CHAPTER I

Review of the Literature

1.1 INTRODUCTION

1.1.1 Poxviruses in history. The Poxviridae is a large family of double-stranded DNA viruses that infect vertebrates (the Chordopoxvirinae) and insects (the Entomopoxvirinae). These viruses are characterised by complex capsid architecture, their ability to replicate in the cytoplasm of infected cells, and their efficiency in de-arming or delaying many different host antiviral responses.

Although vaccinia virus is the type member of the family (Fenner, 1979), probably the most famous member is the agent of the historical smallpox scourge of mankind; i.e. variola virus, which is now eradicated. Smallpox first appeared in China and the Far East at least 2000 years ago, and the Pharaoh Ramses V is believed to have died of the smallpox in 1157 B. C. (Fenner *et al.*, 1988; Hopkins, 1983). The disease was established in Europe in 710 A. D., and was introduced to America by Hernando Cortez in 1520, causing the deaths of at least 3.5 million Aztecs in the next two years. Meanwhile, in the great cities of 18th century Europe, smallpox had reached plague proportions, killing hundreds of thousands of people. At least five reigning European monarchs succumbed to the virus during the 18th century (Hopkins, 1983).

The practice of variolation originated around the 10th century- Asia, Africa and China, whereby pus from a smallpox pustule was deliberately inoculated into the skin, in the hope that a mild, protective and non-transmissible disease would result (Baxby, 1991). Properly done, it did produce a protective lifelong immunity against the naturally spreading disease, but had a mortality of about one in 200. In spite of the risks, variolation was successfully introduced into England by Lady Mary Wortley-Montagu and into America Cotton Mather and Zabdiel Boylston in 1721 (Miller, 1957; Winslow, 1974), providing some control over the spread of the disease. The most important advance in the control of smallpox, however, was the introduction of vaccination (*vacca* = cow) by Edward Jenner. On the 14th of May, 1796, Edward Jenner used cowpox inoculate from a milkmaid, Sarah Nemes, to "vaccinate" eight year old James Phipps, who Jenner later challenged with variola virus, and showed that the boy was protected.

Vaccination had the advantage over variolation that it had no mortality and left a trivial lesion at the vaccination site (Jenner, 1798).

For at least 100 years thereafter, the "vaccine" strains were propagated from arm to arm, with vaccination almost universally adopted worldwide around 1800. Although the virus was eradicated from Europe and North America, it remained endemic in South America, Africa and Asia. In 1967 smallpox was still endemic in 33 countries, and it was the Intensified Smallpox Eradication Program initiative of the World Health Organization, which aimed to achieve global eradication of smallpox within ten years, that finally achieved eradication of smallpox's variola virus. The last case of endemic smallpox occurred in Merka, Somalia, in October 1977. In 1980, the WHO declared smallpox eradicated, and in March 1983, the WHO reported that 155 of its 160 member states and associated members had officially discontinued vaccination (Nakano, 1986).

Fenner regularly used the abovementioned eradication of smallpoxvirus as an example of biological control in his Florey lectures of 1979, in the context that a virus can be used to confer protective immunity against a related but pathogenic one. A classic example of biological control however, is provided by another of the poxvirus family members: myxoma virus causes a mild, benign infection in its evolutionary host, the North American brush rabbit (*Sylvilagus californicus*) or the South American tapeti (*Sylvilagus brasiliensis*), but it causes a rapid systemic and lethal infection, known as myxomatosis, in European rabbits (*Oryctolagus cuniculus*). Mortality rates reach up to virtually 100% (Fenner, 1983). Myxomatosis was deliberately introduced into Australia in 1950 and into Europe in 1952 to control *O. cuniculus*. It was at first spectacularly successful in controlling the rabbit pest (Fenner, 1983), but later proved to be ineffective, due to the combination of increased host resistance in the surviving rabbit populations and genetic attenuation of field virus strains (Fenner and Ratcliffe, 1965; Kerr and Best, 1998). Despite its ultimate failure to control the rabbit population, myxomatosis is an extensively-characterised veterinary disease that provides a well-defined *in vivo* model for the study of virus-encoded virulence factors, including those involved in the evasion of the host immune response (Cameron *et al.*, 1999).

1.1.2 The origins, epidemiology, and distributions of capripoxviruses

"Knopvelsiekte", referring to the characteristic lumps in the hide of infected cattle, is an alternative name for lumpy skin disease (LSD) in southern Africa, and was first described in 1931 under the name of "pseudo-urticaria" by MacDonald in Northern Rhodesia. According to him, the disease had been completely unknown to the experienced cattlemen of the area when the disease made its appearance in Lusaka and Mazabuka in 1929 (MacDonald, 1931; Morris, 1931). Von Backstrom (1945) described the same disease as "Ngamiland cattle disease" in 1945, when it broke out in Bechuanaland, and in 1944 it was reported in the Marico district of Northern Transvaal, by Mare (Thomas and Maré, 1945).

In March of 1945, material collected from the Marico outbreak was taken to Onderstepoort Veterinary Institute, where Thomas, Robinson and Alexander used it in transmission experiments. They injected emulsions of skin nodules into a group of cattle housed in the isolation stables. Five animals reacted from the fourth to the thirty-eighth day after injection. By some unknown manner, the infection spread from the isolation stables, to two other stables located some 80 and 200 yards away, respectively. These stables were being maintained as *Anaplasma centrale* reservoirs. In a period of 77 days after the first case was observed at Onderstepoort, 43 of the 83 animals in the *A. centrale* stables manifested clinical symptoms of LSD, and an unknown number developed inapparent signs. Some of these inapparently-infected cattle (which later proved to be infected) were bled to several farmers, and two of the oxen bled on the 16th of March showed clinical symptoms on the very next day, but it was too late to recall the blood, as most of it had already been used. Twelve days later, a number of cattle inoculated with the infected blood developed the clinical symptoms of lumpy skin disease. This was the first positive proof that the disease could be transmitted by blood inoculation (Henning, 1949).

While the epizootic was at its peak at Onderstepoort, lumpy skin disease broke out on three neighbouring farms separated from the institute by the Apies River. All cattle in the area, including those from the Institute's stables, watered at this river, although the 600

cattle kept on Kaalplaas, a region of the institute located about two miles from the infected stables, did not appear to be affected by the disease outbreak. At the end of May 1945, a number of outbreaks also occurred in the nearby then-village of Pretoria North, but it was not clear if these outbreaks were associated with the Onderstepoort epizootic, or if the disease had been contracted from animals that were brought from the Western Transvaal.

For the rest of 1945 and throughout 1946 the disease spread throughout South Africa, reaching the Cape peninsula in June 1946. During this time, it also entered Swaziland, Natal and Mozambique (MacOwan, 1959). Despite its rapid spread, the disease often jumped for several miles, leaving large areas unaffected. Diesel recorded in 1949 that approximately 8 000000 animals had been infected with the disease. By 1956, lumpy skin disease had been reported from the Belgian Congo, Northern and Southern Rhodesia, South West Africa, and Madagascar (I. B. E. D., 1956). Lumpy skin disease was confirmed in Kenya for the first time in December 1957 in the Nakuru farming district (MacOwan, 1959). From then on it was recorded in Sudan in 1971, Chad in 1973 and Niger in 1973, suggesting an initial northward and then westward extension of the disease (MacOwan, 1959; FAO reports, 1973; Ali and Obeid, 1977). In 1974, an epizootic was recorded in Nigeria's Federal Capital Territory (Nawathe *et al.*, 1978). The first reports of LSD outside of Africa were from Kuwait in 1986 to 1988 (Anonymous, 1988), and the disease was subsequently diagnosed in the cattle of the village of Tel El Kabir, Egypt, in October 1988 (House *et al.*, 1990).

In contrast, the other capripoxvirus family members are endemic in sheep and goats in Central and North Africa, the Middle East, Pakistan, India, Nepal and parts of the Chinese Peoples Republic (FAO, 1985), where they are considered to be of major economic importance in sheep and goat husbandry (Kitching, 1983). Sheep- and goatpox has been endemic to the Yemen Arab Republic and the Sultanate of Oman since the early eighties (Kitching *et al.*, 1986). Sheeppox has also been recorded in recent years in Italy, and goatpox has been reported in the western USA and Sweden (Odend'hal, 1983).

1.1.3 The clinical disease. Clinically, lumpy skin disease can be acute, sub-acute or inapparent in the bovine host. The disease is characterised by fever and the sudden appearance of firm circumscribed skin nodules, 1-5 cm in diameter and larger, which affect the full thickness of the skin. The lesions usually undergo necrosis. Similar lesions may be present in the skeletal muscles and the mucosae of the digestive and respiratory tracts. Lameness resulting from inflammation and necrosis of tendons, and from severe oedema of the brisket and legs are also characteristic of the disease (Weiss, 1968). Superficial lymph nodes draining infected areas of the skin may become enlarged to up to ten times their normal size (Weiss, 1968). While morbidity may vary from 5 to 100 %, mortality rarely exceeds 5% (Woods, 1988). Direct losses due to lumpy skin disease are not the main reason that LSD is considered a disease of major economic importance in Africa. Main economic losses result from indirect causes, for example, secondary infection, sterility, stenosis of the trachea, cachesia, decreased milk yield (up to 50% yield loss) and damage to the hide (Henning, 1949). LSDV is classified as an OIE List A disease (OIE Manual, 1996). A disease closely resembling lumpy skin disease clinically is caused by a completely unrelated virus of the herpes virus group. This so-called "Allerton" virus, also known as pseudo-lumpy skin disease, was described by Weiss in 1963. The virus has since been identified as bovine herpes virus II (Nawathe *et al.*, 1978).

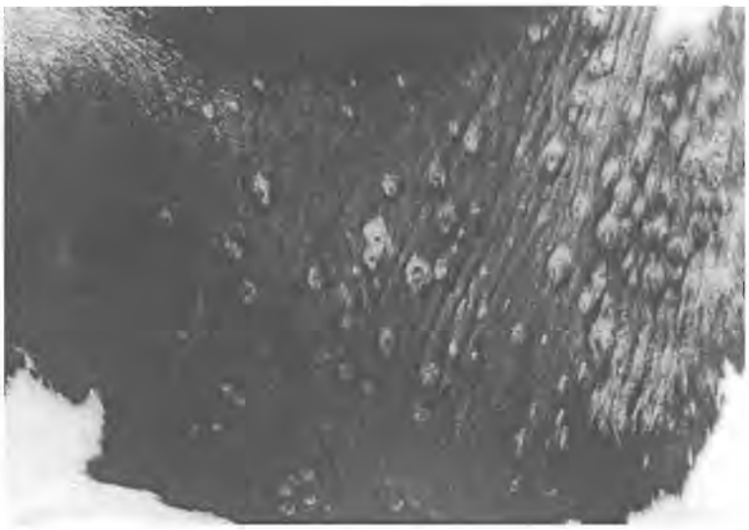

Figure 1: Clinical signs of cow with LSDV (Photograph with courtesy of Dr. P. Hunter, OVI)

The virus causing lumpy skin disease was first isolated in tissue culture by Alexander *et al.* (1957). It was subsequently recovered on numerous occasions from the skin lesions of infected cattle in South Africa as well as Kenya. There is no doubt that this virus, described by Alexander *et al.* (1957) as the Group III virus (prototype Neethling), is the cause of lumpy skin disease (Alexander and Weiss, 1959; Prydie and Coackley, 1959).

1.2 CLASSIFICATION

Smallpox was specifically a human disease, and molluscum contagiosum virus, of the genus molluscipoxvirus, still afflicts only humans. Some poxvirus family members, like the cowpox, monkeypox and vaccinia viruses, have a wide range of potential hosts, while others like variola, camelpox and the lumpy skin disease viruses are restricted to a narrow host range. Capripoxviruses affect only ungulates, usually causing clinical disease in only one species (Weiss, 1968). The factors that determine this host range specificity or lack thereof at the genomic level are of major interest in this study.

Table 1: Classification of the *Poxvirinae*

Subfamily	Genus	Members	Natural Host
Chordopoxvirinae	Avipoxvirus	**Fowlpox virus**	Birds
		Canary pox	
		Junco pox	
		Pigeon pox	
		Quail pox	
		Sparrow pox	
		Starling pox	
		Turkey pox	
	Capripoxvirus	**Sheeppox virus**	Ungulates
		Goatpox virus	Ungulates
		lumpy skin disease virus	**Cattle**
		Cotia virus- unclassified (Ueda *et al.*, 1995)	Unknown
	Leporipoxvirus	**Myxoma virus**	Leporids and Squirrels
		Shope fibroma virus	
		Hare fibroma virus	
		Squirrel fibroma virus	
	Molluscipoxvirus	**Molluscum contagiosum**	**Humans**

	Orthopoxvirus	**Vaccinia virus** Variola virus- variola major - Alastrim variola minor Buffalopox Camelpox Cowpox Monkeypox Rabbitpox Raccoonpox Taterapox volepox ectromelia virus (mousepox)	Very broad
	Parapoxvirus	**Orf virus** Pseudocowpox Chamois contagious ecthyma	Ungulates
	Sutpoxvirus	**Swinepoxvirus**	**Pigs**
	Yatapoxvirus	**Yaba monkey tumor virus** Tanapox	Primates
Entomopoxvirinae	Entomopoxvirus A	***Melolontha melolontha* EPV**	Coleopterans
	Entomopoxvirus B	***Amsacta moorei* EPV**	Lepidopterans
	Entomopoxvirus C	***Chironomus luridus* EPV**	Dipterans
		***Melanoplus sanguinipes* EPV**	Orthopterans

(Cameron *et al.*, 1999; Afonso *et al.*, 2000; Moss, 1990). Protypal members and natural restricted hosts are boldfaced.

1.2.1 Capripoxviruses and their host specificities. Sheep- and goatpox viruses both cause highly contagious diseases, but differ with respect to incubation periods and virulence in their respective hosts. Sheeppox is often fatal in sheep, whereas goatpox is usually less severe. Some authors have distinguished between the host specificities of the Kenyan sheep- and goatpox strains, and those of the Middle East and India (OIE, 1989; Merck Vet. Man., 1986; Odend'hal, 1983; Kitching, 1983). Thus, Kenyan strains have been described which infect both sheep and goats, whereas a virus such as the Sersenk strain of goatpox virus from Iraq is strictly pathogenic for goats (Kitching, 1983). Lumpy skin disease virus is closely related to the sheep- and goatpox viruses (Alexander *et al.*, 1957; Prydie and Coakley, 1959). LSDV isolates from different countries cannot be distinguished serologically from each other, and LSDV cannot be differentiated from the sheep and goat pox viruses by serum neutralization and fluorescent antibody tests (Davies

and Otema, 1981). Despite the close serological relatedness, there is only one type of virus responsible for true lumpy skin disease (Weiss, 1968).

All of the capripoxviruses studied so far have been shown to share a very high (> 95.8%) nucleotide sequence homology (Gershon and Black, 1988). Kitching *et al* (1989) demonstrated that it is possible to study and compare individual virus isolates by the use of simple restriction endonuclease digest patterns. The endonuclease *HindIII* has proved to be the most useful to distinguish between capripoxvirus isolates, as digestion produces approximately 40 fragments with a wide range of sizes (Kitching *et al.*, 1989). This approach was used by Kitching and colleagues (1989) to subdivide the capripoxviruses into four types: Type I isolates were derived from sheep only and type II isolates were derived exclusively from goats. Type III isolates were derived from sheep, goats and cattle but were restricted to Africa, whereas type IV isolates were derived from sheep or goats in the Middle East or Africa (Gershon *et al.*, 1989). It has also been suggested that a type IV isolate may have arisen by genetic recombination between a type II and type III isolate (Gershon *et al.*, 1989), which has implications for the use of recombinant pox virus vaccines, where recombination with a naturally occurring pox virus may occur.

1.2.2 Transmission and spread. Many African and some Middle-Eastern countries support large populations of free-living indigenous wild animals in addition to domestic animals, and their possible role in maintenance and spread of LSD has been investigated: LSD may possibly be transmitted by biting insects (Weiss, 1963; Haig, 1957). In field outbreaks, the resurgence of the disease has been consistently associated with the onset of rains and the emergence of large numbers of the suspected vectors, *Stomoxys calcitrans* and *Culex mirificus* (Woods, 1988). For viral transmission to occur, very close contact between animals may therefore not be necessary. Only cattle have been reported to be affected during outbreaks of LSD, but Davies (1980) reported antibodies in a high proportion in buffalo in East Africa following an epizootic in cattle. Despite this, he failed to detect antibodies in sera from seven other species of wild ruminant examined (Davies, 1982). However, Hedger and Hamblin's (1983) study of free-living wildlife demonstrated antibodies in impala (*Aepyceros melampus*), giraffe (*Giraffa camelopardalis*), kudu (*Tragelaphus strepsiceros*), waterbuck (*Kobus ellipsiprymnus*)

reed buck (*Redunca arundinum*) and springbok (*Antidorcas marsupialis*). In giraffe and reedbuck, virus neutralising titres were of similar levels to those in convalescing cattle, and were assumed to be indicative of past infection. Some ten years prior to the aforementioned surveys, Young *et al.* (1970) investigated the pathogenicity of LSDV for wild species, artificially infected giraffe, impala, African buffalo (*Syncerus caffer*) calves and black wildebeest *(Connochaetes gnou*) with LSDV. Both giraffe and impala died with macroscopic and microscopic lesions characteristic of LSD in cattle, and a significant rise in antibodies was recorded in the impala before death. Neither wildebeest nor buffalo responded clinically or serologically to inoculation of virus.

Since virus isolation from these wild animals has not yet been successful, it has been difficult to place significance on serological titres in random surveys, when only low neutralising antibody titres are recorded. It is known that sera from some animals may contain non-specific inhibitors (Patty, 1970). Due to the close serological relationship between LSD and Kenya sheep- and goatpox viruses, it is possible that some low serum neutralisation titres in wild animals may be due to cross-reactions with other as yet unidentified viruses. Failure of virus isolation and negative or low titres in the above serosurveys would therefore suggest that wildlife might not play a significant part in the maintenance and spread of LSDV (Hedger and Hamblin, 1983).

1.3 THE MOLECULAR BIOLOGY OF POXVIRUSES

1.3.1 Morphology. Munz and Owen (1966) examined negatively stained preparations of purified LSDV under the electron microscope, and observed virus particles that were morphologically similar to the M-forms of VV described by Westwood *et al.* (1964) or the type I VV particles described by Nagington and Horne (1962) and Müller and Peters (1963). Poxvirus virions are oval or brick-shaped bodies about 200 to 400 nm long with axial ratios of 1.2 to 1.7. The external surface contains ridges that may be arranged in parallel rows or, in the case of parapoxviruses, as a single continuous helix. LSDV virions are approximately 295 x 270 nm and are either seen as the the "M" or mulberry form because of a covering of cord-like elements (Kitching and Smale, 1986), or as the "C" or capsule form, also observed by Muntz and Owen (1966), which is surrounded by a membrane. LSDV virus particles measure approximately 3500 Å in length, and 3000 Å in width, with an axis ratio in the region of 1.2 (Munz and Owen, 1966).

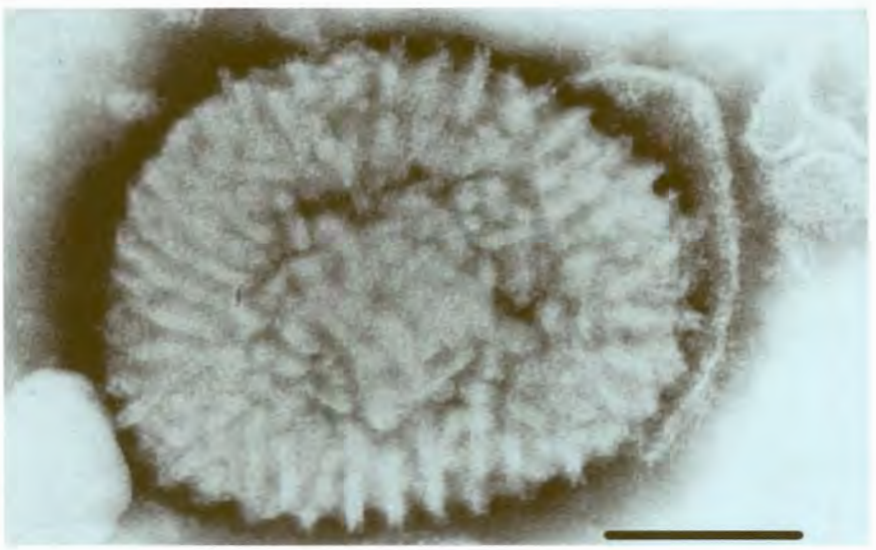

Figure 2: Micrograph of cowpox virus (Dr. Frank Fenner, Australian National University, 1997). Specimen stained with 3% phosphotungstic acid (PTA). The bar represents 100 nm

In cross-section, a pox virion displays a lipoprotein bi-layer, called the outer membrane, surrounding the core. Extracellular forms of poxviruses, released by budding, have been

shown to contain an additional lipoprotein envelope scavenged from the host Golgi apparatus (Dales and Siminovitch, 1961; Medzon and Bauer, 1970; Nagington and Horne, 1962; Peters, 1956). In vertebrate poxviruses, including LSDV, the core appears biconcave with two structures of unknown function, called lateral bodies, nestled in the concavities. In contrast, the core of insect poxviruses may be kidney-shaped with a single lateral body, or brick-shaped, with surrounding material of intermediate density (Westwood *et al.*, 1964; Easterbrook, 1966). The core itself houses a twisted and folded nucleoprotein fiber (Solosky and Holowczak, 1981; Peters, 1956).

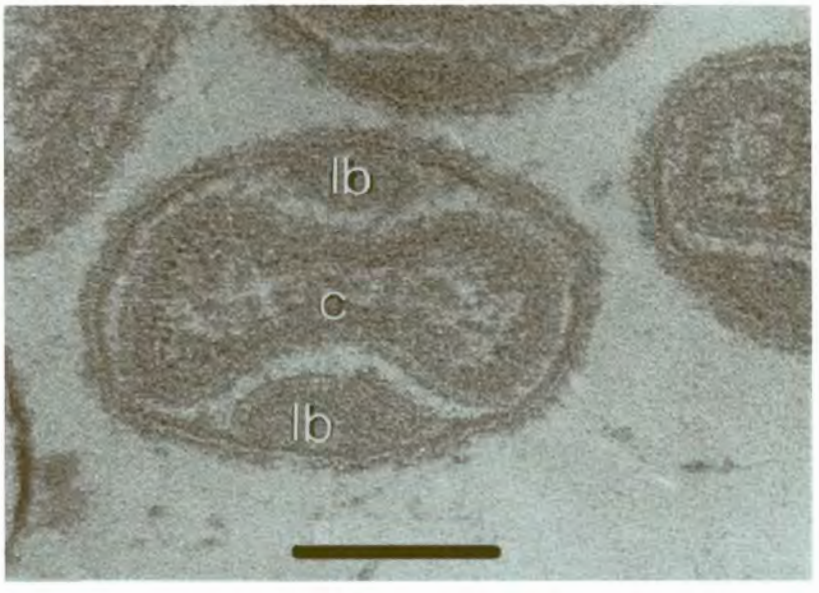

Figure 3: Micrograph of cowpoxvirus- thin section (Dr. Frank Fenner, Australian National University, 1997)
lb- lateral bodies, c-core
the bar represents 100 nm

1.3.2 Genome structure. Poxviruses have the largest known viral genomes (139-288 kbp) each consisting of a linear, double-stranded DNA molecule with covalently linked termini (Baroudy *et al.*, 1982). Cross-linking of poxvirus DNA was first considered because of the rapid renaturation of fowlpox and vaccinia virus DNA, and the inability to separate the two strands (Berns and Silverman, 1970; Jungwirth and Dawid, 1967;

Szybalski *et al.*, 1963). These terminal A+T-rich incompletely base-paired loops of about 100 nucleotides each, were later confirmed by direct sequencing (Baroudy *et al.*, 1982; DeLange *et al.*, 1984). The two ends of the genome inside of the hairpin loops are identical but inverted in sequence (Garon *et al.*, 1978; Wittek *et al.*, 1978). Such inverted terminal repetitions (ITRs) vary in length from 0.7 kbp in the Bangladesh strain of variola, to 12 kbp in the Copenhagen strain of vaccinia (Goebel *et al.*, 1990; Massung *et al.*, 1994). In place of ITRs, variola virus has asymmetrical small arrays in the terminal regions, which distinguishes variola DNA from DNA of many other orthopoxviruses (Fenner *et al.*, 1989; Esposito and Knight, 1985; Antoine *et al.*, 1998; Shchelkunov *et al.*, 1995). Shope fibroma virus terminal inverted repeats seem to be lacking in any repetitive DNA (Willer *et al.*, 1999). In some cases, terminal repeat regions contain expressed genes thought to contribute to phenotypic variability, host-range specificity and pathogenicity (Wittek *et al.*, 1980). The large, conserved central DNA region predominantly contains genes essential for virus structure and replication (Fenner *et al.*, 1989; Buller *et al.*, 1991; Moss *et al.*, 1990; Esposito and Knight, 1985; Esposito, 1991). Gene density for poxviruses is very high, and ORFs are usually separated by only a small number of bases. The fowlpox genome, currently the largest sequenced (288 kbp), contains 260 ORFs (Afonso *et al.*, 2000). The large number of genes is possible due to the apparent absence of introns, the relatively short promoter sequences, and relatively short ORFs. Although both strands of poxvirus DNA are transcribed, overlapping genes are not common. Early and late genes frequently occur in clusters, but can be dispersed across the length of the genome (Moss, 1990).

The genome of capripoxviruses has been partially characterised by hybridisation studies, and shows clear similarities to other poxviruses, notably vaccinia. For example, genome fragments of vaccinia virus and capripox virus cross-hybridise under low-stringency conditions, and a 100-115 kb piece of the genome is co-linear in organisation within the two viruses (Gershon *et al.*, 1989). No cross-hybridisation could be detected between VV DNA and the respective left- and right-hand terminal 8 and 25 kb of capripoxvirus DNA or between capripoxvirus DNA and the respective left- and right-hand terminal 38 and 35 kb of VV. An exception was the hybridisation of the 4.5 kb *HindIII* K-fragment from the

left end of vaccinia to a region near the right end of the capripoxvirus genome. This is a small region that could have translocated between the ends of one or other of the genomes during their divergence (Gershon *et al.*, 1989). From summation of the lengths of restriction endonuclease fragments separated by pulsed field gel electrophoresis, SA-LSDV genome size was determined to be 152 kbp (unpublished data, cited by Fick and Viljoen, 1994).

A high A+T- content of capripoxvirus genomes is indicated by the presence of relatively high frequency of A+T-rich restriction enzyme sites for example *DraI*, and relatively low frequency of G+C-rich sites for example *PstI* and *SalI*. The A+T- content for the capripox thymidine kinase (TK) gene, for example, has been determined as 72.4%, in contrast to 66.4%, 61.6%, and 68.6% for VV, SFV and FPV, respectively (Gershon and Black, 1989). The genomes of the parapoxviruses, which have the same host-range as the capripoxviruses, are of a similar size but have a A+T content of around 37% (Wittek *et al.*, 1979). The high A+T content of capripox DNA may reflect virally controlled aspects of DNA synthesis, such as the characteristics of the capripoxvirus DNA polymerase, or the levels of expression of virally encoded enzymes involved in the biosynthesis of A or T nucleotides, such as TK (Gershon and Black, 1989). In the latter respect, the genomes of poxviruses reported to possess a TK gene are A+T-rich, whereas the presence of a TK gene has not been reported on the G+C-rich genomes of the parapoxviruses (Gershon and Black, 1989). As a result of their widely different G+C contents, the OV and vaccinia (VAC) genomes do not cross-hyridise and homology between these two viruses is evident within the predicted amino acid sequences, but is generally not evident within the DNA sequences (Mercer *et al.*, 1995).

1.3.3 Poxvirus replication. The large dsDNA genomes of poxviruses enable them to have complex virions, and to produce a range of virus-specific enzymes, unlike more limited viruses, which are heavily dependent on cellular enzymes and processes. For example, the enzyme thymidine kinase, which supplies components of eukaryotic DNA synthesis, is only present in actively dividing eukaryotic cells. By producing viral versions of cellular enzymes, the virus ensures that they are available when and where the

virus needs them. The production of infectious virions is dependent on cellular RNA polymerase II (Hruby *et al.*, 1979; Silver *et al.*, 1979; Spencer *et al.*, 1980). Uniquely, for poxviruses and African swine fever virus, their DNA is replicated in the cytoplasm rather than the nucleus, although vaccinia DNA has been demonstrated in the nucleus of infected cells in some systems (La Colla and Weissbach, 1975; Gafford and Randall, 1976, Archard, 1983).

Poxvirus replication occurs in distinct phases. Instead of producing all of the viral proteins simultaneously, poxviruses produce different groups of proteins spread across the genome at particular and appropriate times. The timing is often dependent upon prior viral syntheses, and on specific viral control functions. Different promoters control the transcriptional timing for each group of proteins, which are activated by different but specific stimuli. Most of the current knowledge of poxvirus-gene expression and mRNA synthesis is derived from studies on vaccinia virus (Moss 1990, 1992). Gene expression was initially divided into two phases, namely an early and a late phase, but the discovery of an intermediate class of genes called for a revision of the model for vaccinia virus reproduction, and indicated a more elaborate control of gene expression during the virus infectious cycle (Vos and Stunnenberg, 1988). A cascade model has been proposed, in which early proteins are required for intermediate transcription and proteins encoded by intermediate genes are required for the activation of late transcription (Vos and Stunnenberg, 1988). Furthermore, DNA replication is involved in the regulation of gene expression, a mechanism commonly used among DNA viruses (Tooze, 1981).

a) Entry and uncoating, and transcription of early genes

Both single- and double-membrane forms of poxviruses are infectious. Enveloped virions enter the cell by fusion of the envelope with the cell membrane (Chang and Metz, 1976), although electron micrographs have also demonstrated virus particles fused to membranes within vacuoles caused by surface invagination, which is the entry route for single-membrane forms (Dales and Kajioka, 1964). The epidermal growth factor (EGF) receptor has been implicated as the portal of entry for vaccinia viruses (Eppstein *et al.*, 1965; Marsh and Eppstein, 1987). Soon after penetration, the outer membrane of the

virion is degraded by cellular enzymes, exposing the viral core. After the virus cores have been released into the cytoplasm, a second uncoating event, controlled by the viral genome occurs. A putative 23 kD uncoating protein with trypsin-like activity has been described by Pedley and Cooper (1987), and the uncoating of all cores in the cell is synchronized (Cairns, 1960; Jolik, 1964; Dales, 1965). The nucleoprotein complex passes out through pores in the core walls (Dales, 1965).

Early transcription starts immediately after the virus has entered the cell, and does not require *de novo* synthesis of viral factors or enzymes: all the necessary components, including by the viral polymerase responsible for transcription, are present in the infectious virus particle (Caplan and Holowczak, 1983; Kates and McAuslan, 1967; Munyon *et al.*, 1967). The poxvirus is therefore fully equipped to synthesize fully functional mRNA. Eukaryotic mRNAs typically have a methylated cap structure at their 5' ends. Capping and methylation of vaccinia virus mRNA is accomplished by a 127-kD multifunctional enzyme complex composed of 97-kD and 33-kD subunits (Boone *et al.*, 1977). Polyadenylation of the viral 3'-mRNA tails is catalyzed by the poly(A)polymerase of vaccinia virus, which contains a 55-kD and a 33-kD subunit. Polyadenylation is thought to be important for the stability of mRNA molecules (Moss *et al.*, 1975). Other components of this transcriptional system have been characterised and include DNA topoisomerases, which modify the topological state of DNA via the breakage and rejoining of DNA strands. Poxviruses are the only known eukaryotic viruses that encode their own topoisomerases. The DNA topoisomerase is able to relax both positively and negatively supercoiled DNA in the absence of an energy cofactor, and has the properties of a cellular type I enzyme, except that it is resistant to camptothecin (Bauer *et al.*, 1977; Fogelsong and Bauer, 1984; Shaffer and Traktman, 1987; Shuman *et al.*, 1988). The failure of attempts to construct a mutant vaccinia virus with deletions within the topoisomerase gene suggest that it plays an essential role in the replication cycle of poxviruses (Shuman *et al.*, 1989).

Other components of the proposed transcriptional system include a deoxyribonuclease enzyme with ligase activity, which can catalyze nicking-joining reactions (Lakritz *et al.*,

1985; Pogo and O'Shea, 1977; Merchlinsky *et al.*, 1988; Reddy and Bauer, 1989), a DNA-dependent ATPase thought to play a role in gene expression (De Lange, 1989), and a serine/threonine protein kinase (Kleiman and Moss, 1975; Kleiman and Moss, 1975). Factors required for initiation (Broyles *et al.*, 1988; Broyles and Moss, 1988) and termination (Shuman *et al.*, 1987; Shuman and Moss, 1988) have also been described.

The ability of extracts of vaccinia virus cores to transcribe early genes *in vitro* (Golini and Kates, 1985; Rohrman and Moss, 1985) suggested the presence of a specific transcription factor, which exists in a labile complex with RNA polymerase and the mRNA-capping enzyme (Broyles and Moss, 1987) and which are (late proteins) packaged into newly assembled virus particles (Vos and Stunnenberg, 1988). Highly purified RNA polymerase preparations were unable to initiate RNA synthesis, which prompted a complementation assay for transcription factors. A specific vaccinia virus early transcription factor (VETF) was detected and isolated from core extracts (Broyles *et al.*, 1988). Purified VETF is a heterodimer of two polypeptides of 82 kD and 77 kD, and displays DNA-dependent ATPase activity. The VETF binds to early promoters, the transcriptional start site for early genes, in a manner comparable to the TATA-box factor of eukaryotic RNA polymerase II promoters (Breathnach and Chambon, 1981; Broyles *et al.*, 1988).

Early transcripts, which are less abundant than late transcripts, are of a defined length (Venkatesan *et al.*, 1981). This is due to the fact that early mRNA transcription is terminated by a specific early transcriptional stop signal, TTTTTNT (Rohrman *et al.*, 1986; Yuen and Moss, 1987). Transcription stops 20 to 50 bp downstream of the TTTTTNT termination sequence of many VV early genes (Yuen and Moss, 1987). Sometimes, the terminator is far downstream of the end of the coding region, at other times, inside a late gene; furthermore, it is not uncommon for a single terminator to occur following two or more early genes, suggesting the formation of a 3'-co-terminal nested set of mRNAs. Examples of internal terminator sequences, thought to be non-functional because of secondary RNA structure, have been noted in a few VV genes (Lee-Chen *et al.*, 1988; Luo and Shuman, 1991). Several approaches have been taken to define the

promoter regions of early genes. The most useful have involved the fusion of putative promoter sequence containing the RNA start site to a reporter gene- most commonly encoding easily-assayable enzymes such as chloramphenicol acetyltransferase, β-galactosidease, or herpes simplex virus thymidine kinase. Following this, the chimeric DNA is allowed to recombine into the vaccinia genome. The ability of the recombinant virus to express the reporter gene provides an initial assessment of the promoter function, which could be confirmed by *in vivo* and *in vitro* transcriptional analysis. In this manner, it was shown that the early promoters for the thymidine kinase (TK) and 7.5 kD genes of vaccinia virus extend about 30 bp upstream of the RNA start sites (Cochran *et al.*, 1985; Coupar *et al.*, 1987; Mars and Beaud, 1987; Vasef *et al.*, 1985; Weir and Moss, 1987). In order to pinpoint specific elements essential for promoter function, Davison and Moss (1989) used oligonucleotide synthesis technology to generate a large number of precisely defined mutations in early promoters, and assayed the effect of each mutation on transcription. From their results, they were able to define the early promoter, which may be conceived in terms of three elements: a critical region separated from a seven bp initiation region by an eleven bp spacer region. So far, there is no indication of the involvement of sequences further upstream or downstream in the regulation of early mRNA synthesis. The critical region is defined as that region upstream from the initiation site within which single substitution mutations may have a drastic effect on transcription; it comprises 16bp in the VV 7.5-kD promoter. Like the TATA box of eukaryotic RNA polymerase II promoters, the critical region specifies the start of transcription initiation. Importantly, single nucleotide substitutions in the critical region of the promoter that decrease transcription, also seem to decrease DNA binding by VETF (Yuen *et al.*, 1987).

Early promoter critical region: (Davison and Moss, 1989)

```
CONSENSUS:          AAAAAATGAAAAAAAA
                                  T

STRONG PROMOTER:    AAAAATTGAAAAACTA
                                  T
```

The consensus was derived from alignments of characterised vaccinia virus early promoter sequences (Davison and Moss, 1989)

In contrast with the critical region, which tends to be A-rich, the spacer region in natural promoters tends to be T-rich (Davison and Moss, 1989). The sequence of the spacer region has only a limited effect on promoter strength, since no mutation in this region was found to prevent transcription (Davison and Moss, 1989). Davison and Moss (1989) also proposed a simple model for transcriptional initiation, in which the critical region binds a transcription factor, thus allowing entry of the RNA polymerase complex. The latter might associate with the factor before or after the factor has bound to DNA. Initiation then takes place at a certain distance from the binding site, dictated by the stereochemistry of the complex formed between transcription factor, RNA polymerase and template DNA. The region in which initiation can occur efficiently has a size of 7 bp *in vivo*, but can be considerably larger *in vitro*. The reasons for this difference are unknown, but may depend on the higher order of molecular structure present in the virion than is present in the *in vitro* system. In the study of Davison and Moss (1989), with only one exception, initiation was confined to purines. In fact, initiation did not occur in the initiation region if purines were absent, even if the critical region was present. The sequence at the initiation site has an effect on promoter strength owing to several features: the presence of A and G residues, their locations with respect to the critical region, and their context. Most natural promoters have more than one purine in the initiation region, and the rules governing the relative use of each initiation site appear complex. Beyond the level of promoter sequence, the possibility remains open that the level of expression of a particular early gene may be influenced by its location with respect to adjacent genes (Davison and Moss, 1989). One LSDV early promoter PA8R, from the *BamHI* P-fragment of the Neethling strain has been characterised. PA8R closely resembles critical regions of known poxvirus early promoters (Fick and Viljoen, 1999).

b) DNA replication

Vaccinia DNA replication starts about two hours after infection, quickly reaches a maximum and ceases a few hours before infectious progeny virions are produced. Viral DNA is nicked immediately after uncoating (Pogo, 1977), possibly by the site-specific nuclease described by Lakritz *et al.* (1985), or even the DNA topoisomerase (Sharma *et al.*, 1994). The vaccinia virus DNA polymerase is expressed early in infection, and has a

molecular weight of about 110 kD and an associated 3'-exonuclease activity (Chalberg and Englund, 1979). DNA replication begins near the ends of the molecule, and involves a strand displacement mechanism, with the formation of small DNA fragments covalently linked to RNA primers (Esteban *et al.*, 1977). Only one strand is replicated from either end (Pogo *et al.*, 1981). There is evidence for concatameric forms of replicated DNA (Archard, 1979; Moyer and Graves, 1981; Holowczak, 1983; Baroudy *et al.*, 1983). Cross-linking of the progeny molecules occurs late in the infectious cycle (Esteban and Holowwczak, 1977; Pogo and O'Shea, 1978).

A close examination of the terminal loops at the ends of the VV genome led to the suggestion of an alternative self-priming method of DNA replication, involving the formation of nicks which generate 3' ends (Baroudy *et al.*, 1982; Moyer and Graves, 1981). These 3' ends then allow priming for DNA replication, which is similar to the model proposed for the replication of single-stranded parvovirus DNA (Tattersall and Ward, 1976) and eukaryotic telomeres (Bateman, 1975). This model can account for the regeneration of incompletely base-paired loops and flip-flop sequence rearrangements (Moss, 1990).

Variola and vaccinia viruses encode at least two other enzymes involved in nucleotide metabolism, which enable the high levels of DNA synthesis: thymidine kinase (TK) and ribonucleotide reductase. The vaccinia virus TK was the first poxvirus gene to be mapped by marker rescue (Weir *et al.*, 1982) and expressed by cell-free translation of mRNA (Hruby and Ball, 1982; Weir *et al.*, 1982). Although the TK gene is not required for virus growth in tissue culture cells, it presumably provides a selective advantage in non-replicating cells that have low levels of thymidylic acid, for in animals, TK- mutants are significantly attenuated (Buller *et al.*, 1985). The "Kenya sheep and goat pox" vaccine strain genome fragment containing the TK gene was sequenced by Gershon and Black in 1989. Another virus-encoded enzyme involved in poxvirus replication, ribonucleotide reductase, is induced soon after vaccinia virus infection. This enzyme converts ribonucleotides to DNA precursors, (Slabaugh *et al.*, 1984).

An unusual feature of the shope fibroma virus (SFV) genome, is the partial absence of most of the genes involved in nucleotide biosynthesis (Willer *et al.*, 1999). Variola and vaccinia viruses encode several such genes, as mentioned above, while molluscum contagiosum encode none at all. SFV thymidylate and guanylate kinases are absent, while thymidine kinase is still encoded (Willer *et al.*, 1999). SFV also encodes only the smaller R2 subunit of the ribonucleotide reductase, presumably rendering it dependent upon host cell functions for the synthesis of DNA biosynthetic precursors. A similar situation is observed in human cytomegalovirus except that this virus encodes only the larger R1 subunit (Bankier *et al.*, 1991). Why SFV should retain only one half of an $\alpha 2\beta 2$ protein is unclear. It has been observed that heterodimers composed of a vaccinia R1 subunit and mouse R2 subunit are catalytically active (Hendricks and Mathews, 1998), and it has been suggested that SFV has evolved to exploit the opposite arrangement consisting of the SFV R2 and host R1 subunits (Willer *et al.*, 1999).

c) Transcription of intermediate genes

Intermediate genes are transcribed immediately after DNA replication (which is essential for the de-repression of intermediate genes), and requires positive *trans*-acting transcription factors, which are present in an active state in the infected cells prior to the onset of DNA synthesis. It was suggested by Vos and Stunnenberg (1988) that these transcription factors are encoded by early viral genes. Alternatively, host proteins or viral structural proteins might be involved in the activation of intermediate transcription, as is the case for the structural protein Vmw65 in HSV transcription (O'Hare and Goding, 1988; Preston *et al.*, 1988).

An intermediate promoter for vaccinia virus was characterised by mutational analysis (Hirschmann *et al.*, 1990), and shown to consist of an upstream element from –20 to –9, and a downstream element from +1 to +9, relative to the $TA_{(3\text{-}6)}T$ transcriptional initiation site. These elements are sufficient to direct intermediate transcription both *in vivo* and *in vitro*. The size and position of the upstream element relative to the transcriptional site is comparable to the upstream regions of similar elements in late promoters of VV (Davison and Moss, 1989). The upstream element appears to be essential for intermediate promoter

activity, and the distance between this AAANAA element and the RNA start site is not conserved (Hirschmann *et al.*, 1990), but is generally between 12 and 15 nucleotides long. The mutation of a C residue to an A at position –19 in the vaccinia I3 promoter converts the intermediate promoter into a late promoter, which suggests that the C residue at position –19 in the natural promoter might be detrimental to late promoter activity (Hirschmann *et al.*, 1990). DNA replication is not sufficient for the activation of late transcription either, and it appears that trans-acting transcription factors, encoded by intermediate genes are essential (Wright and Moss, 1987; Vos and Stunnenberg, 1988).

d) Viral transcription of late genes

The switch from early to late transcription in lumpy skin disease virus occurs at approximately nine hours post infection (Fick and Viljoen, 1994). Late messenger sequences are only transcribed after DNA synthesis commences (Oda and Joklik, 1967; Kaverin *et al.*, 1975; Boone and Moss, 1978). The encoded RNA sequence starts within three consecutive A-residues present in a conserved TAAAT element in the non-coding DNA strand. (Rosel *et al.*, 1986; Berthelot *et al.*, 1986; Weir and Moss, 1987; Lee-Chen and Niles, 1988). This poly(A) head is probably generated by slippage (stuttering) of the polymerase complex during transcription initiation (DeMagistris and Stunnenberg, 1988; Schwer and Stunnenberg, 1988). On average, late mRNAs are much longer than those coding for the early proteins, often many-fold longer than the length required to encode their translation product (Cooper *et al.*, 1981; Mahr and Roberts, 1984). This may relate to unusual processing events (Bertholet *et al.*, 1987). Four studies have been done on separate late promoters, in order to precisely define the region necessary for late gene regulation. Each investigation involved linking a late promoter to the coding region for chloramphenicol acetyltransferase, mutating upstream sequences, and assaying promoter activity in a transient expression system, or in recombinant viruses (Cochran *et al.* 1985; Bertholet *et al.* 1986; Weir and Moss,1987). An extensive study of the sequence elements involved in late gene regulation *in vivo,* was later done by detailed mutagenic analysis (Davison and Moss 1989). From the results of these investigations, late promoters were defined by the transcription initiation site TAAAT sequence. This motif often forms part of the initiating methionine codon, by being followed by a G. In strong late promoters,

the TAAAT is preceded by a short spacer of 4-10 nucleotides, and then a T-rich tract of 5-15 nucleotides. The data of Davison and Moss (1989) further indicate that the least detrimental mutations of the TAAAT sequence are CAAAT, TAAAG and TAAAA. A natural example of a functional promoter with the latter of these mutations is demonstrated in the vaccinia A2L gene. Promoters that deviate from the consensus (of Davison and Moss) have been described, and may represent other regulatory classes (Lee-Chen and Niles, 1988; Vos and Stunnenberg, 1988).

Figure 4: Late promoter consensus sequence, derived from the alignment of characterized vaccinia virus late promoter sequences (Davison and Moss, 1989).

A late promoter, PA7L from the *BamHI*-P fragment of the Neethling strain has been characterized. It contains the conserved TAAAT motif, although this is separated from the ATG codon (at position +8) of the ORF P1L, to form a TAAATAAAATG motif (Fick and Viljoen, 1999).

e) Viral assembly and release

The initial stages of virion formation occur in circumscribed, granular electron-dense areas of the cytoplasm apart from cell membranes. In the subsequent stages of development, the immature viral envelopes appear spherical, with a dense nucleoprotein mass embedded in a granular matrix. Electron microscopy studies suggest that the nucleoprotein enters the immature envelopes just before they are completely sealed (Morgan, 1976). Mature virus particles move out of the assembly areas, and are transported to the cell periphery where they may become enveloped by additional membranes (Morgan, 1976). High-voltage electron microscopy of whole infected cells revealed a membrane-like sheath surrounding mature virions, which align themselves with cytoplasmic fibres (Stokes, 1976). Immunofluorescence studies indicated that the

fibres are actin-containing microfilaments of the cell cytoskeleton (Hiller *et al.*, 1981). The double-layered membrane surrounding the mature virus particles is derived from the Golgi apparatus (Hiller and Weber, 1985). The enveloped viral particles fuse with the plasma membrane, resulting in the externalization of the virion, still enclosed by one layer of Golgi membrane (Payne and Kristensson, 1979).

1.3.4 Virus-Host interactions

1.3.4.1 Poxviruses and the evasion of the immune response. A virus able to evade the host immune response, has the advantage to propagate itself over one which is rapidly neutralised and destroyed in the eukaryotic arsenal of antiviral strategies. An important feature of the early inflammatory response to an initial virus challenge, is the influx and activation of leukocytes. Leukocytes help to initiate the earliest phases of antiviral immune activation (Nathanson and McFadden, 1997; Zinkernagel, 1996). Neutrophils, monocytes/macrophages, and Natural Killer (NK) cells in particular (all leukocytes) participate in the first wave of cellular infiltration, but require directional signals in order to migrate to the injured tissues bearing infected cells (Ben-Baruch *et al.*, 1995; Springer, 1994). Poxviruses have been known to actively interfere with the host's immune response by mimicking various host proteins involved in cell signaling and immune regulation (Smith, 1993). This includes interfering with complement function, interference with MHC-I presentation, interference with cytokine production/function, interference with interferon-induced protein kinase, interference with apoptosis and interference with inflammation. (Harper, 1998).

a) Virus-encoded Interferon gamma and receptor homologs

Interferon-γ (IFN-γ) is a cytokine (of which chemokines make up a class), which functions within the immune system as a potent anti-viral and immunoregulatory stimulant (Epstein, 1984; Vilcek *et al.*, 1985; Farrar and Schreiber, 1993). The importance of IFN-γ in response to poxvirus infection has been demonstrated many times. In a murine experimental model, infection with vaccinia virus, which is usually effectively cleared by the immune system, becomes lethal when monoclonal antibodies to IFN-γ are added (Ruby and Ramshaw, 1991). In immuno-deficient mice, where vaccinia

virus infection is lethal, clearance is achieved when a recombinant virus expressing IFN-γ is injected (Kohonen-Corish *et al.*, 1990; Ramshaw *et al.*, 1992). Interferon-γ has also been proven to be critical in the recovery of mice from ectromelia virus infection (mousepox) (Karupiah *et al.*, 1993). In order to infect successfully, poxviruses have evolved several strategies to counteract the pleiotropic effects of IFN-γ. For example, the RNA-dependent protein kinase pathway is responsible for inducing the anti-viral effects of interferons in the host. The E3L gene of vaccinia virus encodes an inhibitor of the double stranded RNA-dependent protein kinase (Chang *et al.*, 1992), while vaccina protein K3L blocks the same pathway at the stage of phosphorylation of the alpha subunit of eukaryotic initiation factor II (Davies *et al.*, 1992). The aforementioned are intracellular processes, but poxviruses are also capable of inhibiting IFN-γ extracellularly, before the cellular receptors are engaged (Mossman *et al.*, 1995). Cells infected by myxoma virus were shown to secrete a 37-kDa virus-encoded protein, which has homology to the ligand-binding domain of the known mammalian interferon-γ receptor, and effectively functions as a soluble interferon-γ receptor homolog (Upton *et al.*, 1992). Putative interferon-γ binding proteins have since been identified in the genomes of shope fibroma virus, vaccinia virus, variola virus, ectromelia virus and swinepoxvirus (Upton *et al.*, 1992; Goebel *et al.*, 1990; Shchelkunov *et al.*, 1993; Massung *et al.*, 1993, 1994; Mossman *et al.*, 1995).

b) Virus-encoded chemokines and chemokine receptors homologs

Chemokines are a family of structurally related cytokines, which selectively induce chemotaxis and activation of phagocytic cells and lymphocytes. They are also able to rapidly trigger integrin-mediated leukocyte mediation (Roitt, 1997). The production and secretion of chemokines can be induced at inflammatory sites, for example, virus-infected cells or tissues. A chemokine receptor homolog produced by the virus on the surface and/or on the inside of the infected cells could therefore result in the inhibition of the release of chemokines and, thus, inhibition of the chemokine-induced inflammatory response against the virus infection. Graham *et al.* (1997) reported a novel class of secreted poxvirus proteins, collectively termed T1/35kDa. These proteins interact with a broad spectrum of chemokines *in vitro*, and retard the extent of leukocyte influx into

virus infected lesions *in vivo*. The mechanisms of these interactions remain to be demonstrated. Based on their primary structure, particularly the conservation of a four-cysteine motif, the chemokine receptor family is divided into two classes: C-C and C-X-C, according to whether the first two cysteinyl residues in the motif are adjacent or separated by an intervening residue (Schall, T. J. 1991). Chemokines signal through a family of serpentine receptors, which span the cell surface seven times, are coupled to G proteins, and are generally class restricted, binding either to CC or CXC chemokines. It has been shown though, that multiple chemokines from within a family can bind and signal through common receptors (Kelvin *et al.*, 1993; Murphy, 1996). Based on amino acid identity, Cao *et al* (1995) demonstrated that the Q2/3L ORF of a capripoxvirus is closely related to the C-C or C-X-C chemokine receptor family. Further structural and topological analysis of the deduced amino acid sequence of ORF Q3/2L revealed that it has key structural characteristics of the G-protein-coupled receptor superfamily; e.g., seven hydrophobic regions (predicted to span the cell membrane) and cysteinyl residues in the first and second extracellular loops, that are implicated in forming a disulphide bond. The sequence identity is highest in the transmembrane-spanning domains. The possibility of there being homologs of this Q2/3L ORF in other capripoxviruses represented by Iraq goat pox, Nigeria sheep pox, and lumpy skin disease, was evaluated by Southern blot analysis of *HindIII* restriction endonuclease digests. ORF Q2/3L homologs were detected only in the equivalent fragments in all of the capripoxviruses examined, except for Iraq goatpox DNA. The equivalent locus in Iraq goatpox DNA is present in a larger fragment (Gershon *et al.*, 1988). Only a single copy of the Q2/3L ORF, located in the left terminal region, is present in the capripoxvirus genome. In contrast, swinepoxvirus contains two almost identical ORFs (C3L and K2R) located within the ITR at each terminus (Massung *et al*, 1993). The equivalent gene was not present in vaccinia virus (Johnson *et al*, 1993).

c) Tumor Necrosis Factor (TNF) inhibitors

Tumor necrosis factors (TNF) -α and -β are two related cytokines named for their cytotoxic effects on certain tumor cells, but also have immunoregulatory functions. Macrophages that have engulfed an impostor are activated by agents such as IFN-γ and

TNF-β, amongst others. Consequently, the macrophages become microbicidal. IFN-γ sets the cell up for destruction by inducing the formation of TNF receptors (Roitt, 1997). A poxviral TNF was identified by Upton *et al.* (1991). Secreted poxviral TNFs would presumably bind to receptors on the infected host cell, preventing the signalling for cell destruction.

d) Viral inhibition of the complement cascade

The complement system is a multicomponent triggered enzyme cascade, which is used to attract phagocytic cells to invading microbes and engulf them (Roitt, 1997). Some twenty proteins are involved in the complex. Along with blood clotting, fibrinolysis and kinin-formation, the complement proteins produce a rapid, highly amplified response to a trigger stimulus. This occurs by a cascade phenomenon, where the product of one reaction is the enzymic catalyst of the next (Roitt, 1997). Briefly, the most abundant component of the complement system, C3, is split by a convertase enzyme formed from its own cleavage product C3b and factor B. The C3b product associates with the microbial surface, which stabilises it against breakdown caused by factors H and I. As it aggregates, C3b becomes covalently linked to the microorganism. The next component, C5, is activated, yielding a small peptide viz. C5a. The residual C5b binds to the surface and assembles the terminal components C6-9 into a membrane attack complex, which is freely permeable to solutes, and can lead to osmotic lysis of the pathogen. C5a is also a potent chemotactic agent for lymphocytes and greatly increases capillary permeability. C3a and C5a act on mast cells, causing the release of further mediators such as histamine, leukotriene B_4 and tumor necrosis factor, with effects on capillary permeability and adhesiveness, and neutrophil chemotaxis; they also activate neutrophils (the major circulating phagocytic polymorphonuclear granulocyte). Following the activation of complement with the ensuing attraction and stimulation of neutrophils, the activated phagocytes bind to the C3b-coated microbes by their surface C3b receptors and ingest them (Reid, 1995; Law and Reid, 1988). The 35-kD major secreted protein of vaccinia virus was recently identified and its sequence revealed that it closely resembled a family of eukaryotic proteins that have 60 amino acid repeats (Kotwal and Moss, 1988). Computer searches indicated that it most closely resembled a complement control protein

that binds to complement C4b. *In vitro* data indicate that the purified protein does inhibit the classic complement pathway (Moss, 1990). Two genes identified in the rabbitpox virus genome have homology with the β chain of complement component C4a, and their roles in viral virulence have been demonstrated (Bloom *et al.*, 1991).

e) Interleukins

Interleukin-I (IL-I) is also a chemokine, and reacts on endothelial cells, fibroblasts and epithelial cells to secrete (another) chemokine, MCP-I, a potent monocyte chemotactic protein that attracts mononuclear phagocytes. It has been reported that a cowpox gene, crmA, encodes an inhibitor of the interleukin-1β converting enzyme, which can inhibit these inflammatory processes (Ray *et al*, 1992). Interleukin I-β receptors have also been identified in other poxviruses (Alcami *et al*, 1992; Spriggs *et al*, 1992).

All of the above-mentioned poxvirus encoded proteins seem to lack the membrane anchor and cytoplasmic domain found in their cellular counterparts. Furthermore, they are secreted from the infected cells, and are incapable of cellular signal transduction. They appear to modify pathogenesis by binding and neutralising the respective cytokine before it can react with its cellular receptor (Moss, 1990).

f) Serpin-like proteins

Serpin homologs have been identified in numerous poxviruses including vaccinia, ectromelia, cowpox, swinepox and myxoma viruses (Massung at al., 1993). In cowpoxvirus, the serpin homolog inhibits the cellular IL-1β converting enzyme (Ray *et al.*, 1992), and interferes with the lipoxygenase pathway of arachidonic acid metabolism (Palumbo *et al.*, 1993). Presumably, the immune response is influenced by the alteration of the synthesis of various arachidonic acid metabolites, including leukotrienes such as the chemoattractant LTB_4, and prostaglandins (Palumbo *et al.*, 1993).

1.3.4.2 Poxvirus genes contributing to host-range specificity and pathogenicity.

Many different types of poxvirus-encoded proteins and their functions have been presented so far. In addition to the genes commonly shared by the many species of

poxvirus, including those involved in DNA replication, nucleotide metabolism, and immuno-regulatory proteins, individual poxviruses also encode proteins that don't seem to have homologs in the other poxviruses. These proteins have been suggested by various authors to be factors which determine pathogenesis and the host range specificity, allowing the respective virus to survive and replicate in a particular animal host.

It has previously been suggested that the essential core of vertebrate poxviruses be defined by the region from vaccinia ORF F9L at the left to near ORF A29L at the right. This represents 102 kbp. The same conclusion for the position at the left end of a postulated central core was reached in an analysis of swinepox virus (Massung *et al.*, 1993). In addition, deletion mutants of vaccinia virus have revealed that all of the open reading frames at the left end of the genome and up to and including ORF F4L are not required for growth in cell culture (Perkus *et al.*, 1991). The deletion mutants also showed that at the right end, ORFs B13R to B29R are nonessential, and a number of open reading frames between ORFs A42R and A56R have been shown to be nonessential for growth in cell culture. These ORFs may therefore play crucial and specific roles during natural infection, to ensure virus survival in the host (Buller *et al.*, 1991; Smith, 1993; Johnson *et al.*, 1993).

The host-interactive proteins include homologs of epidermal growth factor, probably for increasing virus-susceptible cell numbers around infectious centers; a zinc-finger protein, and protein kinases, which have the potential to influence viral gene expression and macromolecular synthesis. Also included are proteins (for example, haemagglutinin protein and serine protease inhibitor analogues) that potentially influence viral spread by altering infected cell membranes, and cellular erythrocyte ankyrin-related proteins that may specify tissue tropism.

a) Ankyrin-like proteins

In 1987, Breedon and Nasmyth reported a ~33 residue repeating motif in the sequence of two yeast cell-cycle regulators, Swi6p and Ccd10p, and in the Notch and LIN-12 developmental regulators from *Drosophila melanogaster* and *Caenorhabditis elegans*

(Breedon and Nasmyth, 1987). Subsequently, the discovery of 24 copies of this sequence in the cytoskeletal protein ankyrin (Lux *et al.*, 1990) led to the naming of this motif as the ankyrin (ANK) repeat. ANK-repeat proteins carry out a wide variety of biological activities and have been detected in organisms ranging from viruses to humans. The motif has now been recognized in >400 proteins (Bork, 1993), including cyclin-dependent kinase (CDK) inhibitors, transcriptional regulators, cytoskeletal organizers, developmental regulators and toxins (Michaely and Bennet, 1992; Axton *et al.*, 1994). These molecules are present in the nucleus, cytoplasm and the extracellular environment. The number of repeats within any one protein is highly variable: the *Drosophila* plutonium protein contains only two repeats (Axton *et al.*, 1994), but 20 or more are not uncommon. Some ANK-repeat proteins consist solely of ANK repeats; others are multi-domain molecules in which ANK repeats are combined with other unrelated structural modules.

Shchelkunov and colleagues' analysis of VAR and VV proteins (1993) demonstrated that both viruses code for a large set of ankyrin-like proteins, and two of these, the M1L and K1L genes of VV were previously found to determine the virus host cell range in cell cultures (Goebel *et al.*, 1990). VV does not propagate on the CHO cell line, although this line is permissive for another orthopoxvirus, CPV. It was shown (Spehner *et al.*, 1988) that the CPV gene coding for the 77 kDa protein, in which three ANK-repeats were identified, is responsible for this property of the virus and, when integrated into the VAC genome, allows the latter virus to propagate in CHO cells. Drillien *et al.* (1978) were furthermore able to demonstrate that, for VV, the infection of non-permissive CHO cells resulted in a rapid inhibition of the protein synthesis both of the host cell and the virus. The mutant of VV with a deleted K1L gene is also incapable of multiplication in cultures of human MRC-5, KB or Hep-2 cells. This process is accompanied by a rapid inhibition of protein synthesis both of the virus and the infected cell (Drillien *et al.*, 1981). The individual repeats or their groups in ankyrin-like proteins are supposed to form binding sites for various integral membrane proteins of the cell cytoskeleton. The first three-dimensional structure of an ANK-repeat molecule, that of 53BP2 bound to the p53 cell-cycle tumor suppressor (Gorina and Pavletich, 1996), was determined almost ten years

after the discovery of the motif. This and subsequent structures have shown that the ANK-repeat motif can be defined as a β-hairpin-helix-loop-helix ($\beta_2\alpha_2$) structure. Clearly, specific types of residues are prevalent at any one position within the consensus. However, ANK repeats can sustain a considerable variety of amino acid substitutions (Michaely and Bennett, 1992; Bork, 1993). In many ANK-repeat molecules, sequence insertions occur most often within that region that connects helix 2 to the first β-strand of the adjacent repeat. The differences in amino acid sequences of each repeat could account for their different specificities (Lux *et al.*, 1990).

The ANK-repeat consensus: (Sedgwick and Smerdon, 1999)

```
1         10                  20                  30
D__G_TPLHA___G___VV_LLL__GADVNA_
```

Binding to target molecules commonly involves contacts with the tips of the β-hairpins and the surface of the helical bundle facing into the ankyrin groove. In general, these residues are not conserved in the ANK consensus. As such, they are not structurally constrained and are ideally located to perform binding roles. The molecular stacking of the repeat motifs allows incorporation of additional non-conserved elements that might contribute both to the structural stability of the domain, and to the binding to the target molecule (Sedgwick and Smerdon, 1999). In VAR, ORF G3R is the most typical ankyrin-like protein among those detected in poxviruses. It contains five ankyrin repeats. (Shchelkunov *et al.*, 1993). A unique characteristic of CPV is that not only does it have the widest host range among the orthopoxviruses, but also the widest set of genes coding for the proteins with multiple ankyrin repeats (Shchelkunov *et al.*, 1997).

b) Kelch-like proteins

The kelch family of proteins is defined by a 50 amino acid repeat that has been shown to interact with actin (Xue and Cooley, 1993). This motif was first identified as a tandemly-arranged repeating sequence in the C-terminal domains of a number of poxvirus ORFS (Goebel *et al.*, 1990; Koonin *et al.*, 1992). However, it was a prior description of the

same motif in the kelch protein of *Drosophila*, as a distinct structural motif, that led to the term "kelch repeat" (Xue and Cooley, 1993). Proteins containing kelch repeats have been identified in a diverse set of organisms, including the A55R and C2L ORFs of vaccinia virus (Goebel *et al.*, 1990; Koonin *et al.*, 1992), the predicted *Ipp* protein and Ectoderm-neural cortex of mice (Chang-Yeh *et al.*, 1993; Hernandez *et al.*, 1997), spe-26 of *Caenorhabditis elegans* (Varkey *et al.*, 1995), α- and β-scruin of *Limulus* (Way *et al.*, 1995), the bovine and human calicins (von Bulow *et al.*, 1995), human NRP/B (Kim *et al.*, 1998), and human NS1-binding protein (Wolff *et al.*, 1991). Direct evidence of actin association has been demonstrated for *Limulus* α-scruin (Sanders *et al.*, 1996), in which the actin binding activity of the protein has been localized specifically to the kelch repeat domains of the molecule. Additional evidence of actin association in members of the kelch family includes co-precipitation and co-localization of mouse ENC-1 in a transfected medulloblastoma cell line (Hernandez *et al.*, 1997), co-localization of actin filaments with *Drosophila* kelch in ovarian ring canals (Xue and Cooley, 1993), co-sedimentation of human IPP fusion proteins (Kim *et al.*, 1999), and the disruption of actin filament localization in the spermatocytes of *C. elegans* spe-26 mutants (Varkey *et al.*, 1995). In addition to the kelch repeat motif, most kelch proteins also possess a ~120 amino acid motif, usually found at the extreme N-terminus of the protein, and referred to as the POZ domain (Bardwell and Treisman, 1994). The POZ domain has been reported to mediate homo- and hetero-dimerization *in vitro* and the formation of multimeric complexes *in vivo* (Bardwell *et al.*, 1994; Robinson and Cooley, 1997).

Metabolism of animal viruses, depending on their peculiarities, suggests the use of structures of cytoplasmic or nuclear protein cell skeleton, as well as cell membranes. In particular, virus-specific cytopathic effect can be caused by a specific reconstruction of the elements of cell skeleton and certain membranes. The aim of this reconstruction is to create the conditions for viral propagation (Luftig, 1982). The cytoskeleton proteins are coded by a large set of genes, which are characterised by tissue-specific expression. This accounts for the difference in the protein contents of the "framework" of various cell types, which affects the function of these cells (Steinert and Roop, 1988). Various cell types may also differ in their membrane contents. The above differences can also affect

such parameters of virus propagation inside the organism such as tissue tropism (Shchelkunov *et al.*, 1993). The detection of genes that determine the tissue tropism of viruses is important for the fundamental understanding of virus-cell interaction, and also for the creation of more safe live vaccines.

c) Virus Growth Factor (VGF)

The vaccinia virus growth factor (VGF) is encoded in the left terminal region, and is thought to be involved in pathogenesis. VGF binds to the epidermal growth factor receptor (EGFR), induces phosphorylation of EGFR, and the proliferation of certain cell types (King *et al.*, 1986; Stroobant *et al.*, 1985; Twardzik *et al.*, 1985). Deletion of this gene from VV results in the attenuation of the virus (Buller and Palumbo, 1991). It has been suggested that in SPV, the absence of this gene, together with the complement-binding C4b binding protein of vaccinia, and a secreted vaccinia protein N1L homolog may be responsible for the highly attenuated SPV phenotype, and the characteristically mild natural infections in swine (Massung *et al.*, 1993).

d) Other proteins and novel poxvirus-encoded proteins

Exclusive or novel genes are difficult to define, because novelty will depend on the size of the poxvirus sequence database and upon the cut-off point of for alignment scores. Furthermore, alignment scores are dependent upon protein size, and it is easy to miss homologous genes if they happen to be very small. Nevertheless, all authors of poxviral genome sequence publications report genes that do not seem to have homologs in any other of the poxvirus species. The most novel variola protein, according to Massung *et al.*, is ORF B22R, which shows poor homology to current database proteins. The protein (Mr = 214K) has six possible transmembrane domains, suggesting that it may locate to virion or infected-cell membranes, and thereby impact infection. Another exclusive VAR gene is ORF D6R, which contains a complete zinc-finger motif not found in a truncated 21.7 kD VV homolog (Kotwal *et al.*, 1988). Variola ORF D11L specifies a homolog of the vaccinia ORF C7L product known to influence viral infectivity for certain cell lines (Perkus *et al.*, 1990), but ORF D11L includes an Arg-Gly-Asp cell-adhesion motif (Rhouslati and Perschbacher, 1986) which could profoundly influence tissue tropism.

Finally, variola ORF J5.5R (of which 45 residues have partial homology to vaccinia ORF A55R), is an analogue of vascular endothelial growth factor having a homolog in the genus *Parapoxvirus* (personal communication to Massung *et al.*, by A. J. Robinson), suggesting that vascular endothelial growth factor is a virulence determinant stemming from a common poxvirus ancestor. Several genes of vaccinia and variola viruses are absent in shope fibroma virus, suggesting that SFV may not use the same infection route(s) as do other Chordopoxviruses. SFV appears to lack homologs of the A-type inclusion proteins (Willer *et al.*, 1999). These genes have been shown to be mutable in a variety of Orthopoxviruses, but most viruses produce at least a truncated form of the protein that may promote infection of phagocytic cells by intracellular mature virions (Ulaeto *et al.*, 1996). The complete absence of any shope fibroma virus gene homolog is unusual and predicts that Leporipoxvirus infections might produce relatively small amounts of intracellular mature virions (Willer *et al.*, 1999). A very interesting example of another unusual protein encoded by a poxvirus, is the *Leporipoxvirus* S127L gene product: this gene encodes a 52.5-kDa product that is homologous to proteins belonging to the family of type II cyclobutane pyrimidine dimer (CPD) photolyases (Willer *et al.*, 1999). Photolyases catalyze a light-dependent reaction that returns CPD adducts back to their component pyrimidines, thus reversing the damage caused by ultraviolet light. Although a similar gene is encoded by molluscum contagiosum virus (Afonso *et al.*, 1999), this is the first DNA repair gene of its kind found to be encoded by a Chordopoxvirus, in fact, any mammalian virus (Willer *et al.*, 1999). The presence of this gene is particularly unusual in that whilst marsupials such as the common opossum (*Monodelphis domestica*) are known to encode photolyases (Yasui *et al.*, 1994), no placental mammal has yet shown to have retained the gene. Because it is generally believed that poxvirus originally acquired most genes from their hosts, this discovery has raised interesting questions concerning the evolutionary history of a New World virus now thought to only infect rabbits (Willer *et al.*, 1999). Preliminary experiments show that the gene can complement an *E. coli phr* (photolyase deficient) mutation *in vivo* (D. Willer, M. Moon and D. Evans, unpublished data) and imply that Leporipoxviruses can repair the damage caused by ultraviolet light exposure either before and/or during viral infection (Willer *et al.*, 1999). Other genes encoded exclusively by Leporipoxiruses

include a 34-kD α 2,3-sialyltransferase nearly identical to the myxoma enzyme recently characterised by Jackson and colleagues (Jackson *et al.*, 1999). Shope fibroma virus sialyltransferase presumably catalyses an identical reaction and likewise promotes viral virulence. Another SFV gene product, S141R also has a previously identified myxoma gene homolog (MA56)(Jackson *et al.*, 1999), encodes a 24-kDA protein that is predicted to contain two transmembrane helices and resembles the immunoglobulin-like domains of a large family of eukaryotic cell-surface proteins (Willer *et al.*, 1999).

Database searches with several SPV ORFs revealed no significant homology to any previously identified poxvirus polypeptides. Of these, the C1L and C3L ORFs are the most interesting. The C1L ORF, located nearest the terminus and encoding a polypeptide of 340 amino acids, shows limited homology to the members of the immunoglobulin superfamily, and also contains the tripeptide RGD, a cell attachment sequence found in a variety of cellular glycoproteins (Rouslahti and Pierschbacher, 1986, Massung *et al.*, 1993). The C3L ORF exhibits 25% identity at the amino acid level with the G-protein coupled receptor family, and is predicted to have six transmembrane domains (Massung *et al.*, 1993). No significant homologs were reported for SPV ORFs C7L and C9L, either (Massung *et al.*, 1993).

As an increasing amount of poxviral genome sequences become available, it is possible, even likely, that homologs of these exclusive genes will be identified. Massung and colleagues (1993a) suggested, against the argument that host range specificity and virulence is determined by function-specific genes, that rather small differences, as presented above, between homologous poxvirus proteins augment the collective contributions of all proteins in determining the systemic or localised nature of infection by these different viruses.

1.4 VACCINES

1.4.1 Lumpy Skin Disease Vaccines Two live attenuated strains of capripoxvirus have been used as vaccines specifically for the control of LSD (Capstick and Coakley, 1961; Carn, 1993): a strain of Kenya sheep pox and goat pox virus passaged 18 times in lamb testes or fetal calf muscle cells, and the South African vaccine strain (Neethling), which originated from the adaptation of a strain of LSDV to propagation in embryonated eggs. The strain of virus selected for these propagation studies was the Neethling type of Group III virus described by Alexander *et al.* in 1957. This virus had been stored in freeze-dried form after four passages in calf kidney, and six passages in lamb kidney tissue cultures. After reconstitution, the virus was seeded on to lamb kidney tissue cultures. On the 5th day, when cytopathogenesis was almost completed, the culture was stab-inoculated into fertile eggs and incubated. Eggs containing dead embryos on succeeding days had their embryos together with the chorioallantoic membrane (CAM) harvested, emulsified, and re-injected into the next series of eggs. At each stage, each inoculum prepared for egg passage was passed to lamb kidney or –testis monolayers to determine the cytopathic effect on tissue culture (the attenuated virus produces no cytopathic effect upon monolayer tissue cultures), and the material was titrated to determine the LD_{50} end-point (where the attenuated virus is not lethal to the chick embryo). The strain was passaged for 45 generations. (van Rooyen *et al.*, 1959). At the 20th passage level, the virus failed to cause a generalized skin eruption or other signs of illness in cattle, and only produced a local swelling at the site of inoculation in about 50% of animals injected. The local reactions disappeared within four to six weeks without any evidence of necrosis. Virus could not be re-isolated from the blood or the local lesions of inoculated cattle, and there was no evidence of reversion of virulence of the attenuated virus by re-passage in lamb kidney tissue cultures (Weiss, 1960). The modern LSDV Neethling vaccine is prepared in much the same way (*Office International des Epizooties* Manual, 1996).

All strains of capripoxvirus examined so far, whether of bovine, ovine or caprine origin, share a major neutralising site, so that animals recovered from infection with one strain are resistant to infection with any other strain. Consequently, it is possible to protect cattle against LSD using strains of capripoxvirus derived from sheep or goats (Coakley

and Capstick, 1961). Protection following vaccination is probably life-long, although as immunity wanes, local capripoxvirus replication will occur at the site of inoculation, but the virus will not become generalised. Both strains of capripoxvirus used routinely as vaccines can produce a large local reaction at the site of inoculation in cattle breeds, which some stock owners find unacceptable. This has discouraged the use of vaccine even though the consequences of an outbreak of LSD are invariably more severe (OIE Manual, 1996).

1.4.2 Pox-vectored vaccines. After the eradication of the smallpox virus, and consequently the lack of reasons to vaccinate against it, vaccinia virus was considered for use in the immunoprophylaxis of other diseases. Live recombinant poxviruses-vectored vaccines have significant advantages over more conventional vaccines, yielding long-lasting protective immunity to more than one disease simultaneously and in this regard, offer huge potential (Weir *et al.*, 1982; Nakano *et al.*, 1982; Mackett *et al.*, 1982, Panicani and Paoletti, 1982).

Vaccinia virus has a capacity for at least 25 000 bp of additional DNA (Smith and Moss, 1983), but the large size of the poxvirus genome makes it necessary to rely on homologous recombination for the insertion of foreign DNA. This procedure depends on infrequent recombination events occurring in infected cells between the poxvirus DNA flanking the foreign gene, and homologous DNA within the poxvirus genome. For efficient expression of foreign genes, vaccinia virus promoters had to be used in the construction of the recombinants. A series of plasmid vectors for vaccinia were constructed, based on the scheme outlined by Mackett *et al.* (1982). Such plasmids contain a DNA segment (with a vaccinia virus transcriptional promoter region and one or more restriction endonuclease sites for insertion of a foreign gene) that is flanked by DNA from a non-essential region of the vaccinia genome. The choice of promoter determines the time and level of expression, whereas the flanking DNA determines the site of insertion. Insertion into the thymidine kinase (TK) gene provided a simple method of selecting recombinant virus that had a TK^- phenotype. The cells were consequently infected with vaccinia virus and transfected with the recombinant plasmid. The virus

progeny were then incubated with a monolayer of TK^- cells, and 5-bromodeoxyuridine is added to the agar overlay to prevent plaque formation of TK+ parental virus. Other screening methods that don't rely on special cell lines have also been developed (Chakrabarti *et al.*, 1985). Alternatively, DNA hybridisation (Panicali *et al.*, 1982) or specific antibody binding (Smith *et al.*, 1983; Paoletti *et al.*, 1984) may be used for primary identification of recombinant plaques. Over the past fifteen or so years, several of these live recombinant vaccines have been produced, and their ability to protectively immunise experimental animals against a variety of viruses has been demonstrated. Some good examples of these recombinant vaccinia vaccines are the hepatitis B surface antigen (Smith *et al.*, 1983; Moss *et al.*, 1984), influenza hemagglutinin gene (Smith *et al.*, 1983, Panicali *et al.*, 1983), herpes simplex type I glycoprotein D (Cremer *et al.*, 1985), hepatitis B surface antigen + herpes simplex D glycoprotein (Paoletti *et al.*, 1984), vesicular stomatitis virus genes (Mackett *et al.*, 1985) and a rabies virus recombinant (Wiktor *et al.*, 1984). Occasional adverse clinical reactions in response to wild type vaccina virus (Lane and Millar, 1971), and its broad host range, however, limits the use of vaccinia virus as a general vaccine vector (Arita and Fenner, 1985). Species-specific poxviruses such as fowlpox virus (Taylor and Paoletti, 1988) and canarypox virus (Taylor *et al.*, 1988), offer certain advantages: they do not spread to non-target species since outside their host-range, and infection leads to incomplete replication. This, nevertheless, appears to be sufficient for the expression of foreign genes and the induction of a protective immune response (Taylor and Paoletti, 1988). A fowlpox recombinant expressing the fusion (F) protein of Newcastle disease virus (NDV) has been generated at OVI, and is currently under evaluation in small animal trials as a vaccine against Newcastle disease (Z. Wang, unpublished data).

Modified live capripox vaccines have achieved promising levels of immunity, and these viruses therefore also have excellent potential as alternative antigen vectors to vaccinia virus for veterinary applications. Such a recombinant capripoxvirus was constructed by Romero *et al.* (1993), and contains a full-length cDNA of the fusion protein gene of rinderpest virus. The gene was inserted in the thymidine kinase gene of the capripox genome under the control of the vaccinia virus major late promoter p11, together with the

E. coli gpt gene in the opposite orientation under the control of the vaccinia early/late promoter p7.5. A vaccine prepared from the recombinant virus protected cattle against clinical rinderpest after a lethal challenge with a virulent virus isolate. In addition, the vaccine protected the cattle against lumpy skin disease.

Because it is restricted to its bovine host, LSDV is another excellent candidate as a viral vaccine vector. OVI is currently investigating the development of its LSDV vaccine strain as a recombinant vector veterinary vaccine. A Rift Valley fever virus (RVFV)/LSDV recombinant, previously generated by the insertion of the RVFV glycoprotein genes, G1 and G2, into an intergenic region in the LSDV genome, has been evaluated in sheep. Sheep receiving the RVFV/LSDV vaccine all sero-converted. All sheep survived a virulent viral challenge, but the RVFV/LSDV-vaccinated animals showed an overall lower temperature profile, shorter period of viraemia and a quicker and elevated neutralising antibody response compared with control groups. This preliminary data suggests that the RVFV/LSDV recombinant vaccine is as good, if not better than the Smithburn vaccine currently in use against RVF (unpublished data).

The terminal regions of the LSDV genome containing the genes for host range specificity and virulence are non-essential for growth in tissue culture, and therefore viral replication and survival. They make ideal sites then, for the insertion of foreign genes. The sequencing of the LSDV genome would enable the identification of these genes, as well as their regulatory elements, as the success of a recombinant viral vaccine is very much dependent on the levels of immunity that is induces in the host, and a strongly-expressed protein under the control of a well-characterised, strong LSDV promoter would be ideal.

1.5 Aims of this study

LSDV is a capripoxvirus causing an economically serious disease of cattle in Africa and the Middle East, resulting in its classification as an OIE List A disease. Despite the availability of vaccines, periodic outbreaks still occur. The South African vaccine strain (also called the Neethling strain) of LSDV is the subject of this study. In order to gain a better understanding of the virus-host interaction, specifically the genetic factors which determine the virulence of the virus (or avirulence in the case of a vaccine strain) and the bovine-specific nature of LSDV infection, the determination of the nucleic acid sequence of the viral genome to identify the responsible genes is necessary. Virtually no nucleotide sequence data is published for LSDV. Once the genome sequence is determined, we would be better equipped to understand the virus-host interaction, and ultimately be able to develop safer and more effective vaccines. Regions of the genome that are non-essential for the replication cycle of the virus are of particular importance, as LSDV shows enormous potential as a recombinant vaccine vector.

The first aim of this study was to clone central and terminal regions of the LSDV genome, specifically those regions expected to encode the genes involved in host range specificity and virulence. From the known sizes and locations of the *PstI* clones in the LSDV genome, the *PstI* -E, -M and -K clones from the left terminal regions of the LSDV SA vaccine strain genome were selected, representing approximately 12 700 bp of the left hand terminal sequence, a variable region in which the pathogenicity and host-range specificity genes of poxviruses are known to occur. Additionally, *PstI* clones -F and -G, representing about 16 800 bp of the central regions of the genome (containing essential genes required for viral replication and survival) were selected for cloning, sequencing and analysis.

The second aim after determining the nucleotide sequences of the clones, was to assemble the fragments into a continuous LSDV DNA sequence. Large-scale laboratories sequencing other poxvirus genomes use a "shotgun" cloning and automated sequencing approach. As a result of limited infrastructure this was not possible here, and so the mapping and cloning approach would be a more systematic one. To distinguish between

similar-sized clones bearing different LSDV fragments, a restriction endonuclease pattern profiling method was used. Computer analysis was used to assemble non-overlapping sub-clone sequences into the correct orders and orientations. This involved using the sequences of homologous poxvirus proteins from the public databases as templates, to assemble the sub-clone amino acid sequences into full-length ORFs. Oligonucleotides designed from LSDV sequences would be used to primer-walk gaps in DNA sequence in certain regions.

The third aim of this investigation was to use computer analysis to translate assembled LSDV nucleotide sequence into amino acid sequences. Open reading frames were identified, and through homology searches of public databases, possible functions were assigned to the putative LSDV proteins. Several proteins with possible roles in host-range specificity and virulence were identified in the sequenced regions.

CHAPTER II

The Restriction Endonuclease Mapping, Sub-cloning and Sequencing of LSDV *PstI* Clones -E, -M, -K, -F and -G

2.1 INTRODUCTION

Poxviruses have the largest known viral genomes (139-288 kbp). Several complete genomic DNA sequences have been deposited in GenBank to date, including a single viral member of the Entomopoxvirinae family, *Melanoplus sanguinipes* entomopoxvirus (MsEPV) (Afonso *et al.*, 1999), as well as those of several *Chordopoxvirinae* family. The latter include the *Orthopoxviruses* vaccinia (VV) (Goebel *et al.*, 1990; Antoine *et al.*, 1998) and variola (VAR) (Massung *et al.*, 1994; Shchelkunov *et al.*, 1995), *Molluscipoxvirus* molluscum contagiosum virus (MCV) (Senkevich *et al.*, 1997) *Avipoxvirus* Fowlpox virus (FPV) (Afonso *et al.*, 2000), and the *Leporipoxviruses* Shope (Rabbit) Fibroma Virus (SFV) (Willer *et al.*, 1999) and myxoma virus (MYX) (Cameron *et al.*, 1999), with new complete genomic DNA sequences well underway to completion. In this study, important regions of the previously-unsequenced genome of the capripox virus, lumpy skin disease virus (LSDV), were selected for DNA sequencing and analysis. It was first necessary to sub-clone large fragments into smaller ones that could be sequenced in the forward and reverse direction, and therefore the locations of infrequent six-base restriction endonuclease sites in the LSDV clones had to be mapped, so that cloning strategies could be devised.

The restriction endonuclease *HindIII* has successfully been used in the molecular characterisation of different capripoxviruses (Kitching *et al.*, 1989) and in addition to the *HindIII* restriction endonuclease maps thereby generated for SA-LSDV, *AvaI*, *SalI* and *PstI* maps have also been compiled (unpublished data presented in the MSc dissertation of Perlman, 1993, UCT). *AvaI* cleaves the 152 kbp genome into seven fragments, varying from 47.35 kbp to 2.1 kbp, whereas *SalI* cleaves the genome into six fragments, ranging in size from 33.15 kbp to 2.8 kbp. The enzyme *PstI* cuts more frequently than the aforementioned, generating 13 fragments ranging in sizes of 42.0 kbp to 2.4 kbp, and was therefore considered to be the most suitable enzyme for cloning the LSDV genome.

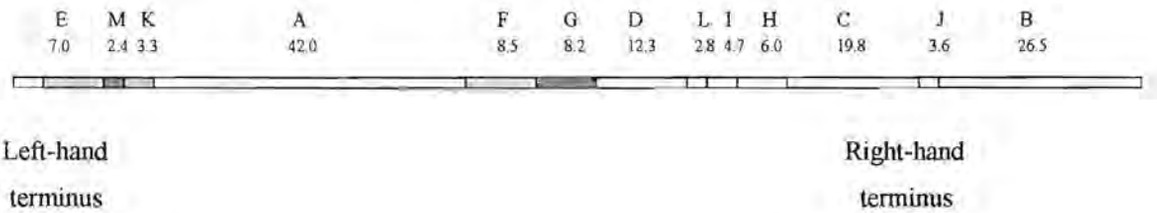

Figure 5: A linear representation of the LSDV *PstI* map (Unpublished data in the MSc dissertation of Perlman, UCT 1993). DNA fragment sizes are given in kilobase pairs, with the regions of importance in highlight. The *PstI*-E clone is 12.0 kb in total; the 7 kb *EcoRI/Pst* fragment is indicated above.

To date, all the *PstI* fragments except the three largest, namely -A, -B and -C, have been cloned into bacterial vector pGEM3Z (F. T. Vreede and D. B. Wallace, OVI). In the cloning of *PstI*-E, the original 12.1 kbp fragment was cleaved at an internal *EcoRI* site, resulting in the cloning of a 7 kbp sub-fragment. It was confirmed to be of the *PstI* E-fragment by hybridization to a Southern blot of a LSDV *PstI* genomic digestion (F. T. Vreede). Furthermore, it was later shown that this *PstI*-E sub-fragment is adjacent to the *PstI*-M fragment i.e., furthest from the left terminus. The remaining 5 kbp *PstI*-E fragment has not yet been cloned. Our laboratory is also involved in cloning *BamHI* fragments overlapping the uncloned *PstI* fragments, and in the cloning of the terminal hairpin loops.

The cloning and sequencing of the terminal hairpin loops of poxvirus genomes has been unsuccessful in most cases. The sequences for the terminal hairpin loops of VV are available (Goebel *et al.*, 1990), which was achieved by the chemical degradation sequencing method of Maxam and Gilbert (1980). In a different approach for the myxoma virus terminal regions, blunt-ended adaptors were ligated to 25 and 29 bp oligonucleotides. The oligonucleotides were designed to be complementary to the upper and lower DNA strands. In a PCR reaction, the terminal sequences adjacent to the hairpin loops were amplified after which the products were sequenced. The actual loop sequence was not determined however (Cameron *et al.*, 1999).

This chapter describes the mapping the LSDV *PstI* clones with restriction endonucleases and the sub-cloning of LSDV DNA fragments into bacterial sequencing vectors. The application of manual and automated sequencing, to determine the linear order of the deoxyribonucleotides from selected regions of the LSDV genome, is also described. It was expected that the LSDV genome would have a high frequency of A+T- nucleotides, as suggested by the high frequency of A+T- rich restriction endonuclease sites in the capripoxvirus genomes (Gershon and Black, 1989). Once the nucleotide sequences were assembled and translated to amino acids (Chapter III), it was expected that several genes involved in pathogenicity and in the restriction of the virus to bovine cell infection would be identified in the selected regions.

2.2 MATERIALS AND METHODS

2.2.1. Restriction endonuclease mapping of the *PstI* clones. Glycerol stocks of a selection of *E. coli* LSDV *PstI*-fragment clones were inoculated into LB medium with ampicillin and grown overnight, with shaking at 37°C. The randomly selected clones from each *PstI* fragment were: M_{72}, M_{94}, M_{102}, E_4, E_5, E_6, K_1, F_3, F_{11}, F_{14}, G_{26}, G_{27}, G_{28} (see Fig. 5 for genomic positions of *PstI*-E, -M, -K, -F and -G). Clones were named according to order of isolation by F. T. Vreede (personal communication). The plasmids were isolated from the pelleted bacterial cells by the modified alkaline lysis, or "miniprep" method (Birnboim and Doly, 1979; Ish-Horowicz and Burke, 1981: Sambrook *et al.*, 1989). Restriction enzyme digests for mapping purposes were performed in 10µl reaction volumes: 10X restriction buffer supplied by the enzyme's supplier, 0.5 U of enzyme, ~ 500ng of plasmid DNA, and distilled water to make up the volume. The reaction was incubated for an hour in a 37°C waterbath, after which RNAse A at a concentration of 10mg/ml was added, and incubated for a further 5 to 10 minutes. 5µl of loading dye (bromophenol blue/ 40 % sucrose) was added, and loaded onto an ethidium bromide-stained (EtBr) 1% agarose gel, with undigested plasmid and a phage-λ DNA *PstI* marker, and electrophoresed at 90-100 volts until the fragments had migrated sufficiently. Restriction enzyme double digests were assembled in 10µl volumes, where enzymes were selected on their optimal requirement for similar buffer components. In cases where the enzymes required totally different buffer conditions, a buffer titration was performed, and the reaction was incubated for an additional half-hour. After electrophoresis, the EtBr-stained DNA bands could be visualised by trans-illumination with long wavelength ultraviolet light (UVP M-20). Images were captured by Lumi-Imager™ (Boehringer Mannheim). Band fragment sizes were estimated according to the *PstI*- Bacteriophage λ DNA ladder.

2.2.2. Cloning

a) Preparation of LSDV insert DNA

~500ng of plasmid DNA was digested in reaction volumes of 25µl or more, with the supplied 10X buffer and 0.5 U of restriction enzyme. The reaction was incubated for 90 minutes in a 37°C water bath, after which 10 µl of RNAse A at a concentration of 10mg/ml was added, and incubated for a further 10 minutes. 20µl of loading buffer was added, and 20µl volumes of the reaction loaded into separate wells of an EtBr-stained 1% agarose gel. A λ-*PstI* DNA marker, and uncut plasmid were also loaded alongside. Electrophoresis at 90-100 V for 45 minutes was performed, or until the fragments had separated fully. Bands of the correct sizes were excised from the agarose gel with a clean, sharp scalpel. DNA was purified from the agarose using one of several commercial DNA purification kits, according to the manufacurers' specifications. These kits included the Talent Cleanmix DNA purification kit™, Gene Clean™, NucleoTrap™ (Machery-Nagel) and QiaQuick Spin™ (Qiagen). DNA fragment concentrations were determined spectrophotometrically (Pharmacia GeneQuant spectrophotometer).

b) Preparation of Vector DNA

One of two vectors were used for cloning of all insert DNA viz. pGEM3ZF+/- (Promega), or pMOSBlue (Amersham). Both are pUC-based and therefore contain commonly- used six-base restriction endonuclease sites in their multiple cloning sites (MCS). pMOS also has a site for *NdeI* in its MCS, which was used to clone fragments generated with *AseI* , since digestion with *AseI* and *NdeI* yields compatible ends.

pGEM or pMOS-plasmid-containing *E.coli* cells were inoculated into LB medium supplemented with ampicillin, and grown with shaking overnight in a 37°C incubator. Vector plasmids were isolated from pelleted bacterial cells by the "miniprep" method, and resuspended in dH_20. Vector plasmid DNA was linearised with the appropriate restriction enzyme/s in 25µl reaction volumes: ~500ng of plasmid DNA, 10X buffer supplied with the restriction enzyme, or a buffer compatible to both in the case of directional cloning strategies, 0.5U DNA of restriction enzyme/s and dH_20 to make up

the volume. The reaction was incubated at 37°C for 60 minutes and loaded onto an EtBr-stained 1% agarose gel, alongside undigested plasmid DNA and the λ-*PstI* marker. This was electrophoresed at 90-100 V for 45 minutes, the linear plasmid bands (3199 bp and 2800 bp for pGEM3Z and pMOSBlue, respectively) were excised from the agarose with a clean sharp scalpel, and purified with one of the aforementioned kits according to the manufacturer's specifications. The linear vector DNA was resuspended in dH_20, and the concentration determined by spectrophotometry (Pharmacia GeneQuant spectrophotometer).

In cases where the cloning would be directional (i.e. the ends of the linearised vector are generated with different restriction enzymes) the vector was used directly for the ligation reaction. In cases where only one enzyme was used to linearise the vector, the ends were dephosphorylated to prevent self-ligation. 0.5 U of Calf-Intestinal Phosphatase (CIP) (Roche) was added to the total volume of purified linear DNA, together with the CIP-10X buffer, and dH_2O, usually to a volume of 40µl. The dephosphorylation reaction was incubated in a 37°C water bath for 20 min. The reaction volume was raised to 200µl by the addition dH_20, and a 3X volume (600µl) of Phenol/Chloroform/Isoamyl alcohol was added, to purify the vector DNA. After vortexing, centrifugation (13000 g's) and removal of the aqueous DNA-containing phase, DNA was ethanol-precipitated and washed before being resuspended in dH_20. Concentration of the DNA was estimated spectrophotometrically (Pharmacia GeneQuant spectrophotometer). Vector DNA was stored at -20°C.

c) Ligation Reactions

Ligation reactions were assembled in 10µl reaction volumes: insert and vector DNA was combined in a molecular ratio of 3:1, respectively, with 0.5 U of T4 DNA ligase enzyme (Roche), the supplied 10X buffer, and dH_20. Ligation reactions were incubated in a water bath of 15°C, overnight, or at 17°C for at least 4 hours. Ligation reactions were stored at -20°C if not used immediately.

d) Competent cells and transformations

Two methods of preparing competent *E. coli* DH5α cells were used: the first was the standard Calcium Chloride method, (Cohen *et al.*, 1972; Sambrook *et al.*, 1989) the second was the High Efficiency Transformation method as described by Inoue *et al.* (1990). The second method was found to yield cells that have superior competency levels to those prepared by the calcium chloride method, and had the added advantage of remaining highly competent after storage in liquid nitrogen at -70°C for longer than 12 months. This method required inoculated DH5α cells to be grown at 18°C, with shaking, to an OD of 0.6 at A600, which took approximately 53 hours. Pelleted cells were then resuspended in TB-buffer, consisting of 10mM Pipes, 55mM $MnCl_2$, 15mM $CaCl_2$ and 250mM KCl. Cells are stored with added DMSO at a final concentration of 7%, in liquid nitrogen.

Cells taken from the liquid nitrogen were allowed to defrost at room temperature, and then placed on ice. 200µl of cells were aliquotted into plastic tubes, and 2µl of the respective ligation reaction added. The cells were incubated in watery ice for 30 minutes, and then heat-shocked at 40°C for 30 seconds. The tubes were returned to the ice for 5 minutes, before 800µl of LB medium (without ampicillin) was added, and the tubes incubated, with shaking, at 37°C, for one hour. 200µl of cells were plated onto LB plates containing ampicillin, and grown overnight at 37°C.

e) Screening for recombinant DNA clones

Individual colonies were picked from agar plates and inoculated into 2ml of LB medium with ampicillin. Cells were grown for a minimum of six hours at 37°C, with shaking. 150µl of the cultures were screened for plasmids with inserts according to the "fast prep" method described by Beuken *et al.* (1998). After electrophoresis, 10µl each of the cell lysates were loaded onto EtBr-stained 1% agarose gel, potential recombinant plasmids were identified on grounds of their increased size and hence their slower migration in the gel. Cells from potential recombinant cultures were pelleted and the plasmids extracted by the "miniprep" method. To confirm a recombinant, restriction enzyme sites flanking

the insertion sites (and/or known to not have a site in the insert itself) in the plasmid were cut with the appropriate enzyme (usually *EcoRI* [Promega] and *HindIII* [Promega]). These restriction enzyme digests were performed in 10µl reaction volumes. Again this included the 10X restriction buffer supplied by Promega, 0.5 U of the respective enzyme, ~ 500ng of plasmid DNA, and dH_2O to make up the volume. The reaction was incubated for an hour in a 37°C water bath, after which RNAse A at a concentration of 10mg/ml was added, and incubated for a further 5 to 10 minutes. 5µl of loading buffer was added, and loaded onto an EtBr-stained 1% agarose gel, with undigested plasmid, an undigested wild-type (pGEM or pMOS) plasmid and a phage-λ DNA *PstI* marker, and electrophoresed at 90-100 volts or until the fragments had migrated sufficiently. Triplicate glycerol stocks were prepared of confirmed recombinant plasmids by adding 200µl of 100% sterile glycerol to 800µl of recombinant cultures, and stored at -20°C.

2.2.3 Sequencing

2.2.3.1 Preparation of DNA template for sequencing. Fresh overnight cultures of the clones to be purified for sequencing were pelleted by centrifugation (13000 x g). Boehringer Mannheim's High Pure Plasmid Isolation Kit™ was used to isolate the recombinant plasmids from the respective clones. In principle, the *E. coli* cells were lysed by the alkaline method. Nucleic acids bound specifically to the surface of glass fibres of the *High Pure* filter tube in the presence of a chaotropic salt. Since the binding process was specific for nucleic acids, the bound plasmid DNA was purified from salts, proteins and other cellular impurities by a washing step, and could then be eluted in a low salt buffer or water. DNA for sequencing purposes was eluted from the spin columns with 100µl of dH_2O. DNA template quantitation is critical for successful sequencing reactions. DNA concentrations were determined by a Pharmacia GeneQuant spectrophotometer: 2µl of purified plasmid was diluted in 98µl of dH_20. This gave a dilution factor of 1 in 50, which was then equated to a DNA concentration in ng/µl.

2.2.3.2 Primers and primer walking. pGEM3Z (Promega) and pMOSBlue (Amersham) are both pUC-based bacterial vectors, and therefore contain binding sites for the M13 forward and reverse primers. The SP6 reverse primer and T7 forward primer were used in some cases, and prime sequences close to the MCS in both vectors. The M13 primers anneal further upstream and downstream of the multiple cloning sites than the T7-promotor or SP6 primer sites, and were therefore used in the majority of the sequencing reactions, as the actual recombination site is important to confirm the correct clone, its orientation in the vector, and for the accurate reading of DNA sequence close to the priming site.

Table 2 : Commercial primers for the sequencing of bacterial vector plasmids

Primer	DNA Sequence
M13+ (GibcoBRL)	5' -TGT AAA ACG ACG GCC AGT- 3'
M13- (GibcoBRL)	5' -CAG GAA ACA GCT ATG AC- 3'
T7- (Promega)	5' -TAA TAC GAC TCA CTA TAG GG- 3'
SP6+ (Promega)	5' -TAT TTA GTT GAC ACT ATA G- 3'

Primer sequence, method of primer synthesis, and approach to primer purification can have a significant effect of the quality of the sequencing data obtained in dye terminator cycle sequencing reactions. The recommendations made for primer design (ABI PRISM™ Protocol) were followed:

1) primers with long runs of a single base were avoided
2) primers were at least 18 bases long to ensure good hybridisation
3) primer melting temperatures were kept above 45°C.
4) a primer content of at least 50% G+C was retained, with a few exceptions, and it was sometimes necessary to extend the primer sequence beyond 18 bases to keep the melting temperature above the recommended lower limit of 45°C

Primers containing palindromic sequences, that would lead to self-ligation, were avoided

Table 3: Primers designed for primer-walking sequencing to sequence gaps

Name	Priming site	DNA sequence	MW	% G+C	Tm (50mM NaCl)
K440	K1200	5'-CCA AAA ACT AGC GAC AGC-3'	6,734.0	50.0%	52.08°C
K875	E2700/ 1450	5'-GCA CAG TTA GTA GCG CC-3'	6,392.0	58.8%	53.48°C
K12/K5P	K500; L	5'-GCA CAT AAG TAG TTT ACA CTG C-3'	8,237.0	40.9%	52.95°C
G68/602P	G6800/ 602	5'-CCG GAA TTG GAG ATG TTT C-3'	7,198.0	47.4%	51.78°C
G68/550P	G6800/ 550	5'-GCA CAT CTT CCC CAT CCT G-3'	7,014.0	57.9%	56.27°C
F66A	F6600	5'-CAT CCC ATA TGG TTG CC-3'	6,327.0	52.9%	50.54°C
F66B	F6600	5'-GGC ACC TAA CCT ACC ATC-3'	6,664.0	55.6%	52.45°C
F29A	F2900	5'-CCA CTG AAA AGA GTG GAG C-3'	7,201	52.6%	53.95°C
F29B	F2900	5'-GGA ACA GGG TAC TAC CG-3'	6,435.0	58.8%	51.43°C
F29C	F2900	5'-CGG GTT CTC TAC CCA AAG C-3'	7,094.0	57.9%	56.05°C
E2150A	F2900	5'-GAT GGA AAA AGT AAC GAC-3'	6,844.0	38.9%	46.27%

Primers were manufactured by Integrated DNA Technologies, Inc., and diluted in dH_20 to a final concentration of 1.5 pmol/µl.

2.2.3.3 The automated DNA sequencing reaction. The ABI PRISM™ Dye Terminator Cycle Sequencing Ready Reaction Kit (Perkin Elmer) was used according to the manufacturer's specifications. The protocol was modified by scaling the reaction mix down to a quarter-reaction.

a) The cycle sequencing reaction

A typical reaction was therefore assembled in 50µl tubes, on ice, as follows:

4-6µl	plasmid DNA template (~600ng)
1µl	primer (1.5pmol/µl)
2µl	kit (quarter-reaction volume)
1.6µl	5x buffer (PE Applied Biosystems, kindly supplied by Joelle vd Walt, UP)
	dH_20

13µl	total volume

Cycle sequencing was performed on a GeneAmp PCR System 2000 thermal cycler™ (Perkin Elmer) under the recommended parameters:

Denaturation: 96°C for 30 seconds/ Annealing: 50°C for 15 seconds/ Elongation: 60°C for 4 minutes (25 cycles).

b) Purification of extension products

The 13µl reaction volume was transferred to a 500µl Eppendorf tube containing 25µl of 96% Ethanol, and 1µl of 3M NaOAC, at room temperature (resuspending very gently with a pipette- no vigorous mixing or vortexing). This was left to precipitate in the dark for 10 minutes. The samples were then centrifuged at 13 000 rpm for 30 minutes, and all the liquid was carefully removed with a pipette, taking care to avoid the area of the (invisible) pellet. 250µl of 70% EtOH was then added (no mixing or vortexing) and centrifuged again on a benchtop centrifuge at 13 000 rpm, for 15 minutes.

All EtOH was carefully aspirated with a pipette, and then vacuum-dried for 10-15 minutes. The tubes containing the dried, pelleted DNA was stored at -4°C until the gel was run, and transported on ice.

c) Gels

Sequencing reactions were analysed on an ABI377 sequencer at either the division of Molecular Biology at OVI, or the Faculty of Natural and Agricultural Sciences, UP. Long Ranger™ gel solution matrix, a ready-to-mix gel solution is used by both laboratories, which is cast and allowed to polymerize. This gel solution was able to give a 25 to 30% increase in read length and an increase in base calling accuracy on automated sequencers when compared with standard polyacrylamide gels. Prior to loading, the dried pellet was resuspended in 4µl of 1:5 blue dextran/ EDTA: deionized formamide, denatured at 95°C and put onto ice. 0.5-1µl was loaded, depending on the size of the gel.

2.2.3.4 The Manual DNA sequencing reaction

a) Denaturing of double-stranded plasmid template

The following reaction was assembled:

8µl plasmid template DNA- of miniprep quality (~500ng)
2µl dH_2O
1µl 2M NaOH/2mM EDTA

After incubation at 37°C in a water bath for 30 minutes, the reaction was stopped by the addition of 0.1 volumes of 3M NaOAc (pH4.5-5.5). The denatured DNA was precipitated by adding 2.5 volumes 100% EtOH, and then incubated at -70°C for 15 minutes. After being centrifuged at 12 000 rpm, room temperature, for 10 minutes, the pellet was air-dried and resuspended in 7µl dH_2O

b) Annealing of template and primer

The reaction for the annealing of the primer to the template was assembled as follows:

7µl denatured DNA template
2µl Sequenase™ 5x reaction buffer
1µl primer (M13+) at 1.5pmol/µl

The total volume of 10 µl was incubated on a 65°C heating block for 10 minutes, followed by incubation in a 37°C water bath for 10 minutes, and then left at room temperature for 10 minutes. The reaction was spun briefly to collect the contents of the tube, and then placed on ice until required.

Four tubes with each respective termination mix (ddATP, ddCTP, ddGTP and ddTTP) were prepared, clearly marked and left at room temperature until use.

c) Radiolabelling of DNA

The reaction for the labelling of the DNA was assembled as follows:

1µl 0.1M DTT
2µl diluted labelling mix
0.5µl [α-^{35}S] dATP
0.25µl Polymerase,
and dH_2O to a total volume of 5.5µl

the above was added to the 10µl of template DNA mix, and incubated for four minutes at room temperature.

d) Termination of the reaction

For the termination reaction, 3.5µl of the radiolabelled DNA was added to each of the respective termination tubes, and incubated for five minutes at 37°C. 4µl of termination stop reagent from the kit was added to each tube to terminate the reactions, and tubes were stored at -20°C overnight, until use. Prior to loading, the samples were boiled for five minutes, and the tubes put onto ice.

e) Gel assembly and electrophoresis

A six percent sequencing gel was prepared as follows: 42g of urea, 15ml of 40% acrylamide, and 10ml of 10X TBE was dissolved in 25ml of dH_20. The volume was adjusted to 100ml with dH_20. 100ml of TEMED and 200ml of freshly prepared 10% APS was added. The gel was injected between thoroughly cleaned and assembled glass plates, and allowed to polymerise overnight.

BIO RAD Sequi-Gen® GT Sequencing Cell apparatus was assembled as per the instruction manual and the buffer tank filled with 0.5X TBE. Stop solution was loaded in some of the wells, and the pre-run at 2000V was started and left for one hour and 40 minutes. 3µl of each sample was then loaded into the wells in the order A-C-T-G, and electrophoresed at 2000V for two hours before a further 3µl volume was loaded and allowed to run for a further two hours. The gel was removed, dried in a Savant slab gel drier (SGD5040) for one hour, and placed in an autoradiography cassette. In a darkroom, film was placed into the cassette, and left overnight for exposure. To develop the autoradiography film, it was taken from the cassette and dipped into the developer for five minutes, before being washed with dH_2O. After it was dipped in fixer for five minutes it was rewashed in dH_2O and hung out to dry. DNA sequence was read from the bottom of the gel upwards, from left to right, in the order A-C-T-G.

2.3 RESULTS AND DISCUSSION

Five LSDV *PstI* fragments between 2400 and 8500 bp in size had to be sub-cloned into fragments small enough to be sequenced. Besides restriction maps for *HindIII*, *PstI*, *AvaI* and *SalI* (Perlman, MSc dissertation, UCT 1993- unpublished data), the locations of RE sites in the LSDV genome were unknown. In order to determine which restriction enzymes cut LSDV DNA at regular intervals, for the generation of fragments with desired sizes, restriction digests with a variety of commonly used enzymes had to be performed on each *PstI*-fragment plasmid. The restriction patterns that were generated were used to compile RE maps for each *PstI* fragment. Four-base cutter enzymes were not used, as they cut more frequently and therefore generate too many smaller fragments to be of use for sub-cloning. The results of the screening of the *PstI* clones for restriction sites follows in Table 4.

Table 4: Summary of restriction endonuclease (RE) sites in LSDV *PstI* clones -E, -M, -K, -F and -G (gels not shown). RE sites occuring in the MCS of pGEM3Z and/ or pMOSBlue are boldfaced. No value indicates that the particular clone was not screened with that RE.

		PstI clone:				
RE	Manufacturer	M	K	E	F	G
AccI	Boehringer Mann.	0	0	0	3	3/4
AseI	Amersham	3	>3	>6	>6	>6
BamHI	Amersham	0	0	1	0	0
BclI	NEB	1				
BglII	Roche	1	0			0
ClaI	Sigma		2			
DpnI	Amersham	>6				
DraI	Amersham	3	4			
EcoRI	Promega	0	0	0	3	0
EcoRV	Amersham	1				
EcoT22I	Amersham	0	0			
HindIII	Promega	1	1	3	3	1
HinfI	Amersham		>1			
KpnI	Amersham	0	0	0	0	0
MluI	Amersham	0	0			0
NcoI	NEB	0	0			0
NdeI	Amersham	0	0			
NheI	Amersham		0			0
NotI	Amersham	0				0
NsiI	Boehringer Mann.	1	0			0
PvuI	Boehringer Mann.	2	0			0
PvuII	Amersham	1	0			0
RsrII	Boehringer Mann.	0	0			
SacI	Roche	0	0	0	0	0
SalI	Amersham	0	0	0	1	1
SfuI	Boehringer Mann.	0				0
SmaI	Amersham	0	0	0	0	0
SpeI	Boehringer Mann.	0	0			
SphI	Amersham	0	0	0	1	
StyI	Boehringer Mann.	0	0			
TaqI	Boehringer Mann.	>6				
XbaI	Amersham	0	0	0	2	0
XhoI	Amersham	0				
XmnI	NEB		1			

2.3.1 The restriction endonuclease mapping and sub-cloning of the *PstI*-M clone

The *PstI* M clone (~ 2 400 bp) was mapped first, due to its small size. Double digests of the parental *PstI*-M plasmid with *AseI/EcoRI*, *AseI/HindIII* and *AseI/PstI* confirmed that the two internal fragments of approximately 480 and 350 bp were flanked by two terminal fragments of approximately 700 bp in size (1a, Fig. 6 on opposite page). The fragment adjacent to the *EcoRI* site in pGEM3Z was named M700, followed by M480, M350 and the opposite terminal fragment, M650.

Neither the pGEM3Z nor the pMOSBlue cloning vectors contain an *AseI* site in their multiple cloning sites, but pGEM3Z contains three internal *AseI* sites (1a, Fig. 6). Digestion with *AseI* produces a TA- overhang which is compatible with the AT-overhang produced by cleavage with *NdeI* in the pMOSBlue multiple cloning site. The two internal fragments, M480 and M350 (1a, Fig. 6), have *AseI* termini and were separately cloned into a pMOSBlue vector with de-phosphorylated *NdeI* ends. Thereby, recombinant plasmid (p) M350 (1d, Fig. 6) and pM480 (1c, Fig. 6) were obtained. In order to clone fragments M650 and M700 directionally, the *PstI*-M clone was double digested with *PstI* and *AseI*. Fragments M650 and M700 are very similar in size, and co-migrated as a single band. Therefore, it was necessary to re-isolate each 700 bp band by two different RE reactions: an *AseI/HindIII* digest to isolate fragment M650, and an *AseI/EcoRI* digest to isolate fragment M700. These were then cloned directionally into *NdeI/HindIII* (1e, Fig. 6) and *NdeI/EcoRI* (1b, Fig. 6) vectors, respectively. These recombinants were confirmed by excising the fragments from the plasmids, and electrophoresing them alongside parental *PstI*-M plasmid digested once again with *AseI/PstI*. Religated *AseI/NdeI* sites have a CATAT sequence that is recognised by neither enzyme, and it was necessary to utilise sites flanking these to excise the inserts. The *PstI*-M fragments are of an appropriate size for automated DNA sequencing (<1 200 bp), and it was not necessary to sub-clone any of them into smaller fragments.

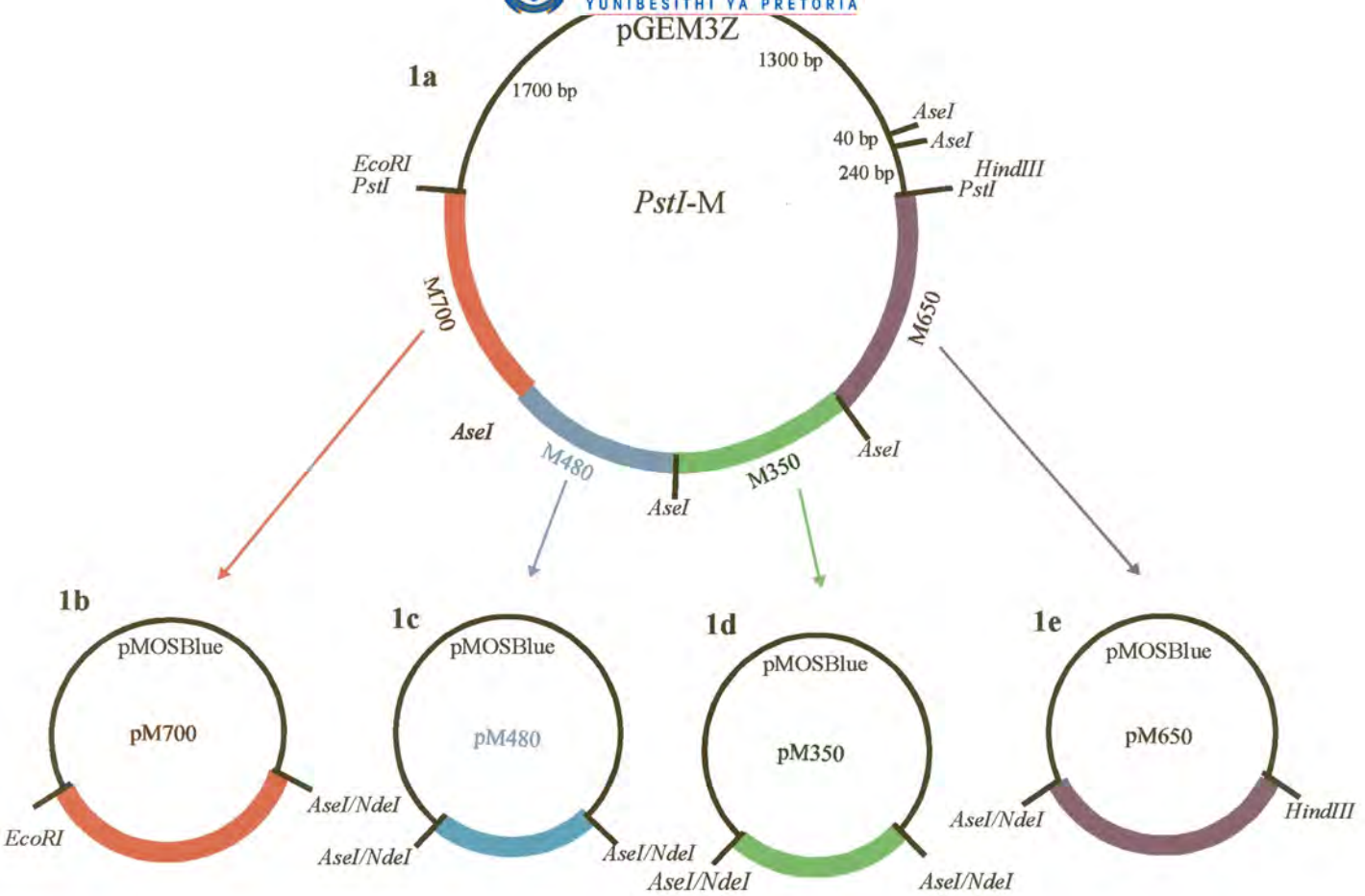

Figure 6: Restriction endonuclease mapping and sub-cloning of the *PstI*-M clone (not drawn to scale)

2.3.2 The restriction endonuclease mapping and sub-cloning of the *PstI*-K clone

The *PstI*- K clone insert is roughly 3300 bp in size. Double digests with *PstI/AseI*, *PstI/HindIII*, and *HindIII/AseI* confirmed the location of a terminal fragment of about 480 bp separated by a *HindIII* site from a fragment of about 650 bp (2a, Fig. 7 on opposite page). An *AseI* digest-generated fragment of approximately 500 bp was found to be adjacent to the 650 bp fragment. On the opposite side of the MCS, a 1200 bp fragment, named K1200, was found to be flanked by *AseI* sites. One of these was situated very close to the pGEM MCS (the distance between the *PstI* site and the *AseI* site was later determined to be eight basepairs by DNA sequencing). Fragment K1200 contains no other internal restriction sites for commonly used enzymes (Table 4). The three smaller fragments were named K480, K650 and K500, from the MCS *HindIII* side.

The cloning of fragments K480, K650 and K500 into vectors proceeded as follows: a digest of the *PstI*-K clone with *HindIII* isolated fragment K480 which was cloned into a vector with *HindIII* termini (2e, Fig. 7); a double digest of the *PstI*-K plasmid *AseI/HindIII* isolated fragment K650 which was cloned directionally into an *NdeI/HindIII* vector (2d, Fig. 7); and an *AseI* digestion isolated the K500 fragment which was cloned into a dephosphorylated *NdeI*-linearised vector (2c, Fig. 7). Recombinant plasmids were sequenced and analysed. Once the LSDV sequences were assembled in a linear order, it became evident that a small fragment was missing from the *AseI* junction of pK650 and pK500 (2a, Fig. 7). The fragment was not visible on agarose gel, and yet there was a gap in the peptide sequence of the predicted ORF. A *ClaI* site located within 15 nt of the *AseI* site of pK500 was identified by sequence analysis (Chapter III), and the distance between this site and the *HindIII* site of the vector was estimated to be approximately 700 bp in size. The K1200 fragment had not been cloned or sequenced at that time, and therefore the possibility other *ClaI* sites occuring within the *PstI*-K plasmid could not be excluded. A double digest with *ClaI/HindIII* produced a fragment which migrated at the appropriate size. The fragment was purified and cloned into pGEM3Z vector linearised with *HindIII* and *AccI*. Digestion with *AccI* produces sticky ends compatible with *ClaI*. Recombinant clones were sequenced and the missing fragment sequence, flanked by *AseI* sites, was discovered to be 71 bp in length (AppendixII).

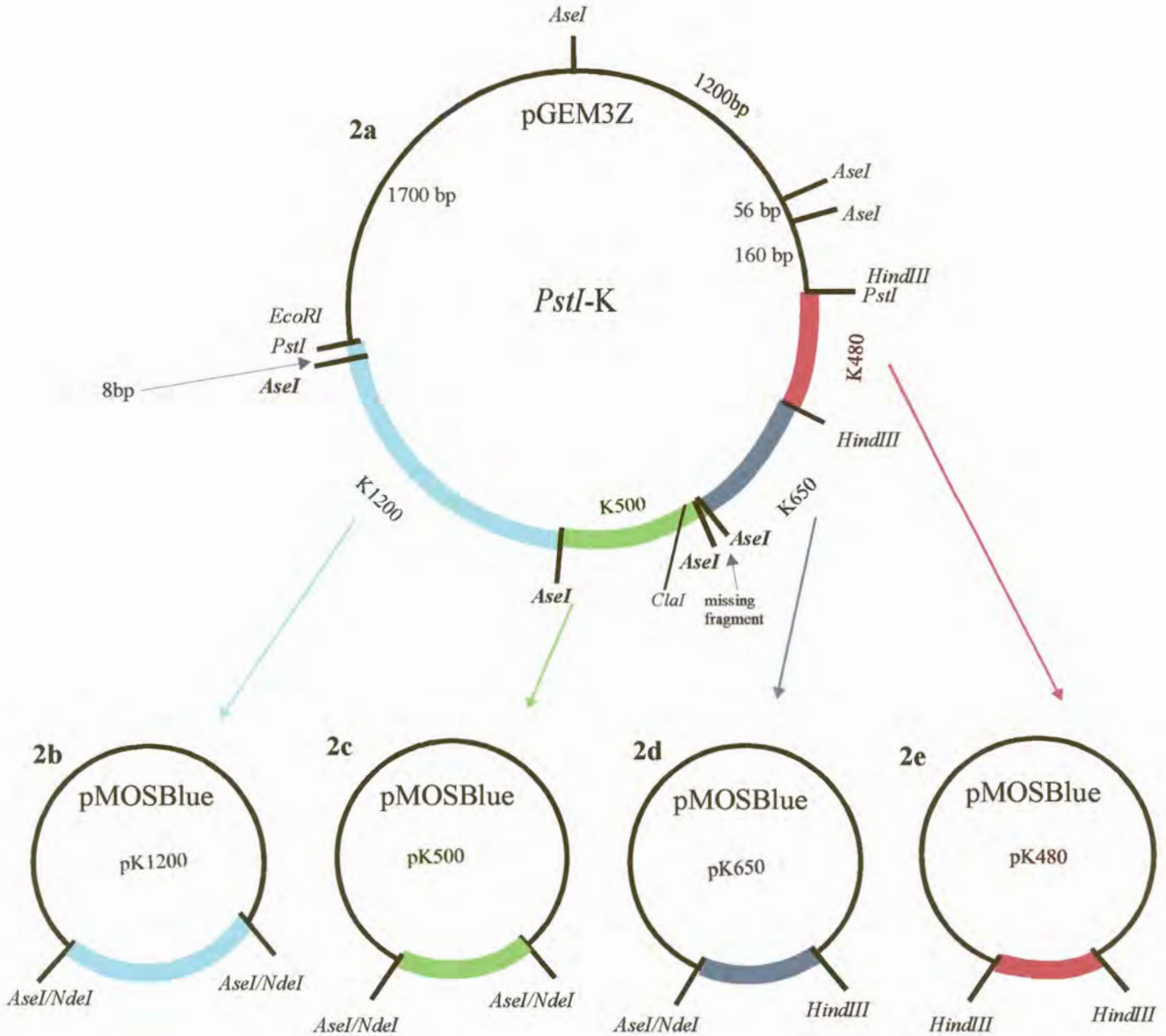

Figure 7: Restriction endonuclease mapping and sub-cloning of the *PstI*-K clone (not drawn to scale)

The K1200 fragment cloning strategy was complicated by several factors: the pGEM3Z vector has three internal *AseI* sites: positions 224, 280 and 1512, which (if the MCS is also cleaved) yield 160-, 56-, 1200- and 1700 bp fragments (2a, Fig. 7). Firstly, fragment K1200 and the 1200 bp vector fragment co-migrate. It was not possible to separate the two electrophoretically, even on a 2 % agarose gel. To avoid this problem, the entire *PstI*-K insert could first be excised, before being digested in isolation with *AseI*. However, because pGEM3Z is 3200 bp in size and the *PstI*-K fragment is 3300 bp in size, they co-migrate. Again, separation by electrophoresis on a 2% agarose gel was unsuccessful, and this strategy was abandoned.

To discriminate between the K1200 fragment and the vector 1200 bp fragment, the pGEM3Z sequence was examined for unique restriction sites. An enzyme was required that would not digest the LSDV DNA, but would cleave the pGEM3Z 1200 bp fragment, allowing the K1200 fragment to be recovered intact. According to a MAPPLOT (GCG) of six-base cutters, the following enzymes with sites in the pGEM3Z 1200 bp fragment were listed: *AlwNI, ApaLI, Bce83I, BsaWI, BsiEI, BsiHKAI, Bsp1286I, BssSI, EciI, HaeI, HaeII, HgiEII, MmeI, MspAIL, SfcI* and *Tth111II*, but since these enzymes are uncommon and expensive, a new strategy was sought. Instead of trying to eliminate the pGEM3Z 1200 bp fragment in order to isolate the K1200 fragment directly, I attempted to fragment the entire pGEM3Z vector in order to recover the entire *PstI*-K fragment (the fragments co-migrate due to similar size). If this was possible, then *AseI* digestions could be done on the purified, isolated *PstI*-K fragment. Restriction endonucleases which only cut in the pGEM3Z vector were sought. Of the commonly available enzymes, *XmnI* cuts pGEM3Z at position 1937, and *NdeI* cuts at position 2509. A double RE digestion was carried out, using *PstI* and *XmnI*, or *NdeI*. While the LSDV fragment would remain undigested at 3300 bp, the pGEM3Z vector would be cleaved into two smaller fragments, allowing the LSDV fragment to be purified from the agarose gel. A restriction endonuclease digest unfortunately proved that there was an *XmnI* site in the *PstI*-K insert, and the approach was abandoned. *Pst/NdeI* digests, followed by purification and *AseI* digests were attempted, but too much DNA was lost in the successive purification steps.

If the co-migrating 1200 bp fragments were "shotgun" cloned, there was possibly a way to distinguish between a pGEM3Z/pMOSBlue recombinant, and a pMOSBlue/LSDV recombinant. The co-migrating 1200 bp fragments with *AseI* termini were excised from the agarose, purified, and cloned into a dephosphorylated *NdeI*-linearised pMOS vector. Thirteen of twenty-four colonies picked were confirmed as recombinants by digestion with *EcoRI/HindIII*. The remaining colonies contained vector that had re-ligated. At this stage, it became evident that *DraI* is an enzyme that frequently cuts the LSDV genome. As a consequence it was decided to digest these recombinants with *DraI* in order to distinguish a possible LSDV-specific clone from a vector-specific clone. Before this was attempted, the *DraI* sites in the *PstI*-K fragment (particularly in the K1200 fragment), had to be mapped. In the subsequent *AseI/DraI* double restriction digests of the *PstI*-K plasmid, the K1200 fragment remained intact, and therefore lacked *DraI* sites. Secondly, the *DraI* map of the vector-specific 1200bp recombinant had to be determined. This was accomplished by RE mapping, and the maps of both orientations are given in Figure 8. The pGEM3Z 1200 bp fragment had one *DraI* site, which was used to distinguish between vector-specific and LSDV-specific recombinants.

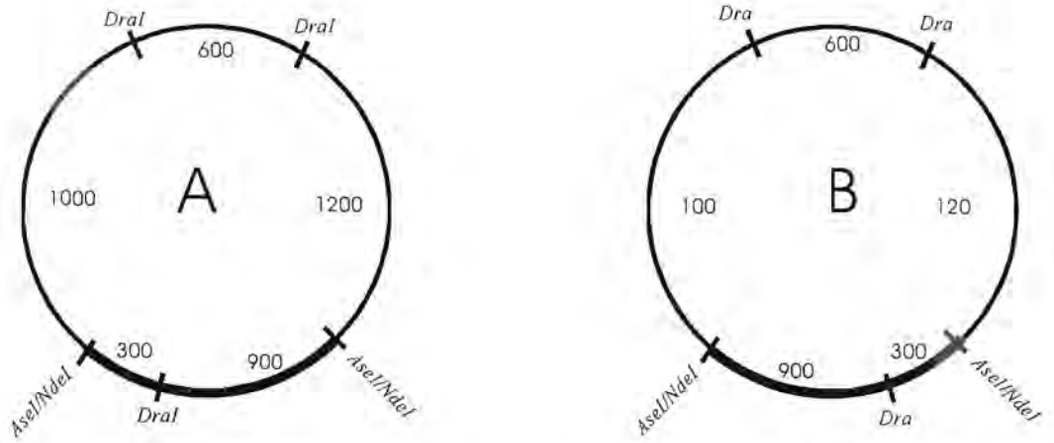

Figure 8: *DraI* RE maps of pGEM/pMOS recombinants in insert orientations A and B. Fragment sizes are given in base pairs.

Bands expected with *DraI* digests (Fig. 8):

A: 2 100 bp + 1 300 bp + 600 bp (4 000 bp)

B: 1 900 bp + 1 500 bp + 600 bp (4 000 bp)

In orientation B, a vector-recombinant, when cleaved with *DraI,* would yield 1500-, 1900- and 600 bp fragments. In the opposite orientation (A), the fragments after *DraI* digestion would be 1300- 2100- and 600 bp in size. Alternatively, *DraI* digestion of an LSDV-specific recombinant would produce the common 600 bp band (from the pMOS vector) and the remainder of the recombinant plasmid (3400 bp) which includes the LSDV 1200 bp fragment, undigested by *DraI*. The thirteen recombinant plasmids were therefore digested with *DraI*, and the results are presented in the agarose gel of Figure 9, which follows:

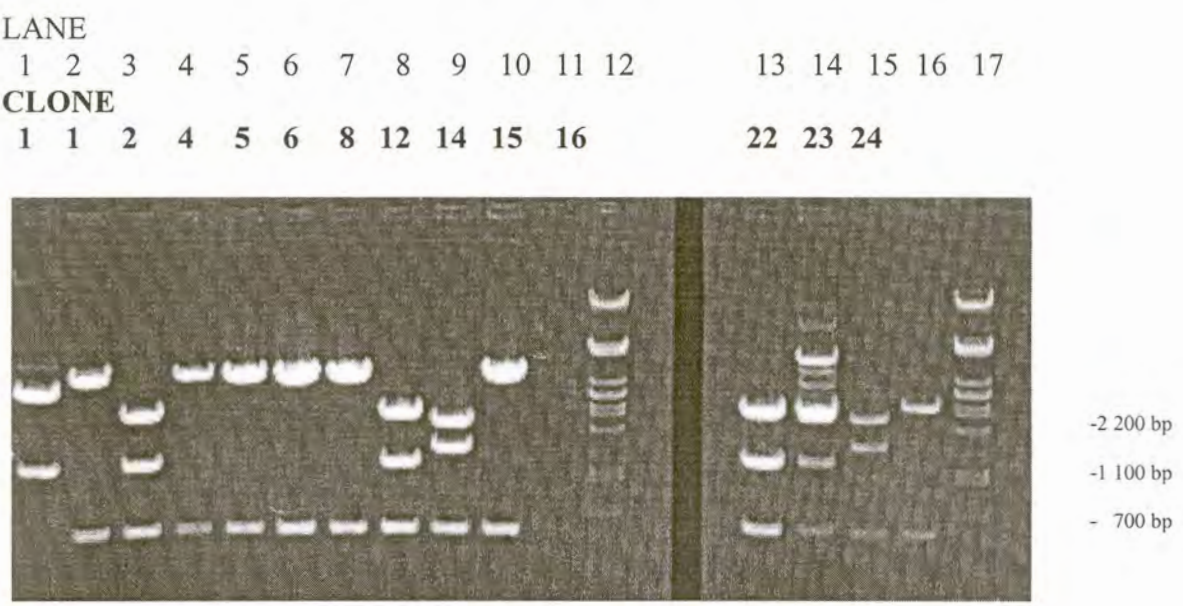

Figure 9: *DraI* digests of recombinant 1 200 bp clones, illustrating pK1200 (clones 1, 4, 5, 6, 8 and 15), recombinant pMOS/pGEM in orientation A (clones 2, 12, 22 and 23) and pMOS/pGEM recombinants in orientation B (clones 14 and 24). Lane 1 is an *EcoRI/HindIII* digest of clone 1, to verify that the sizes of the insert is 1 200 bp; lane 16 is re-ligated pMOS; lanes 12 and 17 are lambda *PstI* DNA markers, with sizes indicated on the right.

All three species were therefore present, and pK1200 (2b, Fig. 7), represented by clones 1, 4, 5, 6, 8 and 15 had therefore been successfully constructed. This technique of *DraI* profiling was consequently used on the *PstI*-E, -F and -G clones, where several co-migrating fragments were of LSDV origin.

2.3.3 The restriction endonuclease mapping and sub-cloning of the *PstI* -E clone

The *PstI*-E clone which is approximately 7 000 bp in size is one of the larger plasmids. The fragment is flanked by an *EcoRI* site, and a *PstI* site, and this *PstI* fragment is the closest to the left hand terminus of the LSDV genome (see Fig. 5 for genomic positions). Many *AseI* sites were expected to occur in the LSDV DNA, and due to this, the plasmid was first sub-cloned into more manageable sizes. Three internal *HindIII* sites were mapped, dividing the *PstI*-E insert into the following fragments: a terminal 800 bp-fragment, 2700 bp fragment, 1300 bp fragment and a 2150 bp terminal fragment (3a, Fig. 10 on opposite page). They were named fragments E4000 (referring to the size of the re-ligated plasmid) E2700, E1300 and E2150, respectively. At this stage, the orders and orientations of E2700 and E1300 relative to each other were not known. This was only determined after sequencing and analysis (Chapter III). The 800 bp terminal fragment lies adjacent to the *EcoRI* site, therefore a single digest with *HindIII* excises fragments E2150, E1300 and E2700 and a 4000 bp fragment containing the original pGEM3Z vector (3200 bp) plus the 800 bp LSDV fragment. All of these were excised from the gel and purified, and the 4000 bp fragment was allowed to re-ligate, hence the naming of the 800 bp fragment (3e, Fig. 10). The E2150, E1300 and E2700 fragments were cloned into dephosphorylated *HindIII*-linearised vectors (3b, 3d and 3c, Fig. 10). Recombinants were confirmed by restriction endonuclease excision of the fragments. pE4000 did not need to be sub-cloned, but the remaining clones were still too large to obtain overlapping sequences from forward and reverse sequencing. Double restriction digests of pE1300 with *HindIII/AseI* located a 700 bp fragment flanked by *AseI* sites (3d, Fig. 10). This fragment is situated approximately 200- and 280 bp from either side of the MCS, respectively. The 700 bp fragment was excised with *AseI* and cloned into a dephosphorylated *NdeI*-linearised vector (3h, Fig. 10). The flanking regions of

approximately 200 and 280 bp in size were not cloned, and these sequences were determined by sequencing in the forward and reverse directions from the parental pE1300. This 700 bp-insert plasmid was named pE1300/700.

Two internal and co-migrating 700 bp fragments, generated by *AseI* were identified in pE2150 (3b, Fig. 10). The two terminal regions are approximately 200- and 250 bp respectively, and are also uncloned. Their sequences were determined from forward and reverse sequencing reactions of the parental E2150 plasmid. The two 700 bp fragments co-migrate, and were shotgun cloned. Recombinant 700 bp-insert clones were distinguished once again by *DraI* profiling: pE2150/702 (3f, Fig. 10) contains no internal *DraI* sites, and yields a 3500 bp band plus the 600 bp band, while pE2150/701 (3g, Fig. 10) has two internal *DraI* sites, producing five bands of 250 bp, 400 bp, 600 bp, 1000 bp, and 1450 bp, approximately (gel not shown). The recombinants were named pE2150/701 and pE2150/702 for convenience. Their orders and orientations relative to each other and the terminal flanking regions were not known at this stage.

pE2700 could be sub-cloned into four smaller fragments after digestion with *AseI* (3c, Fig. 10): E2700/800 (3i, Fig. 10) and E2700/1450 (3l, Fig. 10) were flanking fragments with *HindIII/AseI* termini. The two smaller fragments E2700/450 (3j, Fig. 10) and E2700/300 (3k, Fig. 10) were internal. Recombinants were confirmed by restriction endonuclease digestion.

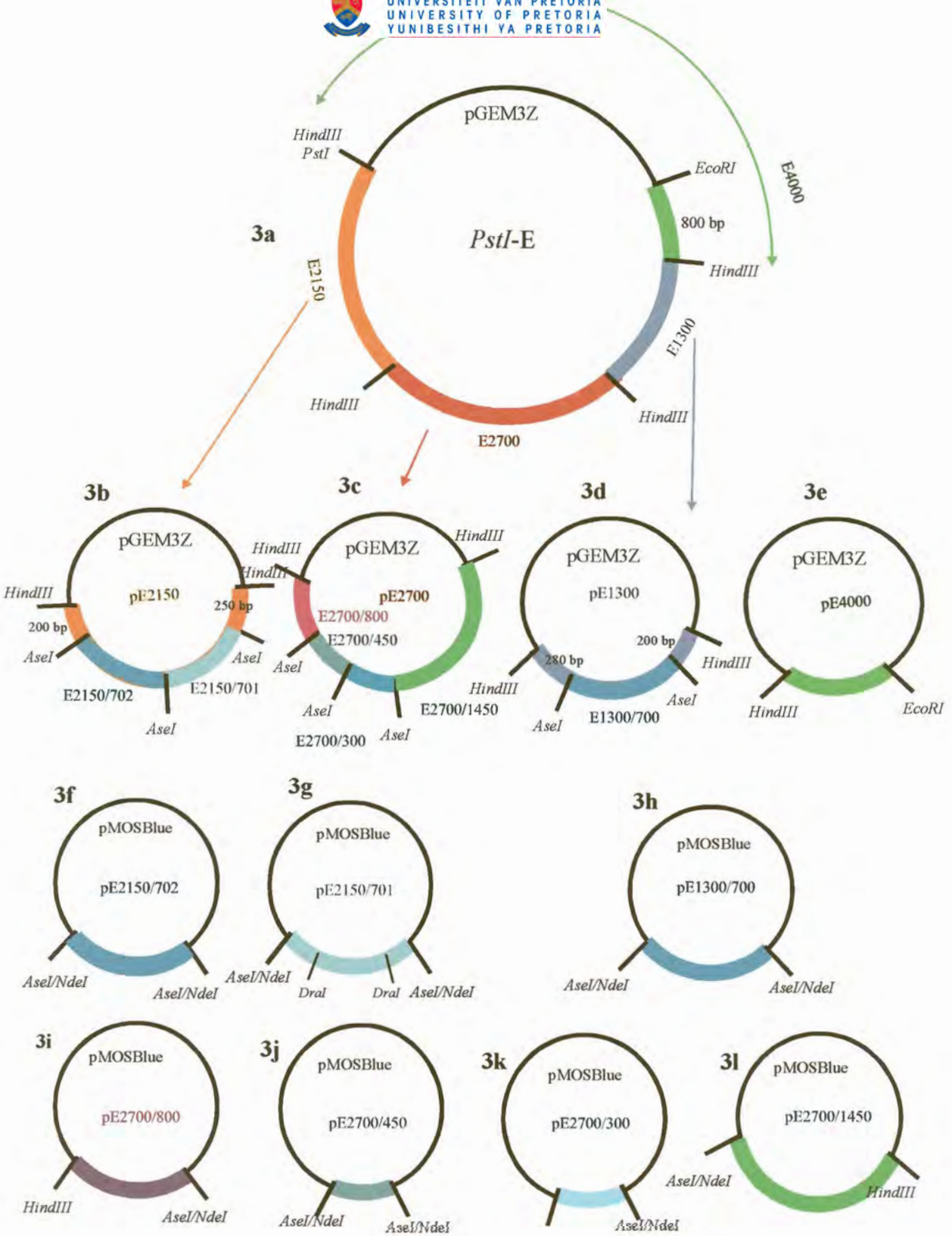

Figure 10: Restriction endonuclease mapping and sub-cloning of the *Pst*-E clone (not drawn to scale)

2.3.4 The restriction endonuclease mapping and sub-cloning of *PstI*-F clone

The *PstI*-F fragment is the largest of the five selected, being 8500 bp in size. The initial *AseI* digest indicated that there were at least twelve internal *AseI* sites in the fragment, and many of these bands co-migrated. To disperse these sites, the *PstI*-F plasmid first had to be sub-cloned into smaller fragments using commonly-used enzymes, before the smaller *AseI*-generated sub-clones were constructed. *PstI*-F is the only fragment of the five that contains *XbaI* sites (Table 4), and mapping revealed two internal sites which divides *PstI*-F into a terminal 1700 bp fragment (pF1700), an internal 2900 bp fragment (pF2900) and a 3400 bp fragment (pF6600) (4a, Fig. 11 on opposite page). After an *XbaI* digest of the *PstI*-F plasmid, the 1700- and 2900 bp bands were purified and cloned into a dephosphorylated *XbaI*-linearised vector (4b and 4c, Fig. 11). The 6600 bp band, which still contained the original "parent" pGEM vector was purified and allowed to self-ligate in an overnight ligation reaction (4d, Fig. 11). All three *PstI*-F sub-clones were confirmed by restriction endonuclease digestion, and separately mapped with double digests of *AseI/XbaI*.

Digestion with *AseI* and *XbaI* cleaved pF1700 (4b, Fig. 11) into two fragments of ~1300 bp and 450 bp, which were cloned to produce pF1700/1300 (4e, Fig. 11), and pF1700/450 (4f, Fig. 11). pF1700/450 borders on pF2900.

Following double digests of pF2900 (4c, Fig. 11) with *XbaI/AseI*, five fragments could be mapped: terminal 950- and 650 bp fragments flanked internal bands of approximately 1000 bp, and two in the region of 300- to 350 bp. The fragments cloned into plasmids and were named pF2900/950 (4g, Fig. 11), pF2900/650 (4k, Fig 11), pF2900/1000 (4j, Fig. 11), pF2900x (for the larger of the two smaller fragments) (4I, Fig. 11) and pF2900y (4h, Fig. 11). All five sub-clones were confirmed by restriction endonuclease digestion.

pF6600 (4d, Fig. 11) could be cleaved into six fragments of estimated sizes in descending order: 1000 bp, 800 bp, 470 bp, ~400 bp, ~380 bp and 300 bp. They were cloned into plasmids and named pF6600/1000 (4q, Fig. 11), pF6600/800 (4o, Fig. 11), pF6600/470 (4n, Fig. 11), pF6600/x (4m, Fig. 11), pF6600/y (4p, Fig. 11) and pF6600/300 (4l, Fig.

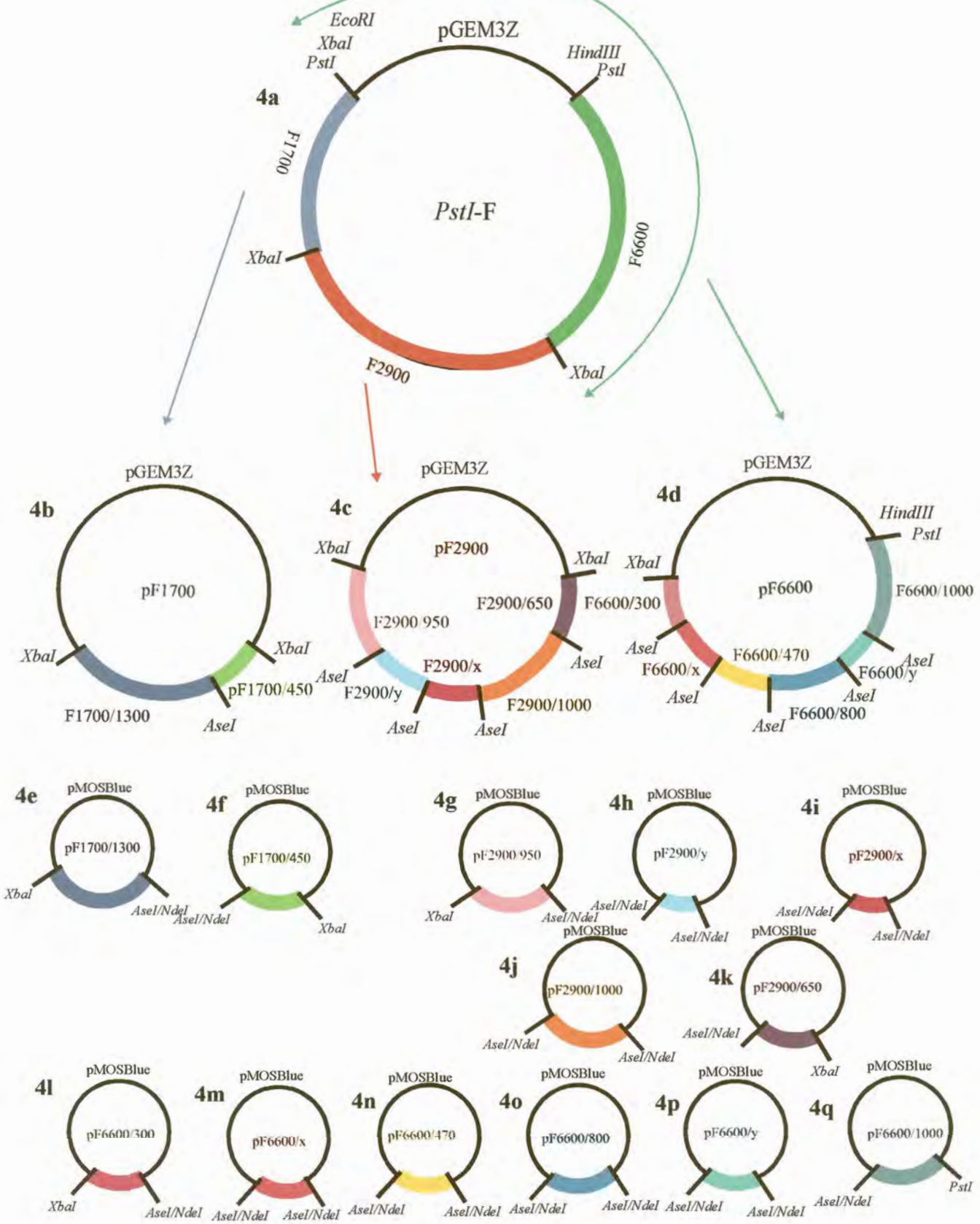

Figure 11: Restriction endonuclease mapping and sub-cloning of the *PstI*-F clone (not drawn to scale)

11), respectively. pF6600/300 contains the terminal fragment bordering on pF2900, and pF6600/1000 contains the opposite terminal fragment, although the orders and orientations relative to each other of the internal fragments were not known at this stage.

2.3.5 The restriction endonuclease mapping and sub-cloning of *PstI*-G clone

RE digestions of the *PstI*-G clone showed it to contain numerous internal *AseI* sites, the majority producing fragments of between 500- and 700 bp in size. The possibility therefore existed that if fragments migrating so close to each other were to be individually excised from an agarose gel, some cross-contamination would result. There is one *SalI* and one *HindIII* site each within the *PstI*-G fragment, where *HindIII* was mapped to cut within 500 bp of the terminus, and *SalI* cut closer to the center, producing a 1400 bp band, and a 6800 bp band (5a, Fig. 12 on opposite page). These fragments were still too large to sequence, but at least some of the *AseI* sites were distributed between them. Although *AccI* does appear to cut about four times in *PstI*-G clone (not shown), it produces degenerate ends and was not used. The *PstI*-G sub-clones were obtained by digesting the plasmid with *SalI*, isolating the 6800 bp band and cloning it into dephosphorylated *SalI*-linearised vector, and allowing the purified 4600 bp fragment (comprising 1400 bp of LSDV plus 3200 bp of pGEM3Z) to self-ligate. These two plasmids were named pG6800 (5c, Fig. 12) and pG4600 (5b, Fig. 12), respectively.

The construct pG4600 contains an internal *AseI*-ended fragment of 650 bp, which was cloned, and named pG4600/650 (5d, Fig. 12). The flanking regions were not cloned, but their sequences were determined by forward and reverse reactions from the parental pG4600 clone. *AseI* was again used in the sub-cloning of pG6800 (5c, Fig. 12). Single 800 bp, 700 bp, 550 bp and 300 bp fragments produced by digestion were isolated from agarose gel, purified and cloned. These clones were confirmed to contain inserts of the correct sizes, and named pG6800/800 (5h, Fig. 12), pG6800/700 (5i, Fig. 12), pG6800/550 (5g, Fig. 12) and pG6800/300 (5e, Fig. 12), respectively. At the approximate position of 500 bp, two co-migrating bands were excised and purified, and "shotgun"

cloned into a dephosphorylated *NdeI*-linearised vector. Once recombinants were identified, *DraI* digests were carried out, and, two different species were identified (bearing in mind that four possible profiles are possible if each insert is able to ligate in both orientations). pG6800/501 (5j, Fig. 12) has no *DraI* sites, while pG6800/502 (5l, Fig. 12) has one, close to its terminus. Similarly, pG6800/601 (5f, Fig. 12) and pG6800/602 (5k, Fig. 12) were cloned from the co-migrating 600 bp band of a pG6800 *AseI* digest. From the intensity of the band, and size estimations from various RE maps of the *PstI*-G clone, I expected to find three different fragments. A third, unique 600 bp *AseI* fragment was never obtained in this manner, although 119 recombinants were profiled. The absence of the third 600-bp fragment was revealed later on, once the assembly of the *PstI*-G fragment sequences (Chapter III) was underway.

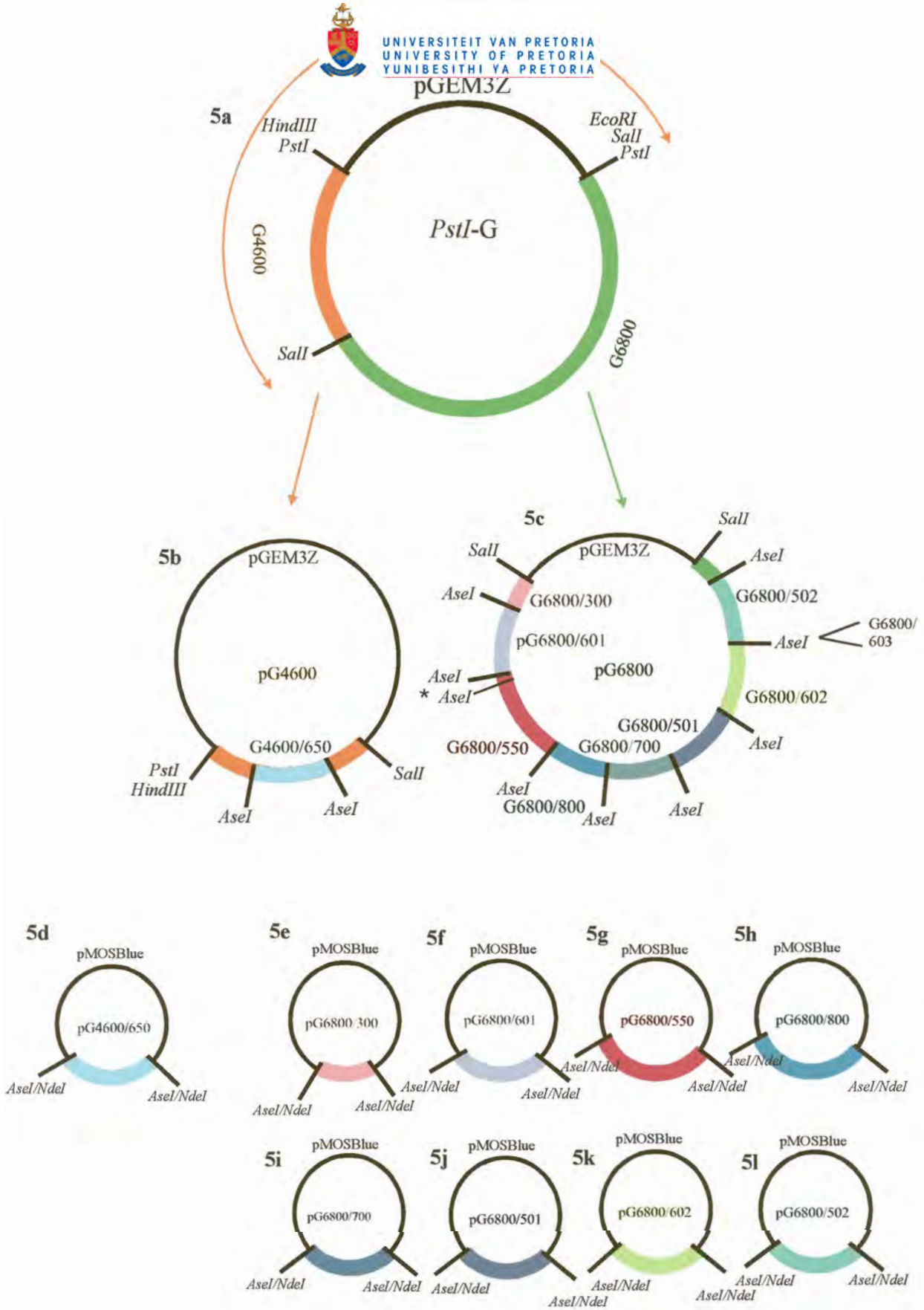

Figure 12: Restriction endonuclease mapping and sub-cloning of the *PstI*-G clone (not drawn to scale)

2.3.6 The mapping and cloning of a fragment in the *PstI* -G clone spanning an *AseI* junction

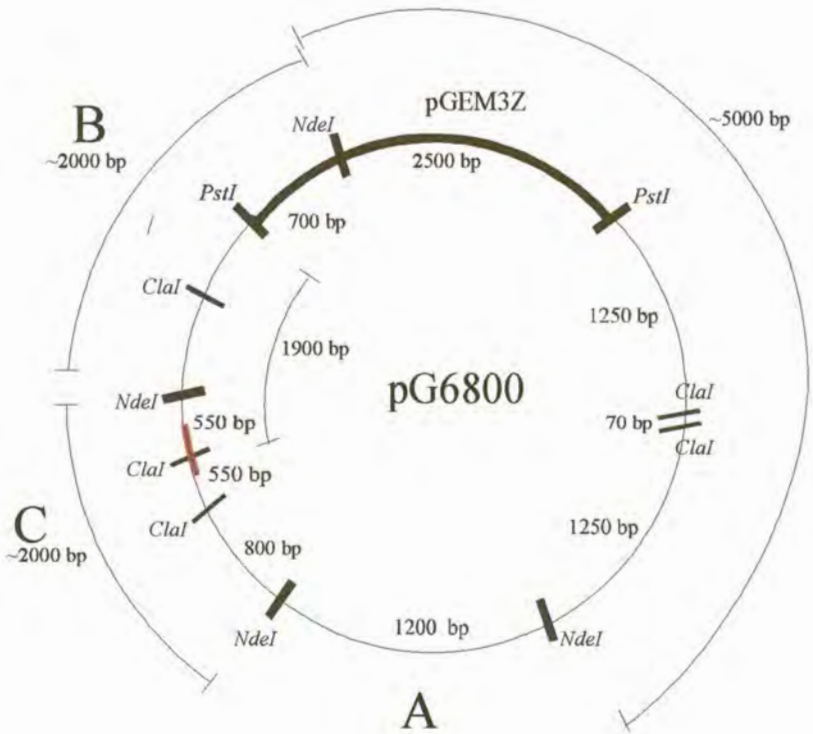

Figure 13: Restriction endonuclease map of clone BC6. The target region is indicated as a red highlighted portion spanning a *ClaI* site.

The sequencing of the pG6800 constructs (5c, Fig. 12), and the analysis of the ORFs they encode, is described in the next chapter. The fragments were assembled into the correct orders and orientations, based on the identified ORFs. By comparison to various other poxviruses sequenced, the absence of the amino terminus of the DNA topoisomerase ORF, and a complete VAR H7 ORF homolog was evident. The missing piece occurred between constructs pG6800/550 and pG6800/601 (5c, Fig. 12, indicated by *). Instead of sequencing the missing fragment by primer-walking sequencing, it was decided to clone a DNA fragment that spanned the *AseI* junction between fragment G6800/601 and fragment G6800/550. GCG MAP analysis (Chapter III) of the partially assembled pG6800 (5c, Fig. 12) sequence identified an *NdeI* RE site located approximately 240 bp upstream of the *AseI* junction between pG6800/550 and the adjacent fragment. Further

RE mapping with both *NdeI* and *ClaI* on pG6800 led to the restriction map of Fig. 13. The fragment spanning the *AseI* junction, including the missing fragment (which was suspected to be between 250 and 500 bp in size), was cloned: RE digestion of pG6800 with *NdeI* yielded four bands (Fig. 13). Co-migrating bands B+C were excised, purified and "shotgun" cloned into a pMOS vector. Clone B recombinants were identified by the additional *PstI* site (determined by RE digests) contributed by the pGEM vector sequence in the fragment. The correct recombinant, containing band C, was identified by the lack of this additional *PstI* site. The clone was named pBC6, and the insert found to be 490 bp in size by sequence analysis (Appendix II).

2.3.7 Sequencing

2.3.7.1 Automated DNA sequencing

Two representative electropherograms are presented in Appendix I

Example 1: G6800/800 REV (sequencing of a cloned fragment with vector sequence primers)

Example 2: F66A (primer walking sequencing)

Forward and reverse automated DNA sequences were edited in Sequence Navigator (Apple Macintosh), which is described in Chapter III: Assembly and Analysis.

2.3.7.2 Primer walking by automated DNA sequencing

EPG output files from the Apple Macintosh were visualized in Chromas (V.1.43, C. McCarthy), which runs in Windows off a PC platform. The DNA sequences were examined for the regions to be inserted, flanked by *AseI* sites. Being relatively short, the sequences were entered manually into the assembled *PstI* fragment sequences, using a text editor program (GCG). The sequences from primer walking sequencing are presented in Appendix II as short nucleotide sequences each flanked by two *AseI* sequences and the name of the adjacent fragment on either side.

The most noticeable feature of an LSDV electropherogram (EPG), is the high frequency of A and T nucleotides. M13 primers were used, which prime further upstream than the T7 and SP6 promotor primers, making it possible to read about 80 bp of the bacterial vector sequence upstream of the LSDV DNA sequence. As shown in the EPG of G6800/800 (Appendix I, Example 1), the peaks continue to be high and sharp for at least 400 base pairs thereafter, and the first N's, or uncalled bases, usually start appearing around 500 bp, although the sequence is still readable and editing possible. Primer walking sequences, using oligonucleotides designed specifically for LSDV sequences, start within 30 bp of the primer annealing site (Appendix I, Example 2- F66A).

It was observed that the read length of quality DNA sequence for quarter- and half-reactions are indistinguishable. Full reactions, with the manufacturer-recommended clean-up, is reported to yield in excess of 700 bp before the first uncalled base is seen,

and those for quarter reactions performed on LSDV were at least 600 bp in length. When designing primers for primer walking, the priming site was chosen to lie at least 50 bp upstream of the target junction site, to compensate for uncalled bases very close to the priming site. Usually, sequences within 30 bp of the priming site could be read. It was also observed that an additional 70% EtOH wash step improved the read length, and it was possible to read a few base pairs closer to the priming site. The recommended annealing temperature for cycle sequencing is 50°C, but the T_ms of the primers used varied from as low as 46.27°C for E2150A to 56.27°C for G68/550P. Despite the deviation from the recommended T_m, good sequences were obtained for all primers.

DNA sequences from primer walking and manual sequencing (Appendix II) were edited into the assembled *PstI* fragment sequences (refer to Table 3 for priming positions) Primer F66B was designed to amplify the region which spans the *AseI* gap between pF6600/800 and pF6600/y in the *PstI*-F clone (4d, Fig. 11), and an additional two *AseI* sites were identified. Two small fragments of 65 and 15 nt were incorporated into the *PstI*-F sequence.

At the junctions sequenced by primers F66A, K12/K5P, E2150A and F29A, no insertions were found. Primers F66A and K12/K5P were designed to read sequences spanning *AseI* junctions in intergenic regions. Primer E2150A was designed to span the intergenic *HindIII* site between E2150 and E2700 (3a, Fig. 10). Primer F29A was designed to sequence a small fragment presumed missing at an *AseI* junction within the RPOL gene. No insertion was found, and the LSDV homolog is therefore accepted to have a natural deletion in this region of the gene. Primer G6800/602 was designed to read the sequence of the missing fragment between G6800/602 and G6800/502 (5c, Fig. 12). Despite the size of the insert, the run was sufficiently long to identify the aforementioned flanking fragment sequences and both *AseI* sites. The new fragment was named G6800/603 (AppendixII).

2.4 Conclusions

In order to determine the DNA sequence of the five *PstI* clones selected for this study, each *PstI*-clone (-E, -M, -K, -F and -G) had to be sub-cloned into smaller plasmids, and sequenced in both directions. Due to limited genomic data, each *PstI*-clone first had to be mapped by double digests with commonly-used restriction enzymes. The RE maps thereby generated were used to develop cloning strategies in which clones smaller than 1000 bp could be constructed and sequenced. These clones did not overlap, and therefore there was a risk of missing very small fragments on the agarose gels from which fragments were excised for cloning. Cloning strategies were often complicated by co-migrating fragments, which led to the development of a *DraI* restriction pattern profiling system. This made it possible to distinguish between different recombinants of the same size, and this approach was used successfully in the cloning of the K1200 fragment, as well as several co-migrating fragments of the large *PstI*-F and -G clones. In total, 46 recombinant plasmids were constructed, containing almost 29000 bp of non-overlapping LSDV DNA. The large majority of these clones were constructed using the restriction endonuclease *AseI*, which recognises ATTAAT sequences in the A+T-rich LSDV genome, and produces ends compatible with *NdeI* in the MCS of bacterial cloning vector pMOSBlue. Once the DNA sequences were assembled and analysed (Chapter III), primers were designed from sequences of the cloned regions, to sequence regions excluded in the sub-clones (particularly at *AseI* junctions) by primer walking sequencing.