiple, defined, central regions. These types of differences have not been identified in the genomes of isolates from other ChPV genera.

The FWPV genome was first shown to exhibit major organisational differences compared to the genome of Vaccinia virus (VACV) using restriction enzyme mapping. It was shown that large segments of the FWPV genome had been reversed and/or translocated relative to VACV although gene content appeared to be largely maintained (Mockett *et al.,* 1992). Sequencing of the FWPV genome and other ChPV genomes allowed for more detailed comparisons which showed that the core region forms a continuous block in all ChPVs except parapoxviruses and avian poxviruses due to various genome rearrangements. It was specifically noted that the core region of avian poxviruses has broken into four segments two of which have been reversed and one of which has been translocated (Gubser et al., 2004; Afonso et al., 2000). Sequencing of the CNPV genome allowed comparison of the regions found between these four segments, to the equivalent FWPV regions, and showed major differences in gene content. At the time, it was unclear if these differences were due to subclade specificities or were a feature of all avian poxviruses.

FP9 was the first clade A isolate noted to be somewhat different to FWPV in the fpv114-126 region with the truncation of fp9 115 and deletion of fp9 125 and fp9 126 (Laidlaw and Skinner, 2004). FeP2 was then noted to have a large deletion of over 10kb and although this region in the PEPV genome was of similar length to FWPV, several inserted and deleted ORFs were noted (Offerman *et al*., 2014). In the TKPV genome, ORF tkpv085 (fpv114) was identified as being affected by genomic rearrangement (Bányai *et al*., 2015).

In this study, alignment of this region in all clade A viruses shows a large variation in length from ~11kb in Fep2, to over 24kb in FWPV and PEPV. In FGPV this region spans 16.5 kb and encodes 16 ORFs (Fig. 3a). If this comparison is expanded to avian poxviruses in other clades, the difference is much larger with a variation in length from ~5kb (encoding seven ORFs) in TKPV to over 42.6kb (encoding 31 ORFs) in CNPV. Four ORFs in this region (fpv117 - fpv120 and the relative equivalents) are conserved among all viruses and syntenic as would be expected, as they are of the 83 genes conserved and considered essential among all ChPVs. What is unexpected is the placement and retention of this pocket of four essential genes in a region of highly divergent gene content and synteny. ORF fpv117 encodes a putative nuclease involved in viral DNA replication (Senkevich *et al*., 2009), fpv118 encodes RNA polymerase subunit RPO7, fpv119 is of unknown function and fpv120 encodes a virion core protein involved in several stages of virion morphogenesis (Mercer and Traktman, 2005). This region also contains several ORFs unique to avian poxviruses.

The second region of difference found in the genome core shows less difference in length among clade A viruses from ~19kb in Fp9 to ~25kb in FGPV but as above, when including viruses in other clades this difference in length increases considerably from ~14kb in TKPV (encoding 17 ORFs) to ~52kb in SWPV1 (encoding 42 ORFs) (Fig. 3b). Only one ORF (fpv148) conserved in all ChPV is present in this region and encodes a virion protein involved in immature virion formation (Szajner *et al*., 2004).

The three regions of difference found closer to the termini (Fig. 4a), b) and c)) are more similar in size across genomes and contain several more conserved ORFs.

Again, this is unusual for poxviruses as we would expect these regions to be less conserved compared to the two, central regions of difference. The three clade B viruses (CNPV, SWPV1 and SWPV2) are more similar to each other in terms of gene content at these locations than the clade A viruses.

Poxviruses have been known to use gene duplication and subsequent, mutationally driven, diversification of paralogues to their advantage to combat host immune responses. Elde *et al*., 2012, specifically looked at the ability of VACV to adapt to growth in human cells where the host range factor K3L is non-functional. E3L functions similarly to K3L and is functional in human cells. It was found that when E3L was deleted, leaving the virus susceptible to host antiviral responses, the K3L gene was recurrently amplified, with each of the paralogues able to explore mutational space until an adaptive substitution was found. Effective copies of the K3L gene were retained and the others lost over generations. It was also noted that duplications other than K3L all occurred in the terminal regions of the VACV genome (Elde *et al*., 2012).

In the case of avian poxviruses, it is interesting to note that several ORFs in the fpv114-126; cnpv141-171 region of difference of CNPV, SWPV1 and SWPV2 are present as repeats/paralogues of gene family proteins: cnpv143-144 = ANK repeat; cnpv150-151 = ANK repeat; cnpv154-155 = B22R; cnpv157-158 = TGF-β; cnpv159-160 = N1R/p28; cnpv161-162 = TGF-β; cnpv166-167 = Ig-like domain; cnpv168-169 = N1R/p28.

In the second region of difference (fpv146-165; cnpv192-238)) found in the central region, alignment of the clade A genomes shows a smaller variation in size from ~19kb in Fp9 to 25kb in FGPV. Again, when the non-clade A genomes are included, this size difference increases considerably to over 38kb. The CNPV genome has 18 copies of N1R/p28 like proteins in this region while SWPV1 and SWPV2 contain differing subsets of these, which may be the result of genomic accordions at work earlier in their evolutionary histories.

The RODs identified in this study are responsible for a significant portion of the differences in genome length between isolates. The TKPV genome is an outlier in

terms of genome length, and resequencing and reassembly would be appropriate to confirm the accuracy of the published sequence. If TKPV is removed from the analysis, the difference in size between the largest and smallest full genomes is 77.6kb and the difference in size between the largest and smallest genomes minus the RODs is 31kb, a reduction of more than half.

4.5 Conclusions

Several important bird species and commercial flocks have been shown to be severely affected by avian poxvirus infection. Genome sequencing and comparative genomics are the gold standards in terms of determining phylogenetic and evolutionary relationships among viral species and explaining differences in host range and pathogenicity. This study provides the genome sequence of two, novel, South African isolates from an African penguin and a lesser flamingo and provides insight into overall genome architecture that appears to be unique to avian poxviruses. Given the relative conservation of the central region of other poxvirus genomes, the regions of difference identified here are particular areas of interest in avian poxvirus genomics, but it is currently unclear why these regions would be so susceptible to rearrangement. The mechanisms responsible for such large-scale rearrangements are also yet to be elucidated. As more avian poxvirus genomes are sequenced, exploration and confirmation of these intriguing differences in these important pathogens can be conducted.

REAGENTS

McIlvains Buffer pH7.4

Solution A: 1.8mM citric acid (2.1g citric acid in 100ml dH_2O)

Solution B: 0.36mM $Na_2HPO_4 \cdot 12H_2O$ (7.2g $Na_2HPO_4 \cdot 12H_2O$ in 100ml dH_2O)

Add 18.17ml solution A and 1.83ml solution B and make up to 1L with dH_2O.

10x Phosphate Buffered Saline (PBS)

80g NaCl

2.0g KCl

14.4g Na_2HPO_4.

2.4g KH_2PO_4.

Dilute in 800ml dH_2O and adjust pH to 7.4.

Make up to 1L with dH_2O.

1x Phosphate Buffered Saline

Add 100ml 10x PBS to 900ml dH_2O.

PBS with Penicillin, Streptomycin and Fungin

1x PBS containing:

500U/ml penicillin

100µg/ml streptomycin

1ug/ml fungin

Physiological Saline

0.85% w/v NaCl

0.85g NaCl

Make up to 100ml with dH_2O.

TE Buffer pH9.0

10mM Tris-HCL

1mM EDTA

Make up to 400ml with dH_2O and adjust pH to 9.

Make up to 500ml with dH_2O.

Lysis Buffer

10N-lauryl sarcosinate

50mM Tris pH7.8

200 mM β-mercaptoethanol (added just before use).

1x Tris-borate-EDTA (TBE) Buffer

89mM Tris (269g)

89mM Boric acid (137.6g)

2mM EDTA (18.63g)

Make up to 25L with dH_2O

10% Buffered Formalin

37-40% Formaldehyde

35.03M $NaH_2PO_4 \cdot H_2O$

21.84M NaH_2PO_4

Make up to 1L with distilled water; ph7.4

REFERENCES

Abdallah, F.M., Hassanin, O., 2013. Detection and molecular characterization of avipoxviruses isolated from different avian species in Egypt. Virus Genes 46, 63–70.

Afonso, C.L., Delhon, G., Tulman, E.R., Lu, Z., Zsak, A., Becerra, V.M., Zsak, L., Kutish, G.F., Rock, D.L., 2005. Genome of Deerpox Virus. J. Virol. 79, 966–977.

Afonso, C.L., Tulman, E.R., Lu, Z., Zsak, L., Kutish, G.F., Rock, D.L., 2000. The genome of fowlpox virus. J. Virol. 74, 3815–3831.

Afonso, P.P., Silva, P.M., Schnellrath, L.C., Jesus, D.M., Hu, J., Yang, Y., Renne, R., Attias, M., Condit, R.C., Moussatché, N., Damaso, C.R., 2012. Biological Characterization and Next-Generation Genome Sequencing of the Unclassified Cotia Virus SPAn232 (*Poxviridae*). J. Virol. 86, 5039–5054.

Allwright, D.M., Burger, W.P., Geyer, A., Wessles, J., 1994. Avian pox in ostriches. J. S. Afr. Vet. Assoc. 65, 23–25.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403–410.

Amano, H., Morikawa, S., Shimizu, H., Shoji, I., Kurosawa, D., Matsuura, Y., Miyamura, T., Ueda, Y., 1999. Identification of the Canarypox Virus Thymidine Kinase Gene and Insertion of Foreign Genes. Virology 256, 280–290.

Anderson, F.E., Swofford, D.L., 2004. Should we be worried about long-branch attraction in real data sets? Investigations using metazoan 18S rDNA. Mol. Phylogenet. Evol. 33, 440–451.

Arai, S., Arai, C., Fujimaki, M., Iwamoto, Y., Kawarada, M., Saito, Y., Nomura, Y., Suzuki, T., 1991. Cutaneous tumour-like lesions due to poxvirus infection in Chilean flamingos. J. Comp. Pathol. 104, 439–441.

Armstrong, J.A., Metz, D.H., Young, M.R., 1973. The Mode of Entry of Vaccinia Virus into L Cells. J. Gen. Virol. 21, 533–537.

Awad, A.M., Abd El-Hamid, H.S., Abou Rawash, A.A., Ibrahim, H.H., 2010. Detection of reticuloendotheliosis virus as a contaminant of fowl pox vaccines. Poult. Sci. 89, 2389–2395.

Babkin, I.V., Babkina, I.N., 2011. Molecular Dating in the Evolution of Vertebrate Poxviruses. Intervirology 54, 253–260.

Babkin, I.V., Shchelkunov, S.N., 2006. Time scale of Poxvirus evolution. Mol. Biol. 40, 16–19.

Bagnall, B.G., Wilson, G.R., 1974. Molluscum contagiosum in a red kangaroo. Australas. J. Dermatol. 15, 115–120.

Bang, F.B. and Murphy, J.S., 1950. Electron microscopy of virus infections of chick chorioallantoic membrane. Bact. Proc. 1950:30.

Bányai, K., Palya, V., Dénes, B., Glávits, R., Ivanics, É., Horváth, B., Farkas, S.L., Marton, S., Bálint, Á., Gyuranecz, M., others, 2015. Unique genomic organization of a novel Avipoxvirus detected in turkey (*Meleagris gallopavo*). Infect. Genet. Evol. 35, 221–229.

Barnett, J., Dastjerdi, A., Davison, N., Deaville, R., Everest, D., Peake, J., Finnegan, C., Jepson, P., Steinbach, F., 2015. Identification of Novel Cetacean Poxviruses in Cetaceans Stranded in South West England. PloS One 10.

Baroudy, B.M., Venkatesan, S., Moss, B., 1982. Incompletely base-paired flip-flop terminal loops link the two DNA strands of the vaccinia virus genome into one uninterrupted polynucleotide chain. Cell 28, 315–324.

Bawden, A.L., Glassberg, K.J., Diggans, J., Shaw, R., Farmerie, W., Moyer, R.W., 2000. Complete Genomic Sequence of the Amsacta moorei Entomopoxvirus: Analysis and Comparison with Other Poxviruses. Virology 274, 120–139.

Becher, P., Orlich, M., Thiel, H.-J.R., 1998. Ribosomal S27a Coding Sequences Upstream of Ubiquitin Coding Sequences in the Genome of a Pestivirus. J Virol 72, 8.

Bidgood, S., Mercer, J., 2015. Cloak and Dagger: Alternative Immune Evasion and Modulation Strategies of Poxviruses. Viruses 7, 4800–4825.

Binns, M.M., Boursnell, M.E.G., Tomley, F.M., Campbell, J., 1989. Analysis of the fowlpoxvirus gene encoding the 4b core polypeptide and demonstration that it possesses efficient promoter sequences. Virology 170, 288–291.

Bolte, A.L., Meurer, J., Kaleta, E.F., 1999. Avian host spectrum of avipoxviruses. Avian Pathol. 28, 415–432.

Boulanger, D., Green, P., Jones, B., Henriquet, G., Hunt, L.G., Laidlaw, S.M., Monaghan, P., Skinner, M.A., 2002. Identification and Characterization of Three Immunodominant Structural Proteins of Fowlpox Virus. J. Virol. 76, 9844–9855.

Bracht, A.J., Brudek, R.L., Ewing, R.Y., Manire, C.A., Burek, K.A., Rosa, C., Beckmen, K.B., Maruniak, J.E., Romero, C.H., 2006. Genetic identification of novel poxviruses of cetaceans and pinnipeds. Arch. Virol. 151, 423–438.

Bratke, K.A., McLysaght, A., 2008. Identification of multiple independent horizontal gene transfers into poxviruses using a comparative genomics approach. BMC Evol. Biol. 8, 67.

Bwala, D.G., Fasina, F.O., Duncan, N.M., 2015. Avian poxvirus in a free-range juvenile speckled (rock) pigeon (*Columba guinea*). J. S. Afr. Vet. Assoc. 86.

Byrd, C.M., Hruby, D.E., 2006. Less Is More: Poxvirus Proteolysis. Microbe Mag. 1, 70–75.

Campbell, J.I., Binns, M.M., Tomley, F.M., Boursnell, M.E., 1989. Tandem repeated sequences within the terminal region of the fowlpox virus genome. J. Gen. Virol. 70 (Pt 1), 145–154.

Carter, G.C., 2005. Entry of the vaccinia virus intracellular mature virion and its interactions with glycosaminoglycans. J. Gen. Virol. 86, 1279–1290.

Carulei, O., Douglass, N., Williamson, A.-L., 2017. Comparative analysis of avian poxvirus genomes, including a novel poxvirus from lesser flamingos (Phoenicopterus minor), highlights the lack of conservation of the central region. BMC Genomics 18.

Carulei, O., Douglass, N., Williamson, A.-L., 2009. Phylogenetic analysis of three genes of Penguinpox virus corresponding to Vaccinia virus G8R (VLTF-1), A3L (P4b) and H3L reveals that it is most closely related to Turkeypox virus, Ostrichpox virus and Pigeonpox virus. Virol. J. 6, 52.

Cary J. Adams, Sanford H. Feldman, Jonathan M. Sleeman, 2005. Phylogenetic Analysis of Avian Poxviruses among Free-Ranging Birds of Virginia. Avian Dis. 49, 601–605.

Chang, A., Metz, D.H., 1976. Further Investigations on the Mode of Entry of Vaccinia Virus into Cells. J. Gen. Virol. 32, 275–282.

Colson, P., De Lamballerie, X., Yutin, N., Asgari, S., Bigot, Y., Bideshi, D.K., Cheng, X.-W., Federici, B.A., Van Etten, J.L., Koonin, E.V., La Scola, B., Raoult, D., 2013. “Megavirales”, a proposed new order for eukaryotic nucleocytoplasmic large DNA viruses. Arch. Virol. 158, 2517–2521.

DaMassa, A.J., 1966. The Role of Culex tarsalis in the Transmission of Fowl Pox Virus. Avian Dis. 10, 57–66.

Delhon, G., Tulman, E.R., Afonso, C.L., Lu, Z., de la Concha-Bermejillo, A., Lehmkuhl, H.D., Piccone, M.E., Kutish, G.F., Rock, D.L., 2004. Genomes of the Parapoxviruses Orf Virus and Bovine Papular Stomatitis Virus. J. Virol. 78, 168–177.

Diallo, I.S., Mackenzie, M.A., Spradbrow, P.B., Robinson, W.F., 1998. Field isolates of fowlpox virus contaminated with reticuloendotheliosis virus. Avian Pathol. J. 27, 60–66.

Diallo, I.S., Taylor, J., Gibson, J., Hoad, J., De Jong, A., Hewitson, G., Corney, B.G., Rodwell, B.J., 2010. Diagnosis of a naturally occurring dual infection of layer chickens with fowlpox virus and gallid herpesvirus 1 (infectious laryngotracheitis virus). Avian Pathol. J. 39, 25–30.

Docherty, D.E., and Slota, P.G., 1988. Use of Muscovy Duck embryo fibroblasts for the isolation of viruses from wild birds. J. Tissue Cult. Meth. 11, 165-70.

Elde, N.C., Child, S.J., Eickbush, M.T., Kitzman, J.O., Rogers, K.S., Shendure, J., Geballe, A.P., Malik, H.S., 2012. Poxviruses deploy genomic accordions to adapt rapidly against host antiviral defenses. Cell. 150, 831–841.

El-Zein, A., Nehme, S., Ghoraib, V., Hasbani, S., and Toth, B., 1974. Preparation of Fowl Pox Vaccine on Chicken-Embryo-Dermis cell culture. Avian Dis. 18(4), 495-506.

Fadly, A.M., Witter, R.L., Smith, E.J., Silva, R.F., Reed, W.M., Hoerr, F.J., Putnam, M.R., 1996. An outbreak of lymphomas in commercial broiler breeder chickens vaccinated with a fowlpox vaccine contaminated with reticuloendotheliosis virus. Avian Pathol. 25, 35–47.

Felsenstein, J., 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evol. Int. J. Org. Evol. 39, 783–791.

Felsenstein, J., 1978. Cases in which Parsimony or Compatibility Methods Will be Positively Misleading. Syst. Zool. 27, 401–410.

Fleischauer, C., Upton, C., Victoria, J., Jones, G.J.B., Roper, R.L., 2015. Genome sequence and comparative virulence of raccoonpox virus: the first North American poxvirus sequence. J. Gen. Virol. 96, 2806–2821.

García, M., Narang, N., Reed, W.M., Fadly, A.M., 2003. Molecular characterization of reticuloendotheliosis virus insertions in the genome of field and vaccine strains of fowl poxvirus. Avian Dis. 47, 343–354.

Gelenczei, E.F., and Lasher, H.N., 1968. Comparative studies of cell-culture propagated avian poxviruses in chickens and turkeys. Avian Dis. 12(1), 142-150.

Gilhare, V.R., Hirpurkar, S.D., Kumar, A., Naik, S.K., Sahu, T., 2015. Pock forming ability of fowl pox virus isolated from layer chicken and its adaptation in chicken embryo fibroblast cell culture. Vet. World 8, 245–250.

Gjessing, M.C., Yutin, N., Tengs, T., Senkevich, T., Koonin, E., Rønning, H.P., Alarcon, M., Ylving, S., Lie, K.-I., Saure, B., Tran, L., Moss, B., Dale, O.B., 2015. Salmon Gill Poxvirus, the Deepest Representative of the Chordopoxvirinae. J. Virol. 89, 9348–9367.

Gough, A.W., Barsoum, N.J., Gracon, S.I., Mitchell, L., Sturgess, J.M., 1982. Poxvirus infection in a colony of common marmosets (*Callithrix jacchus*). Lab. Anim. Sci. 32, 87–90.

Guarino, L.A., 1990. Identification of a viral gene encoding a ubiquitin-like protein. Proc. Natl. Acad. Sci. 87, 409–413.

Gubser, C., 2004. Poxvirus genomes: a phylogenetic analysis. J. Gen. Virol. 85, 105–117.

Gubser, C., Hué, S., Kellam, P., Smith, G.L., 2004. Poxvirus genomes: a phylogenetic analysis. J. Gen. Virol. 85, 105–117.

Gyuranecz, M., Foster, J.T., Dan, A., Ip, H.S., Egstad, K.F., Parker, P.G., Higashiguchi, J.M., Skinner, M.A., Hofle, U., Kreizinger, Z., Dorrestein, G.M., Solt, S., Sos, E., Kim, Y.J., Uhart, M., Pereda, A., Gonzalez-Hein, G., Hidalgo, H., Blanco, J.-M., Erdelyi, K., 2013. Worldwide Phylogenetic Relationship of Avian Poxviruses. J. Virol. 87, 4938–4951.

Halıgür, M., Özmen, Ö., Vural, S.A., Berkin, S., 2009. Pathological, immunohistochemical and electron microscopical examinations on chorioallantoic membrane lesions in experimental Fowl Poxvirus Infection. J. Kafkas Universitesi Veteriner Fakultesi Dergisi. 15.

Hendrickson, R.C., Wang, C., Hatcher, E.L., Lefkowitz, E.J., 2010. Orthopoxvirus Genome Evolution: The Role of Gene Loss. Viruses. 2, 1933–1967.

Henriques, A.M., Fagulha, T., Duarte, M., Ramos, F., Barros, S.C., Luís, T., Bernardino, R., Fernandes, T.L., Lapão, N., da Silva, J.F., Fevereiro, M., 2016. Avian poxvirus infection in a flamingo (*Phoenicopterus ruber*) of the Lisbon zoo. J. Zoo Wildl. Med. Off. Publ. Am. Assoc. Zoo Vet. 47, 161–174.

Hertig, C., Coupar, B.E.H., Gould, A.R., Boyle, D.B., 1997. Field and Vaccine Strains of Fowlpox Virus Carry Integrated Sequences from the Avian Retrovirus, Reticuloendotheliosis Virus. Virology 235, 367–376.

Hillis, D.M., 1996. Inferring complex phylogenies. Nature 383, 130–131.

Hollinshead, M., Vanderplasschen, A., Smith, G.L., Vaux, D.J. 1999. Vaccinia Virus Intracellular Mature Virions Contain Only One Lipid Membrane. J. Virol. 73(2),1503-17.

Huang, C.-Y., Lu, T.-Y., Bair, C.-H., Chang, Y.-S., Jwo, J.-K., Chang, W., 2008. A Novel Cellular Protein, VPEF, Facilitates Vaccinia Virus Penetration into HeLa Cells through Fluid Phase Endocytosis. J. Virol. 82, 7988–7999.

Huan, J., Huang, Q., Zhou, X., Shen, M.M., Yen, A., Yu, S.X., Dong, G., Qu, K., Huang, P., Anderson, E.M., Daniel-Issakani, S., Buller, M.L., Payan, D.G., Lu, H.H., 2004. The Poxviurs p28 Virulence Factor is an E3 Ubiquitin Ligase. J. Biol. Chem. 279, 54110-54116.

Hunter, S., 2014. Avian Poxvirus Infection in a Little Blue Penguin, *Eudyptula minor*. Kokako 21, 26.

Ibrahim, M.S., Antwerpen, M., Georgi, E., Vette, P., Zoeller, G., Meyer, H., 2014. Complete Genome Sequence of the Embu Virus Strain SPAn880. Genome Announc. 2(6).

ICTV 9th Report: Virus taxonomy: classification and nomenclature of viruses: Ninth Report of the International Committee on Taxonomy of Viruses. San Diego: Elsevier; 2012.

Ivanoc, I., Sainova, I., Kril, A., and Simeonov, K, 2001. Propagation of avian pox virus vaccine strains in duck embryo cell line –DEC 99. Exp. Pathol. Parasitol. 4(6), 46-49

Iyer, L.M., Aravind, L., Koonin, E.V., 2001. Common Origin of Four Diverse Families of Large Eukaryotic DNA Viruses. J. Virol. 75, 11720–11734.

Iyer, L.M., Balaji, S., Koonin, E.V., Aravind, L., 2006. Evolutionary genomics of nucleo-cytoplasmic large DNA viruses. Virus Res. 117, 156–184.

Jarmin, S., 2006. Avipoxvirus phylogenetics: identification of a PCR length polymorphism that discriminates between the two major clades. J. Gen. Virol. 87, 2191–2201.

Joklik, W.K., 1962. The purification of four strains of poxvirus. Virology 18, 9–18.

Kane, O.J., Uhart, M.M., Rago, V., Pereda, A.J., Smith, J.R., Van Buren, A., Clark, J.A., Boersma, P.D., 2012. Avian Pox in Magellanic Penguins (*Spheniscus magellanicus*). J. Wildl. Dis. 48, 790–794.

Katoh, K., Misawa, K., Kuma, K., Miyata, T., 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 30, 3059–3066.

Katz, Ehud, Moss, B., 1970. Vaccinia Virus Structural Polypeptide Derived from a High-Molecular-Weight Precursor: Formation and Integration into Virus Particles. J. Virol. 6, 717–726.

Katz, E., Moss, B., 1970. Formation of a vaccinia virus structural polypeptide from a higher molecular weight precursor: inhibition by rifampicin. Proc. Natl. Acad. Sci. 66, 677–684.

Kim, T.-J., Schnitzlein, W.M., McAloose, D., Pessier, A.P., Tripathy, D.N., 2003. Characterization of an avian poxvirus isolated from an Andean condor (*Vultur gryphus)*. Vet. Microbiol. 96, 237–246.

Komander, D., and Rape, M., 2012. The ubiquitin code. Annu. Rev. Biochem. 81, 203-29.

Kotwal, G.J., Abrahams, M.-R., 2004. Growing poxviruses and determining virus titer. Methods Mol. Biol. Clifton NJ 269, 101–112.

Krumsiek, J., Arnold, R., Rattei, T., 2007. Gepard: a rapid and sensitive tool for creating dotplots on genome scale. Bioinforma. Oxf. Engl. 23, 1026–1028.

Kulich, P., Roubalová, E., Dubská, L., Sychra, O., Smíd, B., Literák, I., 2008. Avipoxvirus in blackcaps (*Sylvia atricapilla*). Avian Pathol. J. 37, 101–107.

Kumar, S., Stecher, G., Tamura, K., 2016. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol. Biol. Evol. 33, 1870–1874.

Laidlaw, S.M., Skinner, M.A., 2004. Comparison of the genome sequence of FP9, an attenuated, tissue culture-adapted European strain of Fowlpox virus, with those of virulent American and European viruses. J. Gen. Virol. 85, 305–322.

Le Loc'h, G., Bertagnoli, S., Ducatez, M.F., 2015. Time scale evolution of avipoxviruses. Infect. Genet. Evol. 35, 75–81.

Lefkowitz, E.J., Wang, C., Upton, C., 2006. Poxviruses: past, present and future. Virus Res. 117, 105–118.

Legendre, M., Bartoli, J., Shmakova, L., Jeudy, S., Labadie, K., Adrait, A., Lescot, M., Poirot, O., Bertaux, L., Bruley, C., others, 2014. Thirty-thousand-year-old distant relative of giant icosahedral DNA viruses with a pandoravirus morphology. Proc. Natl. Acad. Sci. 111, 4274–4279.

Legendre, M., Lartigue, A., Bertaux, L., Jeudy, S., Bartoli, J., Lescot, M., Alempic, J.-M., Ramus, C., Bruley, C., Labadie, K., Shmakova, L., Rivkina, E., Couté, Y., Abergel, C., Claverie, J.-M., 2015. In-depth study of Mollivirus sibericum, a new 30,000-y-old giant virus infecting Acanthamoeba. Proc. Natl. Acad. Sci. U. S. A. 112,

Lüschow, D., Hoffmann, T., Hafez, H.M., 2004. Differentiation of avian poxvirus strains on the basis of nucleotide sequences of 4b gene fragment. Avian Dis. 48, 453–462.

Manarolla, G., Pisoni, G., Sironi, G., Rampin, T., 2010. Molecular biological characterization of avian poxvirus strains isolated from different avian species. Vet. Microbiol. 140, 1–8.

Mapaco, L.P., Lacerda, Z., Monjane, I.V.A., Gelaye, E., Sussuro, A.H., Viljoen, G.J., Dundon, W.G., Achá, S.J., 2017. Identification of Clade E Avipoxvirus, Mozambique, 2016. Emerg. Infect. Dis. 23, 1602–1604.

Maruri-Avidal, L., Weisberg, A.S., Moss, B., 2013. Association of the Vaccinia Virus A11 Protein with the Endoplasmic Reticulum and Crescent Precursors of Immature Virions. J. Virol. 87, 10195–10206.

Massung, R.F., Knight, J.C., Esposito, J.J., 1995. Topography of variola smallpox virus inverted terminal repeats. Virology. 211, 350–355.

Mätz-Rensing, K., Ellerbrok, H., Ehlers, B., Pauli, G., Floto, A., Alex, M., Czerny, C.-P., Kaup, F.-J., 2006. Fatal Poxvirus Outbreak in a Colony of New World Monkeys. Vet. Pathol. 43, 212–218.

Mayr, A., 1963. New methods for the differentiation of avian pox viruses. Berl. Munch. Tierarztl. Wschr. 76, 316-324.

Mayr, A., and Kalcher, K. 1960 Comparative studies on the propagation of avian pox viruses in tissue culture. Arch. ges. Virusforsch. 76, 316-324.

McKenzie, R.A., Fay, F.R., Prior, H.C., 1979. Poxvirus infection of the skin of an eastern grey kangaroo. Aust. Vet. J. 55, 188–190.

McLysaght, A., Baldi, P.F., Gaut, B.S., 2003. Extensive gene gain associated with adaptive evolution of poxviruses. Proc. Natl. Acad. Sci. 100, 15655–15660.

Mercer, J., Helenius, A., 2008. Vaccinia Virus Uses Macropinocytosis and Apoptotic Mimicry to Enter Host Cells. Science. 320, 531–535.

Mercer, J., Traktman, P., 2005. Genetic and Cell Biological Characterization of the Vaccinia Virus A30 and G7 Phosphoproteins. J. Virol. 79, 7146–7161.

Meyers, G., Rümenapf, T., and Thiel, H-J., 1989. Ubiquitin in a togavirus. Nature. 341, 491.

Middlemiss, E., 1961. Avian Pox in South Africa. Ostrich. 32, 20–22.

Mockett, B., Binns, M.M., Boursnell, M.E., Skinner, M.A., 1992. Comparison of the locations of homologous fowlpox and vaccinia virus genes reveals major genome reorganization. J. Gen. Virol. 73 (Pt 10), 2661–2668.

Mondal, S.P., Lucio-Martínez, B., Buckles, E.L., 2008. Molecular characterization of a poxvirus isolated from an American flamingo (*Phoeniconais ruber ruber*). Avian Dis. 52, 520–525.

Moore, K.M., Davis, J.R., Sato, T., Yasuda, A., 2000. Reticuloendotheliosis Virus (REV) Long Terminal Repeats Incorporated in the Genomes of Commercial Fowl Poxvirus Vaccines and Pigeon Poxviruses without Indication of the Presence of Infectious REV. Avian Dis. 44, 827.

Moss, B., 2006. Poxvirus entry and membrane fusion. Virology. 344, 48–54.

Najera, J.L., Gomez, C.E., Domingo-Gil, E., Gherardi, M.M., Esteban, M., 2006. Cellular and Biochemical Differences between Two Attenuated Poxvirus Vaccine Candidates (MVA and NYVAC) and Role of the C7L Gene. J. Virol. 80, 6033–6047.

Neilan, J.G., Zsak, L., Lu, Z., Kutish, G.F., Afonso, C.L., Rock, D.L., 2002. Novel swine virulence determinant in the left variable region of the African swine fever virus genome. J. Virol. 76, 3095–3104.

Nichols, R., 2001. Gene trees and species trees are not the same. Trends Ecol. Evol. 16, 358–364.

Niemeyer, C., Favero, C.M., Kolesnikovas, C.K.M., Bhering, R.C.C., Brandão, P., Catão-Dias, J.L., 2013. Two different avipoxviruses associated with pox disease in Magellanic penguins (*Spheniscus magellanicus*) along the Brazilian coast. Avian Pathol. 42, 546–551.

O'Dea, M.A., Tu, S.-L., Pang, S., De Ridder, T., Jackson, B., Upton, C., 2016. Genomic characterization of a novel poxvirus from a flying fox: evidence for a new genus? J. Gen. Virol. 97, 2363–2375.

Offerman, K., Carulei, O., Gous, T.A., Douglass, N., Williamson, A.-L., 2013. Phylogenetic and histological variation in avipoxviruses isolated in South Africa. J. Gen. Virol. 94, 2338–2351.

Offerman, K., Carulei, O., van der Walt, A.P., Douglass, N., Williamson, A.-L., 2014. The complete genome sequences of poxviruses isolated from a penguin and a pigeon in South Africa and comparison to other sequenced avipoxviruses. BMC Genomics 15, 463.

Ouyang, P., Rakus, K., van Beurden, S.J., Westphal, A.H., Davison, A.J., Gatherer, D., and Vanderplasschen, A.F., 2014. IL-10 encoded by viruses; a remarkable example of independent acquisition of a cellular gene by viruses and its subsequent evolution in the viral genome. J. Gen. Virol. 95, 245-262.

Papadimitriou, J.M., Ashman, R.B., 1972. A poxvirus in a marsupial papilloma. J. Gen. Virol. 16, 87–89.

Parsons, N.J., Gous, T.A., Cranfield, M.R., Cheng, L.I., Schultz, A., Horne, E., Last, R.P., Lampen, F., Ludynia, K., Bousfield, B., Strauss, V., Peirce, M.A., Vanstreels, R.E.T., 2018. Novel vagrant records and occurrence of vector-borne pathogens in King Penguins (*Aptenodytes patagonicus*) in South Africa. Polar Biol. 41, 79–86.

Pastoret, P.-P., Vanderplasschen, A., 2003. Poxviruses as vaccine vectors. Comp. Immunol. Microbiol. Infect. Dis. 26, 343–355.

Payne, L.G., 1979. Identification of the vaccinia hemagglutinin polypeptide from a cell system yielding large amounts of extracellular enveloped virus. J. Virol. 31, 147–155.

Perkus, M.E., Goebel, S.J., Davis, S.W., Johnson, G.P., Limbach, K., Norton, E.K., Paoletti, E., 1990. Vaccinia virus host range genes. Virology 179, 276–286.

Philippe, N., Legendre, M., Doutre, G., Couté, Y., Poirot, O., Lescot, M., Arslan, D., Seltzer, V., Bertaux, L., Bruley, C., Garin, J., Claverie, J.-M., Abergel, C., 2013. Pandoraviruses: amoeba viruses with genomes up to 2.5 Mb reaching that of parasitic eukaryotes. Science 341, 281–286.

Pratik Singh, William M. Schnitzlein, Deoki N. Tripathy, 2005. Construction and Characterization of a Fowlpox Virus Field Isolate Whose Genome Lacks Reticuloendotheliosis Provirus Nucleotide Sequences. Avian Dis. 49, 401–408.

Ramos, M.C.C., Coutinho, S.D., Matushima, E.R., Sinhorini, I.L., 2002. Poxvirus dermatitis outbreak in farmed Brazilian caimans (*Caiman crocodilus yacare*). Aust. Vet. J. 80, 371–372.

Rampin, T., Pisoni, G., Manarolla, G., Gallazzi, D., Sironi, G., 2007. Epornitic of avian pox in common buzzards (*Buteo buteo*): virus isolation and molecular biological characterization. Avian Pathol. 36, 161–165.

Reilly, L.M., Guarino, L.A., 1996. The Viral Ubiquitin Gene of *Autographa californica* Nuclear Polyhedrosis Virus Is Not Essential for Viral Replication. Virology 218, 243–247.

Reteno, D.G., Benamar, S., Khalil, J.B., Andreani, J., Armstrong, N., Klose, T., Rossmann, M., Colson, P., Raoult, D., La Scola, B., 2015. Faustovirus, an asfarvirus-related new lineage of giant viruses infecting amoebae. J. Virol. 89, 6585–6594.

Rice, N.R., Hiebsch, R.R., Gonda, M.A., Bose, H.R., Gilden, R.V., 1982. Genome of Reticuloendotheliosis Virus: Characterization by Use of Cloned Proviral DNA. J. Virol. 42, 237–252.

Riper, C. van, Forrester, D.J., 2008. Avian Pox, in: Infectious Diseases of Wild Birds. Wiley-Blackwell, pp. 131–176.

Robinson, F.R., Twiehaus, M.J., 1974. Isolation of the avian reticuloendothelial virus (strain T). Avian Dis. 18, 278–288.

Sainova, I., Kril, A., and Ivanov, I, 2001. Propagation of avian pox virus vaccine strains in a mammalian EBTr cell line. A pilot study. Exp. Path. and Parasitolgy. 4(7), 40-46.

Saitou, N., Nei, M., 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4, 406–425.

Sarker, S., Das, S., Lavers, J.L., Hutton, I., Helbig, K., Imbery, J., Upton, C., Raidal, S.R., 2017. Genomic characterization of two novel pathogenic avipoxviruses isolated from pacific shearwaters (*Ardenna spp.*). BMC Genomics 18.

Satheshkumar, P.S., Olano, L.R., Hammer, C.H., Zhao, M., Moss, B., 2013. Interactions of the Vaccinia Virus A19 Protein. J. Virol. 87, 10710–10720.

Schmelz, M., Sodeik, B., Ericsson, M., Wolffe, E.J., Shida, H., Hiller, G., Griffiths, G., 1994. Assembly of vaccinia virus: the second wrapping cisterna is derived from the trans Golgi network. J. Virol. 68, 130–147.

Schmidt, F.I., Bleck, C.K.E., Helenius, A., Mercer, J., 2011. Vaccinia extracellular virions enter cells by macropinocytosis and acid-activated membrane rupture. EMBO J. 30, 3647–3661.

Schmidt, F.I., Bleck, C.K.E., Reh, L., Novy, K., Wollscheid, B., Helenius, A., Stahlberg, H., Mercer, J., 2013. Vaccinia virus entry is followed by core activation and proteasome-mediated release of the immunomodulatory effector VH1 from lateral bodies. Cell Rep. 4, 464–476.

Schnitzlin, W.M., Ghildyal, N., and Tripathy, D.N., 1988. Genomic and antigenic characterization of avipoxviruses. Virus Res. 10(1), 65-75.

Senkevich, T.G., Koonin, E.V., Moss, B., 2011. Vaccinia virus F16 protein, a predicted catalytically inactive member of the prokaryotic serine recombinase superfamily, is targeted to nucleoli. Virology. 417, 334–342.

Senkevich, T.G., Koonin, E.V., Moss, B., 2009. Predicted poxvirus FEN1-like nuclease required for homologous recombination, double-strand break repair and full-size genome formation. Proc. Natl. Acad. Sci. U. S. A. 106, 17921–17926.

Singh, P., Schnitzlein, W.M., Tripathy, D.N., 2003. Reticuloendotheliosis Virus Sequences within the Genomes of Field Strains of Fowlpox Virus Display Variability. J. Virol. 77, 5855–5862.

Somogyi, P., Frazier, J., Skinner, M.A., 1993. Fowlpox virus host range restriction: gene expression, DNA replication, and morphogenesis in nonpermissive mammalian cells. Virology 197, 439–444.

Stannard, L.M., Marais, D., Dumbell, K.R., Kow, D., 1998. Evidence for incomplete replication of a penguin poxvirus in cells of mammalian origin. J. Gen. Virol. 79, 1637–1646.

Szajner, P., Jaffe, H., Weisberg, A.S., Moss, B., 2004. A complex of seven vaccinia virus proteins conserved in all chordopoxviruses is required for the association of membranes and viroplasm to form immature virions. Virology 330, 447–459.

Tadese, T., Reed, W.M., 2003. Use of Restriction Fragment Length Polymorphism, Immunoblotting, and Polymerase Chain Reaction in the Differentiation of Avian Poxviruses. J. Vet. Diagn. Invest. 15, 141–150.

Terajima, M., Urban, S.L., Leporati, A.M., 2013. The N-terminus of vaccinia virus host range protein C7L is essential for function. Virus Genes 46, 20–27.

Thomas, K., Tompkins, D.M., Sainsbury, A.W., Wood, A.R., Dalziel, R., Nettleton, P.F., McInnes, C.J., 2003. A novel poxvirus lethal to red squirrels (Sciurus vulgaris). J. Gen. Virol. 84, 3337–3341.

Tooze, J., Hollinshead, M., Reis, B., Radsak, K., Kern, H., 1993. Progeny vaccinia and human cytomegalovirus particles utilize early endosomal cisternae for their envelopes. Eur. J. Cell Biol. 60, 163–178.

Toshiaki Terasaki, Mikako Kaneko, Masaji Mase, 2010. Avian Poxvirus Infection in Flamingos (Phoenicopterus roseus) in a Zoo in Japan. Avian Dis. 54, 955–957.

Townsley, A.C., Moss, B., 2007. Two Distinct Low-pH Steps Promote Entry of Vaccinia Virus. J. Virol. 81, 8613–8620.

Townsley, A.C., Weisberg, A.S., Wagenaar, T.R., Moss, B., 2006a. Vaccinia Virus Entry into Cells via a Low-pH-Dependent Endosomal Pathway. J. Virol. 80, 8899–8908.

Townsley, A.C., Weisberg, A.S., Wagenaar, T.R., Moss, B., 2006b. Vaccinia Virus Entry into Cells via a Low-pH-Dependent Endosomal Pathway. J. Virol. 80, 8899–8908.

Tripathy, D.N., Schnitzlein, W.M., Morris, P.J., Janssen, D.L., Zuba, J.K., Massey, G., Atkinson, C.T., 2000. Characterization of poxviruses from forest birds in hawaii. J. Wildl. Dis. 36, 225–230.

Tu, S.-L., Nakazawa, Y., Gao, J., Wilkins, K., Gallardo-Romero, N., Li, Y., Emerson, G.L., Carroll, D.S., Upton, C., 2017. Characterization of Eptesipoxvirus, a novel poxvirus from a microchiropteran bat. Virus Genes 1–12.

Tulman, E.R., Afonso, C.L., Lu, Z., Zsak, L., Kutish, G.F., Rock, D.L., 2004. The Genome of Canarypox Virus. J. Virol. 78, 353–366.

Tulman, E.R., Afonso, C.L., Lu, Z., Zsak, L., Sur, J.-H., Sandybaev, N.T., Kerembekova, U.Z., Zaitsev, V.L., Kutish, G.F., Rock, D.L., 2002. The Genomes of Sheeppox and Goatpox Viruses. J. Virol. 76, 6054–6061.

Tuomi, P.A., Murray, M.J., Garner, M.M., Goertz, C.E.C., Nordhausen, R.W., Burek-Huntington, K.A., Getzy, D.M., Nielsen, O., Archer, L.L., Maness, H.T.D., Wellehan, J.F.X., Waltzek, T.B., 2014. Novel poxvirus infection in northern and southern sea otters (*enhydra lutris kenyoni* and *enhydra lutris neiris)*, alaska and california, usa. J. Wildl. Dis. 50, 607–615.

Ueda, Y., Dumbell, K.R., Tsuruhara, T., Tagaya, I., 1978. Studies on Cotia virus an unclassified poxvirus. J. Gen. Virol. 40, 263–276.

Upton, C., Slack, S., Hunter, A.L., Ehlers, A., Roper, R.L., 2003. Poxvirus Orthologous Clusters: toward Defining the Minimum Essential Poxvirus Genome. J. Virol. 77, 7590–7600.

Vanslyke, J.K., Whitehead, S.S., Wilson, E.M., Hruby, D.E., 1991. The multistep proteolytic maturation pathway utilized by vaccinia virus P4a protein: a degenerate conserved cleavage motif within core proteins. Virology 183, 467–478.

Wei, K., Sun, Z., Zhu, S., Guo, W., Sheng, P., Wang, Z., Zhao, C., Zhao, Q., Zhu, R., 2012. Probable Congenital Transmission of Reticuloendotheliosis Virus Caused by Vaccination with Contaminated Vaccines. PLoS ONE 7, e43422.

Weli, S.C., Nilssen, Ø., Traavik, T., 2005. Avipoxvirus multiplication in a mammalian cell line. Virus Res. 109, 39–49.

Weli, Simon Chioma, Nilssen, Øivind, Traavik, T., 2004. Morphogenesis of fowlpox virus in a baby hamster kidney cell line. Med. Electron Microsc. 37, 225–235.

Weli, S. C., Traavik, T., Tryland, M., Coucheron, D.H., Nilssen, Ø., 2004. Analysis and comparison of the 4b core protein gene of avipoxviruses from wild birds: Evidence for interspecies spatial phylogenetic variation. Arch. Virol.

Wittek, R., Moss, B., 1980. Tandem repeats within the inverted terminal repetition of vaccinia virus DNA. Cell 21, 277–284.

Yang, S., Hruby, D.E., 2007. Vaccinia virus A12L protein and its AG/A proteolysis play an important role in viral morphogenic transition. Virol. J. 4, 73.

Zeigel, R.F., Theilen, G.H., Twiehaus, M.J., 1966. Electron microscopic observations on RE virus (strain T) that induces reticuloendotheliosis in turkeys, chickens, and Japanese quail. J. Natl. Cancer Inst. 37, 709–729.

Zimmermann, D., Anderson, M.D., Lane, E., van Wilpe, E., Carulei, O., Douglass, N., Williamson, A.-L., Kotze, A., 2011. Avian Poxvirus Epizootic in a Breeding Population of Lesser Flamingos (Phoenicopterus minor) at Kamfers Dam, Kimberley, South Africa. J. Wildl. Dis. 47, 989–993.