DNA COMPARISONS OF THE TWO ORTHOPOXVIRUSES MONKEYPOX AND VARIOLA

by

NICOLA JENNIFER PARE

Submitted in fulfilment of the requirements for the degree of MASTER OF SCIENCE (Med) in the department of Medical Microbiology in the Faculty of Medicine at the University of Cape Town.

JUNE 1988

ACKNOWLEDGEMENTS

I would like to thank Prof. K.R. Dumbell for his invaluable guidance and encouragement as my supervisor. Thanks also go to Peggy McDowell for her assistance in the dark room and to everyone in the plasmid unit for their help in the lab and most beneficial discussions. May I also thank Prof. J.W. Moodie for his endless patience behind the computer and his permanent interest in my work. The help of Prof. E.H. Harley and Dr. R. Kirby in providing me with DNA programs is gratefully acknowledged. Thanks are extended to Prof. Dowdle for allowing this thesis to be printed on his printer.

The support of my friends and family has been tremendous. Special thanks go to my parents for providing me with the best possible education.

ABSTRACT

Although smallpox has been eradicated there are animal poxviruses which are closely related. It is desirable to measure the closeness of this relation to assess whether Variola virus could reemerge as a complex mutant of an animal poxvirus. The most likely candidate is Monkeypox, which can produce human infection clinically resembling smallpox. The work in this thesis is the beginning of a detailed comparison of the DNA of Variola and Monkeypox.

A 15.3kb section of the Variola genome was compared with a corresponding 14.4kb region of Monkeypox. This enabled both a comparison of corresponding sequences and the location of a short sequence present only in Variola.

Initially restriction enzyme mapping of the two stretches of DNA showed considerable homology and narrowed down the area containing any nonhomologous Variola sequences to within 2.9kb.

Sequence comparisons show a level of 96% similarity. When the 2.9kb Variola fragment was compared with the corresponding 2.4kb Monkeypox fragment, a 400bp insert was found in Variola flanked by sequences common to both viruses. Analysis of the insert revealed two overlapping open reading frames present on opposite DNA strands. The DNA and putative polypeptide sequences were compared with known sequences, but no significant homology was detected. The presence or absence of this sequence in other Orthopoxviruses is being established, but the expression of these open reading frames *in vivo* and function of the putative polypeptides is still to be investigated.

LIST OF ABBREVIATIONS

bp	Base pairs
^{0}C	Degrees Celsius
Ci	Curie
CIP	Calf Intestinal Phosphatase
cm	Centimetre
CTP	Cytosine triphosphate
d	Dalton
dATP (A)	Deoxyadenosine triphosphate
dCTP (C)	Deoxycytidine triphosphate
dGTP (G)	Deoxyguanosine triphosphate
DNA	Deoxyribonucleic acid
dTTP (T)	Deoxythymidine triphosphate
EDTA	Ethylenediaminetetra-acetic acid
g	Gram
IPTG	Isopropyl-B-D-thio-galactoside
kb	Kilobase pair
kd	Kilodalton
M	Molar
Md	Megadalton
mg	Milligram
ml	Millilitre
mm	Millimetre
mM	Millimolar
MPE	Monkeypox DNA HindIII E fragment cloned into pBR328
N	Normal
nm	Nanometre
^{32}P-ATP	Adenosine triphosphate radioactively labelled with Phosphorous-32 (γ position)

^{32}P-dCTP	dCTP radioactively labelled with Phosphorous-32 (a position)
RNA	Ribonucleic acid
rpm	Revolutions per minute
^{35}S-dCTP	dCTP radioactively labelled with Sulphur-35
SDS	Sodium dodecyl sulphate
TEMED	NNN'N'-tetramethylethylenediamine
Tris	Tris (hydroxymethyl) amino methane
U	Unit
uCi	Microcurie
ug	Microgram
UV	Ultraviolet
V	Volt
VarD	Variola HindIII D fragment cloned into pAT153
v/v	Volume in volume
W	Watt
w/v	Weight in volume
X-GAL	5-bromo-4-chloro-3-indolyl-B-galactoside

TABLE OF CONTENTS

CHAPTER 1 INTRODUCTION

1.1 GENERAL FEATURES OF POXVIRUSES

Poxviruses have been classified into genera on the grounds of morphology, serology and host range (see table 1.1). Most work has been done on Vaccinia, the prototype of the Orthopoxvirus genus, but the biology of other poxviruses is believed to be similar. The family Poxviridae consists of two subfamilies, those infecting vertebrates (Chordopoxviridae) and those with invertebrate hosts, in particular, insects (Entomopoxviridae). The vertebrate poxviruses comprise six genera, viz. Orthopox-, Avipox-, Capripox-, Leporipox-, Parapox-, and Suipoxviruses. Some vertebrate viruses do not fit any of these groups, e.g. Molluscum contagiosum, tanapox and Yaba monkey tumor pox.

Variola (smallpox) and Monkeypox are species of the genus Orthopoxvirus of which Vaccinia is the prototype. Morphologically they are large (250 X 200 um) brick-shaped viruses with an outer surface of tubules or filaments arranged in a characteristic whorled pattern. Within the outer layer is a protein coat, two lateral bodies and a biconcave viral core consisting of DNA bound to well-defined protein subunits. Globular structures (20 to 60 nm in diameter) of supercoiled DNA folded into nucleosome structures with protein have been observed (Soloski and Holowczak 1981). The DNA has a molecular weight of approximately 120 Md.

The Parapoxviruses, viruses of mammals including milker's node virus have a smaller genome (The DNA is approximately 85 X 10^6 daltons.) and are more ovoid in structure. The tubules or filaments are arranged in a regular crisscross pattern probably caused by one continuous filament wound around the virion in a

left-handed sense. The internal components are similar to those of the other poxviruses.

Capripoxviruses, viruses of ungulates, are distinguished by a long, narrow virion.

A varied morphology is found amongst the Entomopoxviruses. Instead of tubular elements, globules are found on the outer surface. Internally, one lateral body may be present instead of two and the core may be kidney shaped. Serologically, these viruses do not cross-react with the vertebrate poxviruses, nor does one genus cross-react with another, except for one nucleoprotein antigen.

Table 1.1 The Poxviruses - continues on the following page (taken from Moss 1985)

Subfamilies	Genera	Representative members	Distinctive features
Chordopox-viridae (vertebrate poxviruses)	Orthopoxvirus	Buffalopox Camelpox Cowpox Ectromelia Monkeypox Vaccinia Variola	Variola cause of smallpox, DNA 120 X 10^6
	Parapoxvirus	Bovine pustular stomatitis Chamois contagious ecthyma Milker's node orf	Viruses of mammals, surface tubules forms regular spiral, DNA 85 X 10^6
	Avipoxvirus	Canarypox Fowlpox Juncopox Pigeonpox Quailpox Sparrowpox	Viruses of birds, arthropod transmission, DNA 200 X 10^6
	Capripoxvirus	Goatpox Lumpy skin disease sheeppox	Viruses of ungulates, virions longer and narrower than Vaccinia, arthropod transmission,
	Leporipoxvirus	Myxoma Rabbit fibroma Squirrel fibroma	Viruses of leporids and squirrels, DNA 150 X 10^6
	Suipox	Swinepox virus	Viruses of swine, limited host range

Table 1.1 The Poxviruses (continued)

Subfamilies	Genera	Representative members	Distinctive features
Entomopox-viridae (insect poxviruses)	A	Melontha melolontha virus	Related viruses isolated from Coeloptera. one lateral body, DNA 200 X 10^6
	B	Amsacta moori virus	Related viruses isolated from other Lepidoptera, indistinct lateral body, DNA 140 X 10^6
	C	Chironimus luridus	Related viruses isolated from other Diptera, two lateral bodies

Having a complex composition, vaccinia contains cholesterol, phospholipid, neutral fat, carbohydrate, copper, riboflavin and biotin as well as DNA and protein. The RNA content of the virus is very small. (Zwartouw 1963). The virion contains more than 100 polypeptides which can be separated two-dimensionally by polyacrylamide gel electrophoresis and isoelectric focussing (Essani and Dales 1979). The virion itself contains enzymes important in establishing virus infection in the cytoplasm. Unlike most viruses, purified poxvirus DNA is not infectious.

Viral entry of the cell

The binding and subsequent entry of poxviruses into the cell takes place in a series of separate events. Specific proteins localized in the outer envelope of the virus are thought to

mediate these events. Viral entry has been proposed to be mediated by a 54kd surface tubule protein which has been isolated and shown to elicit antibodies which neutralize infectivity. Evidence has also been provided for binding to occur via a 37kd envelope component and subsequent penetration to require a small envelope protein. The involvement of a 14kd envelope protein for virus penetration has been demonstrated. (Dales and Pogo 1981).

The virus enters the cell via fusion of its envelope with the plasma membrane. Entry by endocytosis may not lead to further events and Janeczko et al. (1987) have shown that entry for a productive infection does not require a low pH endocytic pathway. Once the virion envelope has fused with the cell membrane the core is released into the cytoplasm of the cell and the lateral bodies remain attached to the envelope. DNA is released from the core through definite breaks in the core coat. Multiple infectious units in a single cell release DNA synchronously, probably after synthesis of the uncoating factor. This factor is probably synthesized by the poxvirus since inhibition of host transcription (with actinomycin D) does not prevent uncoating, but an antiviral substance (interferon) does. (Dales and Pogo 1981).

The poxvirus "factory"

The DNA molecules remain together throughout the replicative cycle in "factories" in the cytoplasm. All events during the virus cycle have been divided into prereplicative (early) or postreplicative (late) phases. The immediate-early functions are associated with transcription from the core; early functions are those occurring after uncoating, but before replication; late functions take place after replication and are related to virion

assembly and maturation; and late-late functions involve the synthesis of virus-specified products which are not components of the virion itself after maturation.

Upon infection vaccinia inhibits host cell protein synthesis. This is possibly brought about through induction of cellular mRNA degradation (Rice and Roberts 1983). Host cell DNA replication in the nucleus is also inhibited - possibly by DNAses encoded by vaccinia (Oligiati et al. 1976).

The early and late mRNA represents approximately 20% and 50% of the viral genome respectively, the switch from early to late expression never being complete and varying in different host cells. The viral DNA is transcribed by an RNA polymerase of 500kd (Dales and Pogo 1981). This enzyme is coded for by the viral genome and the subunit genes are located within the highly conserved central region (Jones et al. 1987).

Both classes of mRNA are polyadenylated after transcription, even though consensus sequences are not always present in the genome (see section 1.2.2.4). The poly(A) tract is synthesized by the core which contains the enzyme riboadenylate transferase (or poly(A) polymerase). Early and late mRNA is also capped and methylated so that binding to the 40s ribosomal subunit can take place for efficient translation. This is brought about by five enzymes which have been identified within the core, viz. polynucleotide 5'-triphosphatase, mRNA guanylyltransferase, mRNA(guanine-7-)-methyltransferase, mRNA methylase (nucleoside-2'-O-methyl transferase) and 5'-phosphate-polyribonucleotide kinase. (Dales and Pogo 1981). The gene coding for the large subunit of the mRNA(guanine-7-)-methyltransferase has been located by Morgan et al. (1984) to a central conserved region of

the vaccinia genome (3.1kb to the left of the HindIII D fragment). Some mRNA is processed, as revealed by the 8 to 12s cleavage products of 20 to 30s mRNA by Paoletti (Dales and Pogo 1981).

Late transcripts do not appear to have defined 3′ ends - as if transcription runs through other genes (Moss 1985). The 5′ ends are also unusual in that they have long leader sequences tagged onto them via a poly(A) stretch. The leader sequence may be transcribed upstream or downstream of the late gene and from either strand (Bertholet et al. 1987; Schwer et al. 1987).

Viral gene expression is probably regulated at many levels. The early and late genes have different promoter sequences (discussed in section 1.2.2.4) which allow specific genes to be transcribed at a particular time in the virion cycle. The mRNA stability has been shown to influence gene expression; and, at the translational level, suppressor proteins are thought to exist (Moss 1985).

Although poxviruses have their "factories" in the cytoplasm, a host-related nuclear function is thought to be required for late gene transcription. The ability of a-amanitin to prevent infectious vaccinia forming relates this function to RNA Polymerase II.

DNA replication

Vaccinia codes for a DNA polymerase enzyme which has a gap-filling and a proof-reading 3′ exonuclease activity. Other enzymes including two DNAses, a DNA ligase and a nicking-closing enzyme have been identified as playing a role in DNA replication. (Moss 1985). Recently Evans and Traktman (1987) identified a gene coding for an 82kd protein which is essential for DNA

replication.

Semiconservative replication is thought to take place by a strand displacement mechanism with initiation and termination occurring at both ends (Pogo et al. 1981). Initially, nicks are formed near the termini, allowing self-priming by the 3' ends. DNA synthesis proceeds bidirectionally, short segments being linked to RNA primers. Finally the segments are joined together (Moss 1985). If only one DNA strand is nicked concatemers are formed, which are subsequently processed into single units. These concatemers are arranged in a head-to-head and tail-to-tail configuration (Moyer and Graves 1981). This model of replication can explain how deletions and insertions at the termini are generated.

The maintenance of a linear minichromosome structure is achieved by the hairpin ends of the genome (Merlinsky and Moss 1986). Circular E.coli plasmids containing inserted hairpin ends were shown to replicate in vaccinia-infected cells in the form of minichromosomes.

Assembly, maturation and release of virions

Virions are assembled out of viral DNA, virus-specified proteins and host derived lipids. The unit membrane is coated by spicules which confer rigidity. The membrane encloses the viral nucleoprotein complex after which the interior reorganizes itself. The lateral bodies become distinguishable from the core, the membrane loses its spicules and surface tubular elements become attached to the membrane. Requirements for virion maturation include DNA replication, late protein synthesis and post translational processing (Dales and Pogo 1981).

Mature pox virions usually remain in the cell, but may be

released in different ways. Firstly, the virions may migrate to the surface of the cell, where they move to the tips of microvilli-like protrusions and are released by breakdown of the extensions (Dales and Pogo 1981). Secondly, the virion may associate with intracytoplasmic membranes and acquire a second membrane. After migration to the cell surface the double-membraned virion fuses with the cell membrane and a virus (with a single envelope) is released from the cell (Payne and Kristenson 1979). The second membrane is probably acquired from the Golgi body (Dales and Pogo 1981).

Cell to cell spread of viruses may be increased via the induction by the virus of haemagglutinin, which attaches to the host cell membrane. This is a late-late function of the virus. (Dales and Pogo 1981). Another means of dissemination of the virion from the host cell involves the formation of proteinacious inclusions ("A"-type). Inclusions of cowpox, ectromelia and fowlpox viruses have been observed, but no other vertebrate poxviruses have these inclusions. The virions require an occlusion factor on their surface in order to migrate out of the factories (B-type inclusions) to form these inclusions. (Dales and Pogo 1981).

1.1.1 DIFFERENTIATION OF ORTHOPOXVIRUSES

Since this thesis is concerned with Variola and Monkeypox, the Orthopoxvirus genus will be discussed. In the past, because Orthopoxviruses could not easily be differentiated by neutralization or other serological techniques, different species were identified by their phenotypic characteristics. These properties include pock morphology on chorioallantoic membranes, growth capacity at different temperatures, types of lesions in rabbit skins, pathogenicity in chick embryos and mice and growth characteristics in tissue culture. (Baxby 1975). More accurate comparisons of different Orhtopoxviruses have been obtained by immunological methods, polypeptide analysis, nucleic acid hybridization, restriction endonuclease analysis, mapping and, more recently, sequencing. Analysis of DNA shall be discussed in a later section.

The similarity of Monkeypox, Variola and Vaccinia has been demonstrated in neutralization, haemagglutination inhibition and immunodiffusion tests. Extensive cross-reactions take place between heterologous viruses and antisera (Marennikova et al. 1972). Monkeypox, serologically, has been shown to be more closely related to Variola that to Vaccinia or Cowpox viruses (Rondle and Sayeed 1972).

Using cross-absorbed antisera, Monkeypox, Variola and Vaccinia may be distinguished from one another by immunodiffusion (Esposito et al. 1977). The common antibodies present in serum were absorbed out by heterologous viral antigens, leaving only those antibodies specific for individual viruses. These monospecific antigens were isolated and found to be a single protein of 73kd in each virus (Esposito et al. 1977). The unique

antigens of Variola, Vaccinia and Monkeypox have also been differenteated by radioimmunoassay (Hutchinson et al. 1977). Vaccinia and Variola viruses have a common antigen not present in Monkeypox; Variola and Monkeypox have a common antigenic group not found in Vaccinia; and Vaccinia and Monkeypox share a common antigen not present in Variola. Additionally, Monkeypox virus contains a unique antigenic group.

The separation of viral polypeptides by polyacrylamide gel electrophoresis has confirmed serological findings and provided a better means of differentiating viruses both between and within different genera. (Esposito et al. 1977; Harper et al. 1979; Turner and Baxby 1979; and Arita and Tagaya 1980).

1.1.2 CHOICE OF MONKEYPOX FOR AN INITIAL COMPARISON WITH VARIOLA

With the eradication of smallpox, human monkeypox, first recognized in 1970 (Ladnyj et al. 1972), is the most important Orthopoxvirus infection of man. One of the requirements for the successful eradication of smallpox was the absence of a natural animal reservoir of the virus. Although a reservoir of smallpox in non-human primates is unlikely, certain species of monkeys have been infected with Variola; and Monkeypox, which resembles Variola very closely, has been detected in captive monkeys (Arita and Henderson 1968). Evidence supporting the absence of a natural reservoir of smallpox in primates includes: the eradication of smallpox from areas such as Panama and the Philippines where large monkey populations exist; the low incidence of poxvirus infection in monkeys; and the inability of smallpox to be transmitted in serial passages in monkeys. (Noble 1970).

Although Monkeypox is enzootic in the tropical rain forest in West and Central Africa, it is not sufficiently transmissible in humans to become established in man. An important feature of monkeypox is its host range. It has been isolated from captive animals including mice, rabbits, monkeys and two zoo-kept anteaters (Gispen 1975). A more major source of the virus may be squirrels. In 1985 a virus was isolated from a wild squirrel and since then a high level of monkeypox-specific antibodies have been found in two species of squirrels in populated areas of Bumba zone of Zaire. Squirrels are thought to be a major source of infection of monkeypox for humans (Khodakevich et al. 1987).

Between 1980 and 1985, 282 cases of monkeypox infection were

detected in Zaire, more than half the patients being under the age of five and more than 90% less than 15 years of age (Jezek et al. 1987). In contrast, only 54 cases were reported from 1970 to 1979 (Arita et al. 1985). This increase of reported incidence may be due to an intensified surveillance program by the World Health Organization. Other causes may be: a change in the virus, making it better adapted for survival in humans; an increase in zoonotic transmission; or an increase in the susceptibility of people for the virus since vaccination was stopped.

Clinically, human monkeypox infections resemble those of smallpox. A feature distinguishing it from smallpox and chickenpox is lymphadenopathy in the early stages. The disease also presents itself differently in vaccinated and unvaccinated patients, no deaths having occurred among vaccinated people, but 11% of unvaccinated cases being fatal. The fatality rate was closer to 15% in younger children. Unvaccinated patients had many more lesions than those vaccinated and the severity of the symptoms were proportional to the number of lesions.

The illness lasts for two to four weeks and goes through two stages - preeruptive and eruptive. The preeruptive stage begins with a fever for 1 to 3 days during which time lymph nodes become enlarged. By the the third day a rash develops. The patient suffers headache, backache, malaise and prostration. Lesions develop simultaneously during the eruptive stage, usually appearing first on the face. They go through stages of macules, papules, vesicles and pustules after which they umbilicate, dry out and desquamate. Complications due to viral activity and secondary bacterial infections are presented. (Jezek et al. 1987).

1.1.3 FOCUS OF INTEREST

The most accurate way of assessing the possibility of Monkeypox producing a genetic variant with the same propensity as Variola is to compare the genetic material (DNA) of the two viruses. This thesis concerns the comparison of a certain region of the two viral genomes conserved within different species. An attempt is being made to identify a gene present in Variola which is absent in Monkeypox. The identification of such a gene would imply that Monkeypox can not, through random mutations, revert to Variola.

Various comparisons previously made between Orthopoxvirus genomes will be discussed.

1.2 POXVIRUS DNA

1.2.1 STRUCTURE OF THE GENOME

Recently more emphasis has been placed on the genome of Orthopoxviruses than the phenotypic properties. Information on the DNA structure rather than its products of expression is a more sensitive means of comparing Orthopoxviruses and may expose mechanisms of gene regulation.

The poxvirus genome is a single double stranded linear molecule varying in molecular weight from 85 X 10^6 daltons (for Parapoxviruses) to 185 X 10^6d (for fowlpox virus) (Moss 1985). Variola, Monkeypox and Vaccinia have molecular weights of 121, 128 and 124 Md respectively (Mackett and Archard 1979); (approximately 181, 192 and 186 kb respectively).

Having an A+T content of 63-68%, the Orthopoxvirus genome resembles that of prokaryotes more so than eukaryotes. Unlike the other poxviruses Parapoxviruses have a G+C content of 63% (Wittek et al. 1979).

The two DNA strands are covalently linked to one another by phosphodiester bonds and the hairpin loops formed at the ends have been sequenced by Baroudy et al. 1982). The sequence, highly A+T rich, is incompletely base-paired and found in two isomeric forms. The two forms, inverted and complementary in sequence would arise from the replication model proposed by Baroudy et al. (1982).

Another distinctive feature of the pox virus genome is the presence of identical sequences at the termini (Wittek et al. 1982). In Vaccinia (strain WR) this inverted repeat element is 10kb and it includes a series of short direct repeats. The short

repeats consist of 70bp and are found tandemly arranged in two blocks of 13 and 17 units separated by an intervening sequence of approximately 435bp. (Wittek and Moss 1980). Upon sequencing of the termini, Baroudy et al. (1982) found another repeat of 125bp and a sequence of 54bp present in eight units. The termini of the Variola genome are slightly different in that a large deletion is present at one end and the inverted repeat is less than 500bp. (Dumbell and Archard 1980).

In contrast to the central region which is highly conserved between different Orthopoxviruses, the termini are hypervariable and undergo deletions and rearrangements (Wittek 1982).

1.2.2 PREVIOUS COMPARISONS OF ORTHOPOXVIRUS DNAs

1.2.2.1 ELECTROPHORETIC PROFILES

Closely related DNA molecules may be compared for sequence homology by analysing cleavage products generated by specific restriction endonucleases. Because DNA fragments are separated according to size, an important assumption made, is that fragments of the same length are found in corresponding sections of the genomes. Patterns generated by single enzymes from DNA of different viruses may give an estimation of the relatedness of the viruses. Similar comparative results from several different enzymes may increase the validity of the comparison. This method of comparing two stretches of DNA may detect small alterations such as single base changes if the change is within a restriction enzyme site.

The closely related Orthopoxviruses Cowpox, Vaccinia, Monkeypox and Variola have been differentiated by restriction enzyme

patterns of their genomes (Esposito et al. 1978). Different species, subspecies and individual strains of virus could be identified by characteristic gel patterns, the enzymes HindIII and SalI allowing species and strain differentiation respectively. HindIII also allowed variola major to be distinguished from variola minor.

Similar comparisons have been made between other poxviruses - Muller et al. (1977) have compared the DNAs of Rabbitpox, Vaccinia, Cowpox, Ectromelia and Fowl pox virus; Gangemi and Gordon Sharp (1976) and McCarron et al. (1978) have compared two different Vaccinia virus strains; and variants of Cowpox virus have been analysed by Archard and Mackett (1979). Different isolates of Orf virus have been compared by Robinson et al. (1982) and Raffii and Burger (1985); six strains of Leporipoxvirus and Vaccinia were compared by Cabirac et al. (1985); Capripoxvirus DNAs have been compared by Black et al. (1986); and Orthopoxvirus DNAs isolated from zoo-kept mammals were characterized by Pilaski et al. (1986).

1.2.2.2 RESTRICTION MAPS

The location of restriction enzyme cleavage sites on the genome allows for a better comparison between viral DNAs. The first two poxvirus genomes to be mapped were those of Vaccinia (strain Lister) and rabbit poxvirus (Wittek et al. 1977). Sites were located by double digestion with HindIII and SstI; and partial digestion followed by complete cleavage by a single enzyme. Because Orthopoxviruses have many DNA sequences in common other genomes have been mapped by comparison with the Lister strain of vaccinia. Immobilized fragments of the genome to be mapped are probed with radioactively labelled fragments of vaccinia DNA.

Those fragments which hybridize to the probe are located to the corresponding area of the genome. Terminal restriction fragments are identified by their ability to renature rapidly after denaturation (because they are cross-linked). Co-electrophoresis of denatured with non-denatured digests identifies co-migrating fragments as the termini.

These techniques have been used by Mackett and Archard (1979) to map sixteen strains of Orthopoxvirus. Strains of the Rabbitpox, Vaccinia, Monkeypox, Variola, Cowpox and Ectromelia species were mapped for HindIII, XhoI and SmaI restriction enzyme sites. The HindIII maps reveal the central region of approximately 30 X 10^6d to be conserved amongst the Orthopoxviruses, although species specific differences are detected. In contrast, the termini are variable, even within species. Generally, Cowpox virus genomes have additional sequences at either terminus; and Variola strains have deletions at one specific terminus. Monkeypox, Variola and Vaccinia may be readily distinguished by their SmaI maps (Mackett and Archard 1979).

A comparison of many Orthopoxvirus DNAs has been done by Esposito and Knight (1985) who mapped 38 Orthopoxviruses for sites of cleavage by HindIII. Genomes of various species and strains of Orthopoxvirus have been mapped by Archard and Mackett (1979); Schumperli *et al.* (1980); Panicali *et al.* (1981); Esposito *et al.* (1981) and Defilippes (1982). Recently genomes of several strains of Orf virus have been mapped (Mercer *et al.* 1987 and Robinson *et al.* 1987). There appears to be more sequence variability in the internal region of these DNAs than is present within the genomes of Orthopoxviruses.

1.2.2.3 HETERODUPLEX ANALYSIS

Because of the extensive similarity of Orthopoxvirus genomes short, non-homologous sequences within conserved areas cannot be easily identified by standard hybridization and mapping techniques. An alternative approach to identifying differences between two sequences is that of heteroduplex analysis. Under partial denaturing conditions heteroduplexed DNA forms bubbles or loops where non-homologous or inserted sequences are present respectively. These can be detected when differences are 100bp or greater.

Variola and Monkeypox sequences within a conserved region of the genome were compared by heteroduplex analysis by Kinchington et al. (1984). Corresponding fragments for comparison were cloned into different vectors so that loops formed between the vectors could be used to orientate the heteroduplexed inserts. A heterogeneous region of 4 to 6 kb was located within the N and left end of the C HindIII restriction fragments of variola and the O and left of the C HindIII restriction fragments of Monkeypox. (See figure 1.2.1 in section 1.2.3 for location of these fragments). The two C fragments differ in length by approximately 2kb, of which most of the Monkeypox insert was detected as a single-stranded loop. The detection of this inserted sequence in the Monkeypox HindIII C fragment implies that the HindIII sites to the right of the C fragments are co-incident. If this is the case an insert is present in the adjacent HindIII D fragment of Variola. Since the detection of a sequence unique to Variola is required for greater confidence in the eradication of the virus (so it cannot reemerge by genetic alterations to Monkeypox) the D fragment of Variola is being compared to the corresponding E fragment of Monkeypox.

1.2.2.4 SEQUENCES

More recently comparisons of poxvirus DNA have been made by sequence analysis; in particular the thymidine kinase gene of different virus species has been compared. Sequencing has been important in confirming the position of genes on the viral genome. Because poxviruses are suitable vectors for vaccination purposes possible sites for insertion of foreign genes have been sequenced. These include the vaccinia virus haemagglutinin gene (Shida 1986) and the thymidine kinase gene (Weir and Moss 1983) which has already been used for insertion of foreign genes.

DNA sequencing has been essential in the determination of gene organization and regulation of gene expression. The early and late genes have differing control signals; and promoter and terminator consensus sequences have been identified.

Both DNA strands of the poxvirus genome are transcribed and open reading frames tend to be closely packed. Within the 16kb HindIII D fragment of vaccinia virus there are 22 open reading frames, each one separated by no more than 42 base pairs, some of which have overlapping ends (Niles et al. 1986). 12 partly overlapping open reading frames have also been observed in another conserved region of 7.6kb in the vaccinia genome (Pluciennicczak et al. 1985).

Promoter regions for both early and late genes, in general, are A/T-rich and have a TATA box. An early vaccinia virus gene sequenced by Venkatesan et al. (1981) had 53 A+T residues within 60 bases upstream of the transcription start site, but no recognition sequences typical of prokaryotic and eukaryotic genes further upstream. Termination of transcription occurred near or within the sequence CTATTC which was tandemly repeated four

times. The polyadenylation signal AATAAA was not identified. A similar DNA structure was found within the inverted terminal repeat of the Vaccinia virus genome within which three early polypeptides are encoded (Venkatesan et al. 1982). As well as the A/T-rich region upstream of the initiation codon for translation and the tandemly repeated CTATTC hexanucleotide near the 3′ end, these genes were preceded by the eukaryotic CAAT consensus sequence (40 to 60 nucleotides upstream of the initiation codon). A sequence complementary to the end of the 18s rRNA was also present near the start of translation. None of the sequences had a polyadenylation signal near the 3′ end.

Esposito and Knight (1984) identified eukaryotic recognition sequences, viz. the CAAT RNA polymerase II promoter sequence at position -100 and some core enhancer sequences upstream of the initiation codon, in the thymidine kinase genes of Variola, Monkeypox and Vaccinia; but in an experiment to determine how many nucleotides upstream of the start site were required for promoter activity, Weir and Moss (1987) demonstrated that no more than 32 base pairsupstream of the initiation codon was required for efficient promoter activity. This result was obtained from transient expression of chloramphenicol acetyl transferase (CAT) upon transfection of vaccinia-infected cells with plasmids containing deletions of the putative thymidine kinase promoter sequence ligated to the gene coding for CAT. Since CAT was only synthesized in cells infected with vaccinia virus the promoter is thought to function upon binding of a Vaccinia virus gene product.

Termination signals of early genes appear to be varied. The thymidine kinase genes have the AATAAA consensus sequence for

polyadenylation (Esposito and Knight 1984), unlike the other early genes described. As previously mentioned the hexanucleotide CTATTC is repeated at the end of some early genes; and the presence of another sequence TTTTTNT near the 3′ end is indicative of termination of transcription of early genes (Upton et al. 1987).

Regulatory sequences required for late gene expression have been shown to reside within 100 base pairsof the mRNA start site (Bertholet et al. 1985). A consensus sequence of TTCATTTa/tGT 34 to 18 nucleotides upstream of the initiation codon for translation has been identified (Shida 1986) and all late genes (except one) contain a characteristic consensus sequence TAAATG which includes the initiation codon of translation (Roseman and Hruby 1987; Rosel et al. 1986; Hirt et al. 1986). Most late genes can be recognized by an extremely short leader sequence and the absence of A or G in the -3 position relative to the start codon of translation (Rosel et al. 1986).

Comparisons of corresponding sequences in different poxviruses have revealed extensive similarities amongst the different viruses. The thymidine kinase gene of Variola, Monkeypox and Vaccinia is almost identical in all three viruses. Of the eight differences in amino acid sequence between Variola and Vaccinia two may alter the kinetic properties of the protein synthesized. (Esposito and Knight 1984). The thymidine kinase gene of the tumorigenic Leporipoxvirus, Shope fibroma virus was sequenced by Upton and McFadden (1986). Although the amino acid sequence is only fairly similar to those of Vaccinia, Variola and Monkeypox, the open reading frames to the 5′ side of the thymidine kinase gene were very similar in all four sequences. A common ancestral origin is proposed to have existed from which Orthopoxviruses and

Leporipoxviruses diverged. Although the thymidine kinase gene of Fowlpox, (an Avipoxvirus) is located in a different position in the viral genome compared with Vaccinia, the DNA sequences in corresponding positions close to the Vaccinia virus thymidine kinase gene is very similar (Drillien et al. 1987). Six open reading frames have been conserved. The Fowlpox thymidine kinase gene has been sequenced by Boyle et al. (1987) and compared to other thymidine kinases. The promoters are identical and the amino acid sequences similar, although the Fowlpox virus thymidine kinase has 6 more amino acid residues than that of Vaccinia virus. The protein was similar to that of the chicken, man and the three other poxviruses, but there was no resemblance with that of herpes simplex virus.

The DNA sequence of vaccinia virus DNA polymerase was compared to other viral DNA polymerase genes (Earl et al. 1986). A conserved area of 14 amino acid residues in the carboxyl-terminal region of pox-, herpes- and adenovirus DNA polymerases is thought to be an essential binding or catalytic site. Sequencing of the Fowlpox DNA polymerase gene revealed extensive similarity in the amino acid sequences of Vaccinia and Fowlpox DNA polymerases. (Binns et al. 1987).

Comparisons of the repeated units at the termini of poxvirus genomes may be a useful means of comparing the relatedness of poxviruses. The similarity of repeat units of 54, 70 and 125 bp points towards evolution by unequal crossing over (Baroudy and Moss 1982). Sequences of the terminal regions of Cowpox virus DNA have been determined by Pickup et al. (1982) and compared with those of Vaccinia virus DNA. The Cowpox repeat elements are arranged in a more complicated manner.

1.2.3 COMPARISON OF THE HindIII RESTRICTION MAPS OF VARIOLA AND MONKEYPOX

The HindIII restriction enzyme maps for the genomes of Variola and Monkeypox were constructed by Mackett and Archard (1979). These two maps are compared in figure 1.2.1. The Monkeypox genome is slightly larger than that of Variola - due to a deletion at the left end of Variola. The HindIII fragment at the extreme left of the Monkeypox genome does not hybridize to the left end fragment of "Variola". The right terminus also shows variability with respect to the locations of HindIII restriction enzyme sites.

The central portion appears highly conserved. There is, however, a HindIII site separating the C and D fragments of Variola; and the C and E fragments of Monkeypox which do not correspond exactly in position; and the position of this site is conserved within the two different species. Heteroduplex analysis of the two C fragments suggests that the two HindIII sites are in corresponding positions. If this is confirmed, a comparison of Variola D and Monkeypox E fragments will be useful in establishing the location of the surplus DNA in the Variola D fragment. These HindIII fragments have been cloned into various vectors by Hamilton et al. (1985) for the purpose of finer characterization of the genomes. The identification of a gene unique to Variola within this region - possibly conferring species specificity for man - would imply that smallpox may not reemerge through the natural reservoir of Monkeypox.

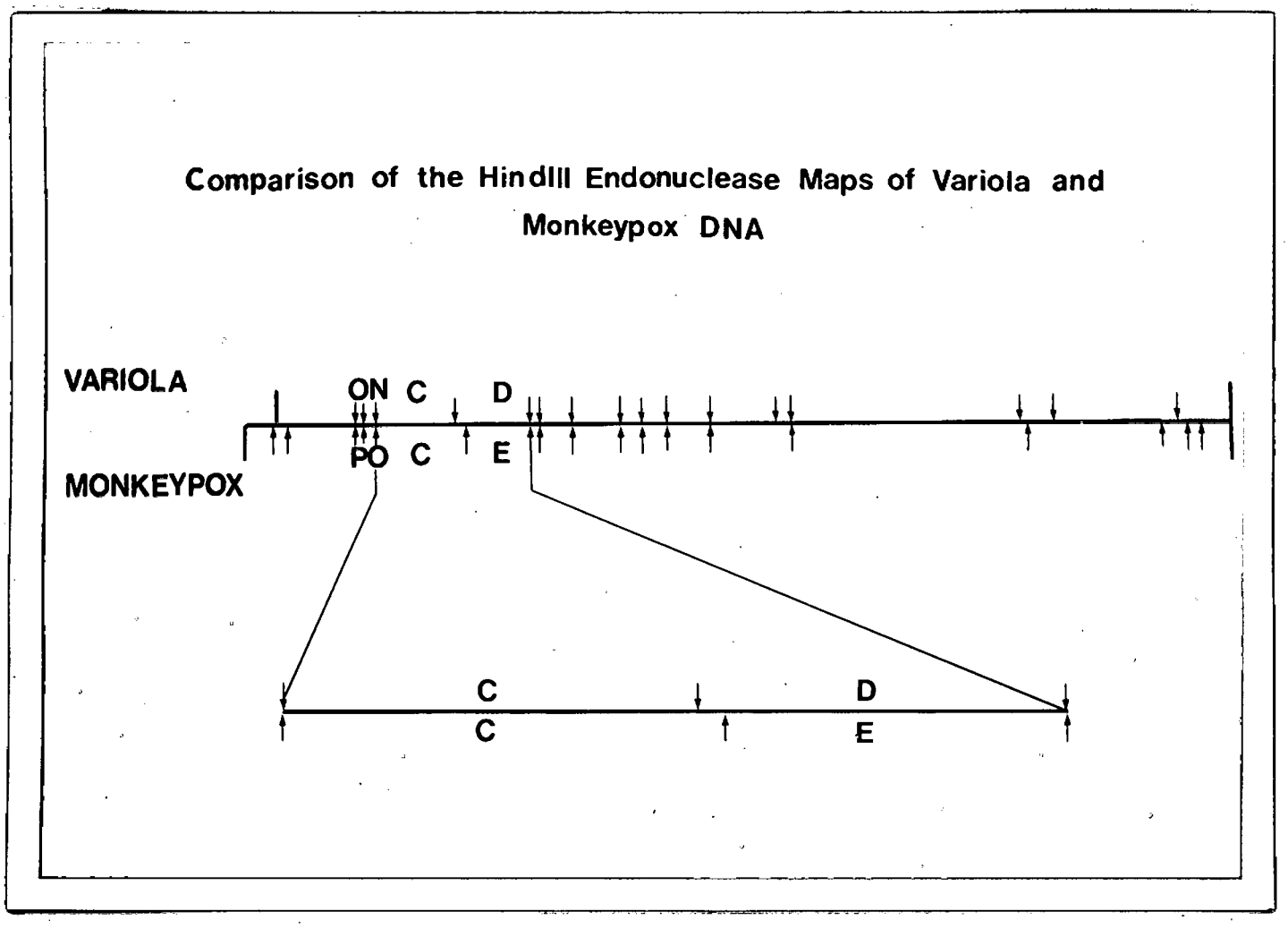

Comparison of the HindIII Endonuclease Maps of Variola and Monkeypox DNA
VARIOLA
ON C D
MONKEYPOX
PO C E
C D
C E

Figure 1.2.1.

CHAPTER 2 MATERIALS

2.1 VECTORS AND HOST CELLS

M13mp19 was supplied by Amersham International plc, Amersham Place, Little Chalfort, Buckinghamshire, England, HP7 9NA.

pUC19 and host cells JM103, JM105 and JM109 were received from Prof. F. Robb, Department of Microbiology, U.C.T., Cape Town.

JM107 cells were obtained from Bethesda Research Laboratories (BRL), Gaithersburg, Maryland 20877.

2.2 ORTHOPOXVIRUS DNA

2.2.1 RECOMBINANT PLASMIDS CONTAINING ORTHOPOXVIRUS SEQUENCES

Recombinant plasmids Variola (Harvey strain) HindIII C, Variola (Harvey strain) HindIII D, Variola (Somalia strain) HindIII D, Monkeypox (Denmark strain) HindIII C, and Monkeypox (Denmark strain) HindIII E were received from Prof. K. Dumbell as frozen transformed HB101 cells.

2.2.2 GENOMIC ORTHOPOXVIRUS DNA

DNA was isolated by Mrs. Madalene Richardson, Department of Medical Microbiology, U.C.T., Cape Town from the following strains of Orthopoxvirus, grown on chick embryo chorioallantois:

Monkeypox: Denmark, Zaire (squirrel 85/218 and human 86/21), Nigeria (71/82) and Sierra Leone (70/266)

Vaccinia: USSR, Dairen, Wyeth and Lister

Cowpox: LB

Camelpox: CMS

2.3 ENZYMES AND KITS

Enzymes were supplied by Amersham; BRL; Boehringer Mannheim GmbH, Biochemica, P.O. Box 310 120, D-6800, Mannheim 31, West Germany; and Pharmacia, Fine Chemicals AB, Uppsala, Sweden.

Amersham supplied Polynucleotide Kinase, ExonucleaseIII, Klenow DNA Polymerase I and restriction enzymes AvaI, BamHI, EcoRI, HincII, HpaI, KpnI, NcoI, PstI, SacI, SalI and XhoI. Boehringer supplied Calf Intestinal Phosphatase (CIP), T4 DNA ligase, S1 nuclease and restriction enzymes SphI and BglI. HaeIII restriction enzyme was obtained from BRL and Pharmacia supplied Mungbean nuclease.

The Southern Cross Hybridization kit was supplied by DuPont, New England Nuclear, 549 Albany Str., Boston, MA02118. Both the Nick Translation kit and the M13 sequencing kit were obtained from Amersham.

2.4 CHEMICALS, REAGENTS AND OTHER MATERIALS

Chemicals supplied by SIGMA, P.O. Box 14508, St. Louis, MO 63178, USA include ammonium persulphate, ATP, bromophenol blue, Lithium Chloride, polyethylene glycol (PEG), SDS, spermidine, sucrose, TEMED and xylene cyanol.

Merck (Frankfurter Strasse 250, D-6100, Darmstadt, Germany) supplied Calcium Chloride, ethidium bromide, formamide, glycerol, hydrochloric acid, 8-hydroxyquinoline, isoamylalcohol, Potassium acetate, Sodium Chloride, Tris and xylol.

BDH, Poole, England supplied acrylamide, boric acid, chloroform, dimethyldichlorosilane, dimethylformamide, EDTA, ethanol, glucose, isopropanol, Magnesium Chloride, 2-mercaptoethanol,

phenol, Sodium citrate, Sodium Hydroxide, triton, urea and Zinc Sulphate.

Cesium Chloride, dithiothreitol and lysozyme were obtained from Boehringer; ampicillin from Beechams; Hybond-N, radionucleotides and M13 reverse sequencing primer from Amersham; and IPTG and X-GAL from BRL. Agarose and Bis (N,N'- methylene bis-acrylamide) were supplied by Bio-Rad laboratories, 32nd and Griffin Ave., Richmond, California, 94804.

Pharmacia supplied Sephadex G-50; FMC Corporation, Marine colloids division, Rocklands, ME 04841, USA supplied low gelling temperature agarose; and DEAE-cellulose was obtained from Schleicher and Schuell, D-3354 Dassel, West Germany. Dialysis tubing was obtained from Spectrum Medical Industries, inc., 60916 Terminal Annex, Los Angeles 90054.

Sodium deoxycholate was obtained from Hopkin and Williams Ltd., Chadwell Heath, Essex, England.

"BLOTTO" is Carnation low fat milk powder.

2.5 MEDIA AND SOLUTIONS

Luria broth and Luria agar were supplied by the department of Medical Microbiology, University of Cape Town.

Luria Agar: 10g Tryptone
5g Yeast extract
10g NaCl
12g Difco agar
pH7.2 with NaOH; total volume 1 litre.

Luria Broth: 10g Tryptone
5g Yeast extract
10g NaCl
pH7.4 with NaOH; total volume 1 litre.

Both the agar and the broth were sterilized at 115^{0}C for 15 minutes and subsequently stored at room temperature.

Antibiotics: Ampicillin (25mg/ml) was stored at -20^{0}C in aliquots and used at a final concentration of 25ug/ml.

Phenol: Phenol was melted at 57^{0}C. 8-Hydroxyquinoline was added to a final concentration of 0,1% (as an antioxidant, a partial inhibitor of RNase and a weak chelator of metal ions). The phenol phase can be identified by the yellow colour. Repeated extractions with an equal volume of 1.0M Tris pH8 were performed until the pH of tne aqueous phase exceeded 7.6. The phenol was stored under 0.1M Tris pH8 and B-mercaptoethanol was added to a final concentration of 0.2%. Phenol solution was stored at 4^{0}C. Molten phenol was stored at -20^{0}C under nitrogen gas.

Chloroform:isoamylalcohol was used at 24:1 v/v and stored at room temperature.

IPTG (100mM) was prepared by dissolving 23.8mg/ml in H_2O and stored at 4^0C.

X-GAL was dissolved in dimethylformamide (2%) and stored at 4^0C.

Solutions for sequencing

The formamide dye mix was prepared by adding a spatula tip of each of EDTA, bromophenol blue and xylene cyanol to 1ml of formamide.

40% acrylamide stock: 38g acrylamide
2g bis
distilled water to 100ml
Stored in the dark at 4^0C.

10 X TBE buffer: 108g Tris base
55g boric acid
9.3g $Na_2EDTA.2H_2O$
Add distilled water to 1 litre.

Top gel: 76.8g urea
24ml acrylamide stock
8ml 10 X TBE
distilled water to 160ml
Stored in the dark at 4^0C.

Bottom gel: 14.4g urea
3.0g sucrose
7.5ml 10 X TBE
4.5ml acrylamide stock
distilled water to 30ml and add a little powdered bromophenol blue
Stored in the dark at 4^0C.

The following deoxynucleotide mixes were made for use with radiolabelled dCTP:

	A^0	C^0	G^0	T^0
0.5mM dATP	1ul	20ul	20ul	20ul
0.5mM dGTP	20ul	20ul	1ul	20ul
0.5mM dTTP	20ul	20ul	20ul	1ul
10mM Tris, pH 8.0, 1mM EDTA	20ul	20ul	20ul	20ul

The sequencing mixes were prepared by adding an equal volume of N^0 and the appropriate dideoxynucleotide solution.

CHAPTER 3 PREPARATION OF POXVIRUS DNA FROM RECOMBINANT VECTORS

3.1 LARGE SCALE PLASMID EXTRACTION

Poxvirus DNA used was in the form of recombinant plasmids. Genomes of the Denmark strain of Monkeypox and the Harvey strain of Variola have been cloned as smaller fragments (Hamilton et al. 1985). Upon HindIII restriction enzyme digestion Variola and Monkeypox DNA generate seventeen and nineteen fragments respectively. Except for the cross-linked termini, all these fragments have been cloned into various vectors. The C and D fragments of Variola are in pAT153 and the C and E fragments of Monkeypox are in pBR329 and pBR328 respectively. Vectors used for cloning purposes include pBR329 and pUC19. All plasmids were prepared in the same way for large scale extractions.

The method used for preparing large quantities of plasmid DNA was based on that described by Greenaway and Dale (1983). The chloramphenicol amplification step was omitted and, for a greater yield, the incubation times were altered.

A single bacterial colony was picked from a Luria-agar plate containing the appropriate antibiotic (ampicillin was used for most of tne preparations) and 10ml of Luria broth containing ampicillin was inoculated. This was incubated at 37^{0}C, shaking, for approximately 8 hours.

All 10ml of the culture was inoculated into 200ml of pre-warmed Luria broth (containing antibiotic) and left shaking at 37^{0}C overnight (+15 hours). To maximize surface aeration a flask of 500ml capacity or larger was used for this step. Shaking in a smaller flask resulted in a lower final plasmid yield.

The cells were harvested by centrifugation at 6000rpm for 15

minutes at 4^{0}C in a JA-10 rotor. All culture supernatant was removed.

The cells were resuspended in 2.6ml of 25% sucrose in 0.05M Tris HCl pH8.0 and transferred to a polypropylene centrifuge tube. The cells were maintained on ice.

A spatula tip of powdered lysozyme was added to the cell suspension which was subsequently swirled on ice for approximately 5 minutes.

2.6ml of 0.25M Na_2EDTA pH8 was added followed by swirling on ice for 5 minutes.

4ml of Triton\DOC solution (1% Triton, 0.4% Sodium deoxycholate in 0.01M Tris, 0.001M Na_2EDTA, pH8) was rapidly added from a 10ml pipette. Even mixing was ensured by drawing up and expelling three times. The mixture was left on ice for 20 to 30 minutes.

Cell debris and chromosomal DNA was pelleted by centrifugation at 15000rpm for 45 minutes at 4^{0}C in a JA20 rotor. Because the bacterial chromosome is attached to the cell membrane it will be pelleted together with the membrane in a detergent solution.

The cleared lysate (supernatant) was decanted into a plastic container and the volume was made up to 10ml with Triton\DOC solution.

Plasmid Purification by Isopycnic Gradient Centrifugation

Isopycnic gradient centrifugation is the separation of macromolecules according to their densities. For DNA separation a Caesium Chloride solution of 1.6g\ml density is made.

9.5g of CsCl was dissolved in the 10ml of DNA solution (0.95g\ml) and 0.2ml of ethidium bromide (10mg\ml) was added. The solution was transferred to two polyallomer tubes suitable for centrifugation in a VTi65 rotor and the tubes were heat sealed. Centrifugation was performed at 50000rpm for 20 hrs at 20^{0}C.

The DNA attained an equilibrium position in the tube at a level corresponding to its density and was visualized under an ulltraviolet light of 375nm wavelength. Usually two bands were visible, the higher band corresponding to chromosomal and nicked plasmid DNA, and the lower band corresponding to covalently closed supercoiled plasmid DNA. Ethidium bromide intercalates into DNA in the supercoiled configuration less efficiently than into DNA in the relaxed configuration; the relaxed form of DNA containing more ethidium bromide is less dense than DNA with less ethidium bromide. This difference in density due to the ethidium bromide intercalation allows supercoiled DNA to be separated from DNA in the relaxed configuration.

The bands were eluted through the bottom of the tube by puncturing both the top and bottom of the tube. To prevent RNA contamination an attempt was made to elute the DNA out from the side of the tube with a syringe, but this caused large losses.

Removal of Ethidium Bromide

Ethidium Bromide was removed by adding an equal volume of isoamyl alcohol, vortexing the tube and separating the two phases by a two-minute centrifugation step. The upper organic layer was discarded and the procedure was repeated until there was no pink colour visible in either phase.

Removal of excess salt (CsCl) and precipitation of DNA

Initially all DNA preparations were dialysed in TE buffer (0.1M Tris, 0.01M EDTA pH8). The buffer was changed twice at four hourly intervals and then left at 4^0C overnight. After removing the DNA solution from the dialysis bags, the volume was determined. Half the volume of phenol was added and well mixed with the DNA solution. Half the initial DNA volume of chloroform:isoamylalcohol was then well mixed with the DNA/phenol solution. The tube was spun to separate the phases and the top aqueous phase was re-extracted with phenol and chloroform:isoamylalcohol. The volume of the aqueous phase was determined and one tenth of the volume of 4M LiCl was added (to make a final concentration of 0.4M LiCl). Two and a half times the volume of absolute ethanol was added and the tube was placed in dry ice (-70^0C) for two hours.

DNA was also precipitated with Na or K acetate but LiCl was preferred as it is more soluble in ethanol and is more selective in precipitating DNA - as opposed to RNA.

The DNA was pelleted by centrifugation in a JA-20 rotor at 5000 rpm for 30 minutes at 4^0C. The pellet was washed with 70% ethanol, spun for 20 minutes at 5000 rpm at 4^0C, dried in a speedivac and resuspended in sterile distilled water.

Not all DNA preparations were dialysed. The omission of the

dialysis step saved time and reduced DNA losses, with no apparent disadvantages. After removal of ethidium bromide with isoamylalcohol a threefold excess of sterile distilled water was added to the DNA preparation, two and a half times the volume of absolute ethanol was added and the DNA was precipitated as described above. No extra salt was added as the CsCl concentration was high enough to precipitate the DNA.

Plasmids were stored at 4^{0}C in sterile distilled water. Water was used instead of TE to prevent alterations in buffer concentrations for specific reactions - such as restriction enzyme digestion with enzymes requiring a low salt buffer. The plasmids remained stable for over eighteen months.

Yields were in the order of 100ug (0.1mg/ml).

3.2 MINIPREPARATIONS OF PLASMID DNA

When constructing recombinant plasmids a quick and easy method of isolating plasmids is required for screening many colonies. The DNA must, however, be pure enough for restriction enzyme digestion. The most successful method used, although not the quickest, was based on that of Birnboim and Doly. Three other methods were attempted, but the results were not reliable as the DNA often appeared denatured and was not always digested properly with restriction enzymes.

3.2.1 MODIFICATION OF THE BIRNBOIM AND DOLY PROCEDURE (1979)

From a 3ml overnight culture of E.Coli cells harboring the plasmid to be isolated, 1.5ml was centrifuged for one minute in a microfuge. The cell pellet was resuspended in 180ul 50mM glucose, 10mM EDTA, 25mM Tris.Cl pH8.0. A spatula tip of powdered lysozyme was added to the tube which was then vortexed. The tube was left at room temperature for 5 minutes.

400ul 0.2M NaOH, 1% SDS was added, the tube gently vortexed and left on ice for 5 minutes. The solution cleared at this stage - as protein and DNA denatured and dissolved.

Chromosomal DNA and protein was then precipitated with 300ul $KOCH_3$ (3M K, 5M acetate) for 10 to 60 minutes on ice.

After a 15 minute centrifugation step 450ul isopropanol was added to and well mixed with 750ul of the supernatant. Plasmid DNA (and RNA) was pelleted by centrifuging for 5 minutes in a microfuge.

The pellet was washed with 70% ethanol by spinning the tube in the opposite orientation such that the pellet migrates through

the 70% ethanol. The supernatant was discarded and the pellet dried in a vacuum.

After dissolving the DNA in 200ul TE + 0.15M NaCl, 200ul phenol was added. The tube was vortexed, 200ul chloroform/isoamylalcohol was added and mixed by vortexing. The phases were separated by centrifuging for two minutes.

The upper aqueous phase was kept and 500ul absolute ethanol was added to it. After 5 minutes incubation in dry ice (-70^0C) the tube was spun for 5 minutes. the DNA pellet was washed with 70% ethanol, vacuum dried and dissolved in 100ul sterile distilled water.

3.2.2 METHOD OF KADU AND LIU (1981)

An overnight culture of cells was pelletted (by microfuging for one minute) and resuspended in 150ul of lysis buffer (4% SDS in 50mM Tris pH12.6).

This was incubated at 100^0C for 30 seconds.

A phenol/chloroform extraction was performed and the phases separated by centrifuging for 5 minutes.

The aqueous phase containing the purified plasmid DNA was retained.

3.2.3 METHOD OF HOLMES AND QUIGLEY (1981)

This method was attempted to isolate DNA for sequencing. The Birnboim and Doly method, however, proved to be better and was therefore used more routinely.

A 10ml overnight culture of cells was centrifuged for 5 minutes at 6000 rpm. The cell pellet was resuspended in 0.7ml STET (50mM Tris-Cl pH8, 50mM EDTA pH8, 8% sucrose, 5% TritonX-100) and transferred to a microfuge tube.

Dry lysozyme was added to the suspension which was subsequently heated for 40 seconds at 100^{0}C. After a 10 minute centrifugation step the supernatant was retained.

50ul RNase A (1mg/ml) was added and the solution left at room temperature for 15 minutes.

A phenol/chloroform extraction was performed (section 3.3) and the DNA precipitated with 10ul 3M NaAcetate and 840ul isopropanol. The solution was left on dry ice for 15 minutes, spun for 15 minutes, dried and resuspended in 360ul H_2O + 40ul 3M NaAcetate. 1ml of ethanol was added and the DNA was precipitated (section 3.3). The plasmid DMA was finally resuspended in 50ul sterile distilled water.

3.2.4 LiCl-BOILING METHOD (Wilimzig 1985)

An overnight culture was microfuged for one minute to pellet the cells harbouring the plasmid. The cells were resuspended in 100ul TELT-buffer (50mM Tris-HCl, 62.5mM EDTA, 0.4% Triton X-100, 2.5M LiCl, pH7.5).

10ul of a 10mg/ml solution of lysozyme, freshly prepared, was added and vortexed. The tube was incubated at 100^{0}C for one minute and then cooled on ice for 5 minutes.

After centrifuging for 8 minutes the DNA in the supernatant was precipitated with salt and ethanol (Refer to section 3.3).

Comments

Method 3.2.1 was the standard method used. The DNA yields were high and the quality good. DNA prepared in this manner was used for restriction enzyme digestion, nick translation, ExonucleaseIII digestion followed by Mung bean nuclease digestion and sequencing by the Sanger/dideoxy method.

Methods 3.2.2 and 3.2.4 were attempted as a short cut in screening for recombinant plasmids; but the first method proved to be more reliable and was used for this purpose, despite its disadvantage of being more time-consuming.

3.3 PHENOL / CHLOROFORM EXTRACTION AND DNA PRECIPITATION

Phenol/ chloroform extractions were performed to purify DNA. The volume of the DNA solution was determined. Half of this volume of phenol was added, the solution vortexed, the same volume of chloroform:isoamylalcohol (ie. half the original DNA volume) was added and the solution was vortexed. The phases were separated by centrifugation - 2 minutes - and the upper aqueous phase was retained. This process was usually repeated at least once.

DNA was precipitated by adding one tenth of its volume 4M LiCl and 2.5 times its volume absolute ethanol. After five minutes on dry ice the tube was centrifuged for 5 minutes. The pellet was washed with 70% ethanol by spinning the tube in its opposite orientation so that the DNA migrated through the 70% ethanol.

The DNA pellet was dried in a vacuum drier and resuspended in sterile distilled water.

3.4 RESTRICTION ENZYME DIGESTION AND GEL ELECTROPHORESIS

INTRODUCTION

Plasmids are identified by their sizes and restriction enzyme profiles. Vectors for cloning purposes are designed to have many unique restriction enzyme sites. Digestion of the vector with one of these enzymes linearizes the plasmid. If the vector carries an insert, for example, a HindIII fragment of poxvirus DNA, the insert is cut out upon HindIII digestion and two linear fragments are generated - vector + insert (See figure 3.4.2).

DNA fragments are separated by agarose gel electrophoresis. Under the influence of an electric current, at neutral pH, the negatively charged DNA will migrate through a slab gel at a rate inversely proportionally to the size of the DNA molecule. Unless there is a string of A's and T's in the DNA chain, the rate of migration is independent of base sequence. A string of A's in one chain and T's in the other causes kinks in the molecule which alter the DNA mobility. Double stranded DNA ranging from 70bp to 800kb can be separated in agarose gels of 3% to 0.1% respectively. DNA of 0.5 to 20kb was separated in 1% or 0.8% agarose gels.

DNA fragments were visualized after ethidium bromide staining on a UV-illuminator (figure 3.4.1). UV light of 302nm is absorbed and re-emitted at 590nm. A source of longer wavelength produces flourescence less efficiently, whereas shorter wavelengths cause photo-nicking, dimerization and rapid bleaching of the DNA-ethidium bromide complex (Sealey and Southern 1982).

Fragment sizes were estimated relative to linear double stranded DNA of known molecular weight run alongside the test fragments.

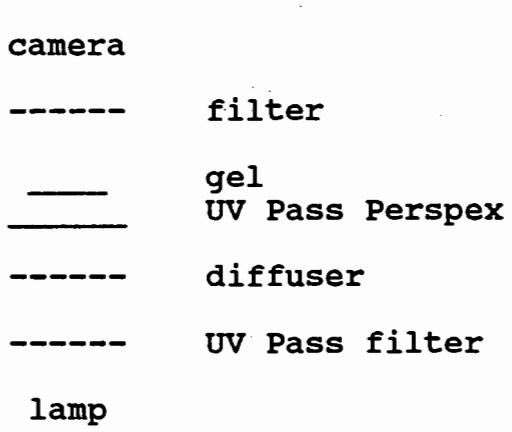


Fig 3.4.1 Diagram of UV-transilluminator

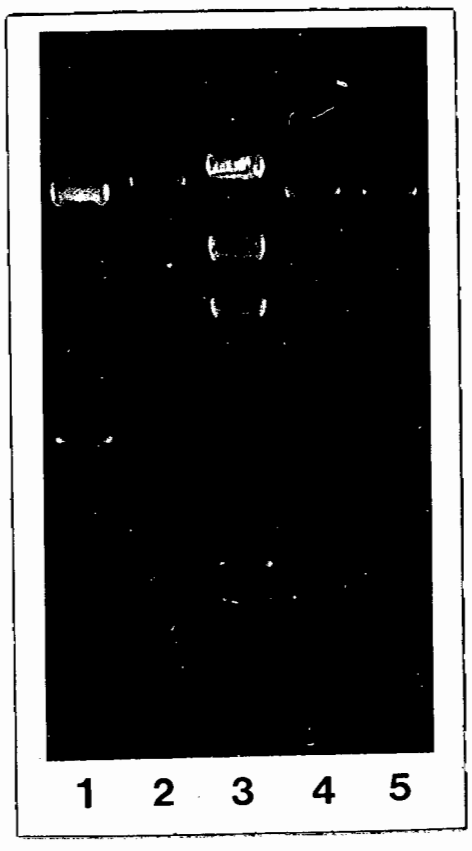


Figure 3.4.2 HindIII digests of recombinant plasmids subjected to 1% agarose gel electrophoresis at approximately 2V/cm. Lane 3 is a size marker of lambda DNA digested with HindIII; lanes 1 and 2 are the C fragments of Variola and Monkeypox respectively; lane 4 is Variola D and lane 5 is Monkeypox E.

3.4.1 RESTRICTION ENZYME DIGESTION

Restriction enzyme digestion was performed as recommended by the suppliers. Care was taken to ensure the correct salt concentrations were used and the glycerol content was low enough to prevent inhibition of digestion. Enzyme volumes were no greater than one tenth the volume of the digestion mix.

Enzymes such as EcoRI and PstI sometimes exhibited star activity. To reduce this effect digestion times were minimal for these enzymes.

For double digestions, two enzymes requiring similar salt conditions were applied to the same tube. If however, the enzymes require very different digestion conditions, the enzyme requiring the lowest salt concentration was used first, the sample was phenol extracted and precipitated (section 3.3) and the second enzyme digestion was performed.

Enzyme reactions were terminated by adding 10 X Stop buffer (50% v/v glycerol, 100mM Na_2EDTA pH8, 1% w/v SDS, 0.1% w/v bromophenol blue, 0.1% cyanol) to the mix. The fragments generated were separated by agarose gel electrophoresis and stained with ethidium bromide.

3.4.2 AGAROSE GEL ELECTROPHORESIS

0.8% or 1% gels were made by dissolving the appropriate amount of agarose in TAE buffer (0.04M Tris-acetate, 0.001M EDTA). This was done in a microwave oven. The gel was poured once the agarose had cooled to 56^{0}C. The gels were poured approximately 3mm thick with the slot former already in position. Once the gel had set - after about 30 minutes - the slot former was removed and the gel placed in a tank of TAE buffer. The gel was totally immersed in buffer.

After the samples were loaded into the wells a potential difference of 2 to 10 volts per cm was applied to the chamber. For good separation of bands a lower voltage was applied for a longer time. A larger voltage was applied for a shorter time for quicker checking gels.

The DNA bands were visualized by staining the gel with a 0.5ug/ml ethidium bromide solution - one drop of a stock solution of 10mg/ml ethidium bromide was added to one litre of TAE buffer. The gel was stained for approximately 30 minutes. If overstained, it was destained in distilled water for a further 30 minutes.
If accurate DNA sizes were not required, a drop of 0.1mg/ml ethidium bromide was added to the gel before pouring it. Ethidium bromide intercalates into the DNA molecule; this interaction may affect the rate of migration of the DNA through the gel - hence the staining of gels after electrophoresis for accurate sizing of fragments.

3.5 DNA ELUTION FROM AGAROSE GELS

INTRODUCTION

Smaller fragments of the cloned poxvirus DNA were often required for purposes such as subcloning or nick translation (chapter 4). Many methods have been established for eluting DNA from agarose gels, possibly because none of them are ideal. Generally, small fragments (<6kb) are eluted quite efficiently, but larger fragments (>10kb) have a low recovery yield. The more conventional method of electroelution proved to be finicky and not very efficient. DEAE-cellulose paper was used to bind DNA which was consequently eluted in high salt buffer. This method was longer, but eluted small fragments more efficiently. The recovery of large fragments in this way was bad. Plasmids were eluted very quickly, giving reasonable yields of smaller fragments, but low yields of larger fragments, by cutting the gel very finely, adding phenol, freezing, and recovering the DNA by centrifugation. The method finally used for eluting fragments of all sizes, were those employing a) low gelling temperature agarose and b) a biotrap apparatus.

3.5.1 ELECTROELUTION (Maniatis et al. 1982)

A dialysis bag was rinsed with distilled water and filled with TAE buffer. The required DNA band was cut out of the gel with a blade and placed into the dialysis tube using forceps. After air bubbles and excess TAE buffer had been squeezed out of the dialysis bag, the tube was closed with a clip. It was placed horizontally in an electrophoresis bath and a potential difference of 100V was applied for two hours. The polarity was reversed for five minutes to prevent DNA sticking to the dialysis tubing, after which the contents of the bag were removed. The DNA

was treated with phenol and precipitated as described in section 3.3.

3.5.2 ELUTION USING DEAE-CELLULOSE FILTERS (Dretzen et al. 1981)

DEAE-cellulose was prepared by soaking small pieces of filter in 2.5M NaCl for 4 hours, followed by several washes in distilled water. Two filters were removed and soaked in TAE buffer.

The gel was cut above and below the band of interest and the filters were placed into the slits. A potential difference of 80V was applied for fifteen minutes during which time the DNA of interest bound to the one filter. The other filter prevented any higher molecular weight fragments from contaminating the required fragment. Using a blade the filter with the bound DNA was shredded and incubated at 37^{0}C in 300ul elution buffer (20mM Tris-HCl (pH7.5), 1mM EDTA, 1.5M NaCl) overnight.

The tube was centrifuged for five minutes and the supernatant kept. 300ul elution buffer was added and the tube was centrifuged again for five minutes. The supernatant was retained, 300ul chloroform:isoamylalcohol added, the sample vortexed and centrifuged for two minutes to separate the phases. The top aqueous phase was treated with salt and ethanol as previously described to precipitate the DNA (section 3.3).

3.5.3 ELUTION BY FREEZING (Seth 1984)

The DNA band required was excised from a gel using a blade and cut up very finely. An equal volume of phenol was mixed with the sample which was then placed at -70^{0}C for at least 10 minutes.

The frozen sample was centrifuged for 10 minutes, the supernatant being retained.

A double phenol/chloroform extraction was performed and the DNA was precipitated with LiCl and ethanol as previously described (section 3.3)

3.5.4 DNA ELUTION FROM LOW GELLING TEMPERATURE AGAROSE (Maniatis et al. 1982)

DNA was electrophoresed through a low gelling temperature agarose in the same way as through the ordinany agarose. The DNA band required was cut out of the gel and a large excess of 20mM Tris-HCl (pH8.0), 1mM EDTA was added. The sample was heated at 65^{0}C until the agarose had melted (approximately 10 minutes). An equal volume of phenol was added, vortexed and centrifuged for 5 minutes. The aqueous layer was re-extracted with phenol and chloroform and the DNA was precipitated as described in section 3.3.

3.5.5 ELECTROELUTION OF DNA INTO A BIOTRAP

This method was performed according to instructions sent with the apparatus supplied by Schleicher and Schuell. For elution of the DNA two types of membrane are implicated - BT 1 and BT 2. BT 1 retains charged molecules of 5000d or larger in an electric field of 10V/cm without affecting the migration of buffer ions. BT 2 has a dual nature; in an active electric field ions pass through it readily, but, without an electric field the buffer is prevented from passing through it.

The biotrap apparatus was placed into an electrophoresis tank. The membranes were placed in positions shown in figure 3.5.1. BT 1 membranes were placed in positions A and C and a BT 2 membrane was placed in position B. A gel slice containing the required DNA band was placed in the apparatus and sufficient TAE buffer to cover the gel was added. A potential difference of 200V was applied for 3.5 hours (10V/cm), after which the polarity was reversed for 20 seconds. The voltage was switched off and the DNA removed from the area between A and B (figure 3.5.1). No phenol extractions were necessary, but sometimes the DNA had to be concentrated by salt and ethanol precipitation (section 3.3).

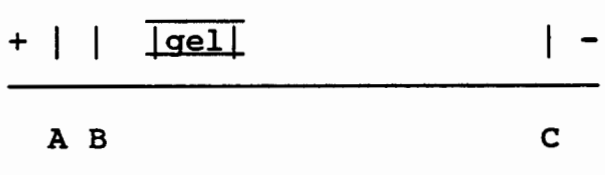


Figure 3.5.1 Diagram of biotrap apparatus. A and C are BT 1 membranes and B is a BT 2 membrane.

CHAPTER 4 MAPPING RESTRICTION SITES ON VARIOLA HindIII D AND MONKEYPOX HindIII E

4.1 CLEAVAGE PATTERNS OF DIFFERENT RESTRICTION ENDONUCLEASES

Variola HindIII D (VarD) and Monkeypox HindIII E (MPE) were mapped for comparison. The two fragments are cloned into the HindIII sites of pAT153 and pBR328 respectively. Maps of the two vectors are shown in appendix 1. VarD had been extensively mapped for restriction enzyme sites (Hamilton et al. 1985), so the same restriction enzymes were used to map the corresponding area in Monkeypox. The number of fragments generated by single restriction enzymes from VarD and MPE have been tabulated in table 4.1.1 and the electrophoretic profiles are displayed in figures 4.1.1 and 4.1.2.

Determination of fragment sizes was by comparison with known fragment sizes. VarD digested with EcoRI generates a ladder of evenly spaced bands and was used as a molecular weight marker. Phage lambda DNA cut with HindIII was also used (see appendix 2 for fragment sizes). Because the VarD marker bands were so well spaced and the total size of the MPE fragment was known, fragment sizes could be accurately estimated by inspection. For more accurate determination of fragment sizes a graph was drawn of $\log_{10}$ molecular weight vs. distance migrated for the markers. The unknown fragment sizes were read off the graph according to their distances migrated from the origin. See appendix 3 for an example.

SphI does not cleave either plasmid. AvaI generates two fragments from MPE (15.5 and 3.9kb); since there is one AvaI site in pBR328, one site is present in the "pox" DNA. No AvaI sites are

present in VarD.

No PstI and SalI sites are present in the Monkeypox HindIII E fragment as opposed to the single PstI and two SalI sites in the Variola HindIII D fragment.

BamHI generates three fragments from MPE, but only two from VarD; the presence of one BamHI site in both vectors means there are two sites in the Monkeypox fragment compared with a single site in the Variola fragment. MPE also has an extra EcoRI site resulting in 10 fragments upon EcoRI digestion as opposed to the nine fragments generated in VarD.

Neither pBR328 nor pAT153 have a HpaI site; HpaI cleaves MPE into three fragments indicating the presence of three sites in the "E" fragment - compared to two sites in VarD.

NcoI generates 8 fragments from MPE; pBR328 has one site so 7 sites must be present in the "E" fragment. 6 NcoI fragments were generated from VarD; pAT153 has no NcoI sites so all the sites are within the "D" fragment.

Table 4.1.1 Restriction Endonuclease Fragments of Recombinant Plasmids Variola HindIII D and Monkeypox HindIII E.(*).

	Enzyme	Fragment sizes (kb)	No. fragments	No. sites in vector	No. sites in "pox" DNA
VarD	AvaI	Not digested	-	0	0
MPE	AvaI	15.5;3.9.	2	1	1
VarD	BamHI	10.0;8.6.	2	1	1
MPE	BamHI	7.25;6.7;5.36	3	1	2
VarD	EcoRI	5.8;4.05;2.85; 1.9;1.3;1.0;1.0; 0.75;0.2.	9	1	8
MPE	EcoRI	5.7;3.2;2.85; 2.7;1.3;1.0;1.0; 0.85;0.75;0.2.	10	1	9
VarD	HindIII	15.3;3.6.	2	1	1
MPE	HindIII	14.4;4.9.	2	1	1
VarD	HpaI	8.6;8.6.	2	0	2
MPE	HpaI	10.7;7.35;1.25.	3	0	3
VarD	NcoI	9.9;2.7;2.55; 1.3;1.3;0.85.	6	0	6
MPE	NcoI	6.6;3.3;2.7;2.2; 1.3;1.3;1.2;0.85.	8	1	7
VarD	PstI	16.0;2.7.	2	1	1
MPE	PstI	19.3.	1	1	0
VarD	SalI	9.2;6.0;3.4.	3	1	2
MPE	SalI	19.3.	1	1	0
VarD	SphI	Not digested	-	0	0
MPE	SphI	Not digested	-	0	0

* The Variola HindIII D fragment is in pAT153 and Monkeypox HindIII E is in pBR328.
Most of the VarD fragment sizes were previously established by Hamilton et al. 1985).

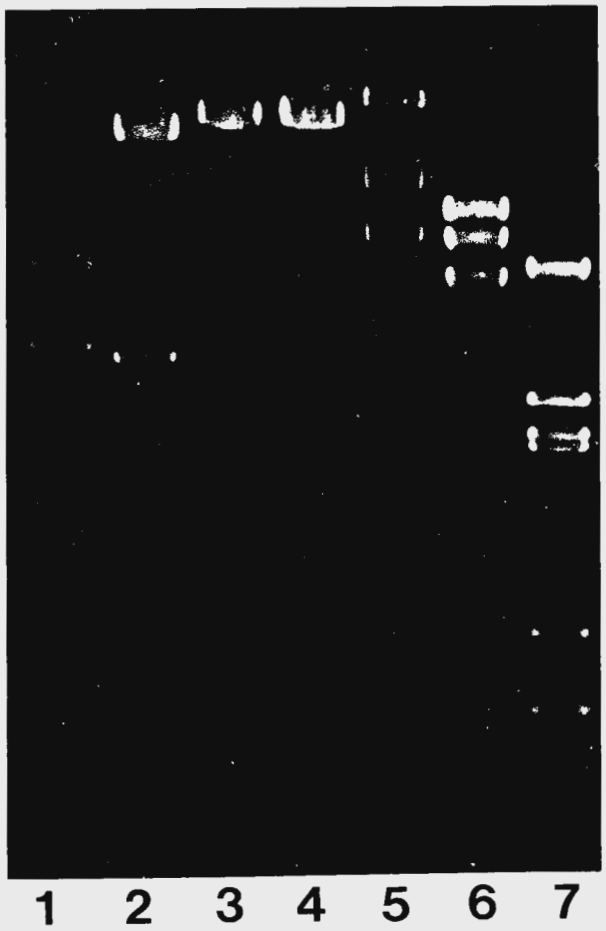


Figure 4.1.1 Electrophoretic profiles of MPE digested with various restriction enzymes. The fragments were separated in a 1% agarose gel at approximately 2V/cm. Lane 1 is VarD digested with EcoRI and lane 5 is lambda DNA digested with HindIII. These were used as size markers. Lanes 2, 3 and 4 are MPE digested with AvaI, PstI and SalI respectively; and lanes 6 and 7 are MPE digested with BamHI and EcoRI respectively.

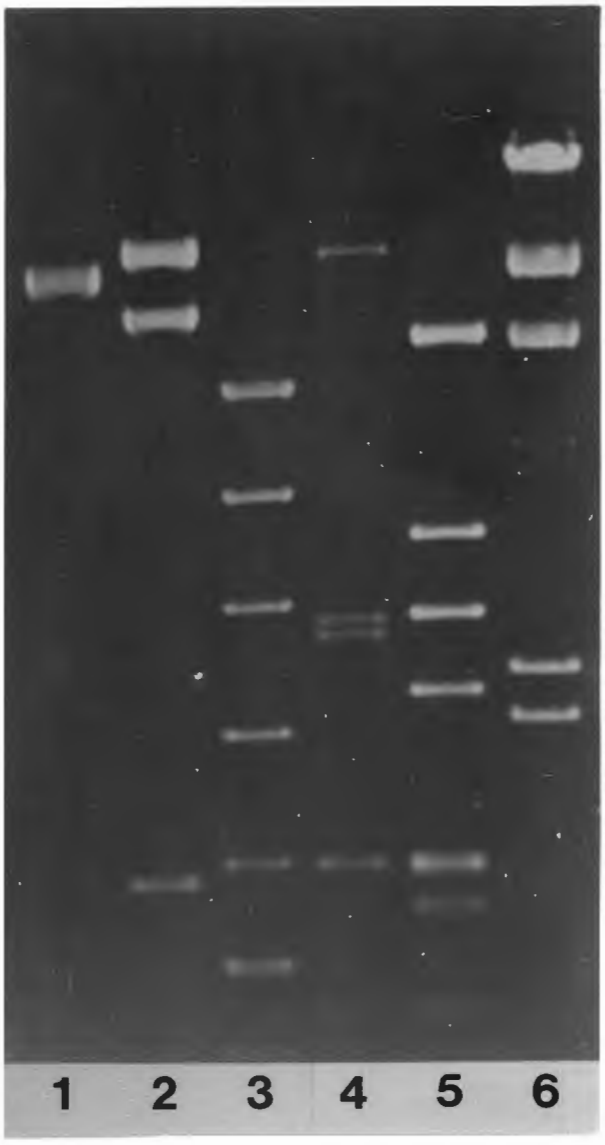


Figure 4.1.2 Comparison of digestion profiles of VarD and MPE upon treatment with HpaI and NcoI. Lanes 1 and 2 are HpaI digests of VarD and MPE respectively; and lanes 4 and 5 are NcoI digests of VarD and MPE respectively. Lane 3 is VarD digested with EcoRI and lane 6 is lambda DNA treated with HindIII. The fragments were separated in a 1% agarose gel at approximately 2V/cm.

4.2 LOCATION OF SITES BY SIZING FRAGMENTS IN SINGLE AND DOUBLE DIGESTS

The map which had already been established for VarD is shown in figure 4.2.1. The various single and double digests performed to map MPE are tabulated in tables 4.2.1 to 4.2.4. Gel photographs are presented in figures 4.2.8 to 4.2.10.

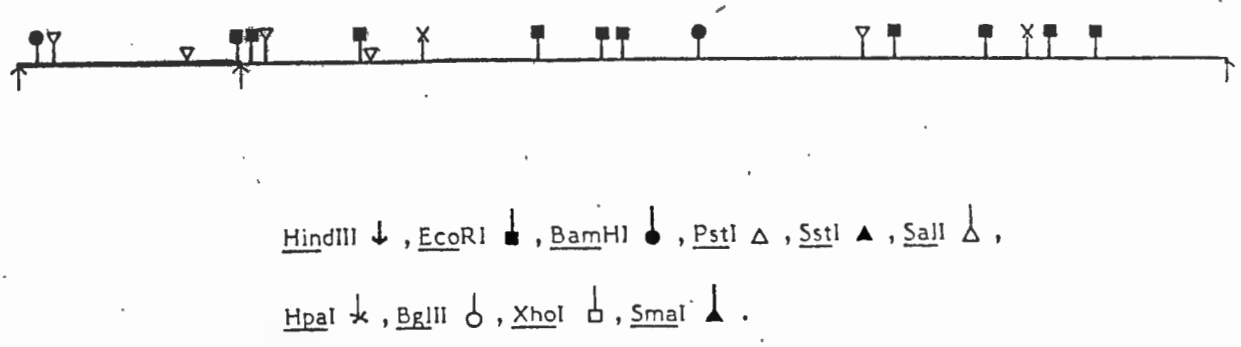


Figure 4.2.1 Restriction enzyme map of VarD established by Hamilton *et al.* 1985. The bolder section represents the cloning vector, pAT153.

4.2.1 LOCATION OF AvaI SITE IN MPE

AvaI generated two fragments of 15.5 and 3.9 kb from the Monkeypox HindIII E plasmid. An AvaI site is present at position 1394 in pBR328. Upon digestion with HindIII + AvaI fragments of 14.0, 3.5, 1.4 and 0.4 kb were generated (table 4.2.1). The 3.5 and 1.4 kb fragments are vector sequences so the AvaI site in the insert can be in one position only - 14.0kb from the one HindIII site and 0.4kb from the other HindIII site; simultaneously it lies 3.9kb from the vector AvaI site (or 15.5kb in the opposite direction), (figure 4.2.2).

Table 4.2.1 Fragment Sizes used in Mapping AvaI Site in MPE

Enzyme used	Fragment sizes (kb)
AvaI	15.5; 3.9.
AvaI + HindIII	14.0; 3.5; 1.4; 0.4.

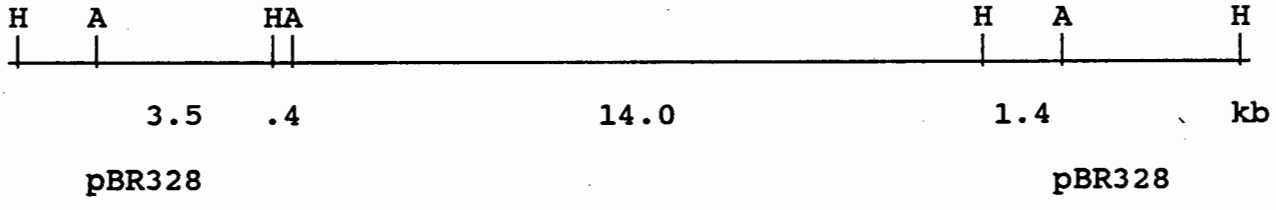


Figure 4.2.2 Location of AvaI sites in MPE. A=Ava I; H=HindIII.

4.2.2 LOCATION OF BamHI SITES IN MPE

BamHI generates three fragments of 7.3, 6.7 and 5.4 kb. When digested with HindIII + BamHI the 6.7kb fragment remains unaltered, but the other two fragments are no longer present. The 7.3kb BamHI fragment is cleaved by HindIII into fragments of 7.0 and 0.3 kb in length. A BamHI site is present 0.3kb from the HindIII site in pBR328; a BamHI site must therefore be present 7.0kb from the HindIII site on the right hand side (figure 4.2.3). The 5.4kb BamHI fragment was cleaved by HindIII into fragments of 4.6 and 0.8 kb. The 4.6kb fragment corresponds to the size of the vector BamHI/HindIII fragment; a BamHI site is therefore present 0.8kb from the left end (figure 4.2.3). The unaltered 6.7kb fragment lies in the central region.

Table 4.2.2 Fragment Sizes used for Mapping BamHI sites in MPE

Enzymes used	Sizes of fragments generated (kb)
BamHI	7.3; 6.7; 5.4.
BamHI + HindIII	7.0; 6.7; 4.6; 0.8; 0.3.

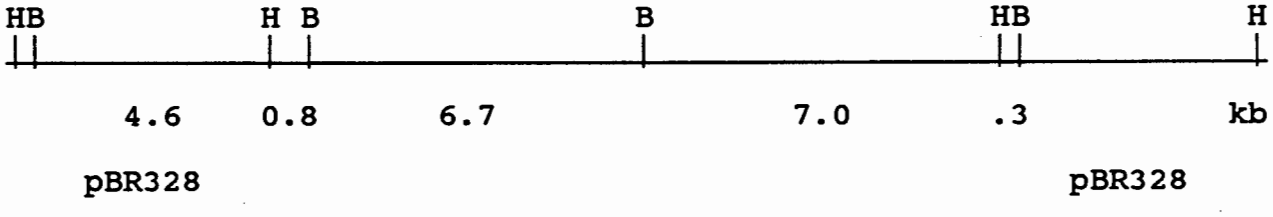


Figure 4.2.3 Location of BamHI sites in MPE. B=BamHI; H=HindIII.

4.2.3 LOCATION OF HpaI SITES IN MPE

Table 4.2.3 gives the sizes of the fragments generated from MPE by the enzymes HpaI, HpaI + HindIII, HpaI + BamHI and BamHI. HpaI generates three fragments of 10.7, 7.35 and 1.25 kb. The HpaI/HindIII double digest shows that the largest HpaI fragment is cut ints three pieces of which the largest 4.9kb fragment is vector. A HpaI site is therefore placed approximately 2.9kb from each end of the insert, leaving the third site 1.25kb internal to either of these first two sites (figure 4.2.4).

BamHI leaves the 1.25kb fragment unaltered, but cleaves the 7.35 and 10.7 kb HpaI fragments. With two HpaI sites established approximately 2.9kb from each end of the insert, the 7.3kb BamHI fragment would be cleaved into two fragments of approximately 3.2kb and 4.1kb (figure 4.2.5). Fragments were generated of 3.15, 3.2 and 4.2 kb. The HpaI site estimated as 2.9kb from the right end is either 2.85 or 2.9kb from the end. Because a fragment of 4.2kb was generated in the double digest - instead of the expected 4.1kb - the HpaI site was positioned 2.85kb from the right end. This location of the HpaI site would account for the 3.15kb fragment generated upon digestion with BamHI + HpaI. (Figure 4.2.5).

The HpaI site 2.9kb from the left end would give rise to a HpaI/BamHI fragment of 2.1kb since the BamHI site divides the HpaI/HindIII fragment of 2.9kb into 0.8 + 2.1 kb.

Fragments from the HpaI/BamHI double digest not accounted for are 3.2 and 1.25 kb. It has been established that the remaining HpaI site is 1.25kb internal to one of the other two HpaI sites (figure 4.2.4). If it were internal to the right HpaI site a BamHI/HpaI fragment of 4.15kb would not be generated. Instead one

would expect fragments of 1.25, 2.85 and 4.45kb. No band of 4.45kb was present. If the HpaI site were 1.25kb internal to to the left HpaI site, fragments of 1.25, 3.2 and 4.15kb would be generated. Fragments of these sizes were observed. The third HpaI site was therefore located 1.25kb internal to the left HpaI site. (Figure 4.2.6).

Table 4.2.3 Fragment Sizes used for Mapping HpaI Sites in MPE

Enzymes used	Sizes of fragments generated (kb)
HpaI	10.7; 7.35; 1.25.
HpaI + HindIII	7.35; 4.9; 2.9; 2.85; 1.25.
HpaI + BamHI	5.4; 4.2; 3.2; 3.15; 2.1; 1.25.
BamHI	7.3; 6.7; 5.4.

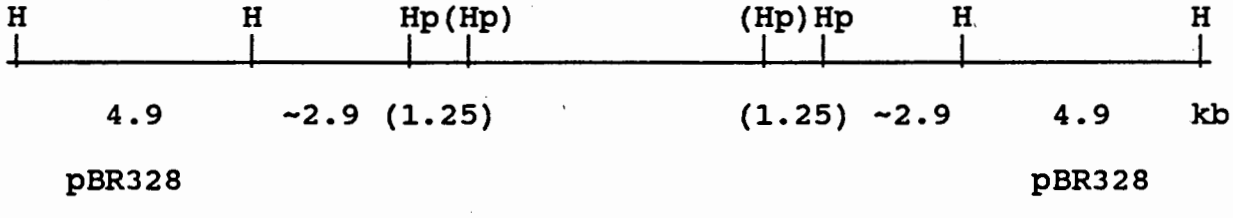


Figure 4.2.4 Mapping of HpaI sites. H=HindIII; Hp=HpaI.

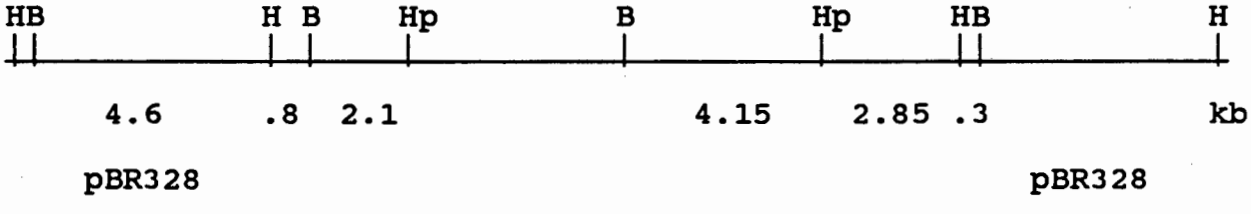


Figure 4.2.5 Mapping of HpaI sites. H=HindIII; Hp=HpaI.

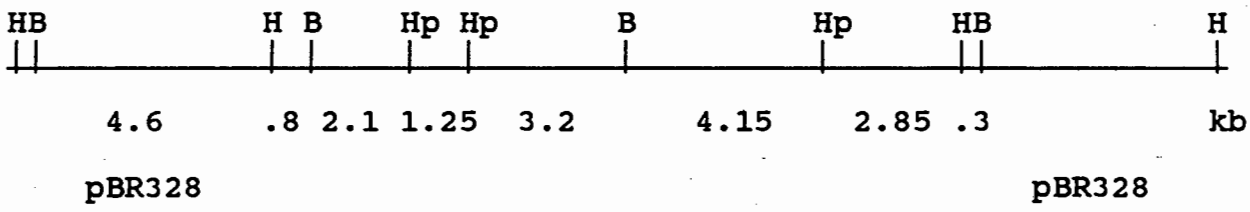


Figure 4.2.6 Location of HpaI sites in MPE. H=HindIII; H=HpaI; B=BamHI.

4.2.4 LOCATION OF TWO EcoRI SITES IN MPE

The fragment sizes used for calculating the positions of some of the EcoRI sites are presented in table 4.2.4. EcoRI generated 10 fragments from MPE. HindIII cleaved the 5.7 and 3.2 kb fragments, generating fragments of 3.65, 2.0, 1.95 and 1.2 kb. The 3.65kb fragment must have arisen from the 5.7kb EcoRI fragment. It corresponds to the size of the EcoRI/HindIII fragment of pBR328. An EcoRI site must therefore be present 5.7-3.65 = 2.05 kb from the right end (figure 4.2.7). A fragment of 2.0kb was produced. The 1.2kb fragment from the EcoRI/HindIII digest corresponds to the other vector EcoRI/HindIII fragment. The EcoRI site at the left end of the insert must be 3.2-1.2 = 2 kb from the HindIII site. A fragment of 1.95kb was produced. Two EcoRI sites can therefore be estimated 2kb from each end.

Table 4.2.4 Fragment Sizes used for EcoRI Mapping of MPE

Enzymes used	Fragment sizes generated (kb)
EcoRI	5.7; 3.2; 2.85; 2.7; 1.3; 1.0; 1.0; 0.85; 0.75; 0.2.
EcoRI + HindIII	3.65; 2.85; 2.7; 2.0; 1.9; 1.3; 1.0; 1.0; 0.85; 0.75; 0.2.

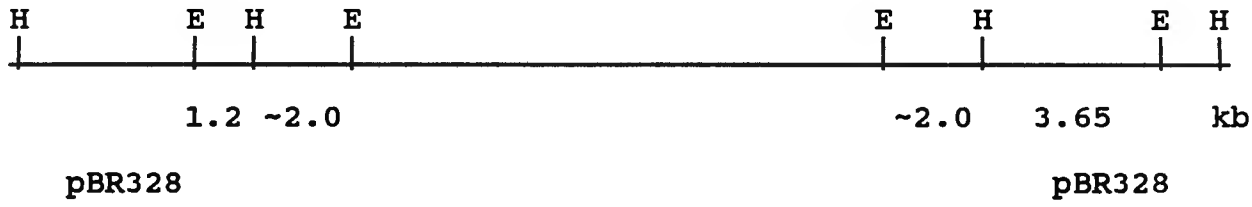


Figure 4.2.7 Location of two EcoRI sites in MPE. H=HindIII; E=EcoRI.

Figure 4.2.8
Gel electrophoresis of MPE digested with AvaI (lane 1) and AvaI + HindIII (lane 2). Lambda DNA cut with HindIII was used as a size marker (lane 3). The fragments were separated at approximately 1.5V/cm in a 1% agarose gel.

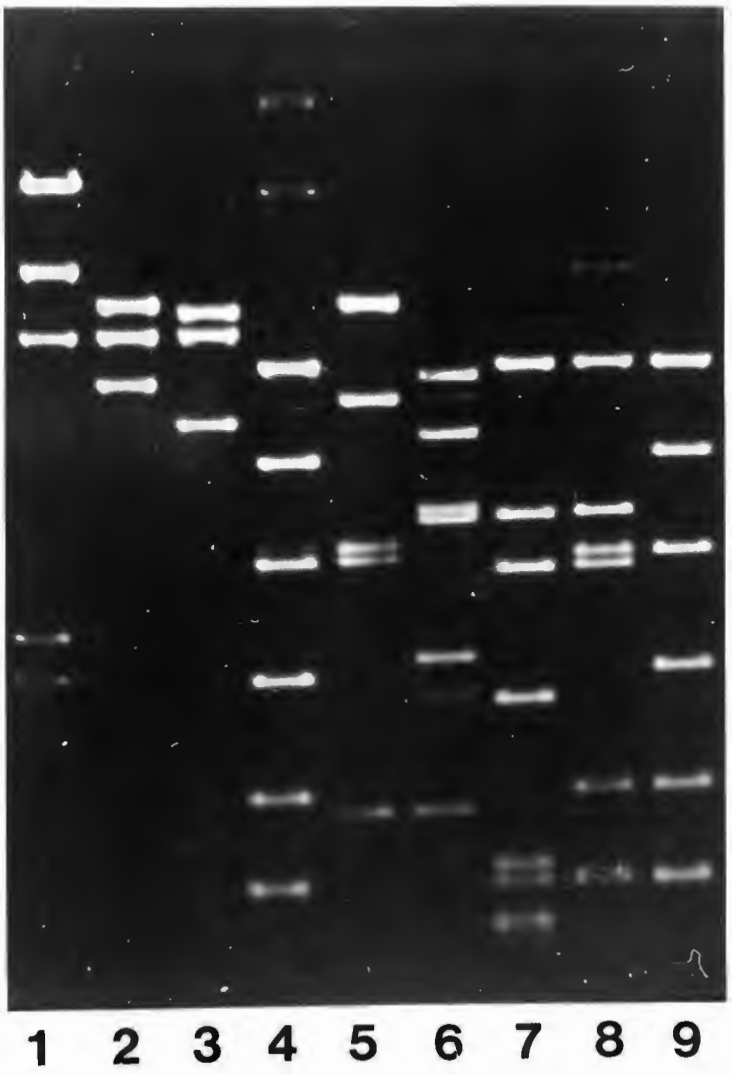


Figure 4.2.9 Various digests of MPE subjected to agarose gel electrophoresis at approximately 2V/cm in a 1% gel. As size markers lambda DNA was cut with HindIII (lane 1) and VarD was digested with EcoRI (lanes 4 and 9). Lanes 2 and 3 are MPE digested with BamHI and BamHI + HindIII respectivily. Other digests of MPE displayed are: HpaI + HindIII - lane 5; HpaI + BamHI - lane 6; HpaI + EcoRI - lane 7 and EcoRI - lane 8.

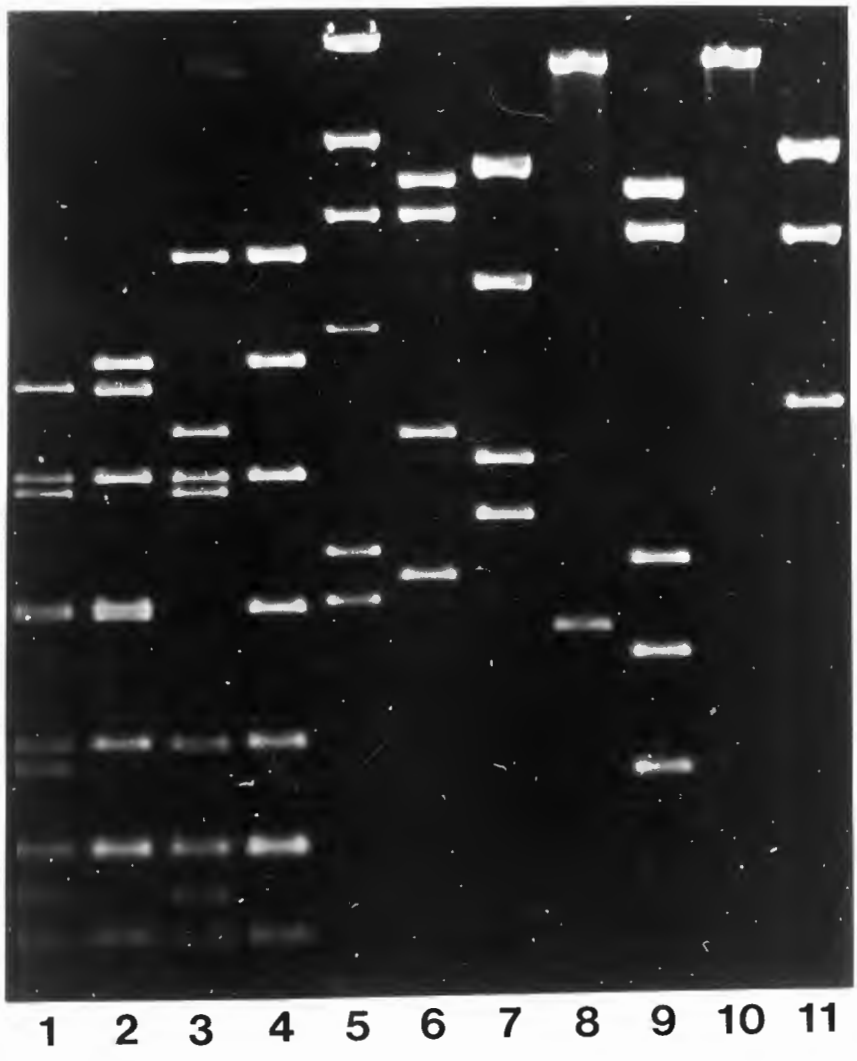


Figure 4.2.10 Comparative digests of MPE and VarD subjected to 1% agarose gel electrophoresis at approximately 2V/cm. Lane 5 is a size marker of lambda DNA digested with HindIII; lanes 1,3,6,8 and 10 are digests of MPE and lanes 2,4,7,9 and 11 are VarD digests. The following enzymes were used: lanes 1 and 2 - EcoRI + HindIII; lanes 3 and 4 - EcoRI; lanes 6 and 7 - BamHI + PstI; lanes 8 and 9 - SalI + PstI; and lanes 10 and 11 - SalI.

4.3 USE OF SOUTHERN CROSS HYBRIDIZATION TO LOCATE THE INTERNAL EcoRI SITES

4.3.1 INTRODUCTION

The presence of many restriction sites makes mapping by double digests followed by gel electrophoresis difficult. The Southern Cross Hybridization method of mapping was designed to map one stretch of DNA for many restriction enzyme sites. Because Monkeypox and Variola DNA map similarly and the EcoRI sites in VarD had already been located, this technique could be used to map the EcoRI sites in MPE.

Two blots were made from preparative gels; one was an EcoRI digest of MPE and the other was an EcoRI digest of VarD which had been radiolabelled with ^{32}P-ATP. The filters were placed on top of one another such that the DNA bands lay perpendicular to one another. Hybridization conditions were applied and the filter containing the MPE digest was washed and autoradiographed.

The pattern of dots which appeared on the autoradiograph allowed one to match up corresponding EcoRI fragments and consequently locate the sites of cleavage in MPE.

4.3.2 METHOD

The protocol supplied by DuPont with the "kit" was used as a guide. The DNA transfer was performed in the manner established in the laboratory, but the method of cross blot construction described in the kit was followed rigidly. MPE will be referred to as the cold digest and VarD as the hot digest. Approximately 4ug of DNA was digested with EcoRI for both hot and cold digests.

Cold Digest

The volume of DNA was made up to 250ul in loading buffer and loaded into a long preparative well of a 0.8% agarose gel. A potential difference of 18 volts was applied for 18 hours(2V/cm), after which the gel was stained and viewed on a UV-transilluminator. The positions of the bands were recorded and measured, and the gel was cut to a size of 15cm square, one of the corners being cut to determine the orientation. The gel was UV-illuminated to break up the DNA fragments.

The DNA was denatured by rocking the gel in 500ml denaturing solution (1.5M NaCl, 0.5M NaOH) for 45 minutes. Neutralization followed by rocking the gel in 1.5M NaCl, 0.5M Tris-HCl pH7.2, 0.001M EDTA for 45 minutes.

The Gene Screen Plus membrane was soaked in 2 X SSC. This membrane has a concave side which binds DNA. The transfer reaction was a modification of the method of Southern. The gel was placed upside down on three pieces of 3MM paper cut larger than the gel and which were soaked in 20 X SSC. The concave side of the Gene Screen Plus membrane was placed on the gel; a piece of 3MM paper the same size as the gel and the membrane was placed on top of the membrane and paper towelling was stacked on top of

the whole lot. The towelling was weighed down with a brick. When the paper towelling became soaked it was replaced with fresh towelling. This was left overnight (approximately 20 hours) at room temperature. The membrane was air dried and kept at room temperature until required.

Hot Digest

After EcoRI digestion, VarD was phenol extracted and ethanol precipitated as described in section 3.3. The DNA was to be end labelled with ^{32}P-ATP.

Polynucleotide Kinase requires a dephosphorylated 5′ protruding end. EcoRI generates phosphorylated 5′ protruding ends, so the sample was treated with alkaline phosphatase to remove the 5′ phosphate group.(See section 4.4.1.1.)

The polynucleotide kinase reaction was performed as described by Greenaway and Dale (1983).

The DNA was resuspended in 50ul Kinase buffer (0.05M Tris-Cl pH7.6, 0.01M $MgCl_2$, 5mM dithiothreitol, 0.1mM spermidine, 0.1mM EDTA) to which was added 40uCi ^{32}P-ATP and 20 units T4 Polynucleotide Kinase. The reaction proceeded at 37^{0}C for 60 minutes.

Unincorporated triphosphate was removed by spinning the sample through a 1ml syringe of Sephadex G-50 which had been equilibrated with 10mM Tris-HCl, 1mM EDTA pH7.5.

100ul TE was spun through the column.

15ul of the sample was loaded into a single well in a 0.8% agarose gel. The remaining sample was made up to 250ul with stop buffer and loaded into a long preparative well next to the single

small well. The DNA bands were separated by electrophoresis at 2V/cm and visualized on a UV-transilluminator after ethidium bromide staining.

The gel was cut into 15cm square. As a check on the labelling reaction the sample lane was cut out, dried and exposed to X-ray film. Denaturation and neutralization were performed as described for the cold blot. (The DNA bands were not UV-irradiated as the fragments were to remain in tact.)

The DNA was transferred to a GeneScreen membrane, of which the two sides were identical in their ability to bind DNA. The method of transfer was as described for the cold blot. Once the DNA had transferred, the blot was not allowed to dry (so that the labelled DNA may be transferred to the cold blot upon hybridization). Prewashing was started before blotting of the labelled fragments was complete.

Cross Blot Construction

The cold membrane was prewashed at 65^{0}C shaking for a few hours (3.5 hrs) in hybridization buffer A (10ml hybridization buffer concentrate + 10ml H_2O + 0.2ml 10% SDS). It was rewashed for 1 hour in hybridization buffer B (25ml hybridization buffer concentrate + 25ml H_2O + 0.5ml 10% SDS + 0.5ml 10% Napyrophosphate). This prewashing was done to reduce the nonspecific binding of DNA to the membrane.

The hot blot was soaked briefly (less than 1 minute) in 2 X SSC (it was not allowed to dry) and placed "face-up" on top of a glass plate in the container provided. The cold membrane (from buffer B) was placed "face-down" in an orientation such that the DNA bands were perpendicular to one another. Any air bubbles present were carefully removed using a glass roller. A piece of

blotting paper was soaked in hybridization buffer C (12.5ml hybridization buffer concentrate + 12.5ml formamide + 0.25ml 10% SDS) and placed on top of the membranes. 5 to 8mls of buffer C was poured over the entire stack, especially near the edges of the membranes. A glass plate was placed on top of the stack and the sandwich clamped together. After standing at room temperature for 1 to 2 hours, the sandwich was incubated at 37^0C overnight. During this time complementary DNA strands hybridize to one another.

Removal of Nonspecific Binding

The blotting paper and hot membrane were removed (kept only in case of troubleshooting). The cold membrane was washed three times in 50ml 2 X SSC, 0.1% SDS for 20 minutes at room temperature, shaking. It was then washed twice ot 50^0C for 40 minutes in 0.1 X SSC, 0.1% SDS.

The membrane was air dried, mounted on 3MM paper and autoradiographed.

4.3.3 RESULTS AND INTERPRETATION

Figures 4.3.2 and 4.3.3 show the EcoRI profiles of VarD and MPE. The single lane was cut off each gel before DNA transfer to the membranes. The single VarD lane was dried and autoradiographed. 7 bands appeared on the X-ray film.

Figure 4.3.4 shows an autoradiogram exposed to the membrane for 60 hours at -70^0C. Another X-ray film was exposed for two weeks (figure 4.3.5). From these two autoradiographs it can be seen that the 2.85kb, 1.3kb, both 1.0kb and the 0.75kb EcoRI fragments are very similar in MPE and VarD. These fragments do

not cross hybridize with fragments of other sizes, indicating that the EcoRI sites are coincidental.

Variola 5.8kb EcoRI fragment, consisting of pAT153 + 2kb of the left end of the insert, hybridizes to Monkeypox 5.7kb and 3.2kb EcoRI fragments. This is expected as the 5.7kb fragment consists mainly of vector sequences which are similar in pAT153 and pBR328 and the 3.2kb fragment consists mainly of 2kb of the left end of the Monkeypox insert. This result indicates that the left end of the two inserts are similar.

The Variola 4.05kb EcoRI fragment hybridized to the Monkeypox 2.7kb and possibly also to the 0.85kb fragment. The 0.85kb EcoRI fragment may lie adjacent to the 2.7kb fragment and the two fragments together correspond in position to the 4.05kb EcoRI fragment of Variola.

The Variola 1.9kb EcoRI fragment, positioned at the right end of the insert hybridized to the Monkeypox 5.7kb EcoRI fragment. This was expected as an EcoRI site was placed approximately 2kb from the right end of the Monkeypox insert, the 5.7kb being made up with vector DNA.

Bands not accounted for are the 0.2kb fragments in Monkeypox and Variola - no signal could be detected in either sample, suggesting that the small fragments may not have transferred efficiently from the gel onto the filter. The two 0.2kb fragments were assumed to correspond to one another.

From this data comparative EcoRI maps may be drawn as in figure 4.3.6.

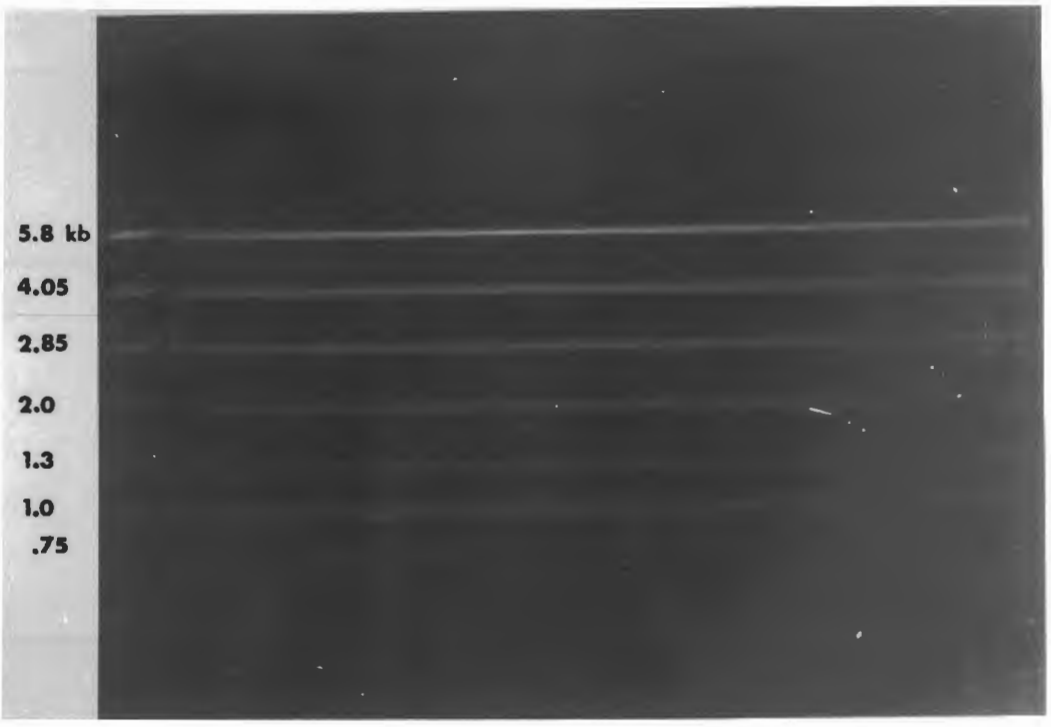
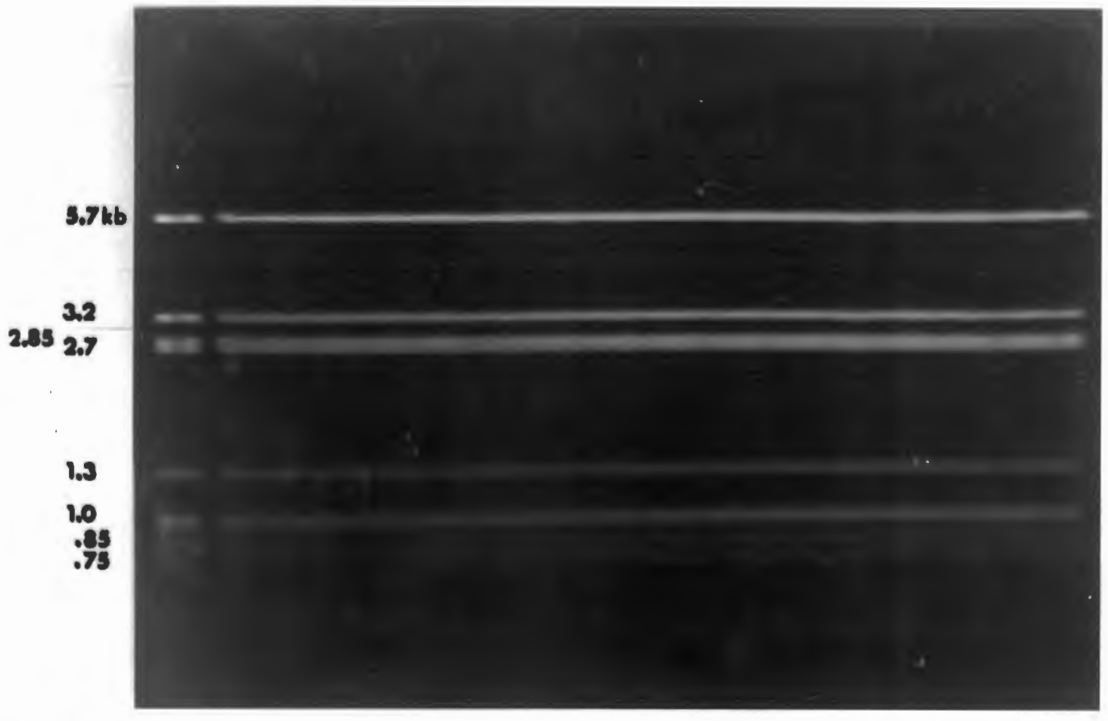


Figure 4.3.2 VarD digested with EcoRI, end-labelled and subjected to 0.8% agarose gel electrophoresis at approximately 2V/cm.

Figure 4.3.3 MPE digested with EcoRI followed by electrophoresis in a 0.8% agarose gel at approximately 2V/cm.



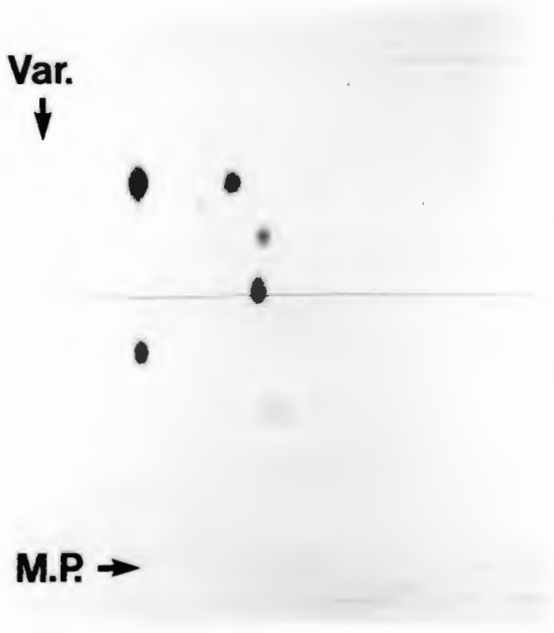


Figure 4.3.4
An autoradiogram exposed to the blot for 60 hours. The VarD bands run vertically and the MPE bands run horizontally.

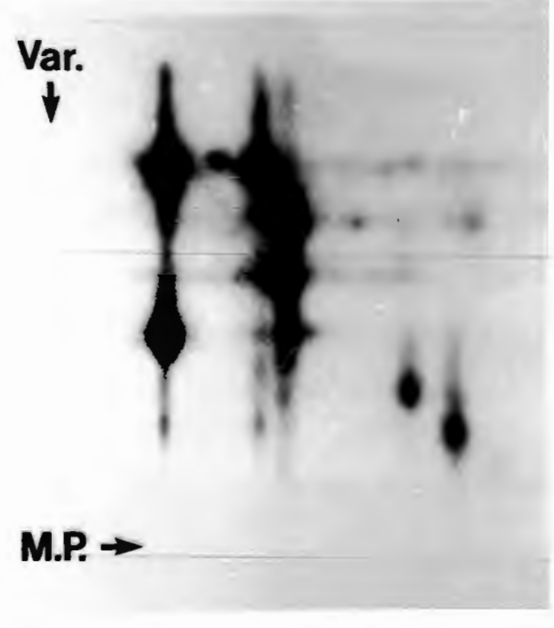


Figure 4.3.5
An autoradiogram exposed to the blot for two weeks.

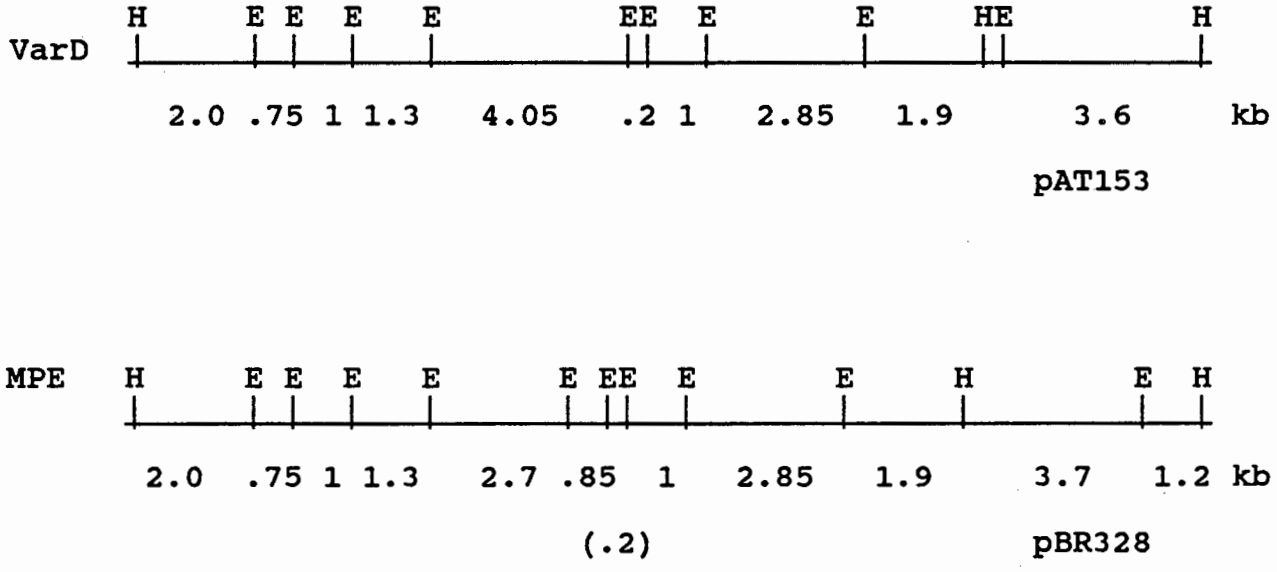


Figure 4.3.6 Location of EcoRI sites in MPE. H=HindIII; E=EcoRI.

4.4 CONFIRMATION OF SOME INTERNAL EcoRI SITES BY SUBCLONING AND HYBRIDIZATION

The locations of some of the EcoRI sites were confirmed by using cloned EcoRI fragments as probes against various digests. The EcoRI fragments of both MPE and VarD were shotgun cloned into the EcoRI site of pUC19. A probe was made of the 2.85kb EcoRI fragments of VarD to confirm the position of the EcoRI site near the right end of MPE.

The location of the 0.85kb EcoRI fragment of MPE was also confirmed by hybridization, making use of subclones of the cloned 2.7kb EcoRI fragment of MPE and the cloned 4.05kb EcoRI fragment of VarD. Both the Monkeypox 2.7kb fragment and the variola 4.05kb fragment are cleaved by BamHI into two fragments. The four EcoRI/BamHI fragments were cloned into pUC19 and the recombinant plasmid containing a 1.15kb insert of Variola DNA was used as a probe against Monkeypox DNA.

4.4.1 CLONING

The vector used for cloning was pUC19. This plasmid has a number of advantages over other vectors - such as pBR322 and others related to it. It is a small, high copy number plasmid, has a multiple cloning site consisting of many single restriction enzyme sites and selection of recombinant plasmids is quick and easy.

The pUC plasmids contain some pBR sequences as well as phage (M13) sequences. The B-lactamase gene (from pBR) conferring ampicillin resistance upon the host cell allows for selection of transformed cells on plates containing ampicillin. When an insert is ligated into the multiple cloning site cells containing

recombinant plasmids may be recognized as white colonies grown in the presence of X-GAL and IPTG. pUC contains a portion of the B-galactosidase gene as well as a multiple cloning site within the gene. In the presence of the lac operon inducer IPTG (isopropyl-B-D-thiogalactopyranoside) B-galactosidase is produced. This enzyme hydrolyses X-GAL (5-bromo-4-chloro-3-indolyl-B-galactoside) to produce bromochloroindole, which is a blue dye. When grown on Luria agar containing ampicillin, X-GAL and IPTG, host cells harboring pUC will appear as blue colonies. If foreign DNA has been inserted into the multiple cloning site a functional B-galactosidase enzyme will not be made and the colonies will be white.

Strains used for transformation by pUC19 are E.coli K-12 derivitives (the JM series). Usually JM107 cells were used. They have the genotype endA1, gyrA96, thi, hsdR17, supE44, relA1, λ^-, $\triangle$(lac-proAB) [F', traD36, proAB, lacI^{q}Z$\triangle$M15] and modify but do not restrict foreign DNA (r-m+). In particular, these cells carry an F plasmid which contains the portion of the B-galactosidase gene complementing that present in pUC (lac I^{q}Z$\triangle$M15). The proAB sequences ensure selection of the plasmid on glucose/minimal medium - as there is a chromosomal pro deletion. The Tra D36 mutation in the plasmid reduces conjugation and is present to comply with US(NIH) and UK(GMAG) guidelines.

SHOTGUN CLONING OF VarD EcoRI FRAGMENTS

4.4.1.1 PREPARATION OF VECTOR AND INSERT

The vector (pUC19) was linearized with a restriction enzyme generating the appropriate sticky ends for ligation. When the EcoRI fragments of VarD and MPE were cloned, pUC19 was linearized with EcoRI.

A vector digested with one enzyme only is capable of recircularizing. The complementary sticky ends hydrogen bond to one another and under ligation conditions the plasmid recircularizes. To prevent recircularization of pUC19 digested with EcoRI, the linearized plasmid was treated with alkaline phosphatase. Calf Intestinal Phosphatase removes the phosphate group from the 5' end of the DNA molecule necessary for the joining reaction.

Alkaline Phosphatase treatment of Vector

The linearized vector was phenol/chloroform extracted to remove the EcoRI enzyme, alcohol precipitated and redissolved in 17ul sterile distilled water. 2ul 10 X CIP buffer (100mM Tris-Cl pH 9.2, 1mM EDTA) was added and the reaction started with 1 unit of Calf Intestinal Phosphatase. After 20 minutes incubation at $37^{0}C$ the sample was treated with phenol and chloroform:isoamylalcohol to remove the enzyme. Ethanol precipitation followed and the dried pellet was finally dissolved in sterile distilled water. The timing of the reaction was found to be critical as the alkaline phosphatase tends to exhibit exonuclease activity if incubated long enough.

Insert preparation

VarD was treated with EcoRI as recommended by the suppliers. The enzyme was removed by phenol/chloroform extraction and the sample precipitated and washed with 70% ethanol as previously described (section 3.3). The dried pellet was dissolved in sterile distilled water.

4.4.1.2 LIGATION

Before ligating the vector and insert a fraction of the two DNA samples was run on a checking gel to estimate their relative DNA concentrations. Six different molar ratios of vector:insert were calculated for pUC19:VarD to determine the optimum ratio for ligation. These ratios were 5:1; 2:1; 1:1; 1:2; 1:5; and 1:10, approximately 5ng of pUC19 being used in each reaction, the quantity of insert DNA varying from 7ng to 350ng (VarD is approximately seven times larger than pUC19).

To make a total of 10ul, 1ul of 10 X ligase buffer (660mM Tris-HCl pH7.2, 10mM EDTA, 100mM $MgCl_2$, 100mM dithiothreitol, 1mM ATP) and 1ul T4 DNA ligase (1 unit) were added to the vector + insert DNA. Ligation proceeded at 12^0C for 4 hours. This temperature is low enough to permit hydrogen bonding between the sticky ends of DNA and high enough for the enzyme to be active. For other ligations the temperature was varied to 15^0C, the altered conditions making no difference to the results. Most subsequent reactions proceeded overnight - giving no change in results.

As ligation controls, linearized pUC19 not treated with alkaline phosphatase was incubated in the presence of T4 ligase (positive) and in the absence of the enzyme (negative). An alkaline

phosphatase control was also set up in which case the vector treated with alkaline phasphatase was incubated under ligation conditions.

4.4.1.3 PREPARATION OF COMPETENT CELLS (Maniatis et al. 1982)

10ml of Luria broth was inoculated with a single JM107 colony grown on minimal medium. This was incubated at 37^0C overnight with shaking.

1ml of the overnight culture was added to 100ml of pre-warmed Luria broth (in a 500ml flask) and incubated for 2 to 4 hours - until log phase had been attained. The culture was left on ice for 10 minutes. The cells were pelletted by centrifugation at 5000rpm for 5 minutes at 4^0C as two volumes of 30ml.

The supernatant was discarded and the cells carefully resuspended in 30ml (2 volumes of 15ml) cold 50mM $CaCl_2$, 10mM Tris-HCl pH8. (50mM $CaCl_2$ was later found to work just as well).

The suspension was left on ice for 15 minutes and then centrifuged for 5 minutes at 5000rpm at 4^0C.

The cells were resuspended in two volumes of 2ml of ice-cold 50mM $CaCl_2$ (and Tris-HCl pH8).

If the cells were to be used within 4 days they were stored at4^0C and 200ul aliquots were used per transformation.

Competent cells were stored at -70^0C. These were resuspended in 15% v/v glycerol.

4.4.1.4 TRANSFORMATION

Competent cells stored at -70^{0}C were thawed slowly (30 minutes) on ice before transformation. Cells transformed by the ligation mixes for shotgun cloning of VarD EcoRI fragments were stored at 4^{0}C.

All of the ligation mix was added to 200ul competent cells and placed on ice for at least 40 minutes. The cells were heat-shocked at 42^{0}C for 2 minutes. 1ml of Luria broth was added and the cell suspension was incubated at 37^{0}C for 1 hour - for expression of the B-lactamase gene encoding ampicillin resistance.

The suspension was spread onto Luria agar plates containing ampicillin, X-GAL and IPTG. After the liquid had been absorbed the plates were incubated at 37^{0}C overnight in an inverted position.

As a transformation control, circular pUC19, having undergone the same ligation treatment, was also used to transform competent cells.

Plasmid DNA was isolated from transformed cells as described in section 3.2.

4.4.1.5 RESULTS

Transformation by the circular pUC19 resulted in hundreds of blue colonies. Transformation frequencies varied from 10^5 to 0.25 X 10^8 colonies per microgram of pUC19. Linearized pUC19 religated gave rise to many blue colonies, in contrast to the negative control (no ligase added) which did not transform at all. pUC19 which had been treated with alkaline phosphatase gave rise to some blue colonies, but about 10 times fewer than the untreated vector.

The vector:insert ratio affected the number of colonies produced as well as the type of recombinant plasmid transformants carried. A 1:1 ratio resulted in very few transformants. As the relative vector concentration increased from a ratio of 2:1 to 5:1 the number of white colonies generated decreased (from about 80 to 50 colonies). With an increase in insert concentration (ratios of 1:2, 1:5 and 1:10) the number of colonies generated increased from approximately 150 to 300 to 400 colonies respectively. Some of these colonies were amplified and plasmid DNA was extracted (section 3.2.1). The insert was cleaved out by EcoRI digestion and subjected to agarose gel electrophoresis (figure 4.4.1).

Ratios of 1:2 and 1:5 generated recombinant plasmids containing single VarD EcoRI fragment inserts. A ratio of 1:10, however, generated recombinant plasmids with two or three EcoRI fragments inserted into a single vector - as shown in the EcoRI digests of figure 4.4.2. A ratio of 1:5 was therefore chosen as the optimal ratio for ligation as it gave rise to the highest number of transformants containing single inserts. This ratio was used for subsequent ligation experiments.

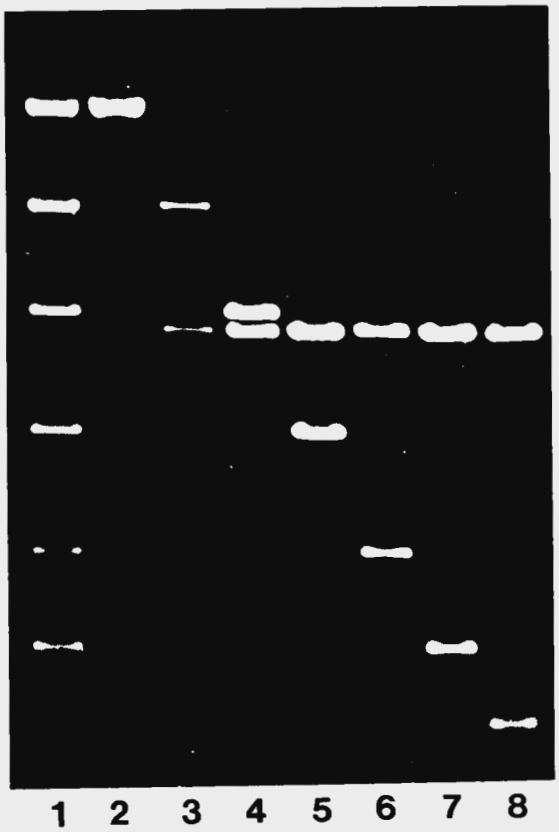


Figure 4.4.1 EcoRI fragments of VarD cloned into pUC19 were cleaved out with EcoRI and subjected to 1% agarose gel electrophoresis at approximately 2V/cm. Lane 1 is VarD digested with EcoRI and lanes 2 to 8 are the various EcoRI fragments cleaved out of pUC19.

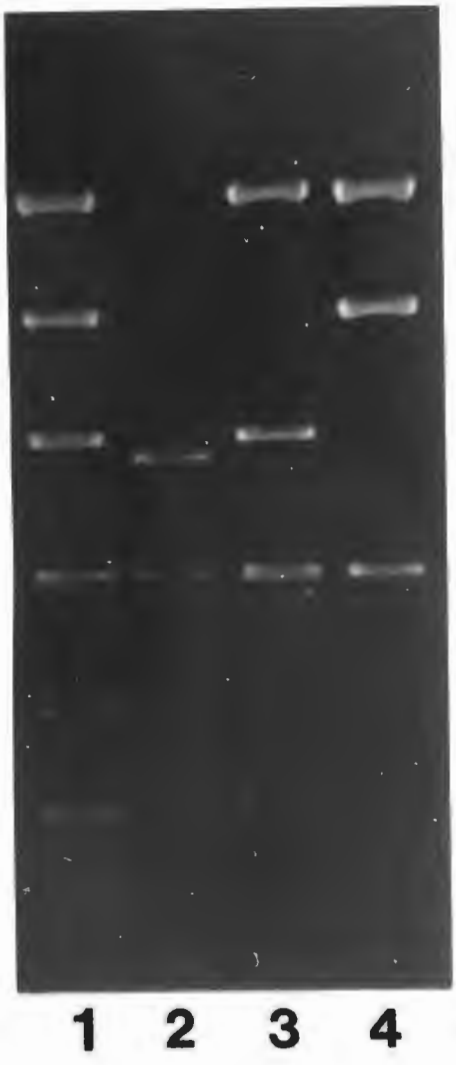


Figure 4.4.2
Plasmids containing multiple inserts of EcoRI fragments of VarD were cut with EcoRI and subjected to 1% agarose gel electrophoresis at approximately 2V/cm. Lane 1 is VarD digested with EcoRI; lane 2 shows two inserts cleaved from pUC19; and lanes 3 and 4 contain multiple inserts in pAT153 (the vector in which VarD was cloned).

The EcoRI fragments of MPE were shotgun cloned as described above using a vector:insert ratio of 1:5 for the ligation step. The inserts were released from the vector by EcoRI digestion (figure 4.4.3). The 2.7kb fragment is the same size as pUC19 and therefore cannot be identified by EcoRI digestion. The recombinant plasmid was a)treated with BamHI to generate two fragements, their sizes depending on the orientation of the insert and b)digested with EcoRI to generate two fragments of 2.7kb (figure 4.4.4).

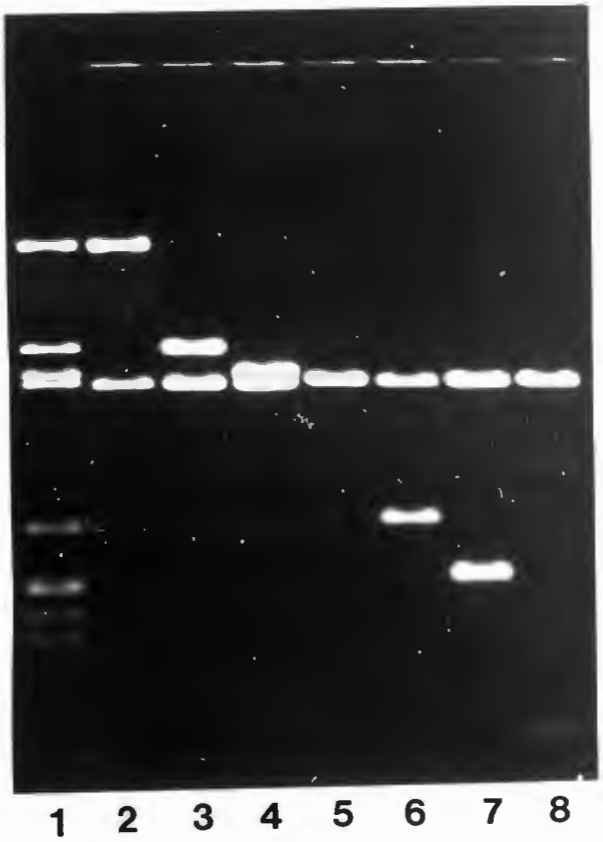


Figure 4.4.3
EcoRI fragments of MPE cloned into pUC19 were cleaved out with EcoRI and subjected to 1% agarose gel electrophoresis at approximately 2V/cm. Lane 1 is MPE digested with EcoRI and lanes 2 to 8 are various EcoRI fragments cleaved out of pUC19.

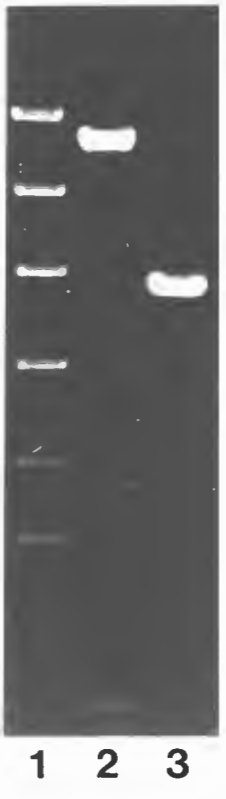


Figure 4.4.4
The 2.7kb EcoRI fragment of MPE cloned into pUC19 was identified by BamHI digestion (lane 2) and EcoRI digestion (lane 3). VarD digested with EcoRI was used as a size marker (lane 1). The fragments were separated at approximately 2V/cm in a 1% agarose gel.

The cloned 4.05kb fragment of VarD and 2.7kb EcoRI fragment of MPE (figure 4.4.5) were double digested with EcoRI and BamHI. (The location of these fragments can be seen in figure 4.3.6). Both plasmids generated two insert fragments as well as the vector having one BamHI sticky end and one EcoRI sticky end. No additional vector was required and ligation and transformation was performed as described in sections 4.4.1.2 to 4.4.1.4. Recombinant plasmids containing inserts of 2.9kb and 1.15kb of Variola DNA and 2.4kb of Monkeypox DNA were isolated. They were recognized by EcoRI + BamHI double digestion as well as linearization with EcoRI followed by agarose gel electrophoresis (figure 4.4.7).

The 0.3kb EcoRI/BamHI fragment of Monkeypox was not isolated as a recombinant plasmid. This fragment was therefore eluted from a gel containing the cloned 2.7kb EcoRI fragment of Monkeypox digested with EcoRI and BamHI. Ligation and transformation was performed and the recombinant plasmids isolated as described (sections 4.4.1.2 to 4.4.1.4). The required recombinant plasmids were recognized by EcoRI + BamHI digestion (figure 4.4.6).

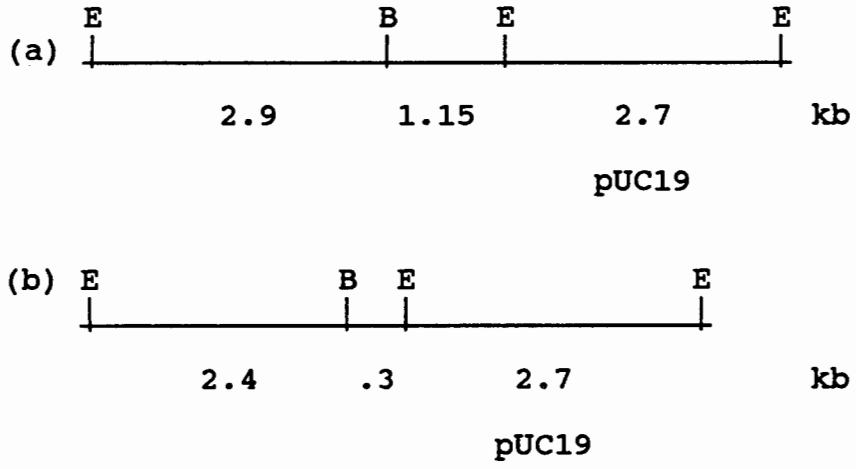


Figure 4.4.5 Diagrams of the cloned 4.05kb fragment of VarD (a);and cloned 2.7kb EcoRI fragment of MPE (b). E=EcoRI; B=BamHI.

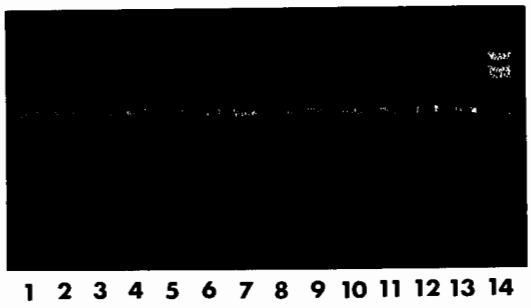


Figure 4.4.6 1% Agarose gel electrophoresis of recombinant plasmids cleaved with EcoRI + BamHI. Lane 1 is pUC19 and lanes 2 to 13 are minipreps of the cloned 0.3kb EcoRI/BamHI fragment of MPE. Lane 14 is lambda DNA cleaved with HindIII. The fragments were separated at approximatnely 4V/cm.

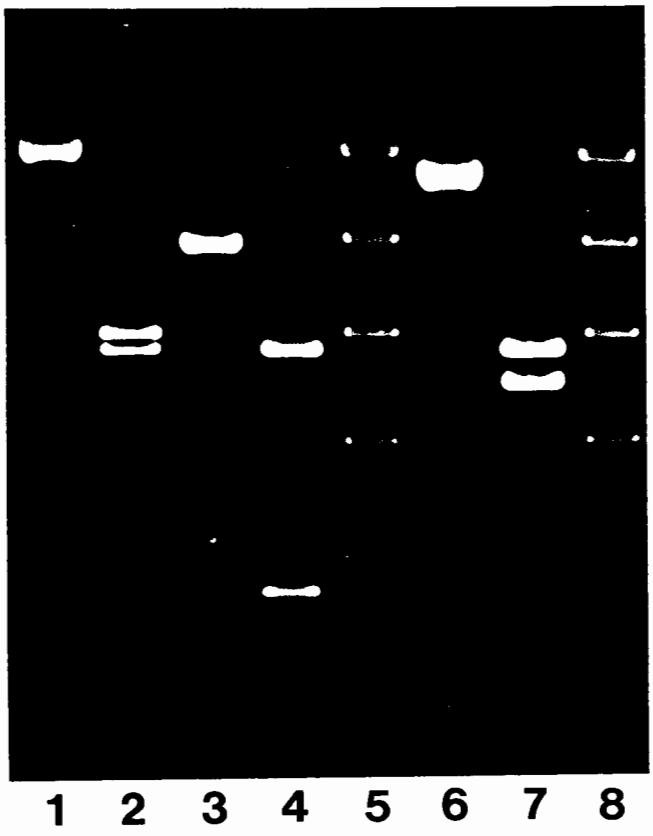


Figure 4.4.7 Identification of recombinant plasmids containing EcoRI/BamHI inserts by gel electrophoresis at approximately 2V/cm in a 1% agarose gel. Lanes 1,3 and 6 are EcoRI digests of plasmids containing the 2.9kb VarD, 1.15kb VarD and 2.4kb MPE EcoRI/BamHI inserts respectively. Lanes 2,4 and 7 are EcoRI + BamHI digests of the same plasmids, the inserts being released in each case. Lanes 5 and 8 are size markers of VarD digested with EcoRI.

4.4.2 HYBRIDIZATION

DNA separated by agarose gel electrophoresis was denatured and transferred onto a Hybond-N membrane by a modified method of Southern (1975). The membrane was probed with DNA radiolabelled with ^{32}P-dCTP and autoradiographed. DNA bands homologous to the probe lit up on the X-ray film. By probing different digests and determining the size of the bands which hybridize, restriction sites can be mapped.

4.4.2.1 DNA TRANSFER TO HYBOND-N MEMBRANES

DNA subjected to agarose gel electrophoresis in a 0.8% agarose gel was "broken up" by UV-illuminating the gel for five minutes. It was denatured by shaking the gel in an excess volume (500ml) of denaturing buffer (1.5M NaCl, 0.5M NaOH) for 45 minutes. Neutralization followed by shaking the gel in 500ml neutralizing buffer (1.5M NaCl, 0.5M Tris-HCl pH7.2, 0.001M Na_2EDTA) for at least 45 minutes.

Figure 4.4.8 shows the DNA transfer set-up. Three pieces of 3MM paper cut larger than the gel were soaked in 20 X SSC (1 X SSC is 0.15M NaCl, 0.015M sodium citrate pH 7.0) and placed onto some saran wrap on a flat bench top. The gel was placed upside down on the 3MM paper and a piece of Hybond-N cut to the same size as the gel was placed on top of the gel. Care was taken to prevent air bubbles forming between the gel and the membrane. If two blots of the same gel were required, a piece of Hybond-N was placed underneath as well as on top of the gel.

A single piece of 3MM paper was cut to the same size as the gel and placed on top of the Hybond-N membrane. A stack of paper towelling was put on top of the 3MM paper and weighed down with a

brick. As a variation, incontinent sheets (from Groote Schuur Hospital) or Kimbies nappies (from S.A.A.) were used as well as paper towelling to absorb the buffer.

The paper towelling was changed at least once and the transfer was allowed to proceed overnight.

The Hybond-N was air-dried, wrapped in saran wrap and UV-illuminated for 15 minutes (DNA side down) to irreversibly bind the DNA to the membrane.

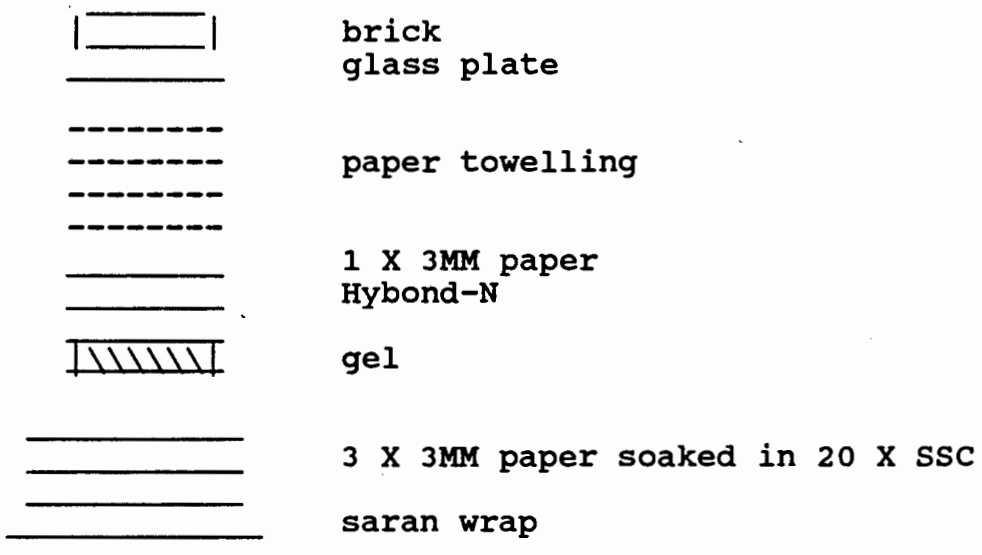


Figure 4.4.8 Diagram of DNA transfer apparatus.

4.4.2.2 RADIOLABELLING DNA BY NICK TRANSLATION

DNA was radiolabelled as described in the manual supplied with the Nick Translation Kit (Amersham No.N5000). Approximately 1ug of DNA was used to which was added 20ul of nucleotide / buffer solution containing the deoxyribonucleotides required, 5ul dCTP($\underline{a}$-^{32}P), sterile distilled water to make the volume up to 90ul and 10ul enzyme solution containing DNA Polymerase I. The contents of the tube were mixed and incubated at 15^{0}C for 1.5 hours. The reaction was terminated with 5ul Stop buffer.

Purification of Probe by Sephadex Chromatography

a) Using a pipette

A piece of siliconized glass wool was inserted at the end of a plastic pipette. The pipette was packed with sephadex G-50 to the 5ml mark. The sample was loaded and the sephadex was kept immersed in nick translation buffer (150mM NaCl, 50mM Tris-Cl pH 7.5, 10mM EDTA, 0.1% SDS).

Once the xylene cyanol dye had migrated half way down, fractions were collected, three drops per microfuge tube. After 20 tubes had been collected the radioactivity of each tube was monitored with a geiger counter. Five tubes containing the peak of racioactivity were pooled and kept for hybridization.

b) Using a syringe

A piece of siliconized glass wool was inserted into a 1ml syringe. The syringe was filled with sephadex G-50 and placed into a plastic tube. The syringe was centrifuged at 1500rpm for 5 minutes.

The sample was loaded, 200ul nick translation buffer was added and the syringe was recentrifuged for 5 minutes. The sample was

collected in a microfuge tube.

Before hybridization the probe was denatured by heating it for 7 minutes at 100^{0}C.

4.4.2.3 HYBRIDIZATION

Hybridization was performed as described by Johnson et al. (1984). To reduce nonspecific binding of the probe to the membrane, the Hybond-N was prehybridized in 20ml of 0.25% nonfat dry milk (0.05 X BLOTTO), 6 X SSC. The membrane and prehybridization fluid were sealed in a plastic bag and incubated at 42^{0}C for at least two hours.

The fluid was discarded and an equal volume (usually 5ml) of prehybridization fluid and formamide was mixed and, together with the probe, added to the membrane in the bag. The bag was sealed and incubated at 42^{0}C overnight, shaking. The formamide ensured denaturation of the probe.

After hybridization the membrane was washed to remove nonspecific binding of the probe to the membrane. Two thirty-minute washes were performed in 500ml 0.05 X BLOTTO, 2 X SSC, 0.1% SDS with shaking. Another two washes were performed at 55^{0}C in 500ml 0.1 X SSC, 0.1% SDS.

The blot was then wrapped in saran wrap and autoradiographed.

4.4.2.4 AUTORADIOGRAPHY

An X-ray film was exposed to the blot for a few hours or overnight at -70^{0}C. Sometimes it was left for a few days.

Under the illumination of a red light only the film was developed for 3 minutes in GBX developer. It was then placed in a 2% acetic acid stop bath for 1 minute and finally in fixer containing hardener for 2 minutes. The film was then washed under tap water for approximately 15 minutes and air-dried.

4.4.2.5 RESULTS AND INTERPRETATION

Figure 4.4.9 shows a gel which was transferred to Hybond-N membrane and probed with the cloned 2.85kb EcoRI fragment of VarD. Lane 1 is MPE digested with EcoRI + BamHI; lane 2 - MPE digested with EcoRI; lane 3 - MPE digested with EcoRI + HpaI and lane 4 - VarD digested with Eco RI. VarD digested with Eco RI was used both as a positive control and as a molecular weight marker. Sites already mapped are shown in figure 4.4.10 with the EcoRI site being confirmed in brackets.

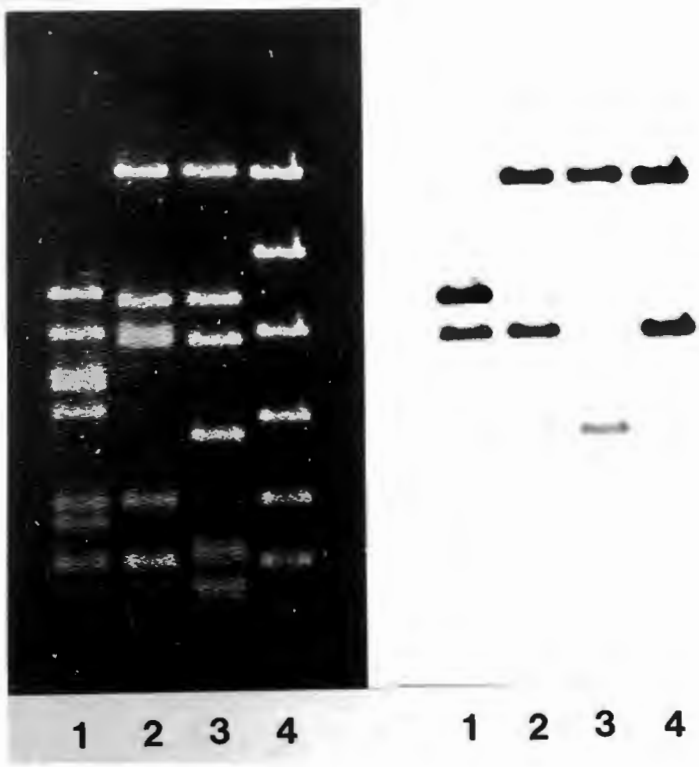


Figure 4.4.9 A - 0.8% Agarose gel subjected to approximately 2V/cm and subsequently blotted. MPE was digested with EcoRI + BamHI (lane 1), EcoRI (lane 2) and EcoRI + HpaI (lane 3). Lane 4 is VarD digested with EcoRI.
B - Autoradiogram of the blot after probing with the cloned 2.85kb EcoRI fragment of VarD.

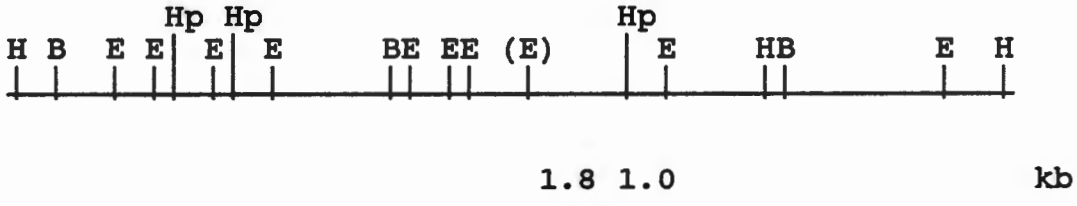


Figure 4.4.10 Map of MPE. The EcoRI site being confirmed is bracketed. H=HindIII; B=BamHI; E=EcoRI; Hp=HpaI.

Two bands of the EcoRI/BamHI double digest lit up - 2.85kb and 3.3kb. The 3.3kb band corresponds to vector sequences. (Appendix 1 shows the maps of pUC19 and pBR328). The 2.85kb fragment must correspond to the 2.85kb EcoRI fragment of VarD. It was not cut by BamHI. Lane 2 (the EcoRI fragments of VarD) also shows that the MPE 2.85kb fragment is homologous to the probe. The 5.7kb band contains vector sequences.

In lane 3 (MPE digested with EcoRI and HpaI) three bands lit up. The 5.7kb band contains vector sequences. The other two bands (1.8kb and 1.0kb) must be the 2.85kb fragment cleaved by HpaI. A Hpa I site is located 1.0kb internal to the right EcoRI site. Another EcoRI site must therefore be present 1.8kb from the HpaI site.

Figure 4.4.11 shows restriction enzyme profiles which were probed with the 1.15kb EcoRI/BamHI fragment of VarD cloned into pUC19. VarD digested with EcoRI was used as a molecular weight marker as well as a positive control (lane 1). Lambda DNA cleaved with HindIII was also run as a molecular weight marker (lane 4). Lane 2 is MPE treated with EcoRI and lane 3 is the 0.3kb EcoRI/BamHI fragment of MPE cloned into pUC19 and cleaved out with EcoRI and BamHI.

The probe hybridized to the 0.3kb EcoRI/BamHI fragment of MPE. This fragment lies at the right end of the 2.7kb EcoRI fragment of MPE (figure 4.4.12). The probe also hybridized to a fragment of 0.85kb in the EcoRI digest of MPE. This means the 0.85kb EcoRI fragment lies on the right of the 2.7kb EcoRI fragment.

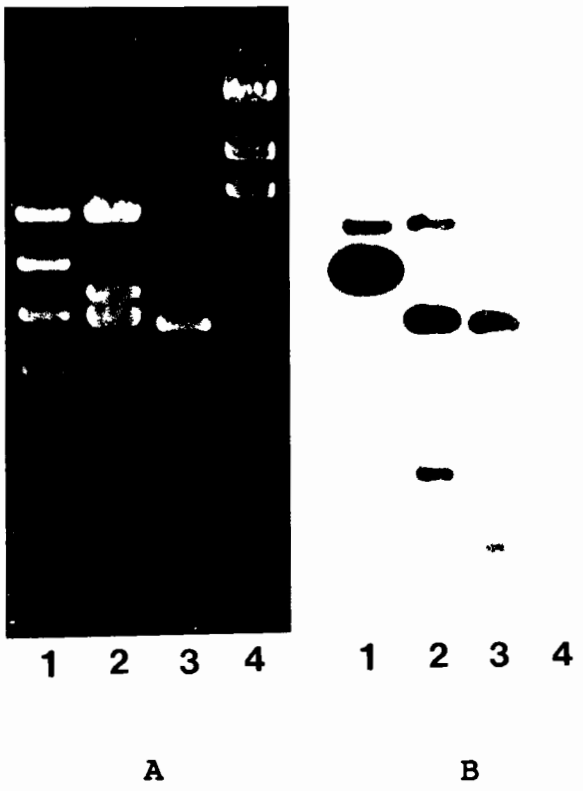


Figure 4.4.11 Digests of cloned Monkeypox DNA probed with the cloned 1.15kb EcoRI/BamHI fragment of VarD.
A - The following digests were subjected to approximately 2V/cm in a 0.8% agarose gel: lane 1 - VarD digested with EcoRI; lane 2 - MPE digested wwith EcoRI; lane 3 - the cloned 0.3kb EcoRI/BamHI fragment of MPE treated with EcoRI + BamHI; and lane 4 - lambda DNA cleaved with HindIII.
B - Autoradiograph of the membrane onto which the fragments were transferred and subsequently probed with the 1.15kb EcoRI/BamHI fragment of VarD.

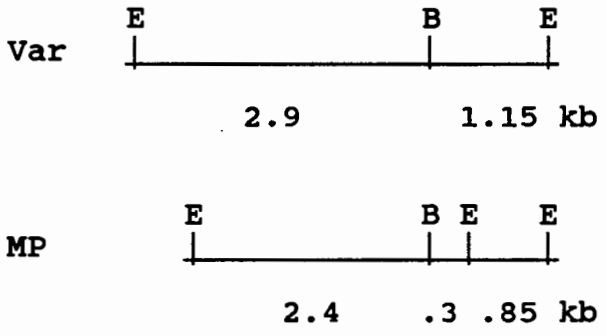


Figure 4.4.12 Comparative maps of the internal regions of VarD and MPE. E=EcoRI; B=BamHI.

4.5 COMPARISON OF RESTRICTION ENZYME MAPS OF VarD AND MPE

Figure 4.5.1 shows the restriction enzyme maps of VarD (labelled Harvey) and MPE. The Variola HindIII D fragment is cloned into pAT153 and the Monkeypox HindIII E fragment is cloned into pBR328. The "pox" sequences lie between the two left HindIII sites (represented as arrows), the Variola sequence being slightly longer than that of the Monkeypox. The maps are orientated as they would appear in the whole viral genome from left to right.

Although similar, the restriction enzyme maps of these two fragments are not identical. MPE has an extra AvaI and an extra BamHI site near the left end. More centrally it has an extra HpaI site and an extra EcoRI site. In the central region VarD has a SalI site not present in MPE. At the right end VarD has an extra PstI site and an extra SalI site.

Sites present in both fragments in corresponding positions include seven EcoRI sites, two HpaI sites and one BamHI site. This similarity is remarkable, as these fragments are from two different strains of Orthopoxviruses having contrasting phenotypic properties.

Each of the sites mentioned represents a sequence of 6 base pairs. If there is an alteration in one base pair the presence or absence of a site may be affected. Single base changes may not necessarily be of great importance - due to the degeneracy of the genetic code - since an altered DNA sequence may not give rise to an altered amino acid sequence. The different restriction enzyme sites are therefore not necessarily of great significance as the difference may only be in a single base pair. If, however, one is looking for differences, the areas where the restriction enzyme

pattern varies would be investigated. An aim of this project was to locate differences in the Variola genome (compared to that of Monkeypox). With this in mind the left end 2kb HindIII/EcoRI fragment, the right end 2kb EcoRI fragment together with its neighboring HpaI/EcoRI fragment and the central 4kb EcoRI fragment were compared to the corresponding regions in the monkeypox DNA.

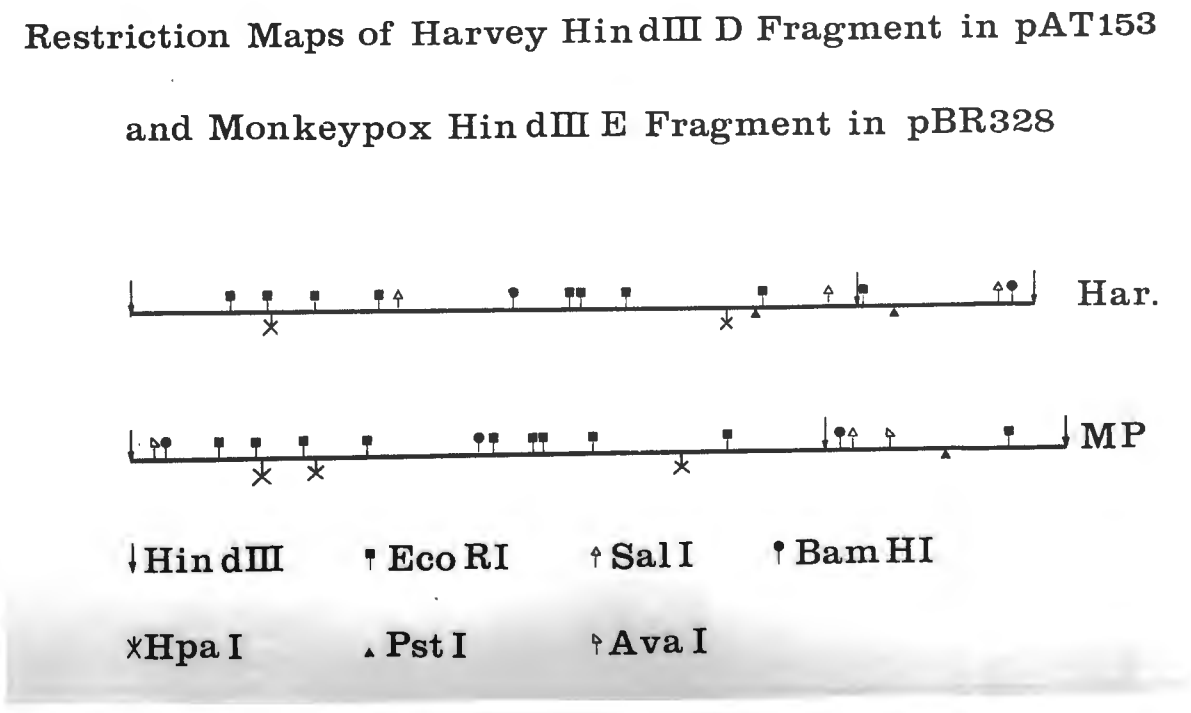


Figure 4.5.1 Comparison of Restriction Enzyme Maps of VarD and MPE.

CHAPTER 5 LOCATION OF HETEROLOGOUS REGIONS IN VarD

5.1 HYBRIDIZATION

In comparison with the corresponding area of the Monkeypox genome, heterologous VarD DNA sequences were established. The search for these differing areas can be divided into four sections, viz. the left end, a general comparison of homology, the right end and the central region.

5.1.1 THE LEFT END OF VarD

Initially it was important to establish whether the two HindIII sites at the left of VarD and MPE were coincident or not. (See figure 1.2.1). This was uncertain as Variola HindIII C is smaller than Monkeypox HindIII C and VarD is larger than MPE.

The 2kb HindIII/EcoRI fragment at the left of VarD was eluted from a gel and cloned into the HindIII/EcoRI sites of pUC19. This plasmid was radiolabelled with ^{32}P-dCTP and used to probe Variola HindIII C, VarD, Monkeypox HindIII C and MPE. VarC was a negative control and VarD was a positive control. Monkeypox HindIII C did not hybridize to the probe, but MPE did. (Figure 5.1.1). This means that Monkeypox sequences to the right of the HindIII site separating MPC from MPE correspond to VarD sequences. The HindIII sites can therefore be said to be coincident.

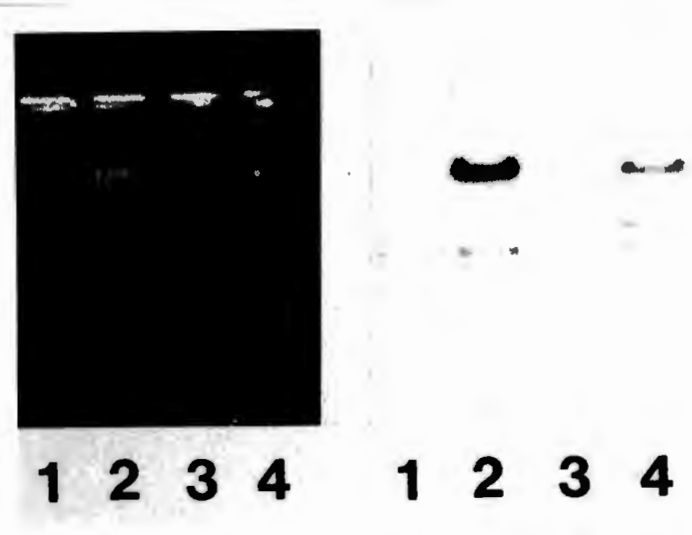


Figure 5.1.1 A - Gel electrophoresis of HindIII digests of recombinant plasmids in a 0.8% agarose gel at approximately 2V/cm. Lane 1 - VarC; lane 2 - VarD; lane 3 - MPC and lane 4 - MPE.
B - Autoradiogram of membrane probed with the cloned 2kb EcoRI/HindIII fragment at the left end of VarD.

5.1.2 A GENERAL COMPARISON OF HOMOLOGY

In an attempt to detect a difference in DNA sequence between VarD and MPE, VarD was digested with EcoRI + SalI, electrophoresed through a 0.8% agarose gel, transferred to a Hybond-N membrane and probed with ^{32}P-dCTP labelled MPE. The resulting autoradiograph displayed all the bands of VarD (figure 5.1.2). This means that most of the DNA sequence is very similar in VarD and MPE. Two bands appear distinctly fainter than the others - those of 3.5kb and 1.5kb. The 3.5kb fragment lies within the 4.0kb EcoRI fragment in the central region of VarD; and the 1.5kb sequence lies within the 1.9kb EcoRI fragment at the right end. These two EcoRI fragments were analysed in more detail to locate any differences.

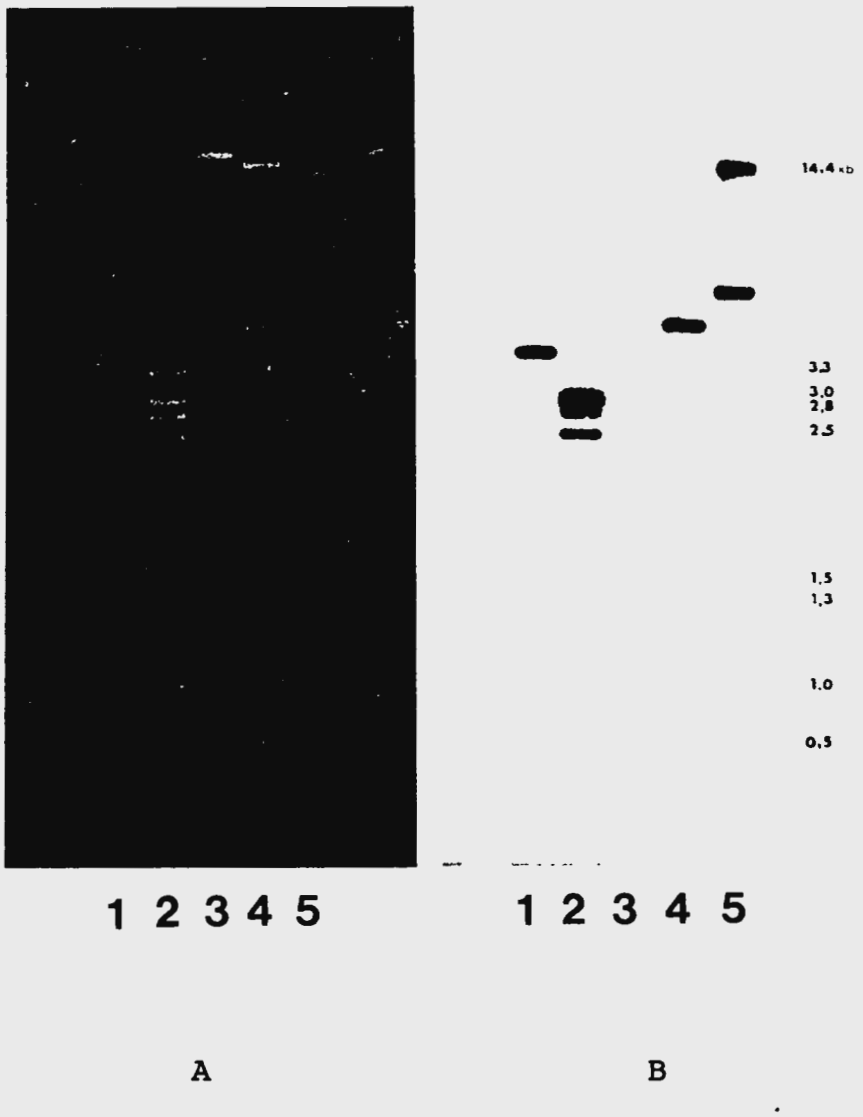


Figure 5.1.2 A - Electrophoresis of restriction fragments in a 0.8% agarose gel at approximately 2V/cm. Lane 1 is VarD digested with HindIII; lane 2 - VarD digested with EcoRI + SalI; lane 4 - MPC digested with HindIII; and lane 5 - MPE digested with HindIII. B - Autoradiogram of the membrane probed with MPE.

5.1.3 THE RIGHT END OF VarD

The cloned 1.9kb and 2.85kb EcoRI fragments of VarD were employed for analysis of the right end. The plasmid containing the 1.9kb insert was digested with SalI + EcoRI; and the 2.85kb fragment cloned into pUC19 was digested with HpaI and EcoRI (see figure 4.5.1). In each case a fragment of 2.7kb consisting of vector sequences was generated. Together with VarD digested with EcoRI as a molecular weight marker the digests were subjected to electrophoresis in a 0.8% agarose gel. The DNA was transferred to Hybond-N membrane and probed with MPE (see figure 5.1.3).

The probe hybridized to all the fragments comprising the 2.85kb EcoRI fragment. In contrast, only the 1.5kb band of the 1.9kb EcoRI fragment hybridized (and this signal was relatively faint). The 0.4kb EcoRI/SalI fragment could not be visualized on the X-ray film.

To check that the DNA had transferred onto the membrane, the same membrane was probed with the 0.4kb EcoRI/SalI fragment eluted from a gel. A strong signal was emmitted at the position of the 0.4kb band. The DNA had therefore been transferred onto the membrane.

Attention must be drawn to the fact that the 0.4kb fragment was the smallest fragment present. Kinetically, larger fragments would hybridize to a probe more frequently than smaller fragments.

Because of the apparent nonhomology located within the 1.9kb EcoRI fragment of VarD, this fragment, together with the corresponding region of MPE was sequenced. (See chapter 6).

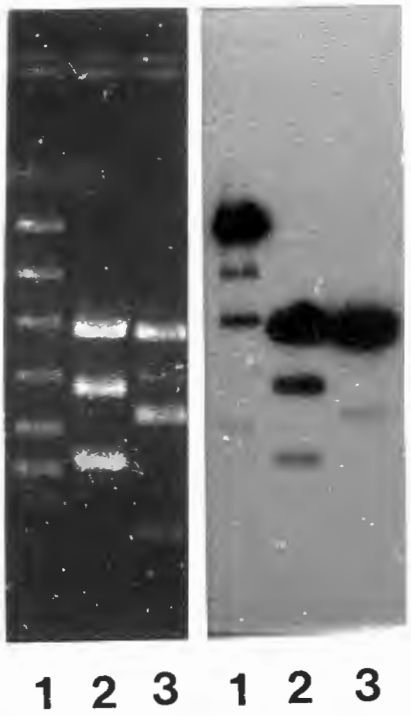


Figure 5.1.3 A - Electrophoresis of restriction fragments at approximately 2V/cm in a 0.8% agarose gel. Lane 1 is VarD digested with EcoRI; lane 2 is a HpaI + EcoRI digest of the 2.85kb EcoRI fragment of VarD cloned into pUC19 and lane 3 is a SalI + EcoRI digest of the 1.9kb EcoRI fragment of VarD cloned into pUC19.
B - Autoradiogram of the membrane probed with MPE.

5.1.4 THE CENTRAL REGION OF VarD

The 4.05kb EcoRI fragment of VarD cloned into pUC19 was digested with EcoRI + SalI + BamHI. (See figure 4.5.1). Additional to the 2.7kb vector fragment, three fragments were generated. These were separated on a 0.8% agarose gel, transferred to a Hybond-N membrane and probed with MPE (figure 5.1.4). All the fragments hybridized to MPE.

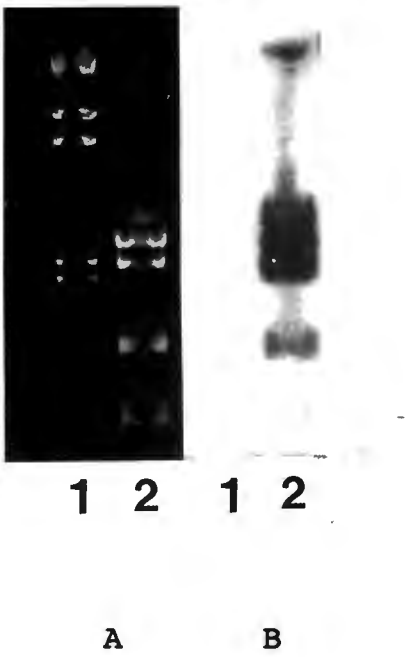


Figure 5.1.4
A - Electrophoresis of restriction fragments at approximately 5V/cm in a 0.8% agarose gel. Lane 1 is lambda DNA digested with HindIII and lane 2 is VarD digested with EcoRI + SalI + BamHI.
B - Autoradiogram of the membrane probed with MPE.

It has been established that within this 4.05kb EcoRI fragment - and in particular, within the larger EcoRI/BamHI fragment - there is an extra 0.5kb sequence relative to the Monkeypox fragment (figure 4.5.1). The Variola EcoRI /BamHI fragment is 2.9kb, whereas the Monkeypox EcoRI/BamHI fragment is only 2.4kb. These two fragments were mapped for other restriction enzyme sites in an attempt to locate the extra 0.5kb sequence in the Variola DNA.

5.2 RESTRICTION ENZYME MAPPING OF THE 2.9kb EcoRI/BamHI FRAGMENT OF VarD AND THE 2.4kb EcoRI/BamHI FRAGMENT OF MPE

5.2.1 NcoI MAPPING OF THE 2.9kb EcoRI/BamHI FRAGMENT OF VarD

NcoI cleaves the recombinant plasmid once. There is no NcoI site in the vector (pUC19), so there must be one site in the insert. Digestion with NcoI + HindIII generated two bands of 2.7 and 2.9kb. Since the vector is 2.7kb and has no NcoI sites, an NcoI site is either a) 0.2kb from the EcoRI site (the 2.7kb band consisting of insert sequences and the 2.9kb band consisting of the vector + 0.2kb of insert sequences) or b) right next to the EcoRI site (2.9kb=insert only; 2.7kb=vector only). (See figure 5.2.1).

A SalI site is present 0.5kb from the EcoRI site. Upon digestion with SalI + NcoI fragments of 2.9, 2.4 and 0.3kb were generated. The 2.9kb fragment must consist of the vector, which is 2.7kb; the additional 0.2kb means the NcoI site must be 0.2kb from the EcoRI site. Positioning of the NcoI site here would give rise to a 0.3kb NcoI/SalI fragment which was produced. The 2.4kb band was the SalI/BamHI stretch of the insert extended to the SalI site of the vector.(Figure 5.2.2).

Table 5.2.1 Fragment sizes used for NcoI mapping of the cloned 2.9kb EcoRI/BamHI stretch of VarD

Enzymes used	Sizes of Fragments generated (kb)
NcoI	5.6.
NcoI + HindIII	2.9; 2.7.
NcoI + SalI	2.9; 2.4; 0.3.

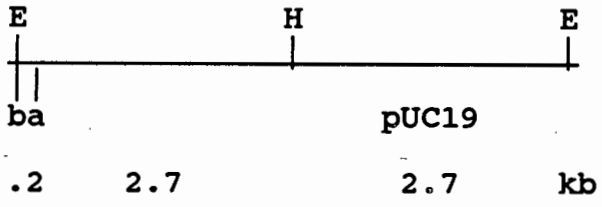


Figure 5.2.1 Mapping of NcoI site. E=EcoRI; H=HindIII. Possible positions of the NcoI site are depicted as a and b.

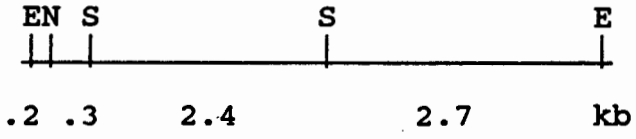


Figure 5.2.2 Location of the NcoI site. E=EcoRI; N=NcoI; S=SalI.

5.2.2 NcoI MAPPING OF THE 2.7kb EcoRI FRAGMENT OF MPE CLONED INTO pUC19

The orientation of the insert was determined by cleaving the plasmid with BamHI (see figure 4.4.4). A BamHI site is present in the multiple cloning site of the vector and 0.3kb from the right end of the insert (as orientated in VarD). BamHI generated fragments of 5.1 and 0.3kb; the BamHI site is therefore 0.3kb from the vector BamHI site. If in the opposite orientation, bands of 3.0 and 2.7kb would have been produced. (Figure 5.2.3).

NcoI generated two fragments of 3.2 and 2.2kb. Two sites are therefore present in the insert. The 3.2kb fragment must contain all of the vector sequences and should therefore be cleaved twice by EcoRI. NcoI + EcoRI digestion gave rise to 4 fragments of 2.7, 2.2, 0.35 and 0.15 kb. The 2.7kb fragment is pUC19 and the 2.2kb fragment must be the unaltered internal NcoI fragment. NcoI sites are therefore present 0.35 and 0.15 kb from either end.

The location of these sites was determined by NcoI + SalI double digestion. SalI cleaves the plasmid once only (in the multiple cloning site). Fragment sizes of 2.85, 2.2 and 0.35 kb were generated. The 2.85kb fragment consists of pUC19 + 0.15kb insert sequences. An NcoI site is therefore present 0.15kb from the left EcoRI site (figure 5.2.4). The 2.2kb band is the internal 2.2kb NcoI fragment, and the 0.35kb band is the right end NcoI/SalI fragment.

Table 5.2.2 Fragment sizes used for mapping NcoI sites in the cloned 2.7kb EcoRI fragment of MPE

Enzymes used	Sizes of fragments generated (kb)
NcoI	3.2; 2.2.
NcoI + BamHI	5.0; 0.35.
NcoI + EcoRI	2.7; 2.2; 0.35; 0.15.
NcoI + SalI	2.85; 2.2; 0.35.

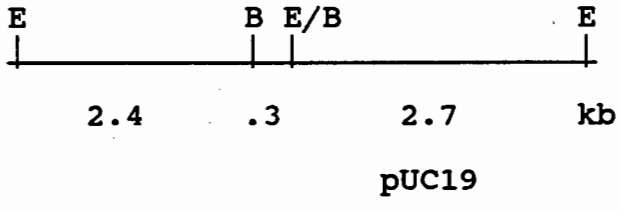


Figure 5.2.3 Orientation of the 2.7kb EcoRI insert in pUC19. E=EcoRI; B=BamHI.

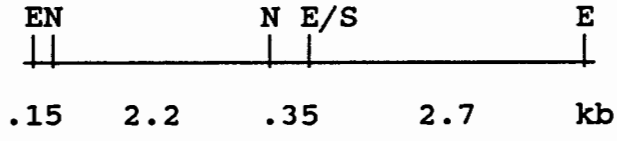


Figure 5.2.4 Location of the NcoI sites in the 2.7kb EcoRI fragment of MPE. E=EcoRI; N=NcoI; S=SalI.

5.2.3 MAPPING OF THE HincII SITES IN THE 2.9kb EcoRI/BamHI FRAGMENT OF VarD AND THE 2.4kb EcoRI/BamHI FRAGMENT OF MPE

HincII cleaves pUC19 once near the BamHI site in the multiple cloning site. When the 5.6kb " Variola" recombinant plasmid was treated with HincII two fragments of 3.4kb and 2.2kb were generated. The 3.4kb fragment consists of 2.7kb pUC19 + 0.7kb insert. A HincII site is therefore present 0.7kb from the EcoRI site of pUC19 (figure 5.2.5). The 2.2kb HincII fragment is the remaining insert DNA.

The 5.1kb "Monkeypox" recombinant plasmid gave rise to two fragments of 3.4kb and 1.7kb upon HincII digestion. Since the 3.4kb fragment must include 2.7kb of vector sequences, a HincII site is located 0.7kb from the EcoRI site. The 1.7kb fragment is the remaining insert sequence (figure 5.2.6).

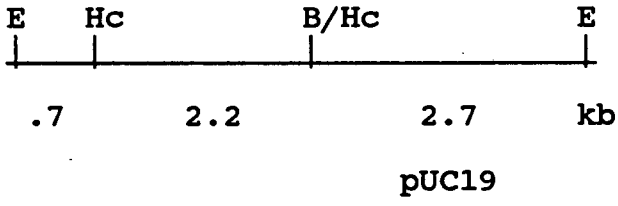


Figure 5.2.5 Location of the HincII site in the VarD 2.9kb EcoRI/BamHI fragment. E=EcoRI; B=BamHI; Hc=HincII.

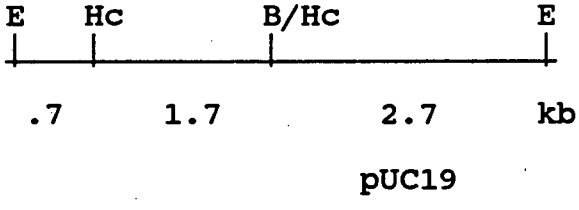


Figure 5.2.6 Location of the HincII site in the MPE 2.4kb EcoRI/BamHI fragment. E=EcoRI; B=BamHI; Hc=HincII.

5.2.4 MAPPING FOR HaeIII SITES

HaeIII cleaves pUC19 in many positions. In particular there are HaeIII sites near both ends of the multiple cloning site. When the 2.9kb "Variola" recombinant was subjected to HaeIII digestion two bands of 1.7kb and 1.2kb were generated in addition to many smaller fragments which were derived from pUC19. One HaeIII site is therefore present in the insert 1.2kb from one end and 1.7kb from the other end. The position was confirmed by double digestion with HincII + HaeIII. As well as many smaller fragments, one fragment of 1.7kb was generated. The HaeIII site must therefore be 1.7kb from the BamHI site. If it were 1.7kb from the EcoRI site HincII would have cleaved the 1.7kb fragment and left the intact 1.2kb fragment. Instead HincII cleaved the 1.2kb fragment, leaving an intact 1.7kb fragment (figure 5.2.7).

When the "Monkeypox" recombinant plasmid was digested with HaeIII one large fragment of 2.4kb was generated together with the many smaller fragments. The insert is 2.4kb; therefore no HaeIII sites are present in the insert.

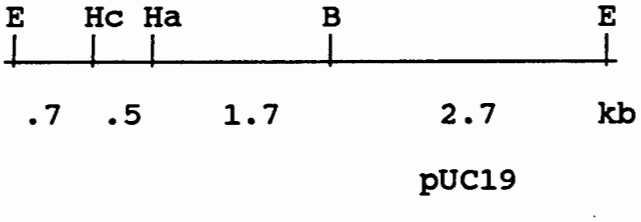


Figure 5.2.7 Location of the HaeIII site in the 2.9kb EcoRI/BamHI fragment of VarD. E=EcoRI; B=BamHI; Hc=HincII; Ha=HaeIII.

5.2.5 LOCATION OF KpnI SITES IN THE 2.9kb EcoRI/BamHI FRAGMENT OF VarD AND THE 2.4kb EcoRI/BamHI FRAGMENT OF MPE

KpnI linearized both the recombinant plasmids as well as pUC19. During the EcoRI/BamHI cloning procedure the KpnI site was removed; one KpnI site is therefore present in each of the inserts. The position of this site was established by double digestion with KpnI + BamHI and confirmed by KpnI + EcoRI double digestion. KpnI + BamHI generated two fragments from each of the recombinant plasmids. A doublet of 2.8kb from the "Variola" recombinant; and fragments of 2.3kb and 2.8kb from the "Monkeypox" recombinant. In each case the 2.8kb fragment consists of 2.7kb of pUC19 and 0.1kb of insert. A KpnI site is therefore present 0.1kb from the EcoRI site in both recombinant plasmids (figure 5.2.8). The remaining insert DNA was the other 2.8kb fragment for "Variola" and the 2.3kb fragment for "Monkeypox". Double digestion with KpnI + EcoRI confirmed the position of the KpnI site by generating bands of almost full size recombinant plasmids. The 0.1kb fragment, due to its small size, was barely visible.

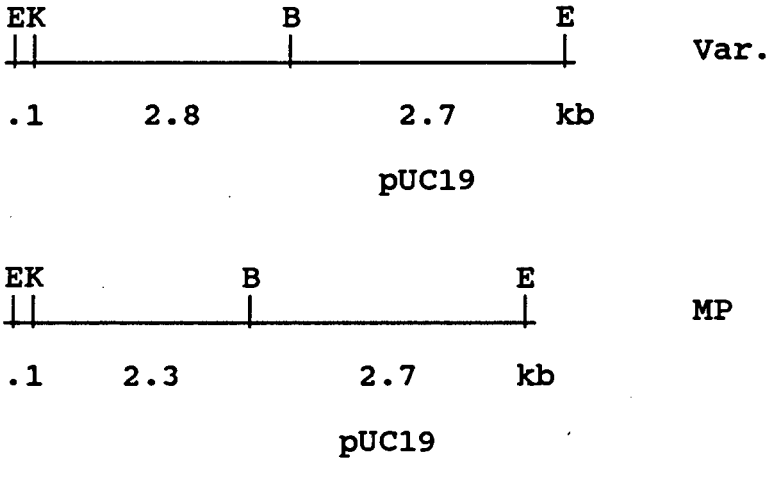


Figure 5.2.8 Location of KpnI sites in the EcoRI/BamHI fragments of VarD and MPE. E=EcoRI; B=BamHI; K=KpnI.

XhoI and SacI did not cleave either of the two recombinant plasmids. pUC19 has no XhoI site and the single SacI site present was cleaved out upon cloning into the EcoRI + BamHI sites. BglI generated 2 fragments from pUC19. It also generated two fragments from both the "Variola" and "Monkeypox" recombinant plasmids. No BglI site was destroyed during the cloning precedure; therefore neither the 2.9kb VarD fragment nor the 2.4kb MPE fragment contain a BglI site.

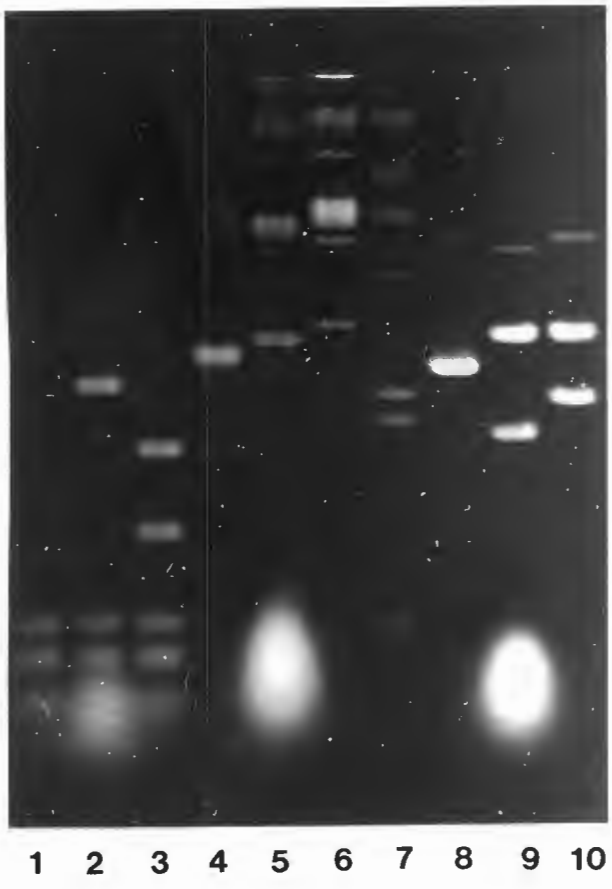


Figure 5.2.9
Electrophoresis of restriction fragments in a 1% agarose gel at approximately 2V/cm. Lanes 1,2 and 3 are HaeIII digests of pUC19, the recombinant 2.4kb EcoRI/BamHI fragment of MPE and the recombinant 2.9kb EcoRI/BamHI fragment of VarD respectively; lanes 4,5 and 6 are XhoI digests of the same plasmids respectively; and lanes 8,9 and 10 are HincII digests of the same plasmids respectively. Lane 7 is a size marker of lambda DNA digested with HindIII.

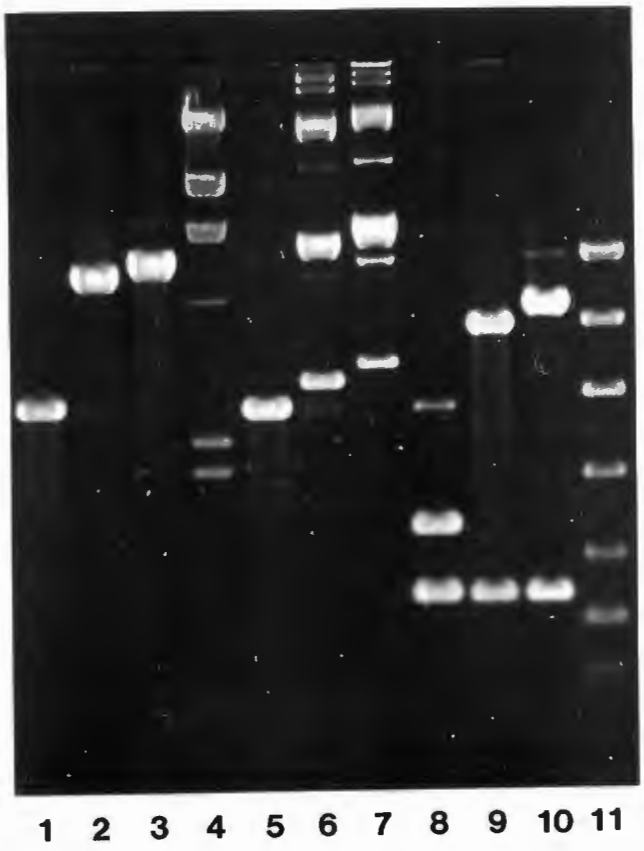


Figure 5.2.10 Electrophoresis of restriction fragments in a 1% agarose gel at approximately 2V/cm. Two size markers were used, lambda DNA digested with HindIII (lane 4) and VarD digested with EcoRI (lane 11). Lanes 1,2 and 3 are KpnI digests of pUC19, the recombinant 2.4kb EcoRI/BamHI fragment of MPE and the recombinant 2.9kb EcoRI/BamHI fragment of VarD respectively; lanes 5,6 and 7 are SacI digests of the same plasmids respectively; and lanes 8,9 and 10 are BglI digests of the same plasmids respectively.

Figure 5.2.11
Electrophoresis of restriction fragments in a 1% agarose gel at approximately 4V/cm. Lane 1 is lambda DNA digested with HindIII; lane 2 is the cloned 2.9kb EcoRI/BamHI fragment of VarD digested with HincII + HaeII; and lane 3 is pUC19 digested with HincII + HaeIII.

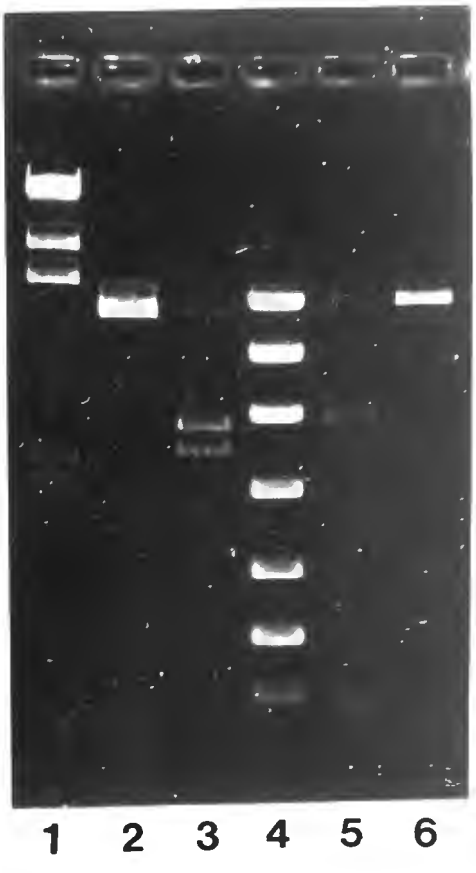


Figure 5.2.12
Electrophoresis of restriction fragments in a 1% agarose gel at approximately 4V/cm. Lanes 1 and 4 are size markers of lambda DNA digested with HindIII and VarD digested with EcoRI respectively; lanes 2 and 6 are KpnI + EcoRI digests of the recombinant 2.4kb EcoRI/BamHI fragment of MPE and the recombinant 2.9kb EcoRI/BamHI fragment of VarD respectively; and lanes 3 and 5 are KpnI + BamHI digests of the "Monkeypox" and "Variola" plasmids respectively.

5.2.6 COMPARISON OF THE 2.9kb EcoRI/BamHI FRAGMENT OF VarD WITH THE 2.4kb EcoRI/BamHI FRAGMENT OF MPE

Figure 5.2.13 compares the restriction enzyme maps of the corresponding 2.9kb and 2.4kb EcoRI/BamHI fragments of VarD and MPE respectively. The left end of these 2 fragments appear similar as they have the same restriction enzyme sites viz. KpnI, NcoI, and HincII. There is not enough evidence to postulate whether the two fragments are similar or not to the right of the HincII site. There is, however, an extra 0.5kb sequence in "Variola" (Har.). This sequence probably lies to the right of the HincII site.

A more precise location of this sequence - possibly unique to Variola - was needed for it to be sequenced (to minimize the length of DNA to be sequenced). For this purpose heteroduplexes were made of the corresponding MPE and VarD fragments (the 2.4 and 2.9 kb EcoRI/BamHI fragments respectively) and mapped with SI nuclease.

Comparison of Harvey 2,9kb Eco RI/Bam HI Fragment with

the Corresponding 2,4kb Region in Monkeypox

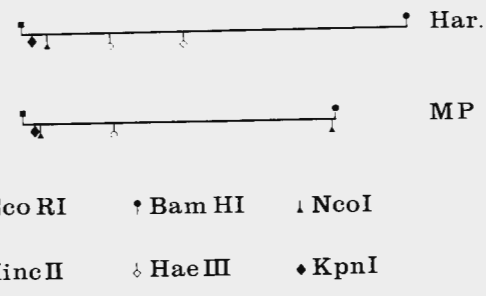


Figure 5.2.13.

5.3 SI MAPPING OF HETERODUPLEX DNA

The location of the extra sequence in Variola was determined by hybridizing the Variola DNA to the Monkeypox DNA, treating the heteroduplex formed with SI nuclease - which will digest any single stranded DNA - and separating the fragments by agarose gel electrophoresis. This experiment was performed starting with the cloned 2.9kb EcoRI/BamHI fragment of VarD and the cloned 2.4kb EcoRI/BamHI fragment of MPE linearized with 1)BamHI and 2)KpnI. Initially EcoRI was used to linearize the two fragments, but this enzyme exhibited star activity, making interpretation of the multiple bands complicated. KpnI cleaves both the Monkeypox and the Variola DNA very close to the EcoRI site; this enzyme was therefore used to linearize the plasmids. (The KpnI site of pUC19 was removed upon cloning into the EcoRI + BamHI sites).

After the two fragments were linearized, denatured, renatured, treated with SI nuclease and subjected to agarose gel electrophoresis, certain banding patterns were expected. The 2.9kb Variola fragment ligated to pUC19 would generate a homoduplex of 5.6kb and the 2.4kb Monkeypox insert would generate a homoduplex of 5.1kb (pUC19=2.7kb). Another two bands would be expected, the sum of the two being 5.1kb (this is assuming the "unique" sequence is one continuous stretch). Because pUC19 is 2.7kb the one band should be greater than 2.7kb; and because the Monkeypox insert is 2.4kb, the other band expected should be less than 2.4kb. Out of the 4 bands expected, the smaller two fragments generated should indicate where the nonhomology between the Variola and Monkeypox DNA is.

Instead of SI nuclease, Mung bean nuclease can also be used. Results obtained using Mung bean nuclease were identical to those

when SI nuclease was used.

5.3.1 ESTABLISHMENT OF CONDITIONS FOR SI NUCLEASE DIGESTION

8 Units of SI nuclease; variable incubation times

Approximately 1.5ug of each of the Variola and Monkeypox inserts was digested with BamHI (in the same tube). The enzyme was heat inactivated at 65^0C for 5 minutes, after which the DNA was precipitated (section 3.3).

The sample was resuspended in 80ul SI buffer (30mM KOAc pH4.6, 250mM NaCl, 1mM $ZnSO_4$, 5% glycerol) and divided into four tubes of 20ul each. Each tube was to be incubated for a different time length during the SI nuclease reaction - 5, 10, 30, and 60 minutes. The DNA was denatured by heating the sample to 95^0C for 5 minutes. Renaturation then took place by placing the tubes in a 65^0C water bath which was switched off. The tubes were removed after 3 hours.

8 Units of SI nuclease were added to each tube and the mixes were incubated at 37^0C. After 5, 10, 30 and 60 minutes respectively the reaction was stopped in each tube by adding 1ul of 0.5M EDTA and storing the tubes on ice.

2ul 10 X Stop buffer was added and the samples loaded onto a 1% agarose gel. After electrophoresis at 2V/cm the gel was stained and photographed.

Results

Figure 5.3.1 shows the effect of increasing the time of SI nuclease incubation. Lane 1 is VarD digested with EcoRI and lane 6 is lambda DNA digested with HindIII. These were both used as molecular weight markers. Lanes 2 to 5 contain the same amount of

DNA treated with 8 units of SI nuclease. The incubation times of the samples were: lane 2 - 5 minutes; lane 3 - 10 minutes; lane 4 - 30 minutes; lane 5 - 60 minutes.

The bands present in each lane are the same, but are much fainter in lanes 4 and 5. From lane 2 to lane 5 there is a progressive increase in degradation of the DNA.

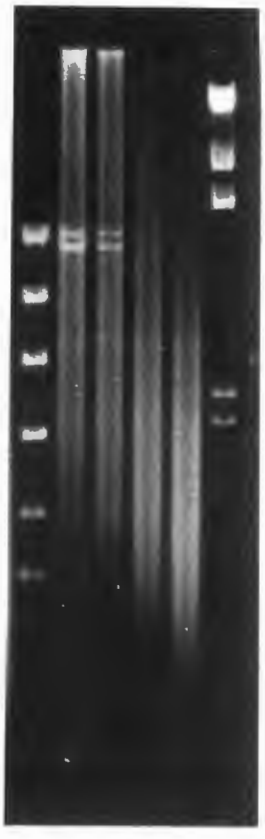
4 Units of SI nuclease; variable incubation times

The above experiment was repeated using 4 units of SI nuclease instead of 8 units. The incubation times were 5, 10 and 20 minutes. As controls homologous plasmid was treated in the same way for both the "Monkeypox" and "Variola" recombinants.

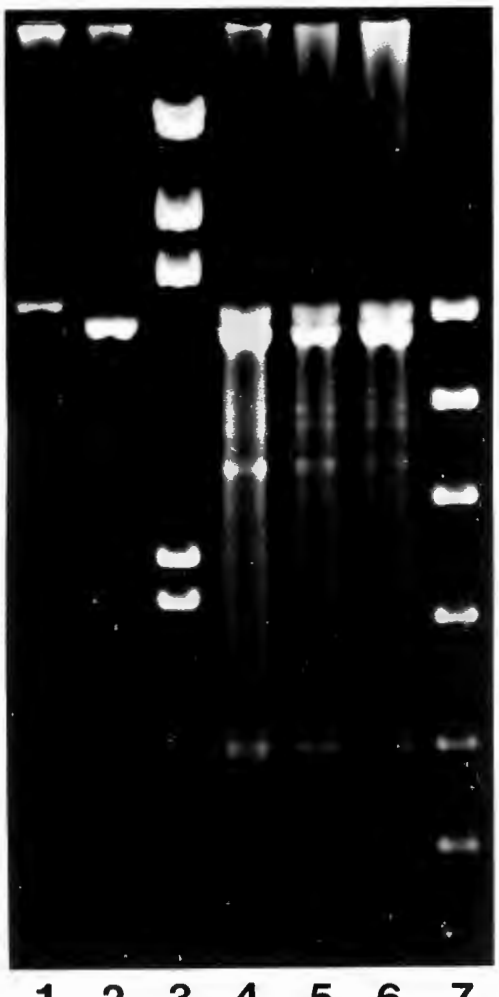
The results of this experiment are shown in figure 5.3.2. Lanes 3 and 7 are molecular weight markers - lambda DNA digested with HindIII and VarD digested with EcoRI respectively; lanes 2 and 1 are the homologous "Monkeypox" (5.1kb) and "Variola" (5.6kb) recombinants respectively; and lanes 4, 5 and 6 are the heterologous plasmids incubated for 20, 10 and 5 minutes respectively with SI nuclease.
Present in the control lanes was one band the size of the linearized plasmid. This means that bands present in the test lanes resulted from SI digestion of the heteroduplex molecules, and not from star activity of the restriction enzyme.

All of lanes 4, 5 and 6 gave interpretable results. Conditions of 4 units of SI nuclease and an incubation time of 10 minutes was used for future reactions.

Figure 5.3.1
Electrophoresis of DNA fragments in a 1% agarose gel at approximately 2V/cm. The samples were linearized with BamHI, denatured, renatured and treated with 8 units of S1 nuclease. Lanes 1 and 6 are size markers of VarD digested with EcoRI and lambda DNA digested with HindIII respectively. Lanes 2,3,4 and 5 show samples treated for 5,10,30 and 60 minutes respectively with S1 nuclease.



1 2 3 4 5 6



1 2 3 4 5 6 7

Figure 5.3.2
Electrophoresis of DNA fragments in a 1% agarose gel at approximately 2V/cm. The samples were digested with BamHI, renatured and treated with 4 units of S1 nuclease for 20 (lane 4), 10 (lane 5) and 5 (lane 6) minutes. Lanes 1 and 2 are homoduplexes of recombinant "Variola" and "Monkeypox" plasmids respectively; and lanes 3 and 7 are size markers of lambda DNA digested with HindIII and VarD digested with EcoRI respectively.

Interpretation of Figure 5.3.2

Present in the test lanes are multiple bands. The two high molecular weight fragments correspond to the homoduplexes of 5.6 and 5.1 kb. A set of three bands larger than 2.7kb are present - approximately 3.9, 3.7 and 3.2 kb; this implies there is nonhomology from approximately 0.5 to 1.2 kb from the EcoRI site (see figure 5.3.3). Another fragment of 1.3kb was generated. This means 1.3kb of DNA is homologous and mismatching probably occurs 1.3kb from the BamHI end of the insert.

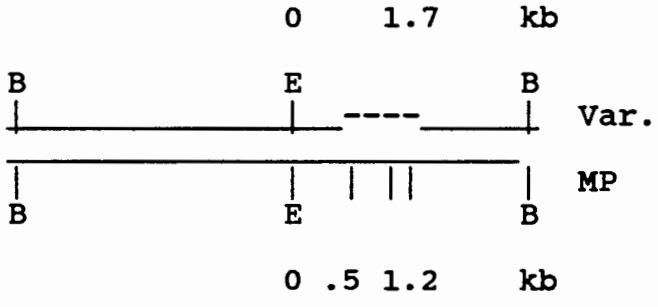


Figure 5.3.3 Interpretation of results from SI nuclease digestion of heteroduplexes. B=BamHI; E=EcoRI; Var.= 2.9kb EcoRI/BamHI fragment of VarD ligated into pUC19; MP = 2.4kb EcoRI/BamHI fragment of MPE ligated into pUC19; --- = predicted area of nonhomology.

5.3.2 LOCATION OF THE "UNIQUE" SEQUENCE IN VARIOLA DNA

The "Monkeypox" and "Variola" plasmids were linearized with KpnI (KpnI cleaves both plasmids very close to the EcoRI site). Three tubes were prepared, one of the two plasmids mixed and the other two of each plasmid separate. The samples were denatured, renatured and treated with 4 units of SI nuclease for 10 minutes as described. Figure 5.3.4 shows the samples after electrophoresis at 2V/cm through a 1% agarose gel.

Lanes 1 and 5 are molecular weight markers - VarD digested with EcoRI and lambda DNA digested with HindIII respectively. Lanes 2 and 3 are the respective "Variola" and "Monkeypox" homoduplexes; and lane 4 is the heteroduplex sample. The homoduplex controls generated the expected single bands of 5.6kb for "Variola" and 5.1kb for "Monkeypox". Bands at the same positions were present in the mixed sample (lane 4) together with other bands. One other fragment larger than 2.7kb was generated. The presence of this 4.2kb fragment means there is a difference between the pox inserts 1.5 kb from the BamHI site, or 0.9kb from the EcoRI site of the "Monkeypox" insert. A cluster of fragments of 0.9, 1.0 and 1.1 kb was generated. 2 bands of about 1.4 and 1.5 kb were also present. This means there is base mismatching around these positions (as measured from the EcoRI site).

The presence of extra bands is not unexpected as SI nuclease will cleave DNA at positions of base-mismatching. When considering all the fragments generated it seems clear that there are two regions very similar in the two DNA strands viz. 0.5kb from the EcoRI site and 1.2kb from the BamHI site. The extra sequence in the Variola DNA is therefore very likely to lie between 0.5 and 1.7kb from the EcoRI site.(See figure 5.3.3).

This region, together with sequences approximately 0.9kb from the EcoRI site in the "Monkeypox" plasmid was sequenced.

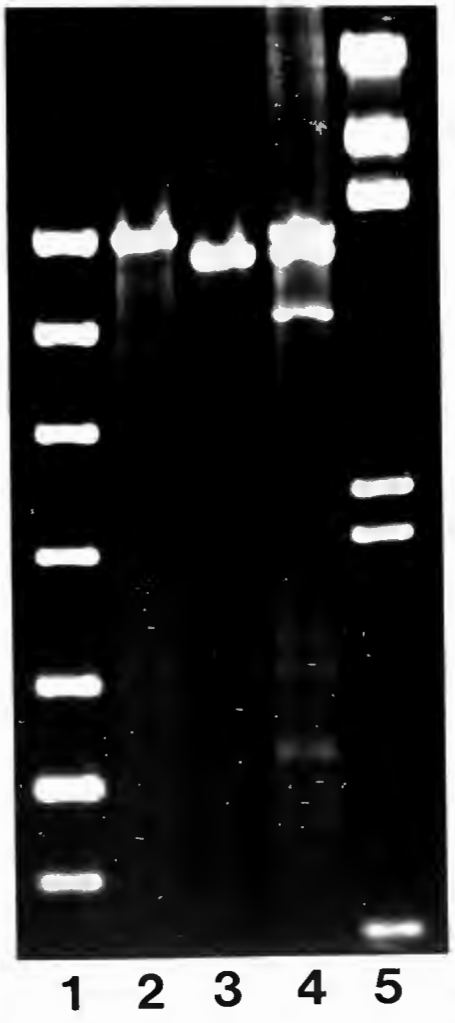


Figure 5.3.4 Electrophoresis of DNA fragments in a 1% agarose gel at approximately 2V/cm. Samples were digested with KpnI, denatured, renatured and treated with 4 units of S1 nuclease for 10 minutes. Lanes 2 and 3 are the homoduplexes of "Variola" and "Monkeypox" respectively; and lane 4 shows the heteroduplexes formed. Lane 1 is VarD digested with EcoRI and lane 5 is lambda DNA digested with HindIII.

CHAPTER 6 DNA SEQUENCING

6.1 GENERAL METHOD

The Sanger-dideoxy method of sequencing (1975; 1977) was chosen instead of the Maxam and Gilbert method (1977; 1980). The reactions are quicker and harmful chemicals are not required. In principle, this method relies on premature termination of DNA synthesis by incorporation of nucleotide analogues into the newly synthesized DNA chains.

The basic requirement is a single-stranded template containing a primer-binding site immediately upstream of the DNA to be sequenced. The single-stranded templates used were in one of two forms a) as single-stranded M13 phage DNA or b) as denatured plasmid DNA. Two primer-binding sites were employed, one on either side of the multiple cloning site of pUC19 or phage M13 (the sites are identical in the two vectors). (See appendix 1). Commercially available forward and reverse primers were used to bind to these sites.

After binding of the primer, DNA synthesis proceeded. The mixture was divided amongst 4 tubes labelled A, C, G and T. Present in the appropriate tube was a specific concentration of the dideoxynucleotide together with all the other deoxynucleotides and the Klenow fragment of DNA Polymerase I. The newly synthesized DNA chain was radiolabelled with ^{32}P-dCTP or ^{35}S-dCTP. The prematurely terminated DNA strands were separated on a denaturing polyacrylamide gel, which was subsequently dried and autoradiographed. DNA sequences were read from 5′ to 3′ from the bottom to the top. (See figure 6.1.1 for a flow diagram of the reaction.)

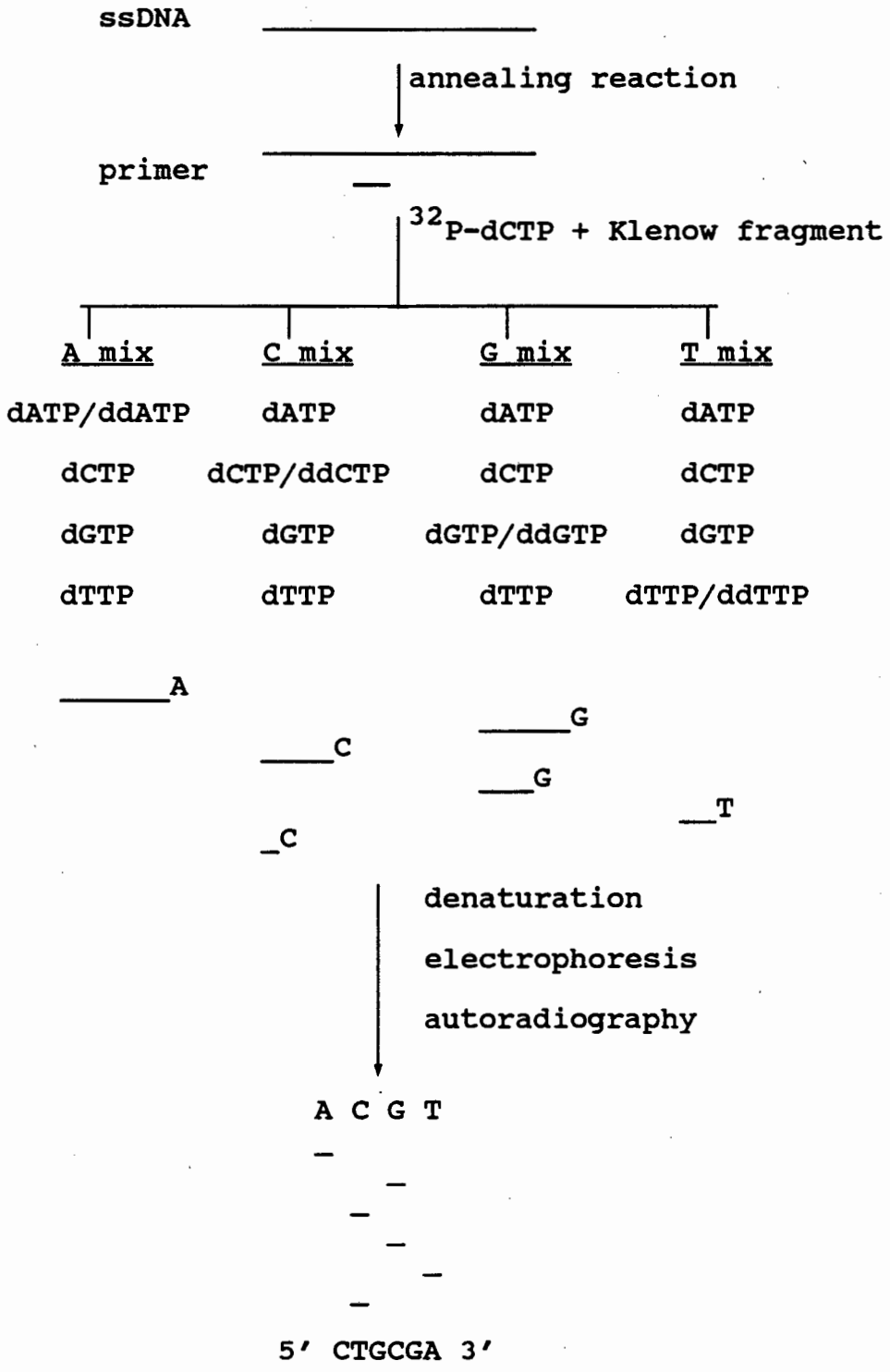


Figure 6.1.1 Flow diagram of the sequencing reaction.

M13 is a single stranded filamentous phage which can enter E.Coli F' cells by way of the F pilus. Once in the cell, the phage is in a circular single-stranded form with no protein coat. A complementary strand is synthesized to give rise to a double-stranded replicative form (RF). This RF DNA replicates to give more than 100 copies per cell. By means of the rolling circle method, single stranded M13 phage DNA is synthesized and subsequently packaged into coat protein. The virus (>200 per cell) is extruded out of the cell without causing cell lysis. The single stranded DNA can be isolated and sequenced.

Single-stranded DNA for sequencing was prepared by cloning the DNA of interest into the double stranded replicative form of M13. A competent host was transformed with this recombinant and single stranded DNA was isolated.

The selection of recombinants was by colour, the principal being similar to selection of pUC recombinants (section 4.4.1). Inserts were cloned into the multiple cloning site present within the B-galactosidase gene. In the presence of IPTG and X-GAL cells harboring intact M13 appear blue, whereas those containing foreign DNA inserted into M13 appear colourless.

The growth of cells infected with phage is retarded. These cells are therefore identified as plaques on a lawn of uninfected cells. For amplification of recombinant M13 phage, colourless plaques were picked out, grown up in larger quantities and DNA was isolated.

6.1.1 M13 CLONING

The technique for cloning into M13 and subsequent sequencing was based on that described in the handbook provided by Amersham upon delivery of the sequencing kit (cat. no. N4502).

The vector M13mp19 (Messing 1983) was cleaved with the appropriate restriction enzyme and treated with Calf Intestinal Phosphatase as previously described (section 4.4.1). The DNA fragment to be sequenced was treated with the same enzyme and ligated into the vector as described in section 4.4.1, an excess of insert molecules being present. The relevant controls were set up.

Transformation

5ul (~5ng of vector) of DNA was added to 300ul of competent JM107 cells (prepared as in section 4.4.1). These were stored on ice for 40 minutes in a 6ml tube. The cells were heat-shocked at 42^0C for 3 minutes and returned to ice.

During the heat shock step a mixture was prepared of 200ul JM107 cells (fresh exponential culture), 40ul IPTG (100mM) and 40ul X-GAL (2% in dimethylformamide). The exponential culture was prepared by inoculating 10ml of Luria broth with 200ul of an overnight culture and incubating at 37^0C with shaking for 1 hour.

280ul of fresh cells/X-GAL/IPTG mix was added to each tube of heat-shocked cells. 3ml of 2 X YT Sloppy agar prewarmed at 42^0C was added, the tube was inverted and the contents poured onto a prewarmed (37^0C) IST plate. The plates were left at room temperature (approximately 30 minutes) to set after which they were incubated at 37^0C in an inverted position overnight.

Results

The test plates contained 5 to 10 colourless plaques as well as some blue ones. Both the intact phage DNA and the vector DNA digested but not treated with alkaline phosphatase gave rise to approximately 200 blue plaques. In comparison, the vector digested and treated with CIP generated about 20 blue plaques. The "CIPped" vector also generated one colourless plaque, possibly due to a deletion mutant. In each case approximately 5ng of vector was used for transformation.

6.1.2 PREPARATION OF SINGLE-STRANDED TEMPLATE

20ml of Luria broth was inoculated with 200ul of an overnight culture of JM107 cells. 1.5ml aliquots were dispensed into 5ml tubes which were subsequently inoculated with a colourless plaque.

The tubes were shaken vigorously at 37^{0}C for 4.5 hours.

The suspensions were transferred to microfuge tubes and centrifuged for 5 minutes. After pouring the supernatant (which contains single-stranded phage) into a fresh tube it was recentrifuged to remove all the cells.

The supernatant was added to 200ul 20% PEG, 2.5M NaCl, shaken and left at room temperature for 15 minutes. After centrifuging for 5 minutes and discarding the supernatant, the tube was recentrifuged for two minutes. During this step the virus was precipitated. All traces of PEG were removed with a drawn out pasteur pipettte and the pellet was dissolved in 100ul TE.

50ul Phenol was added, vortexed and left at room temperature for 15 minutes. 50ul chloroform:isoamylalcohol was added and

vortexed. The phases were separated by centrifugation, the top aqueous layer retained and the DNA precipitated as described in section 3.3.

The DNA pellet was redissolved in 50ul sterile distilled water and 5ul was subjected to agarose gel electrophoresis.

Results

DNA isolated from one blue plaque and 11 white plaques was run on a 0.7% agarose gel at 1.5V/cm. Two of the samples contained inserts. The others may be deletion mutants of M13 (see figure 6.1.1).

The RF was prepared from the cell pellet by means of a standard plasmid preparation.

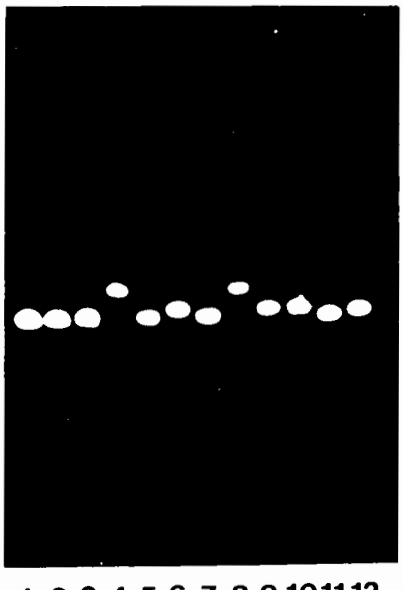


Figure 6.1.1 Electrophoresis of single stranded phage DNA at 1.5V/cm in a 0.7% agarose gel. DNA in lanes 1 to 11 was isolated from white plaques and lane 12 is M13 phage DNA isolated from a blue plaque. Lanes 4 and 8 show samples containing inserts.

6.1.3 THE SEQUENCING REACTION

The primer was annealed to the template by mixing 7.5ul single-stranded DNA with 2ul M13 forward primer (1.2ug/ml) and 1.5ul 10 X Klenow reaction buffer (100mM tris, pH8.0, 50mM $MgCl_2$), heating the mixture for 5 minutes at 90 - 100^0C and allowing the test tube to cool to room temperature for 30 to 45 minutes.

1ul ^{32}P-dCTP (>400 Ci/mmol) was added followed by 1ul of Klenow fragment (1 unit/ul). The contents were spun to the bottom of the tube. 2.5ul of the annealed template/primer/label/enzyme mix was added to each of 4 tubes labelled A, C, G and T. The mix consisted of an equal volume of N^0 (see Materials) and ddNTP. The ddNTP concentrations used were: ddATP - 0.15mM; ddCTP - 0.03mM; ddGTP - 0.1mM; ddTTP - 0.8mM.

The reaction proceeded for 15 minutes at 40^0C, followed by a chase reaction at room temperature which required the addition of 2ul of 0.5mM dCTP.

4ul Formamide dye mix was added to stop the reaction. Before loading the samples onto a gel they were denatured by heating at 90^0C for at least 3 minutes.

6.1.4 POLYACRYLAMIDE GEL ELECTROPHORESIS

Oligonucleotides of different lengths were separated on a 6% polyacrylamide gel under denaturing conditions (7M urea). The temperature of the gel was kept high to prevent secondary structures forming within the DNA molecules. To maximize the number of bases determined from a single gel, buffer gradient gels were poured (Biggin et al. 1983). The use of ^{35}S-dCTP instead of P-32 labelled nucleotides also increased resolution of bands (discussed later). A higher buffer concentration at the bottom of the gel than at the top causes retardatation of the faster migrating smaller DNA fragments. The spacing of the bands produced is more even with the result that approximately 200 nucleotides can be read from one reaction (using P-32 labelled nucleotides).

Two systems of sequencing gel electrophoresis were used; the one was a BRL Model SO and the other was a Hoefer sequencer, SE 1500.

Siliconisation of glass plates

Siliconising the glass plates before use has two advantages, a) it aids pouring of the gel and b) it prevents adhesion of the gel to the plates when the plates are separated.

The plates were well washed with detergent and rinsed with boiling water. After treatment with an organic solvent such as xylol, they were cleaned with ethanol. To siliconise the plates dimethyldichlorosilane was poured onto the larger plate, the smaller plate was slid over the larger one and the pair was left for approximately 10 minutes. The plates were parted and well washed with detergent followed by rinsing with boiling water.

The plates were used approximately 10 times before being resiliconised. After each use they were washed with detergent and rinsed with boiling water.

Preparation of the gel

The two plates were clamped together as recommended by the suppliers.

Preparation of a buffer gradient gel requires two acrylamide solutions (see Materials). The "top gel" is less concentrated than the "bottom gel". A dye is included into the bottom gel which enables one to assess the gradient.

12ml of bottom gel was added to 12ul freshly prepared 25% ammlnium persulphate and 12ul TEMED. In a separate beaker 50ml of top gel was added to 75ul 25% ammmonium persulfate and 75ul TEMED. Approximately 10ml of top gel was drawn up in a syringe. Using the same syringe the bottom gel was drawn up and slightly mixed so that a gradient formed in the syringe. This mixture was poured between the glass plates, care being taken to avoid air bubbles forming. The remaining top gel was poured between the plates and a comb inserted into the top of the gel where the samples were to be loaded. The gel was allowed to set for at least one hour.

The comb was removed and the gel clamped onto the apparatus. Prewarmed TBE was poured into the buffer tank(s) of the apparatus and the well at the top of the gel was washed out with buffer to remove unpolymerised acrylamide. The BRL model SO apparatus had two buffer tanks - one at the top and one at the bottom. The "Poker face" sequencer, being an improved model, had one large tank, the small glass plate forming the one wall of the chamber.

Allowing the whole area to be in contact with buffer ensures even distribution of the heat and reduces the "smiling" effect at the edges of the gel.

A sharkstooth comb was inserted into the well at the top of the gel. This formed wells into which the samples could be loaded.

Running the gel

Before loading the samples the gel was prerun. The voltage and current were variable and the power kept constant. The BRL apparatus was run at 67W and the "Poker face" gel at 85W. A higher voltage could be applied to the "Poker face" apparatus due to the heat distribution over the whole gel. If the BRL apparatus had too high a voltage applied to it, the heat generated would cause the glass plates to crack.

The time of prerunning did not prove to be critical, as long as the gel was warm before loading the samples. This was achieved by prewarming the TBE buffer.

Before loading the samples the wells were washed out with TBE to remove urea which leaches out of the gel. The presence of urea in a well prevents a sharp band forming upon loading the gel.

The samples were heated for at least 3 minutes at 90^{0}C before being loaded. 3ul were loaded using a Gilson's pipette. A Hamilton syringe and a drawn out capillary tube were also used for loading samples, but the Gilson's pipette proved to be adequate and easier.

The samples were run for 2 to 5 hours. For determining the maximum sequence from one sample two loadings were performed, one allowed to run for 2 hours and the other for approximately 4 hours.

Drying the gel

After electrophoresis the two glass plates were prized apart, the gel sticking to one plate only. A piece of 3MM paper the size of the large glass plate was placed on top of the gel. The gel stuck to the 3MM paper and was peeled away from the glass plate. Saran wrap was laid over the gel and the 3MM paper containing the gel placed on a SE 1150 slab gel dryer (Hoefer Scientific Instruments). The gel was dried at 80^{0}C under a vacuum for approximately 3 hours.

Finally, the dried gel was autoradiographed, the time of exposure varying from a few hours to a few days.

Figure 6.1.2 Autoradiogram of a sequencing gel. Phage DNA was used as a template and the newly synthesized DNA was labelled with ^{32}P-dCTP.

6.2 SEQUENCING OF PLASMID DNA

Instead of cloning the DNA to be sequenced into M13 - phage the technique of sequencing directly from pUC19 was instigated (Zagursky, R. - personal communication through Stein, L.M. 1984). To obtain single-stranded template the plasmid was denatured by an alkaline treatment. The sequencing reaction was much the same as that described for M13 sequencing.

2ul 2M NaOH was added to 8ul plasmid DNA resuspended in sterile distilled water. The DNA was denatured at room temperature for 5 minutes.

3ul Potassium acetate (3M K; 5M acetate), 2ul 4M LiCl, 5ul primer (10ng/ul) and 75ul absolute ethanol were added to the sample, which was subsequently placed on dry ice for 30 minutes. The tube was often left on dry ice overnight. The tube was centrifuged for 5 minutes to pellet the DNA. After a 70% ethanol wash the sample was vacuum dried and resuspended in 10ul annealing buffer (100mM Tris, pH8.0, 50mM $MgCl_2$).

The annealing reaction proceeded at 40^0C for 15 minutes. During this time the tubes for the sequencing reactions were prepared. 3ul of each dNTP/ddNTP mix was pipetted into a sterile microfuge tube and preheated at 40^0C. The ddNTP concentrations used for preparing the mixes were : ddATP - 0.15mM; ddCTP - 0.1mM; ddGTP - 0.1mM; ddTTP - 0.4mM.

To the annealing reaction 1ul DNA Polymerase I - Klenow fragment (5U/ul) and 4ul radiolabelled dCTP (>400Ci/mM) were added. The tube was microfuged for 1 second to mix the contents and 3.5ul was dispensed into each dNTP/ddNTP reaction tube. To save radiolabelled dCTP only 2ul was sometimes added in which case 3ul

of the mixture was dispensed into each reaction tube.

The reaction tubes were microfuged for 1 second and incubated at 40^{0}C. If P-32 was used to label the DNA the reaction proceeded for 15 minutes; for incorporation of S-35 label 20 minutes was allowed for the reaction.

1.3ul Chase solution containing 0.5 mM dATP, dCTP, dGTP and dTTP was added to each tube and the reaction continued at room temperature for 15 minutes.

4 to 6 ul formamide dye mix was added to each tube.

Before loading the samples onto a 6% polyacrylamide gradient gel they were denatured by heating at 90^{0}C for at least 3 minutes. Elecrophoresis was performed as described in section 6.1.4 followed by drying and autoradiography.

6.3 REDUCING THE SIZE OF INSERTS

DNA stretches of approximately 2000 base pairs had to be sequenced. But only 200 bases could be read from one clone. To obtain different size DNA clones the original recombinant plasmid was treated with Exonuclease III followed by SI nuclease or Mung bean nuclease and subsequently religated (Henikoff 1984). This targetting of deletion breakpoints makes sequencing much quicker than random subdivision and cloning of the DNA. Random cloning may allow for sequencing errors in repetitive DNA. A nonrandom method uses BAL31 to obtain a set of deletion clones. This technique includes a gel purification step which probably makes it more time-consuming than the Exonuclease III method.

Another method of obtaining overlapping clones makes use of a transposon mediated deletion vector (Ahmed 1985; Peng and Wu 1986). This technique requires neither in vitro enzymatic manipulations for obtaining deletions nor subcloning. Specific restriction sites within the plasmid are not required. The deletions arise via integration of the vector - containing a portion of the Tn9 transposon and a truncated gal operon - into the gal genes. Deletion mutants are isolated by positive selection for galactose resistance. (Cells with functional gal genes convert galactose into UDP-galactose, which kills them). The deletions occur naturally and extend from a fixed site at the transposon terminus to variable sites on adjoining DNA. The required deletions are isolated by separating plasmids from a preparation from many colonies by agarose gel electrophoresis. Plasmids are eluted, a competent strain of cells is transformed and the plasmids are isolated for sequencing. Peng and Wu (1986) describe how single-stranded DNA can be obtained for sequencing, making use of an M13 helper phage.

The enzymatic manipulations using Exonuclease III and Mung bean nuclease were used for obtaining deletion clones as the method was established before the above method was known to exist. Exonuclease III digests DNA at a uniform rate from the 3′ terminus. Digestion only proceeds from an intruded or blunt-ended 3′ end, 3′ protruding ends being protected. Both SI nuclease and Mung bean nuclease cleave single-stranded DNA, but Mung bean nuclease is thought to be more specific than SI nuclease for single-stranded DNA.

From the inserts to be sequenced an ordered set of deletion clones were obtained by first digesting the recombinant plasmid with two restriction enzymes - one which generates a protective 3′ overlapping end preventing digestion of the vector DNA; and another which generates a 5′ overhang allowing digestion into the inserted DNA. After exonuclease digestion for strict time periods, SI nuclease digestion, Klenow DNA Polymerase I treatment and ligation, competent cells were transformed. The deletion clones were isolated by minipreps (section 3.2).

Method

Approximately 5ug of an insert of ~2kb cloned into pUC19 was double digested. The one enzyme generated a 5′ protrusion adjacent to the insert from which deletions were to be made and the other enzyme left a 4 base 3′ protrusion protecting the vector. After digestion the DNA was precipitated as described in section 3.3.

The pellet was dissolved in an appropriate volume of Exo buffer (66mM Tris-HCl pH8, 0.66mM $MgCl_2$) e.g. 60ul. Before the next step 12 tubes were prepared with 15ul Exo stop buffer in each (0.2N NaCl, 5mM EDTA pH8).

Approximately 1000 units (a large excess) of Exonuclease III was added to the DNA, mixed rapidly and incubated at 37^0C. Aliquots of 5ul were removed and mixed with Exo stop buffer at 27 second intervals after an initial lag of 20 to 30 seconds. Exonuclease III proceeds at 445 bases/minute; after each 27 second interval an extra 200 bases should have been deleted. The first aliquot was removed immediately after mixing, the second one after 47 seconds and subsequent aliquots were taken at 27 second intervals. The 11th sample should have digested 2kb of the one strand and sample number 12 should have digested 200 bases into pUC19.

The Exonuclease III enzyme was inactivated by heating the samples for 10 minutes at 70^0C. Precipitation of the DNA followed as described in section 3.3.

Each pellet was dissolved in 50ul SI mix (30mM KOAc pH 4.6, 250mM NaCl, 1mM $ZnSO_4$, 5% glycerol, 67 Vogt units/ml SI nuclease or 150 units/ml Mung bean nuclease). After 30 minutes at room temperature the reactions were stopped by addition of 6ul SI stop (0.5M Tris pH8, 0.125M EDTA). Each sample was extracted with phenol and chloroform and precipitated as described in section 3.3.

The next step was included in the first of the experiments done, but was subsequently omitted. It is a filling-in reaction to ensure that all the molecules are blunt ended. Each pellet was dissolved in 10ul Klenow mix (10mM Tris pH8, 5mM $MgCl_2$, 10 units/ml large fragment DNA Polymerase I), incubated at 37^0C for abour 2 minutes before adding 1ul dNTPs (0.5mM each of dATP, dCTP, dGTP, dTTP) and incubated again for about 2 minutes at 37^0C.

40ul Ligase mix (66mM Tris pH 7.6, 6.6mM $MgCl_2$, 10mM DTT, 100ug/ml BSA, 1mM spermidine, 0.2mM ATP, 25 units/ml T4 DNA ligase) was added and ligation proceeded overnight at 15^0C.

Transformation of competent cells followed as described in section 4.3.1.

Plasmid DNA was isolated as in section 3.2 and the colonies containing the required size inserts were kept for sequencing. All samples to be sequenced were treated with RNase.

Comments

Sometimes plasmids of varying sizes were generated from one sample. Initially the enzyme was thought to be limiting in the reaction - which could explain the presence of some larger fragments - but even using the large excess of Exonuclease III different size fragments were generated. Many of the plasmids were smaller than expected. This could be due to pre-existing nicks in the double stranded plasmid or to breakage during SI nuclease digestion.

A possible limitation of this method is the requirement for two restriction enzyme sites with the appropriate overhanging ends. This drawback has been overcome with the construction of better vectors containing many restriction sites suitably located in the multiple cloning site, for example bluescript and the PMTL series.

6.4 AREAS CHOSEN AND PREPARATION OF SAMPLES FOR SEQUENCING

The 1.9kb stretch at the right end of both MPE and VarD were sequenced. The VarD fragment had been cloned as a 1.9kb EcoRI fragment into pUC19 (section 4.3.1). PstI and BamHI were used to protect the vector and allow Exonuclease III digestion respectively. Samples were obtained as described in section 6.3 and are shown in figure 6.4.1 linearized with EcoRI. The first lane is the full 1.9kb insert in pUC19 linearized with PstI and lane 10 is pUC19 linearized with EcoRI.

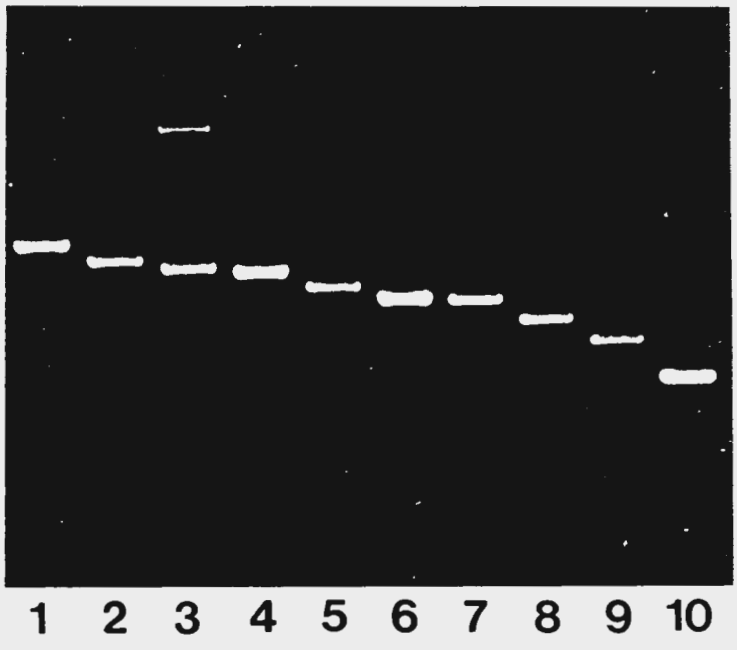


Figure 6.4.1 Electrophoresis of linearized plasmids in a 1% agarose gel at approximately 2V/cm. Lane 1 is the 1.9kb EcoRI fragment of VarD cloned into pUC19 and lane 10 is pUC19. Lanes 2 to 9 are progressive deletions of the 1.9kb insert in pUC19. All plasmids were linearized with EcoRI, except for the full 1.9kb insert which was digested with PstI.

The right end of MPE was cloned as an EcoRI/HindIII fragment into pUC19. (Figure 6.4.2 shows the insert cleaved out with EcoRI and HindIII as well as the linearized plasmid of 4.6kb). EcoRI and HindIII lie at the two extremes of the multiple cloning site. This means there are no sites available for protection of the vector from exonuclease digestion. To overcome this problem the 1.9kb EcoRI/HindIII fragment was cloned into pBR328 cut with EcoRI and HindIII and, together with a 0.3kb HindII/BamHI fragment of pBR328, recloned into the EcoRI + BamHI sites of pUC19 (see figures 6.4.3 and 6.4.4).

Exonuclease III digested the inserted DNA from the BamHI site, but the vector was protected by PstI digestion. Samples isolated for sequencing are shown in figure 6.4.5 linearized with HindIII. Lane 1 shows the linearized 2.2kb insert in pUC19 (4.9kb) and lane 12 contains linear pUC19.

Figure 6.4.2
Electrophoresis of restriction fragments in a 1% agarose gel at approximately 2V/cm. Lane 1 is a size marker of VarD digested with EcoRI; lane 2 is the 1.9kb EcoRI/HindIII fragment of MPE cloned into pUC19; and lane 3 is the insert cleaved out with EcoRI + HindIII.

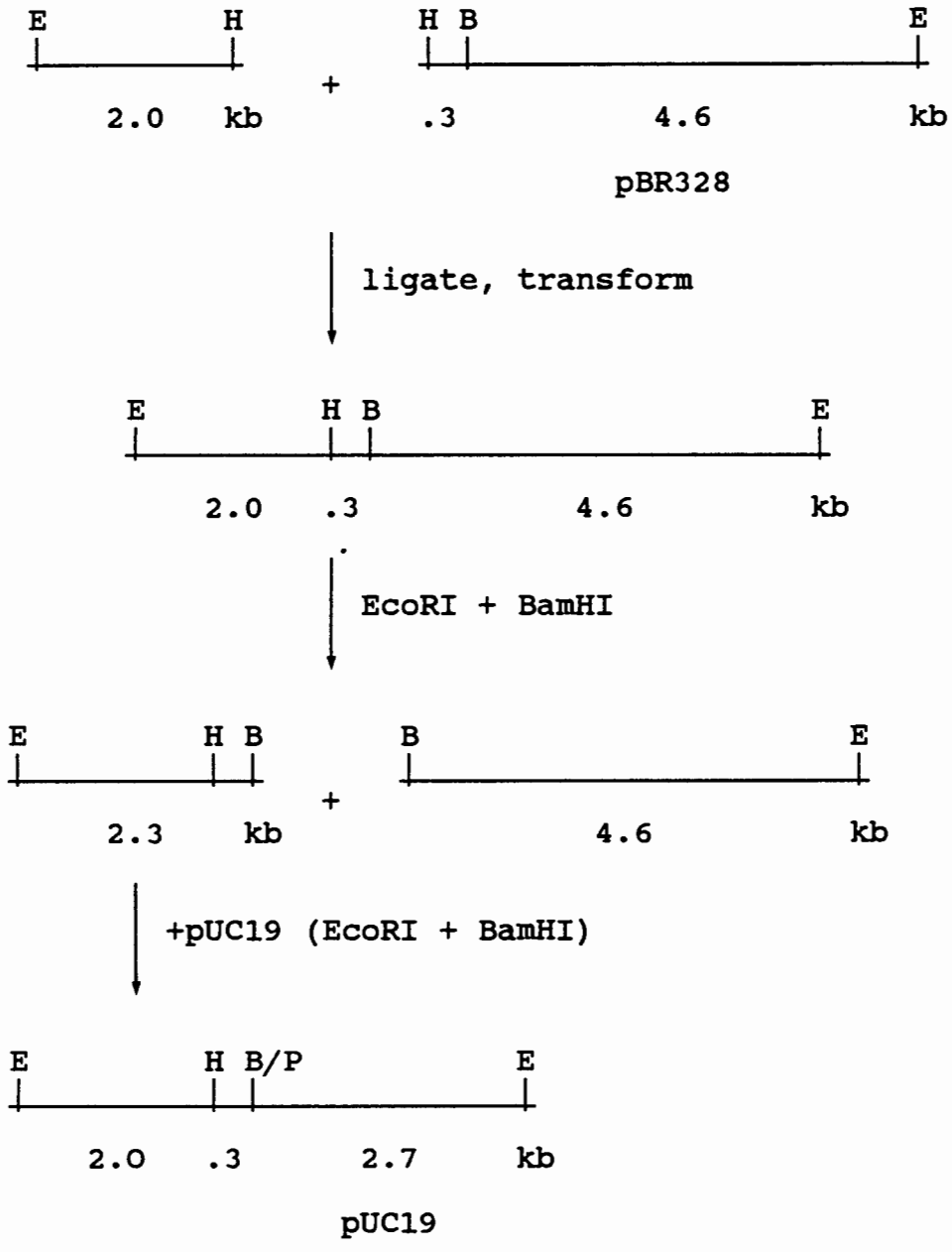


Figure 6.4.3 Cloning strategy employed prior to Exonuclease III treatment of the right end 2kb fragment of MPE. E=EcoRI; H=HindIII; B=BamHI; P=PstI.

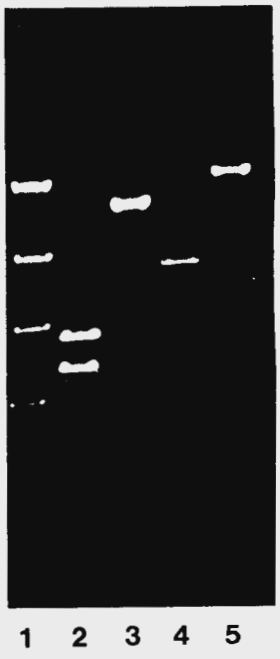


Figure 6.4.4
Electrophoresis of restriction fragments in a 1% agarose gel at approximately 2V/cm. Lane 1 is VarD digested with EcoRI as a size marker; lanes 4 and 5 are the cloned 1.9kb EcoRI/HindIII fragment of MPE in pBR328 digested with EcoRI + HindIII and EcoRI respectively; and lanes 2 and 3 are the cloned 2.2kb EcoRI/BamHI fragment in pUC19 digested with EcoRI + BamHI and BamHI respectively.

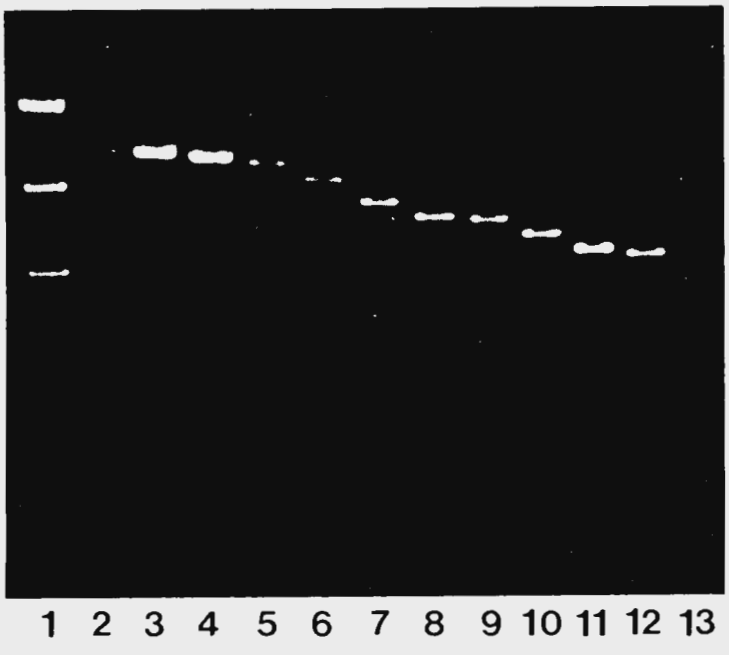


Figure 6.4.5 Electrophoresis of plasmids used for sequencing at approximately 2V/cm in a 1% agarose gel. Lane 1 is a size marker of VarD digested with EcoRI; lane 2 is the 2.2kb EcoRI/BamHI fragment in pUC19; lane 13 is pUC19; and lanes 3 to 12 are progressive deletions of the 2.2kb EcoRI/BamHI fragment cloned into pUC19. The plasmids were linearized with HindIII.

A small stretch of DNA to the right of the central BamHI site in both MPE and VarD was sequenced (figure 4.5.1). The 0.3kb EcoRI/BamHI fragment of MPE cloned into pUC19 was sequenced and compared to the corresponding sequence in the 1.1kb EcoRI/BamHI fragment of VarD.

Finally, the "heterologous" sequence of VarD was located by sequencing slightly less than 1kb in the 2.9kb EcoRI/BamHI fragment cloned into pUC19. 462 base pairs in the corresponding region of MPE (2.4kb EcoRI/BamHI fragment cloned into pUC19) was sequenced. Slightly downstream of this region approximately 150bp were sequenced in both the Variola and Monkeypox DNA.

Samples were obtained for sequencing the specific area required by treating the cloned 2.9kb EcoRI/BamHI fragment of VarD and the 2.4kb EcoRI/BamHI fragment of MPE with Exonuclease III for the appropriate length of time. Cleavage with PstI protected the vector whilst Exonuclease III digested the insert DNA from the BamHI site. Reactions were performed as described in section 6.3 and DNA was isolated. DNA clones sequenced are shown in figure 6.4.8. They have been digested with EcoRI + HindIII to release the insert DNA from the vector. Lanes 2 to 5 are "Variola" samples, lane 7 is a "Monkeypox" sample and lanes 1 and 6 are molecular weight markers (VarD digested with EcoRI and lambda DNA digested with HindII respectively).

Another sample sequenced (in the opposite direction) was the subcloned 1.7kb HincII fragment from the cloned 2.4kb EcoRI/BamHI fragment of MPE. (See diagram 6.4.6.)

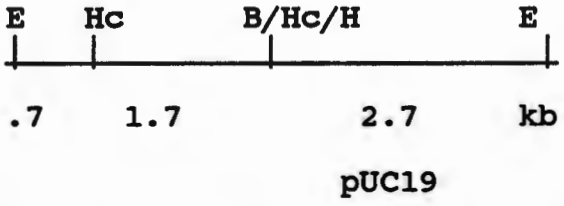


Figure 6.4.6 Diagram of the cloned 2.4kb EcoRI/BamHI fragment of MPE showing the positions of the HincII sites. E=EcoRI; Hc=HincII; B=BamHI; H=HindIII.

Figure 6.4.7 shows two recombinant plasmids with the 1.7kb HincII fragment cloned into pUC19 in opposite orientations (a and b). In both orientations HincII will cleave the 1.7kb insert out of the 2.7kb vector (lanes 1 and 2). BamHI sites are present both in the multiple cloning site of pUC19 and at the end of the insert. Digestion with BamHI would therefore generate either two fragments of 2.7kb and 1.7kb (a); or one large fragment of 4.4kb (b).

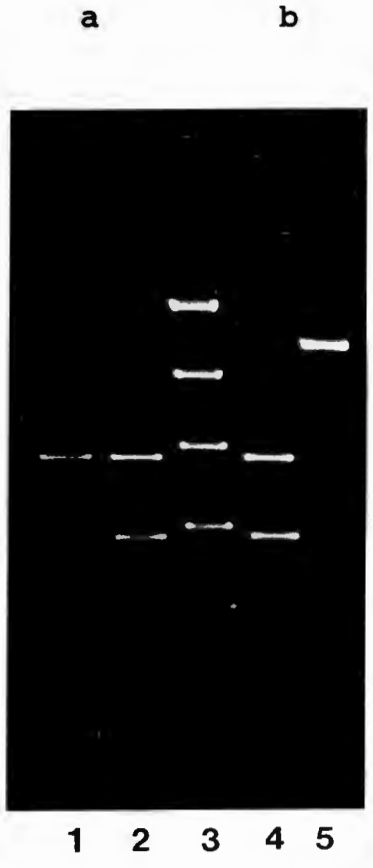


Figure 6.4.7
Determination of insert orientation by 1% agarose gel electrophoresis at approximately 2V/cm. The 1.7kb HincII fragment was subcloned from the 2.4kb EcoRI/BamHI fragment of MPE in pUC19. Lanes 2 and 4 show two recombinant plasmids digested with HincII; lanes 2 and 5 show the same plasmids respectively digested with BamHI; and lane 3 is a size marker of VarD digested with EcoRI.

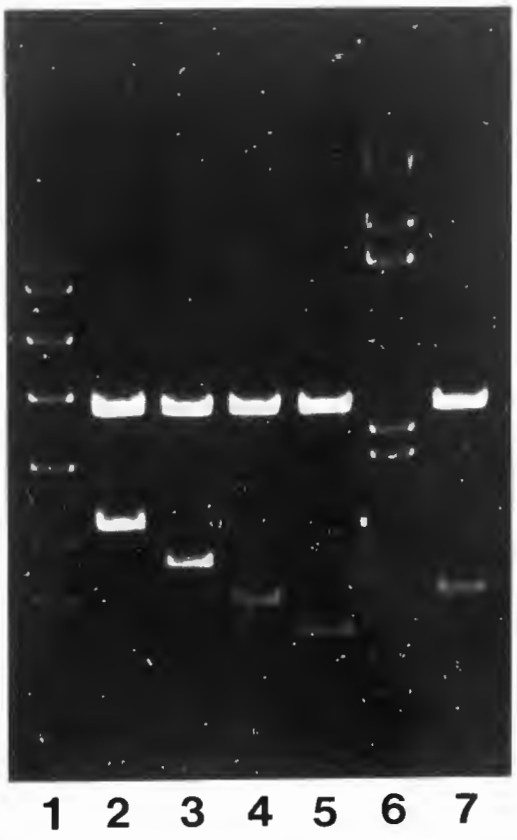


Figure 6.4.8 EcoRI + HindIII digests of plasmids used for sequencing subjected to 1% agarose gel electrophoresis at approximately 4V/cm. Deletions were made in the cloned 2.9kb and 2.4kb EcoRI/BamHI fragments of VarD and MPE respectively. Lanes 2 to 5 are "Variola" deletions; lane 7 is a "Monkeypox" deletion; and lanes 1 and 6 are size markers of VarD digested with EcoRI and lambda DNA digested with HindIII respectively.

The other two samples sequenced were approximately 1.8kb and 1.4kb from the EcoRI site in the recombinant "Variola" and "Monkeypox" plasmids respectively. Figure 6.4.9 shows the plasmids digested with EcoRI and HindIII. Present in lane 1 is the 1.8kb "Variola" sample; and in lane 3 the 1.4kb "Monkeypox" insert has been cleaved out of pUC19. Lane 2 is a molecular weight marker of VarD digested with EcoRI.

Figure 6.4.9 Electrophoresis of corresponding sections within VarD and MPE at approximately 4V/cm in a 1% agarose gel. Lane 1 shows an insert of approximately 1.8kb of Variola DNA in pUC19 and lane 3 shows an insert of approximately 1.4kb of Monkeypox DNA in pUC19. The inserts were cleaved out with EcoRI + HindIII. Lane 2 is VarD digested with EcoRI as a size marker.

6.5 RESULTS

The DNA sequences determined are shown in figures 6.5.1 to 6.5.6, a total of 5930 bases being sequenced. Most of the reactions were performed at least twice, but the complementary strands of DNA were not always sequenced. Only at the ends of each particular stretch were the complementary sequences determined.

On average 150 to 200 bases were read off one set of reactions using P-32 labelled dCTP. In comparison, the use of S-35 labelled dCTP allowed up to 400 bases to be determined from each sample. S-35 labelling resulted in sharper bands and, due to the lower radiation dose, larger fragments could be resolved in autoradiography. (See figures 6.5.7 and 6.5.8 for comparison of P-32 and S-35 labelling).

Although P-32 was used for much of the sequencing S-35 proved to be more advantageous on the whole. With a half-life six times longer than that of P-32 (87.2 days as opposed to 14.3 days), S-35 could be stored for a longer time and handled more easily. Once the reactions had been performed, S-35 labelled samples could be stored for some time at -20^{0}C. P-32, however, destroys DNA and samples could not be stored. If stored for longer than 24 hours before gel electrophoresis, fuzzy bands were generated or additional bands were formed.

The β-particles emitted by S-35 are of lower energies than those of P-32. This means less scatter is produced around the source of emission, resulting in better resolution of bands on an autoradiograph of a sequencing gel.

P-32 has the advantage of giving a quicker result. As opposed to the two days required for exposure of S-35 labelled samples, P-32

labelled samples required only a few hours of autoradiography. When less P-32 was used gels were autoradiographed overnight. P-32 was therefore more useful than S-35 for establishing the correct conditions.

To sequence further into one particular clone the ratios of dNTP/ddNTP were altered and the time of electrophoresis was increased. The ddNTP concentration was reduced by a factor of 4 so that longer DNA chains were synthesized before termination.

The 1.9kb stretches at the right end of VarD and MPE were sequenced to establish a possible heterologous region. Two areas of the Variola sequence were not connected, each gap consisting of approximately 100 bases. These gaps were not filled in as comparison of the 1784 bases sequenced with the corresponding sequence in Monkeypox showed extensive similarity. (See section 7.3). 1907 bases at the right end of MPE, reading from the EcoRI site towards the HindIII site, are shown in figure 6.5.1. The corresponding sequence in VarD is shown in figure 6.5.2, slashes representing the two unconnected regions.

To establish whether the central BamHI site of MPE and VarD were coincidental or not, DNA was sequenced to the right of these two sites. The sequences are displayed in figure 6.5.3. The two sequences are almost identical, implying that the BamHI sites are coincidental.

Figure 6.5.4 shows another stretch of DNA almost identical in the two viral genomes. This region was sequenced during the search for the "unique" stretch of Variola DNA.

Finally, figures 6.5.5 and 6.5.6 show the corresponding areas of MPE and VarD within which the extra sequence in VarD lies. The

beginning and the end of the "Variola" sequence show a high degree of similarity with the whole of the "Monkeypox" sequence; but in the middle of the "Variola" sequence an extra 389 bases are present. (See section 7.3). This stretch of DNA, present in VarD but absent in MPE, was further analysed.

EcoRI
5'

---------- ACTGTTATCT TAGTGAAATC TTTAACAAAC AGCAAGGGTT CGTCAAAGAC

ATAAAACTCA TTGTTTACGA TCGAAATAGA CCCCCTATCA CACTTAAAAT AAAAAATATC

CTTATCCTTT ACCACCAAAT AAAATTCTGA TTGGTCAATG TGAATGTATT CACTTAACAG

TTCCACAAAT TTATTTATTA ACTCCGAGCC ACAGACATCG TCGGTATTTT TTATGACAAA

CTTTACTCTT CCAGCATCCG TTCTAAAAAA ATATTAACGA GTTCCATTTA TATCATCCAA

TATTATTGAA ATGACGTTGA TGGACAGATG ATATAAATAA GAAGGTACAG TACCTTTGTC

CACCATCTCC TCCAATTCAT ACTCTATTTT GTCATTAACT TTAATGTGTG AAAACATCTT

CCATGACAGT GTGTAACACT TTGGATACAA AATGTTGACA TTAGTATAAT TGTCCAAGAC

TGTCAATCTA TAATAGATAG TAGCTATAAT ATATTCTATG ATGGTATTGA AGAAGATGAC

AACCTTGGCA TATTGATCAT TTAACACAGA CATGGTATCA ACAAATAGCT TAAATGAAAG

AGAATCAGTA ATTGGAATAA GCGTCTTCTC GATGTAGTGT CCGTATACCA ACATGTCTGA

TATTTTGATG TATTCCATTA AATTATTTAG TTTTCTTTTA TTCTCGATTA AACAGAATTT

CTGTCATGGA CCCCAACATC GTTGACCTAT TAAGTTTTGA TTGAATTTTT CCTGGTAAGG

CGTATCTAGT CAGATCGTAT AGCCTATCCA ATAATCCATA GTCTGTGCGT AGATCACATG

GTACACTTTT TAATTTTCTA TAGAAGAGTG ACAGACATCT GAAGCAATTA CAGACAGCAA

TTTCTTTATT CTCTACAGAT GTAAGATACT TGAAGACATT CCTATGATGA TGCAGAATTT

TGGATAACAC AGGTATTGAT GGTATCTGTT ACCATAATTC CTTTGACTGA TAGTGTCAAA

GTACAAGATT TCCAATCTTT TGCAATTTTC AGTACCATTA TCTTGTTTTG ATATCTATAT

Figure 6.5.1 DNA sequence of the right end of MPE (continued on the next page).

CAGACAGCAT GGTACGTCTG ACAACACAGG GATTAAGACG GAAAGATGAA AGATGAAATG

ATTCTCTCAA CATCTTCAAT AGATACCTTG CTATTTTTGG ATATGTATAT TGGAAGAATT

CCTTAAAGCT AGTATCCTTG TGATATGGCT CTATGACATG GACGTCTAAC CTCTATCTAT

CACAGTGCAT GCTGATTGTC TCTTCTTTAT CATGATGTAA TCTCTAAATT CATCGGCAAA

TGTCTATATC TAAAATCATA ATATGAGTGT TTACCTCTAC AAATATCTGT CGTCCATGTT

AGAGTATCTA TATCAGTTTT GTATTCCAAA TTAAACATGG CAACGGATTT AATTTTATAT

TCCTCTATTA AGTCCTCGTC GATAATAACA GATTGTAGAT AATCATTTAA TCCATCGCTA

ACATGGTTGG AAGATGCCGT GACAAAACCT TTAATTCTCC TGATGAAGGT GGGACTATAT

CTAACATCCT GATTAATAAA ATTTATACAA TGTCCATAGG ATACTTTGTA ACTAGTTTAT

ACACATGTGT TCATTGGTAA GTTTATACAG AATATCGTGA ACAGGTGGTA TATTATATTC

ATCAGATATA CGAAGAATAA TGTCCAAATC TATATTCTTT AATATATTAT ATAGATGTAG

TGTAGCTCCT ACAGGAATAT CTTTAACTAA GTCAATGATT TCATCAACAG TTAGATCTAT

TTTAAAGTTA ATCATATAGG CATTGATTTT TAAAAGGTAT GTAGCCTTGA CTACATTCTC

ATTAATTAAC CATTCCAAGT CACTGTGTGT AAGAAGATTA TATTCTATCA TAAGCTT--

3′

HindIII

Figure 6.5.1 DNA sequence of the right end of MPE (continued).

EcoRI

5′

GAATTCTACC CCCGTTACAT TTGTATAATC TTTAACAAAC AGTAAGGGCT CGTCAAATAC

GTAAAATCTA TTACTTACGA TAGAAATAGA TCCCTTATCA CACTTAAAAT AAAAAATATC

CTTATCCTTT ACCACCAAAT AAAATTCTGA TTGGTCTATG TGAATGTATT CACTTAACAA

TTGCACAAAT TTATTTATTA ACTCCGAGGC ACATACATCG TCGGTATTTT TTATGACAAA

CTTTACTCTT CCAGCATCCT TTTCTAAAAA AATATTAACG AGTTCCATTT ATATCATCCA

ATATTATTGA AATGACGTTG ATGGACAGAT GATGCAAATA AGAAGGTACA ATACCTTTGT

CTACCATCTC CTCCAATTTA TGCTCTATTT TCTCATTAAC TTTAATGTGT GAAAACAGAT

Figure 6.5.2 The right end of VarD. Gaps in the sequence are represented as /. The sequence continues on the next page.

TCCATAACAT GTGGTAACAC TTTTAATACG AAATGTCTGA TGGCGACATA ATTGTCCAAA

ACTGCCAATC TATAATAGAT GGTGGCCACA ATATATTCGA TGATGGTATT GAAGAAGATA

ACACCTTGGC ATATTGATCA TTTAGTACGG ACATGGTATC ACAGATACGT TGAATGAAAG

AGACTCAGTA ATCGGAATAA GTGTTTTCTC GATAGAGTGT CCGTATACCA ACATGTCTGA

TATTTTGATG TATTTCATCA AATTGTTTAG TTTT/TTTCC ATGTAAGGCG TATCTAGTCA

GATCGTATAA TCTAGCCAAC AATCCATCGT CGGTGTTTAG GTCACATCGT ATACTTTTTA

ATTCTCTATA TAATATATAC ACACACCTGG AACAATTACA GACAGTAATT TCTTTATTAT

CTACAGATGG AAGATACTTG AAGACATTCC TATGATGATG TAGAATTTTT GATAACACTG

TATTGATGGT ATCTGTTACC ATAATACCTT TTATGGCGGA CAGTGTCAGA GCACAGGATT

TCCAATCTTT TGCAATTTTC AGTACCATTA TCTTGTTTTG ATATCTATAT TAGATAAGCA

AAGATAAGCA TGGTACGTCT GACAACACAG GGATTAAGAC GAAAAGATGA AA/TATCCTT

TTTGATAATA GTCGATCTAA GTGACATGGG ACGTCTAAAG CTTCTTATTT TATCACTGTT

TGCATGGTGA TTTGTCTGCT TTCTTTTAAC ATGATGTAAT CTCTAAATTC ATTGGCAAAT

TGTCTATATC TAAAATCATA ATATGAGATG TTTACCTCTA CAAATATCTG TTCGCCAATG

TTAGAGTATC TACATCAGTC TTGTATTCCA AATTAAACAT GGCAACGGAT TTAATTTTAT

ATTCCTCTAT TAGGTTCCTT GTCTACAATT AACAATAGTG GCAGATAATC AATTAATTCC

GTCTGACAAT GGTTGGAAGA TGCGGATTAA CAAAATCTTT AATTGTCTTG ATGAAGGTGG

GACTATATCT AACATCTTGA TTAATAAAAT TTATAACATT TTCCATAGGA TACTTTTGTA

ACTAGCTTTA TACACATCTC TTCATCTGTA AGTTTAGACA GATATCGTGA CAGGTGGCTA

TATTATATTC ATCAGATATC GAGATATGTC AATCTAGATT GTGGAATATA TTATATAGAT

GTAGTGTAGC TCCTACAGGA ATATCTTTAA CTAAGTCAAT TATTTCATCA ACCGTTAGAT

CTATTTTAAA GTTAATCATA TAGGCATTGA TTTTTAAAAG GTATGTAGCC TTGATTACAT

TCTCATTAAT TAACCATTCC AAGTCACTAT GTGTAAGAAG ATTATATTCT ATGTAA---

3'

HindIII

Figure 6.5.2 1784 bases of the right end of VarD (continued). Gaps in the sequence are represented as /.

```
      5'

MP    ATAGAGAACG GAATACTAAA GAAATTTAAA TCTATAAAAT CCAAGAGTAG
                             * ***
Var.  ATAGAGAACG GAATACTAAA GAAACTGTCA TCTATAAAAT CCAAGAGTAG

MP    ACGGCTGAAC TTGTTTAGCA AAAACATTTT AAAATATTAT TTAGACGGAC
                  *
Var.  ACGGCTGAAC TTCTTTAGCA AAAACATTTT AAAATATTAT TTAGACGGAC

MP    AATTGGCTCG TCTAGGTCTT GTGTTAGATG ATTATAAAGG AGACTTATTA
                                                      *
Var.  AATTGGCTCG TCTAGGTCTT GTGTTAGATG ATTATAAAGG AGACTTGTTA

MP    GTTAAAATGA TAAACCATCT CAAATCTGTG GAGGATGTAT CCGCATTCGT
                            * *          *      *
Var.  GTTAAAATGA TAAACCATCT CAAGTTTGTG GAGGATCTAT CAGCATTCGT

MP    TAGATTTTCT ACAGATAAAA ACCCTAGTAT TCTTCCATCG CTAATCAAAA
                            *       *                    **
Var.  TAGATTTTCT ACAGATAAAA TCCCTAGTGT TCTTCCATCG CTAATCACCA

MP    CTATTTTAGC TAGTTATAAT ATTTCCATCA TCGTCTTATT TCAAAGGTTT
            **                          *
Var.  CTATTTACGC TAGTTATAAT ATTTCCATCA TTGTCTTATT ----------

MP    TTAAGAGAT

Var.  ---------

            3'
```

Figure 6.5.3 Comparison of DNA sequence to the right of the central BamHI site in MPE and VarD. Base differences are marked with a *.

```
      5'
MP    AATTTTTCAT CTTGTGTTAG TTTATGCACG GTTTCCTTGT ACGACTGTAT
                                      *                *
Var.  ---------- ---------- ----TGCACT GTTTCCTTGT ACGATTGTAT

MP    TGTGTTACTA AACGATAAAT AATTCTTAGT CAATGCTTTA ACGTACAACG
        *                                                *
Var.  TGCGTTACTA AACGATAAAT AATTCTTAGT CAATGCTTTA ACGTATAACG

MP    TGGGTGCATC AAATGCGTGT CGTAAAACCG CTTCGTATAC TTTACAGCAT
                                                       *
Var.  TGGGTGCATC AAATGCGTGT CGTAAAACCG CTTCGTACAC TTTACATCAT

MP    CGTACCATAT GTATG--

Var.  CGTACCATAT GTATGAA
                      3'
```

Figure 6.5.4 Comparison of corresponding DNA sequences within the 2.4kb EcoRI/BamHI fragment of MPE and the 2.9kb EcoRI/BamHI fragment of VarD.

```
5'

TAACAAAGTC AATGCACTCT TAAAGATTCT CAAAGTATTT ATTTACCATT TCTTTACACA

GTATCTTAGC CACAGGCATA CGAACAGGCA TACATGATTC AACTATATCA TGAACAATGT

TGGATGCATT CTATAACAAA TGAATCTATG GTCAACTCCA AACAAATTGG GAAACATGTA

AAGCAGCCGA GCATATATCA TCTTACTCGG ATTCTGTGAT GCTTTGCCAA TGACTGAAAT

GTACGAGAAT AGTTTATAAG TTTTATGATT ATTTGGAAAT AAGTTGGATG TTGATAGTCT

ATTCTAGAGT AGTAGTCCTA ATCATTCCCT TAAATTTTAT ATATCCCAGT TTCACTGTCT

CGTATTGAGT TTGTGGTGCT CTTATCGTCT GATTTATTTC TTTTACCATT TTGACTCTAT

TCTGAAACTT TGTCATCGTC TTATCCATTT TCATTGTTGA C
                                              3'
                                          HincII
```

Figure 6.5.5 DNA sequence to the right of the HincII site within the 2.4kb EcoRI/BamHI fragment of MPE.

5′

CAATTTGAAT TAGTATTGGA GAAAGATGAA ATGCCTATTC ACTAACAAAG TCAATGCACT

CTTGAAGATT CTTACCAATA ATATGCAGCG GATTCTCAAA GTATTTATTT ACCATTTCTT

TACACAGTAT CTTAGCCACT GGCCTACGAA CAGGCATACA TGGTTCAACT ATATCATGAA

CAATGTTTAA TTTTAGTTTA TCCAATCCAA TGAATCCATT TTCTTCGTCA TCATCGTATC

TGTACGCTTT GTATTCTCCA GCTCTTGAAC AATCTCTAAA CAGTATCTTC AACACTTTTT

CCACCAATGC CTGGGATGTC ACACATTGGT CTACCATCTG TTTGATGAAG AGGGTATCTT

CTTCAGATTT TCCCTTGATT CTATATATGG TCTTAGGCGA TACCTTGCAT CTATTGTATA

TTTATCAGAT TCAACTATCA ACAATTTTAT CAACTTGTGT TCCAATTATT CTCCATTTAT

TAGATCTGCA TTCATTATTA TAGAATCTTT CTTCCACTAA TATTCTAATA AGGATTAAGT

TTGAAAGGAG AGAAGATCTT GTGTTTGATT TTACTCATTG GAAACATTCT ATAAAGAATG

AATCTATGGT CATCTCCAAA CAAATTGGGG AACATGTAAA GCAGTAAAGC ATATGCCATT

TTACTCGGAT TCTGTGATAC TTTGCCAATG GCTGAAATGT CCGAGAATAG TTTATAAATT

TTCTGAATAT TTGGAATAGA TGTTGTTAGT ATCTTCTAAA GTAGTAGTCC TAATCAATTT

CTTAAATTTT ATATATCCTA GTTTCACTGT CTCGTAATGA GTTTGTCCTG CTCTTATTGT

CTGATTTATT TCTTTTACCA TTTTGGCTCT ATTTTGACAC TTTGTCCTCT TCTTATCCAT

TTTCATTGT

3′

HincII

Figure 6.5.6 DNA sequence to the right of the HincII site within the 2.9kb EcoRI/BamHI fragment of VarD.

Figure 6.5.7 Autoradiograph of a sequencing gel in which the DNA was labelled with ^{32}P-dCTP.

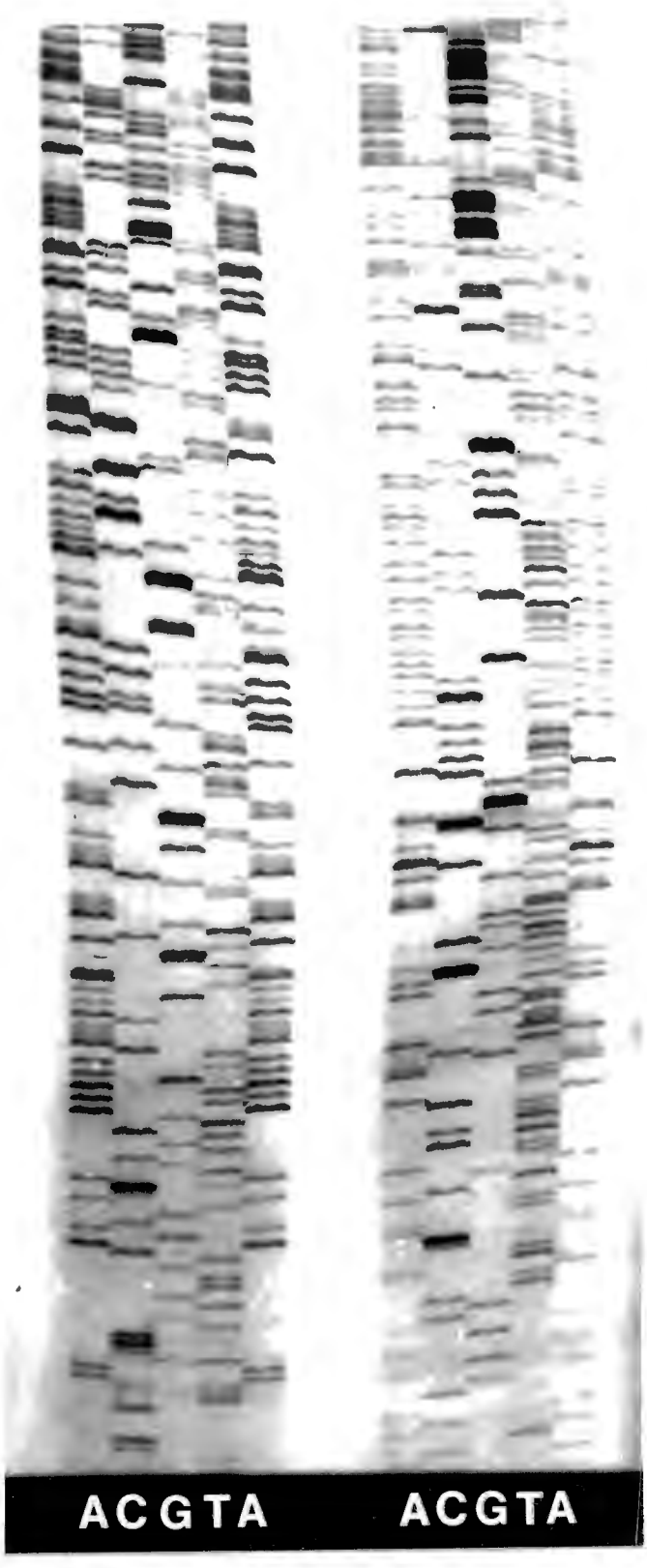


Figure 6.5.8 Autoradiograph of a sequencing gel in which the DNA was labelled with ^{35}S-dCTP.

CHAPTER 7 ANALYSIS OF DNA SEQUENCES

Most of the sequencing was performed in one direction only. The two viral DNAs sequenced proved to be almost identical, this duplication being a "semi"-confirmation of the sequences. The remarkable similarity in DNA sequences may be noted with confidence, but any comparison of derived amino sequences must allow for possible errors in DNA sequence. Within this chapter reference has been made to putative polypeptides. Expression of these polypeptides has not been investigated - they are merely amino acid sequences derived from DNA sequences.

7.1 PROGRAMS USED

The DNA sequences were analysed using two sets of programs. The one program and its hardware requirements are described by Harley (1986). Functions performed include: determination of base composition; conversion of a sequence into its complementary sequence; mapping of restriction enzyme sites; comparison of two DNA strands for homology; identification of open reading frames; translation of DNA sequence into amino acid sequence and analysis of polypeptides for hydrophobicity and ionic charge. The other program used, Genepro; Version 3.1; Riverside Scientific (1986) Serial No. 163, was received from Dr. R. Kirby (department of Microbiology, U.C.T., Cape Town). It had most of the functions already mentioned, but, in addition, it contained a gene bank of DNA sequences established before 1986. Another genebank released in October 1987 was made available (by Dr. R. Kirby) and was used to determine any similarity between some of the sequenced DNA and other known sequences. Derived amino acid sequences of two open reading frames were compared with polypeptide sequences in a 1987 protein data bank.

7.2 DNA COMPOSITION

Pox DNA is known to be AT-rich (Moss 1985). The DNA stretches sequenced were shown to have approximately 67% A+T residues, complying with published reports on poxvirus DNA composition. Table 7.2.1 displays the percentage A+T in each of the DNA sequences established.

Table 7.2.1 Percentage A+T in each DNA stretch sequenced. The DNA stretches are labelled according to the number of bases sequenced.

DNA stretch	% A+T
MP 1907	67.8
Var. 1784	68.2
Var. 909	67.3
MP 462	66.8
MP 309	69.5
Var. 290	67.5
MP 165	63.6
Var. 143	62.9

7.3 COMPARISON OF CORRESPONDING VARIOLA AND MONKEYPOX DNA SEQUENCES

Figures 6.5.3 and 6.5.4 give a good indication of the similarity of the two shorter stretches of DNA sequenced. The respective figures of 17/290 and 5/141 base alterations can be converted to 5.9% and 3.6% mismatching respectively (or 94.1 - 96.4 % similarity).

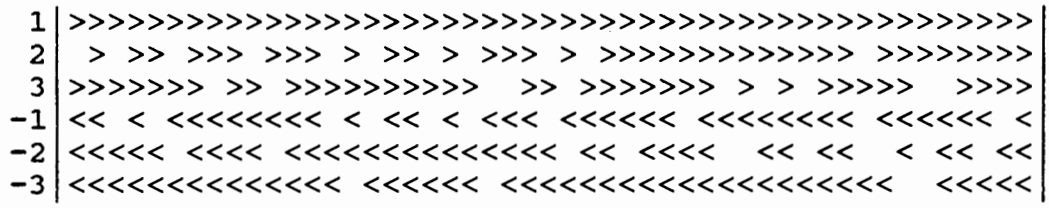
7.3.1 THE SEQUENCES TO THE RIGHT OF THE CENTRAL BamHI SITE

Figure 7.3.1 displays the open reading frames to the right of the central BamHI site a) in VarD and b) in MPE. All six reading frames are represented, frames 1, 2 and 3 of the one DNA strand depicted as 1, 2 and 3 respectively; and the three frames of the complementary strand represented by -1, -2 and -3 respectively. The arrows represent the direction of transcription and any spaces in the sequences correspond to termination codons.

The open reading frames appear to be similar in the two stretches. A single open reading frame in frame 1 of both VarD and MPE appears to code for the same protein. These amino acid sequences are displayed in figure 7.3.2. The variola sequence is translated in (a), (b) showing the corresponding MPE sequence translated. Of the 96 corresponding amino acid residues 8 are different, giving a 92% level of similarity at the protein level of a small section of VarD and MPE. The differences resulted from single base changes in four cases, double base changes in two cases and, for the first two adjacent alterations, four out of the six nucleotides differed. In addition, five nucleotide changes resulted in no alteration of amino acid sequence. Significant differences resulted from two of the single base changes - phenylalanine in the Variola sequence is replaced by

serine in the Monkeypox sequence and isoleucine in "Variola" is replaced by asparagine in "Monkeypox"; the two double base changes code for threonine in place of lysine and tyrosine in place of leucine in "Variola" and "Monkeypox" respectively. The first two amino acid differences may also be important in that leucine and serine in "Variola" are replaced by phenylalanine and lysine in "Monkeypox". These alterations mentioned are amino acid changes which may influence the properties of the putative polypeptide. Single and double base changes may be due to erroneous sequencing data, but the section of four base changes within six nucleotides giving rise to an alteration in two amino acid residues, is probably real.

a) VARBAM,[290]5;5H



b) MPBAM, [309]5;5H

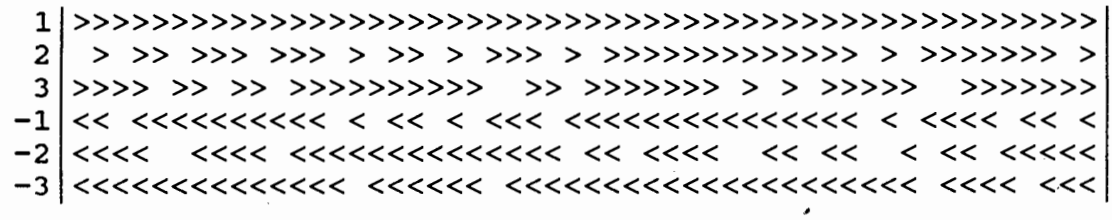


Figure 7.3.1 Open reading frames to the right of the central BamHI site a) in VarD and b) in MPE. The six different reading frames are displayed, negative values representing analysis of the complememtary DNA strand. The arrows point in the direction of transcription and gaps indicate the presence of termination codons.

a)

Ile Glu Asn Gly Ile Leu Lys Lys Leu Ser Ser Ile Lys Ser Lys Ser
Arg Arg Leu Asn Leu Phe Ser Lys Asn Ile Leu Lys Tyr Tyr Leu Asp
Gly Gln Leu Ala Arg Leu Gly Leu Val Leu Asp Asp Tyr Lys Gly Asp
Leu Leu Val Lys Met Ile Asn His Leu Lys Phe Val Glu Asp Leu Ser
Ala Phe Val Arg Phe Ser Thr Asp Lys Ile Pro Ser Val Leu Pro Ser
Leu Ile Thr Thr Ile Tyr Ala Ser Tyr Asn Ile Ser Ile Ile Val Leu

b)

Ile Glu Asn Gly Ile Leu Lys Lys Phe Lys Ser Ile Lys Ser Lys Ser
Arg Arg Leu Asn Leu Phe Ser Lys Asn Ile Leu Lys Tyr Tyr Leu Asp
Gly Gln Leu Ala Arg Leu Gly Leu Val Leu Asp Asp Tyr Lys Gly Asp
Leu Leu Val Lys Met Ile Asn His Leu Lys Ser Val Glu Asp Val Ser
Ala Phe Val Arg Phe Ser Thr Asp Lys Asn Pro Ser Ile Leu Pro Ser
Leu Ile Lys Thr Ile Leu Ala Ser Tyr Asn Ile Ser Ile Ile Val Leu
Phe Gln Arg Phe Leu Arg Asp

Figure 7.3.2 Translation of frame 1 of a and b above. The amino acid residues which differ between the two sequences are underlined.

7.3.2 THE RIGHT END OF VarD AND MPE

Figure 7.3.3 shows a similarity plot of the corresponding sequences at the right end of MPE and VarD. The x-axis represents the Monkeypox sequence of 1907 nucleotides and the y-axis represents 1784 nucleotides of the Variola sequence. A window size of 8 means that 8 bases at a time were compared, a dot being plotted each time the bases were precisely aligned. In this manner a series of dots were generated, any similarity between the two sequences appearing as a diagonal straight line.

The two end fragments proved to be very similar. The two larger gaps in the diagonal line generated in figure 7.3.3 correspond to positions 694 and 1061 in the Variola sequence, which is where the sequencing was not completed (ie. the horizontal displacement of the diagonal line). Other gaps, however, within the diagonal line, represent some mismatching in the DNA sequences. These regions of dissimilarity are very small and were not considered

to be of great significance. The 0.5kb stretch which was expected to be heterologous (from hybridization experiments) lies at the end of the sequence (from the SalI site at position 1329 in the variola sequence to the HindIII site at the end). Except for the small heterologous region around position 1350 in the Variola sequence the two sequences are almost identical.

From the termination codon maps displayed in figure 7.3.4 an open reading frame was found in frame 1 reading from the 3′ end to the 5′ end at the right end of the Monkeypox DNA sequence. Figure 7.3.6 shows the similarity of the translation products of this sequence together with a much smaller open reading frame found in the corresponding region of the Variola sequence. Except for one (leu for ile) the amino acid residues are identical starting from the second methionine residue of the "Monkeypox" sequence. The possibility of errors in sequencing being the cause of apparent early termination in the Variola sequence may not be excluded, as sequencing was only performed in one direction and the substitution of TTGA for TTTGGA caused the termination. These open reading frames are probably the end portion of a larger coding sequence which starts outside of the area sequenced.

The other open reading frame (longer than 240 bases) in the Monkeypox sequence has been translated in figures 7.3.7. A valid comparison of this Monkeypox sequence and any corresponding Variola open reading frames cannot be made, since "gaps" are present in the Variola sequence.

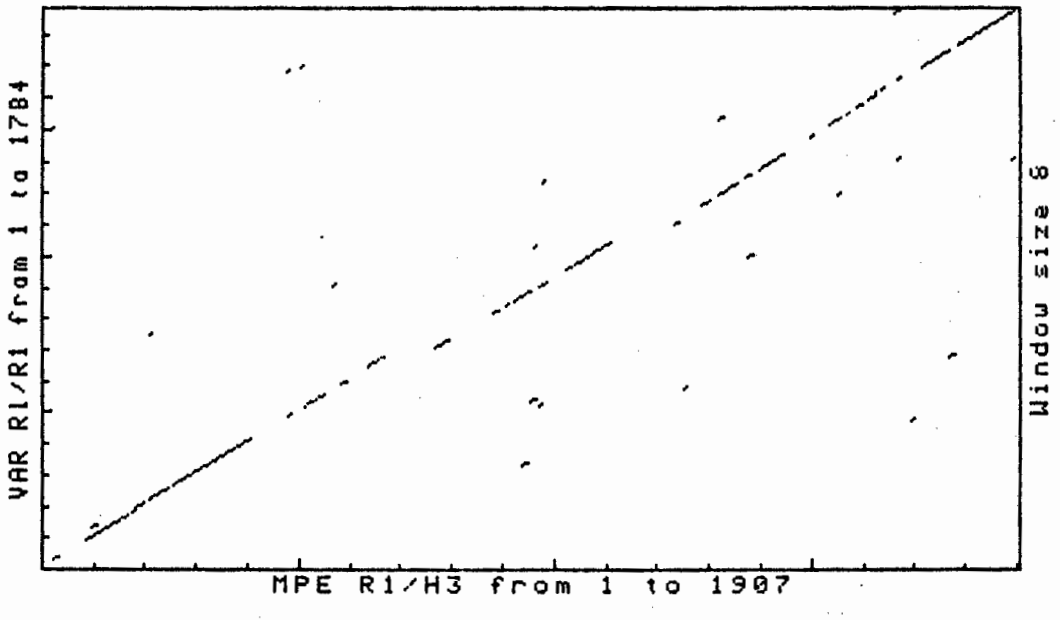


Figure 7.3.3 A plot showing the similarity between the Monkeypox (x-axis) and Variola (y-axis) DNA sequences at the right end of the D and E HindIII fragments respectively.

Search for open reading frames in MPE R1/H3

Open reading frames (longer than 240 bases) found :

ame	From Base	To Base	Length	Pos. of AUG	Length AUG-ter	AUG in frame?
3'-5')	716	266	450	668	405	Yes
3'-5')	1910	1604	306	1901	300	Yes

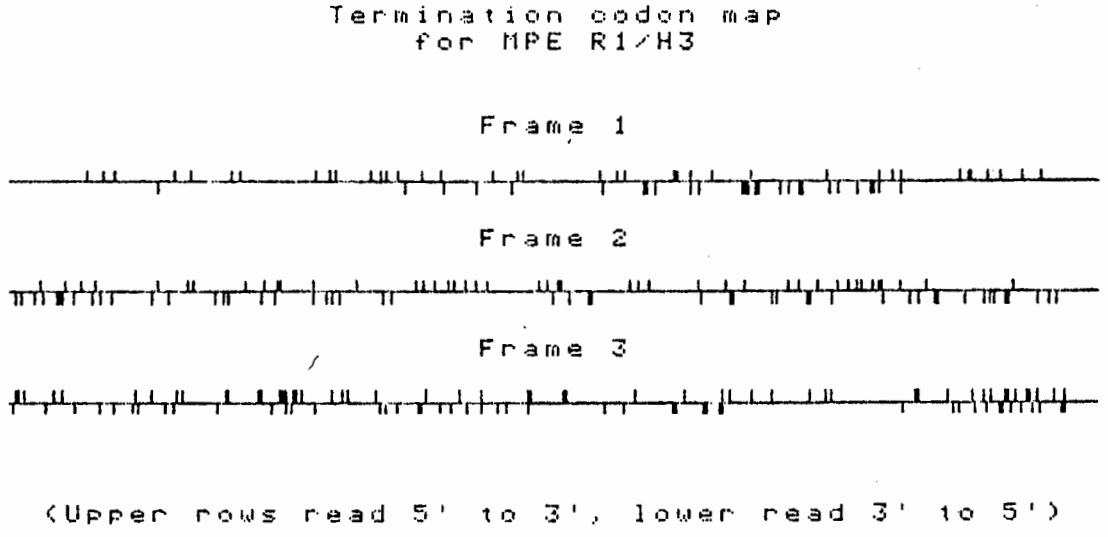


Figure 7.3.4 Termination codon map used to find open reading frames in the Monkeypox 1.9kb sequence. Prof. E.H. Harley's program was used for this analysis.

Search for open reading frames in VAR R1/R1

Open reading frames (longer than 240 bases) found :

ame	From Base	To Base	Length	Pos. of AUG	Length AUG-ter	AUG in frame?
3'-5')	291	2	289	257	258	Yes
3'-5')	1011	564	447	987	426	Yes

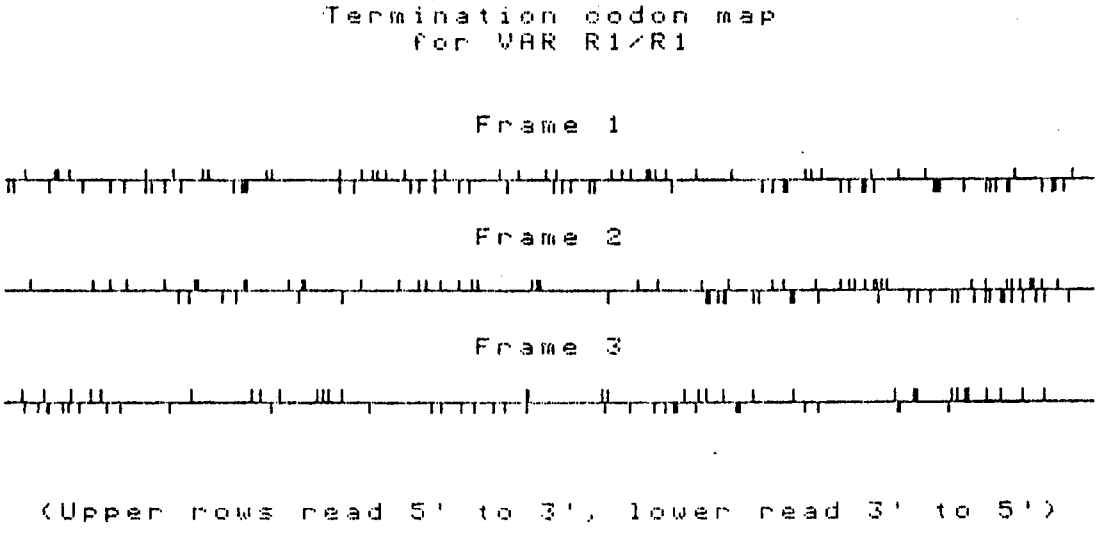


Figure 7.3.5 Termination codon map used to find open reading frames in the corresponding Variola sequence.

Translation of MPE R1/H3 sequence

Reading frame read from 1901 to 1604 in frame 1

```
.Ile.Glu.Tyr.Asn.Leu.Leu.Thr.His.Ser.Asp.Leu.Glu.Trp.Leu.Ile.Asn.Glu.Asn.Val.
 ATA GAA TAT AAT CTT CTT ACA CAC AGT GAC TTG GAA TGG TTA ATT AAT GAG AAT GTA

.Lys.Ala.Thr.Tyr.Leu.Leu.Lys.Ile.Asn.Ala.Tyr.Met.Ile.Asn.Phe.Lys.Ile.Asp.Leu.
: AAG GCT ACA TAC CTT TTA AAA ATC AAT GCC TAT ATG ATT AAC TTT AAA ATA GAT CTA

.Val.Asp.Glu.Ile.Ile.Asp.Leu.Val.Lys.Asp.Ile.Pro.Val.Gly.Ala.Thr.Leu.His.Leu.
 GTT GAT GAA ATC ATT GAC TTA GTT AAA GAT ATT CCT GTA GGA GCT ACA CTA CAT CTA

.Asn.Ile.Leu.Asn.Asn.Ile.Asp.Leu.Asp.Ile.Ile.Leu.Arg.Ile.Ser.Asp.Glu.Tyr.Asn.
 AAT ATA TTA AAC AAT ATA GAT TTG GAC ATT ATT CTT CGT ATA TCT GAT GAA TAT AAT

.Pro.Pro.Val.His.Asp.Ile.Leu.Ser.Lys.Leu.Thr.Asn.Glu.Glu.Met.Cys.Ile.Asn.
 CCA CCT GTT CAC GAT ATT CTG TCT AAA CTT ACC AAT GAA GAG ATG TGT ATA AAC
```

Translation of VAR R1/R1 sequence

Reading frame read from 1687 to 1579 in frame 1

```
t.Ile.Asn.Phe.Lys.Ile.Asp.Leu.Thr.Val.Asp.Glu.Ile.Ile.Asp.Leu.Val.Lys.Asp.Ile.
 ATT AAC TTT AAA ATA GAT CTA ACG GTT GAT GAA ATA ATT GAC TTA GTT AAA GAT ATT

.Val.Gly.Ala.Thr.Leu.His.Leu.Tyr.Asn.Ile.Leu.Asn.Asn.Leu.Asp.
 GTA GGA GCT ACA CTA CAT CTA TAT AAT ATA TTA AAC AAT CTA GAT
```

Figure 7.3.6 Amino acid sequence of the "Monkeypox" open reading frame at the extreme right of the DNA sequenced and a smaller open reading frame in the corresponding position in "Variola". The Monkeypox sequence equivalent to the Variola sequence is underlined and the single alteration in amino acid residues is enclosed in a square.

Translation of MPE R1/H3 sequence

Reading frame read from 668 to 266 in frame 1

```
t.Glu.Tyr.Ile.Lys.Ile.Ser.Asp.Met.Leu.Val.Tyr.Gly.His.Tyr.Ile.Glu.Lys.Thr.Leu.
'G GAA TAC ATC AAA ATA TCA GAC ATG TTG GTA TAC GGA CAC TAC ATC GAG AAG ACG CTT

e.Pro.Ile.Thr.Asp.Ser.Leu.Ser.Phe.Lys.Leu.Phe.Val.Asp.Thr.Met.Ser.Val.Leu.Asn.
'T CCA ATT ACT GAT TCT CTT TCA TTT AAG CTA TTT GTT GAT ACC ATG TCT GTG TTA AAT

sp.Gln.Tyr.Ala.Lys.Val.Val.Ile.Phe.Phe.Asn.Thr.Ile.Ile.Glu.Tyr.Ile.Ile.Ala.Thr.
aT CAA TAT GCC AAG GTT GTC ATC TTC TTC AAT ACC ATC ATA GAA TAT ATT ATA GCT ACT

e.Tyr.Tyr.Arg.Leu.Thr.Val.Leu.Asp.Asn.Tyr.Thr.Asn.Val.Asn.Ile.Leu.Tyr.Pro.Lys.
'C TAT TAT AGA TTG ACA GTC TTG GAC AAT TAT ACT AAT GTC AAC ATT TTG TAT CCA AAG

's.Tyr.Thr.Leu.Ser.Trp.Lys.Met.Phe.Ser.His.Ile.Lys.Val.Asn.Asp.Lys.Ile.Glu.Tyr.
GT TAC ACA CTG TCA TGG AAG ATG TTT TCA CAC ATT AAA GTT AAT GAC AAA ATA GAG TAT

u.Leu.Glu.Glu.Met.Val.Asp.Lys.Gly.Thr.Val.Pro.Ser.Tyr.Leu.Tyr.His.Leu.Ser.Ile.
AA TTG GAG GAG ATG GTG GAC AAA GGT ACT GTA CCT TCT TAT TTA TAT CAT CTG TCC ATC

sn.Val.Ile.Ser.Ile.Ile.Leu.Asp.Asp.Ile.Asn.Gly.Thr.Arg.
AC GTC ATT TCA ATA ATA TTG GAT GAT ATA AAT GGA ACT CGT
```

Figure 7.3.7 The other open reading frame in the "Monkeypox" sequence translated.

7.3.3 ANALYSIS OF THE HETEROLOGOUS SEQUENCES IN VARIOLA

The corresponding DNA stretches sequenced within the EcoRI/BamHI fragments of VarD and MPE are compared in figure 7.3.8. The sequence begins approximately 1.2kb to the left of the central BamHI site (see figures 4.5.1 and 5.2.13) and continues to the HincII site, which is 0.7kb from the EcoRI site. The length of DNA sequenced in VarD was 909 nucleotides. In comparison the corresponding length of MPE was only 462 nucleotides.

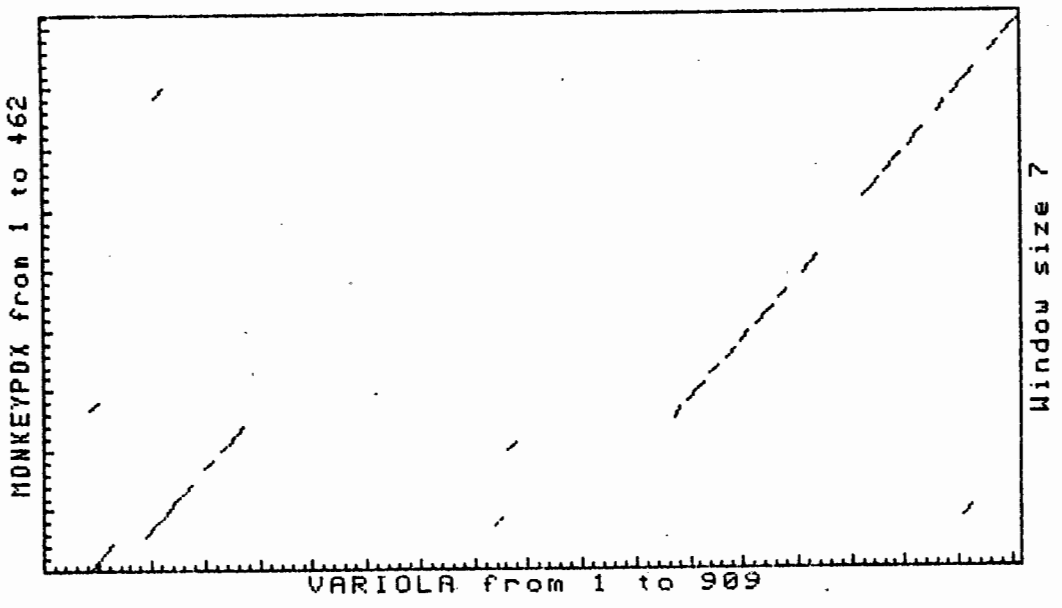


Figure 7.3.8 Comparison of corresponding DNA sequences in Variola and Monkeypox.

Figure 7.3.8 shows the sequences to be very similar up to position 187 in "Variola" and from position 579 to the end. In the "Monkeypox" sequence the first 121 nucleotides correspond to the beginning of the "Variola" sequence and the nucleotides from position 122 onwards correspond to those from position 579 in "Variola". The Variola genome therefore has an insert of 392 nucleotides relative to that of Monkeypox at position 122 in the Monkeypox DNA sequenced.

The DNA sequences (Monkeypox 462 and Variola 909) were searched for open reading frames. The Monkeypox sequence did not have any longer than 240 bases, but the Variola sequence had two relatively long open reading frames, both of which may be of interest. (See figure 7.3.9). At position 183 (in frame 1) an AUG start codon is present in the Variola sequence. From this position 318 nucleotides code for amino acid residues before a termination codon is reached at position 501. The other open reading frame of potential importance is one in the -3 frame from positions 337 to 37. This open reading frame is found on the complementary DNA strand and reads in the opposite direction from the first one mentioned. The DNA sequences of the two open reading frames overlap considerably and are largely composed of the "unique" sequence (positions 187 to 579). This confinement of two open reading frames to one stretch of DNA (possibly unique to Variola) is interesting and may be a relevant difference between the two viral genomes of Variola and Monkeypox.

In comparison, the Monkeypox sequence codes for no equivalent polypeptide to that coded for in frame 1 of the Variola sequence from position 183 to 318. The 303 nucleotide open reading frame in frame -3 of Variola has a counterpart in the Monkeypox sequence of 72 nucleotides from position 91 to 22 in the -3

frame. The amino acid sequences coded for by these sequences are displayed in figure 7.3.10.

Search for open reading frames in VARIOLA

Open reading frames (longer than 240 bases) found :

ame	From Base	To Base	Length	Pos. of AUG	Length AUG-ter	AUG in frame?
5'-3')	177	501	324	183 (	318	Yes
3'-5')	424	37	387	337	303	Yes

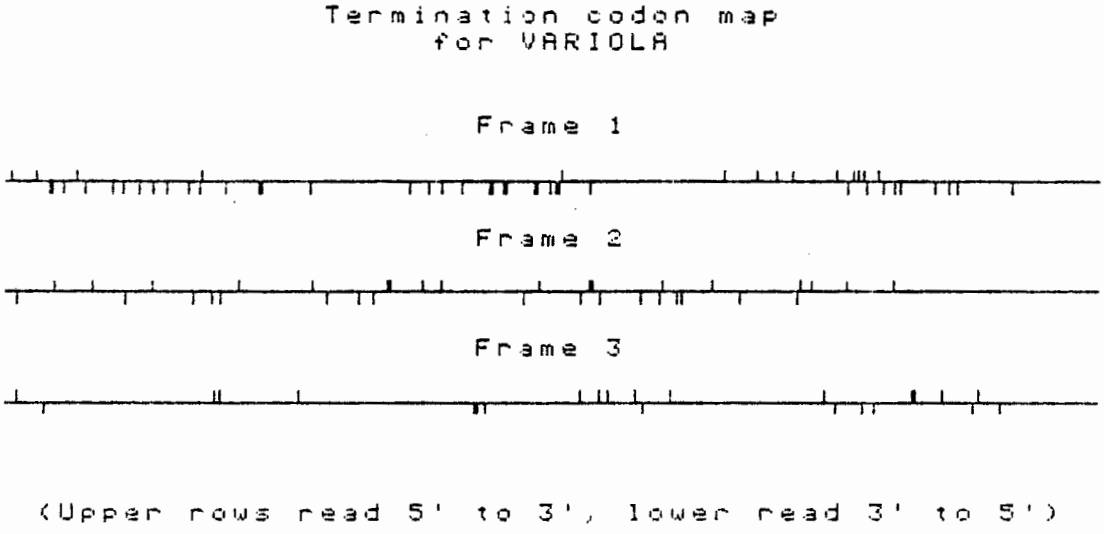


Figure 7.3.9 Termination codon map for the 909 nucleotides of Variola and open reading frames longer than 240 bases within this sequence.

a)

Translation of VARIOLA sequence

Reading frame read from 183 to 501 in frame 1

```
Met.Phe.Asn.Phe.Ser.Leu.Ser.Asn.Pro.Met.Asn.Pro.Phe.Ser.Ser.Ser.Ser.Ser.Tyr.Leu.
ATG TTT AAT TTT AGT TTA TCC AAT CCA ATG AAT CCA TTT TCT TCG TCA TCA TCG TAT CTG

Tyr.Ala.Leu.Tyr.Ser.Pro.Ala.Leu.Glu.Gln.Ser.Leu.Asn.Ser.Ile.Phe.Asn.Thr.Phe.Ser.
AC GCT TTG TAT TCT CCA GCT CTT GAA CAA TCT CTA AAC AGT ATC TTC AAC ACT TTT TCC

Thr.Asn.Ala.Trp.Asp.Val.Thr.His.Trp.Ser.Thr.Ile.Cys.Leu.Met.Lys.Arg.Val.Ser.Ser
ACC AAT GCC TGG GAT GTC ACA CAT TGG TCT ACC ATC TGT TTG ATG AAG AGG GTA TCT TCT

Ser.Asp.Phe.Pro.Leu.Ile.Leu.Tyr.Met.Val.Leu.Gly.Asp.Thr.Leu.His.Leu.Leu.Tyr.Ile
TCA GAT TTT CCC TTG ATT CTA TAT ATG GTC TTA GGC GAT ACC TTG CAT CTA TTG TAT ATT

Tyr.Gln.Ile.Gln.Leu.Ser.Thr.Ile.Leu.Ser.Thr.Cys.Val.Pro.Ile.Ile.Leu.His.Leu.Leu
TAT CAG ATT CAA CTA TCA ACA ATT TTA TCA ACT TGT GTT CCA ATT ATT CTC CAT TTA TTA

Asp.Leu.His.Ser.Leu.Leu.*a*.
GAT CTG CAT TCA TTA TTA TAG
```

b)

Translation of VARIOLA sequence

Reading frame read from 337 to 37 in frame 1

```
Met.Val.Asp.Gln.Cys.Val.Thr.Ser.Gln.Ala.Leu.Val.Glu.Lys.Val.Leu.Lys.Ile.Leu.Phe.
ATG GTA GAC CAA TGT GTG ACA TCC CAG GCA TTG GTG GAA AAA GTG TTG AAG ATA CTG TTT

Arg.Asp.Cys.Ser.Arg.Ala.Gly.Glu.Tyr.Lys.Ala.Tyr.Arg.Tyr.Asp.Asp.Asp.Glu.Glu.Asn.
AGA GAT TGT TCA AGA GCT GGA GAA TAC AAA GCG TAC AGA TAC GAT GAT GAC GAA GAA AAT

Gly.Phe.Ile.Gly.Leu.Asp.Lys.Leu.Lys.Leu.Asn.Ile.Val.His.Asp.Ile.Val.Glu.Pro.Cys.
GGA TTC ATT GGA TTG GAT AAA CTA AAA TTA AAC ATT GTT CAT GAT ATA GTT GAA CCA TGT

Met.Pro.Val.Arg.Arg.Pro.Val.Ala.Lys.Ile.Leu.Cys.Lys.Glu.Met.Val.Asn.Lys.Tyr.Phe.
ATG CCT GTT CGT AGG CCA GTG GCT AAG ATA CTG TGT AAA GAA ATG GTA AAT AAA TAC TTT

Glu.Asn.Pro.Leu.His.Ile.Ile.Gly.Lys.Asn.Leu.Gln.Glu.Cys.Ile.Asp.Phe.Val.Ser.Glu.
GAG AAT CCG CTG CAT ATT ATT GGT AAG AAT CTT CAA GAG TGC ATT GAC TTT GTT AGT GAA
```

c)

Translation of MONKEYPOX sequence

Reading frame read from 91 to 22 in frame 1

```
Met.Pro.Val.Arg.Met.Pro.Val.Ala.Lys.Ile.Leu.Cys.Lys.Glu.Met.Val.Asn.Lys.Tyr.Phe.
ATG CCT GTT CGT ATG CCT GTG GCT AAG ATA CTG TGT AAA GAA ATG GTA AAT AAA TAC TTT

Glu.Asn.Leu.
GAG AAT CTT
```

Figure 7.3.10 Derived amino acid sequences of a) Variola from nucleotide position 183 to position 501; b) Variola from position 337 to 37 and c) Monkeypox from nucleotide position 91 to position 22. The sequence in Variola equivalent to that of the Monkeypox sequence displayed is underlined

7.3.3.1 THE VARIOLA OPEN READING FRAME CODING FOR 106 AMINO ACID RESIDUES (position 183 to 501)

The flanking sequences of the "unique" Variola open reading frames were searched for known regulatory elements. (Figure 7.3.11 shows the DNA sequence). The characteristic TAAATG consensus sequence including the initiation codon of translation for late poxvirus (Moss et al. 1986) genes was not present. Other methionine codons in the sequence were not preceded by TAA. The third nucleotide upstream of the ATG codon is C which is also very uncharacteristic of late genes, which have A or T in this position (Moss et al. 1986). If expressed, this gene is probably transcribed and translated early in the virus cycle (before replication).

Approximately 100 nucleotides upstream of the first AUG codon the consensus sequence CAAT, typical of eucaryotic genes but also present in early poxvirus genes (Esposito and Knight 1984), was found. The TATA promoter sequence is located from positions -13 to -10 with respect to the initiation codon of translation. No eucaryotic cap or polyadenylation signals are present, but there is a hexanucleotice of AATAAG 3' to the termination codon which is similar to the polyadenylation consensus sequence of AATAAA. The repeated unit of CTATTC found at the 3' end of many early poxvirus genes is not present. The closest resemblance to this sequence is ATATTC found at position 520.

VARIOLA sequence from 1 to 909

```
      10          20          30          40          50          60          70
ATTTGAAT TAGTATTGGA GAAAGATGAA ATGCCTATTC ACTAACAAAG TCAATGCACT CTTGAAGATT
      80          90         100         110         120         130         140
TACCAATA ATATGCAGCG GATTCTCAAA GTATTTATTT ACCATTTCTT TACACAGTAT CTTAGCCACT
     150         160         170         180         190         200         210
CCTACGAA CAGGCATACA TGGTTCAACT ATATCATGAA CAATGTTTAA TTTTAGTTTA TCCAATCCAA
     220         230         240         250         260         270         280
AATCCATT TTCTTCGTCA TCATCGTATC TGTACGCTTT GTATTCTCCA GCTCTTGAAC AATCTCTAAA
     290         300         310         320         330         340         350
GTATCTTC AACACTTTTT CCACCAATGC CTGGGATGTC ACACATTGGT CTACCATCTG TTTGATGAAG
     360         370         380         390         400         410         420
GGTATCTT CTTCAGATTT TCCCTTGATT CTATATATGG TCTTAGGCGA TACCTTGCAT CTATTGTATA
     430         440         450         460         470         480         490
TATCAGAT TCAACTATCA ACAATTTTAT CAACTTGTGT TCCAATTATT CTCCATTTAT TAGATCTGCA
     500         510         520         530         540         550         560
CATTATTA TAGAATCTTT CTTCCACTAA TATTCTAATA AGGATTAAGT TTGAAAGGAG AGAAGATCTT
     570         580         590         600         610         620         630
GTTTGATT TTACTCATTG GAAACATTCT ATAAAGAATG AATCTATGGT CATCTCCAAA CAAATTGGGG
     640         650         660         670         680         690         700
CATGTAAA GCAGTAAAGC ATATGCCATT TTACTCGGAT TCTGTGATAC TTTGCCAATG GCTGAAATGT
     710         720         730         740         750         760         770
GAGAATAG TTTATAAATT TTCTGAATAT TTGGAATAGA TGTTGTTAGT ATCTTCTAAA GTAGTAGTCC
     780         790         800         810         820         830         840
ATCAATTT CTTAAATTTT ATATATCCTA GTTTCACTGT CTCGTAATGA GTTTGTCCTG CTCTTATTGT
     850         860         870         880         890         900         910
GATTTATT TCTTTTACCA TTTTGGCTCT ATTTTGACAC TTTGTCCTCT TCTTATCCAT TTTCATTGT.
```

Figure 7.3.11 The "unique" Variola segment and its flanking sequences. An open reading frame extends from position 183 to 501. Start and stop codons and consensus sequences for initiation of transcription are underlined.

7.3.3.2 THE VARIOLA OPEN READING FRAME CODING FOR 101 AMINO ACID RESIDUES (position 337 to 37)

Figure 7.3.12 shows the complementary sequence to that of figure 7.3.11. The sequence reads in the opposite direction, so the open reading frame coding for 101 amino acids (position 337 to 37 reading from the 3′ to the 5′ end in figure 7.3.11) stretches from position 573 to position 873 in the complementary sequence and reads from the 5′ end to the 3′ end.

The absence of TAAATG at the origin of translation and the presence of C three nucleotides upstream of the origin reduces the chances of this open reading frame coding for a late polypeptide. A TATA consensus sequence is present 50 nucleotides upstream of the initiation codon of translation and CAAT is found approximately 80 nucleotides upstream of the origin of translation. No consensus sequences for termination were identified.

```
mplementary sequence to VARIOLA read from position 910 to 1

        10         20         30         40         50         60         70
CAATGAAAA TGGATAAGAA GAGGACAAAG TGTCAAAATA GAGCCAAAAT GGTAAAAGAA ATAAATCAGA
        80         90        100        110        120        130        140
ATAAGAGC AGGACAAACT CATTACGAGA CAGTGAAACT AGGATATATA AAATTTAAGA AATTGATTAG
       150        160        170        180        190        200        210
CTACTACT TTAGAAGATA CTAACAACAT CTATTCCAAA TATTCAGAAA ATTTATAAAC TATTCTCGGA
       220        230        240        250        260        270        280
TTTCAGCC ATTGGCAAAG TATCACAGAA TCCGAGTAAA ATGGCATATG CTTTACTGCT TTACATGTTC
       290        300        310        320        330        340        350
CAATTTGT TTGGAGATGA CCATAGATTC ATTCTTTATA GAATGTTTCC AATGAGTAAA ATCAAACACA
       360        370        380        390        400        410        420
ATCTTCTC TCCTTTCAAA CTTAATCCTT ATTAGAATAT TAGTGGAAGA AAGATTCTAT AATAATGAAT
       430        440        450        460        470        480        490
AGATCTAA TAAATGGAGA ATAATTGGAA CACAAGTTGA TAAAATTGTT GATAGTTGAA TCTGATAAAT
       500        510        520        530        540        550        560
ACAATAGA TGCAAGGTAT CGCCTAAGAC CATATATAGA ATCAAGGGAA AATCTGAAGA AGATACCCTC
       570        580        590        600        610        620        630
CATCAAAC AGATGGTAGA CCAATGTGTG ACATCCCAGG CATTGGTGGA AAAAGTGTTG AAGATACTGT
       640        650        660        670        680        690        700
AGAGATTG TTCAAGAGCT GGAGAATACA AAGCGTACAG ATACGATGAT GACGAAGAAA ATGGAATTCA
       710        720        730        740        750        760        770
GGATTGGA TAAACTAAAA TTAAACATTG TTCATGATAT AGTTGAACCA TGTATGCCTG TTCGTAGGCC
       780        790        800        810        820        830        840
TGGCTAAG ATACTGTGTA AAGAAATGGT AAATAAATAC TTTGAGAATC CGCTGCATAT TATTGGTAAG
       850        860        870        880        890        900        910
ATCTTCAAG AGTGCATTGA CTTTGTTAGT GAATAGGCAT TTCATCTTTC TCCAATACTA ATTCAAATTG
```

Figure 7.3.12 The complementary sequence of figure 7.3.11. Positions 573 and 873 correspond to the start and stop codons respectively of an open reading frame. Consensus sequences and initiation and termination codons are underlined.

7.4 COMPARISON OF THE "UNIQUE" OPEN READING FRAMES WITH DNA AND PROTEIN DATA BANKS

The two overlapping open reading frames within the "unique" Variola segment were compared with DNA sequences established before October 1987. Any sequence containing 33 or more identical nucleotides within a 60bp stretch was identified. (Figure 7.4.1 displays sequences containing more than 35 nucleotides identical to the open reading frame of 303 nucleotides). For both open reading frames many different sequences were picked up, but no consecutive sequences of 60bp were recognised as being very similar. Sequences similar to the "unique" Variola sequence included many exons, large genes such as the albumin and globin genes, various enzymes, immunoglobulin genes (in particular the d-j-c region), transposable elements and, most similar of all, the mitochondrial DNA dimer replication initiation region. Because the areas of similarity were relatively short (repeated homologies were not found in any one particular sequence) and such a large number of similar sequences were detected, the poxvirus DNA could not be identified as any known sequence.

The derived amino acid sequences of the two open reading frames were compared with known polypeptide sequences. The open reading frame from position 183 to position 501 was translated and any sequence containing 8 identical amino acid residues within a stretch of 10 amino acids was identified. A single polypeptide sequence was picked up - cytochrome P-450 (see figure 7.4.2). The similar region was one highly rich in serine residues, the Variola sequence being Asn, Pro, Phe, Ser, Ser, Ser, Ser, Ser, Tyr, Leu. The two differing residues were Phe and the fifth Ser. If the Variola polypeptide is expressed this serine-rich region is probably a necessary functional domain. Many polypeptide

sequences containing 6 out of 10 identical amino acid residues were identified and the similar regions consisted predominantly of this serene-rich sequence.

The translated open reading frame from position 337 to 37 was also compared with known polypeptide sequences. Stretches of 7 out of 10 identical amino acid residues were identified. The last 10 amino acids were similar to a region within the troponin C molecule of bovine, human and rabbit. A stretch of amino acids from the eighth residue through to the 20th residue resembled enzyme sequences. Enzymes with similar sequences were aspartate aminotransferase (mitochondria), alanyl-tRNA synthetase (E.Coli) and carbamoyl-phosphate synthetase: arginine specific (yeast). A common feature of these enzymes is their ability to bind ATP.

Although polypeptide sequences similar to the derived amino acid sequences were found, no definite function could be assigned to the putative polypeptides. Because the sequence of interest is in Variola, an eradicated virus, no in vivo tests can be done to investigate the expression of these open reading frames.

```
TGACATCCCAGGCATTGGTGGAAAAAGTGTTGAAGATACTGTTTAGAGATTGTTCAAGAG
*** * * *** **      ** *  ********** ** **** * * ** * *      36
TGAAACCTCAGCCAAGACCAGACAGGGTGTTGAAGATGCTTTTTACACACTGGTAAGAGA
```

Human N-ras mRNA and flanking regions [2436]

```
TGGATTCATTGGATTGGATAAACTAAAATTAAACATTGTTCATGATATAGTTGAACCATG
*   ** ***    *  *** * ** ******* ***  * ** ** ** **  ** **   36
TTATTTTATTTATATAAATATATTATAATTAAATATTTATAATAATTTATTTTTACAATT
```

Paramecium species 7,227 mt dna dimer: replication init. region [930]

```
CAAAGCGTACAGATACGATGATGACGAAGAAAATGGATTCATTGGATTGGATAAACTAAA
  *  ** *** *  ******    *******  ** * * ****  *** *** ****  38
GGATCCGGACAAAATCGATGAAACGGAAGAAATCCGAGTGAATGGAAAGGACAAAATAAA
```

Petunia hybrida chloroplast DNA with autonomous replication activity (ARSB) [1335]

```
CGAAGAAAATGGATTCATTGGATTGGATAAACTAAAATTAAACATTGTTCATGATATAGT
***** *  *  ***  * * ***  ***** ** *  *  * **  ** ** **** *  36
CGAAGGAGTTTTATTTTTAGTATTTTATAAAATATATATTTATATGATTAATAATATTAT
```

Yeast (S.cerevisiae) mitochondrial cytochrome oxi1 gene and flanks [3804]

```
AAATGGATTCATTGGATTGGATAAACTAAAATTAAACATTGTTCATGATATAGTTGAACC
 * ** *** *  * * ** *  ************    * * * ** *** *  ** *  37
TACTGAATTTAAAGAAATGAAAGAACTAAAATTAAAAGAAGCTAAAGAGATACTAAAAGC
```

P.stuartii PStI restriction and modification genes, complete [3889]

Figure 7.4.1 Sequences containing 36 or more identical nucleotides within any stretch of 60 nucleotides in the open reading frame of 303 nucleotides (position 337 to 37).

```
11   NPFSSSSSYL    8
     ** **** **
80   NPASSSSLYL   108
```

Cytochrome P450 [355]

Figure 7.4.2 An amino acid sequence within cytochrome P450 containing 8 out of 10 identical residues to the derived amino acid sequence of the 318 nucleotide open reading frame.

CHAPTER 8 CONCLUSION

A comparison has been made between a section of the genomes of Variola and Monkeypox. The DNA sequence chosen is one which appears to be conserved within different species - judging by previously constructed restriction enzyme maps (Mackett and Archard 1979). Since the HindIII D fragment of Variola is slightly larger than the corresponding HindIII E fragment of Monkeypox these two DNA sequences were compared in an attempt to locate a gene unique to Variola. By hybridization the HindIII sites separating the fragments of interest from the adjacent C fragments were shown to be co-incident.

MPE was mapped for restriction enzyme sites. No PstI or SalI sites were present, but fifteen other sites were located (3 HpaI, 9 EcoRI, 2 BamHI and 1 AvaI). This was performed by sizing restriction enzyme fragments; mapping by the Southern Cross technique and by subcloning fragments for use as probes in hybridization experiments. The constructed MPE map was compared with the VarD map established by Hamilton et al. (1985). Although not identical, the maps are similar. Differences can be seen at the right end and an insert appears to be present in the central region of VarD.

Hybridization experiments showed VarD and MPE to be very similar. This similarity proved to be quite remarkable upon sequencing approximately 3kb of each fragment. (~95% similarity). The differences were mainly single base changes.

For location of the "unique" sequence in Variola, the corresponding 2.9kb and 2.4kb EcoRI/BamHI fragments of VarD and MPE respectively were mapped for four enzymes. These results were not sufficient for locating the insert in VarD, so the two

sequences were heteroduplexed and mapped by SI analysis. The position of the insert was established and the DNA sequenced. The corresponding Monkeypox sequence was identical to the sequences flanking the 400bp insert.

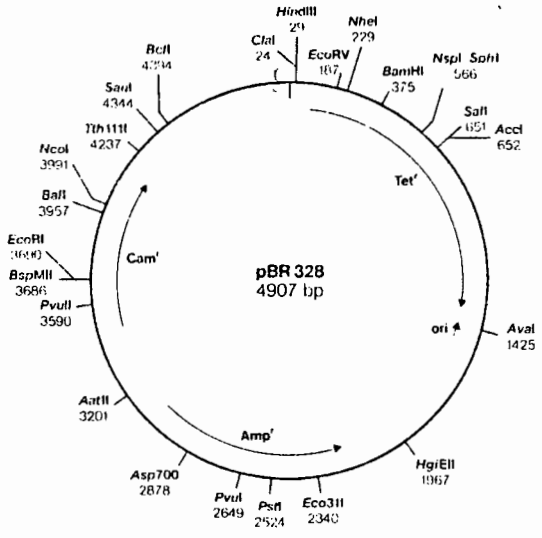
Two overlapping open reading frames of 303 and 318 nucleotides were identified within this "unique" Variola sequence reading in opposite directions. To determine whether or not this sequence is present in other Orthopoxviruses a 100bp section of the "unique" area was subcloned into pUC19 and used to probe DNA from other strains of Orthopoxviruses. The cloned HindIII D fragment of the Somalia strain of Variola (variola minor) hybridized to the probe. Five strains of Monkeypox were shown not to have this sequence, but DNA from Vaccinia, Cowpox and Camelpox hybridized to the probe. This sequence is therefore not unique to Variola. It is absent in Monkeypox; but the presence of it in Vaccinia, Cowpox and Camelpox means that expression of the open reading frames may be investigated in vivo. If mRNA is detected in Vaccinia-infected cells future work may involve characterization of any translation products. Antibodies may be made to a synthetic oligonucleotide to test the level of polypeptide synthesis.

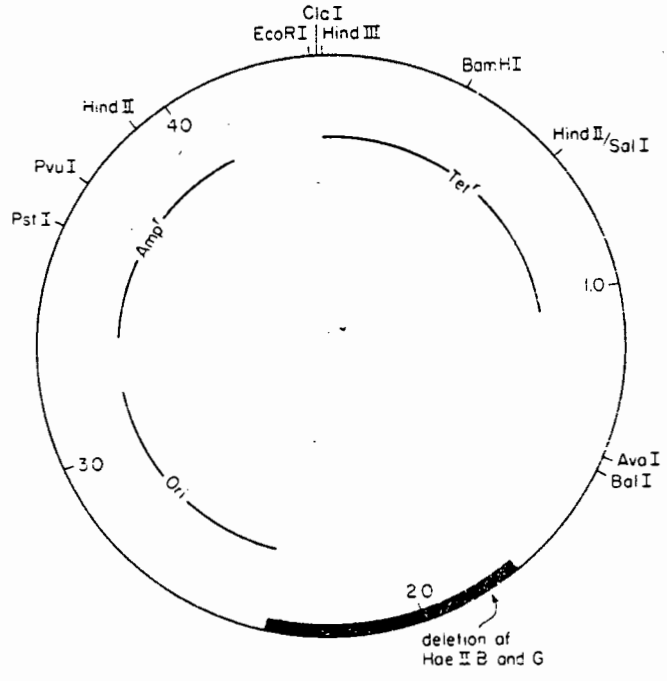
The presence of a sequence in Variola and other Orthopoxviruses excluding Monkeypox is not completely unexpected since a specific antigen (va) has been identified in strains of Variola, Vaccinia, Rabbitpox, Buffalopox and camelpox, but not in Monkeypox (Gispen and Brand-Saathof 1974). Since the DNA sequence of interest is conserved within different species it is probably important for virus propagation. How Monkeypox survives without this sequence may not be established before determining the

function of it. If this sequence is required for maintenance in certain hosts only, man being one, then Monkeypox may be regarded as less of a threat to mankind - since the sequence is absent.

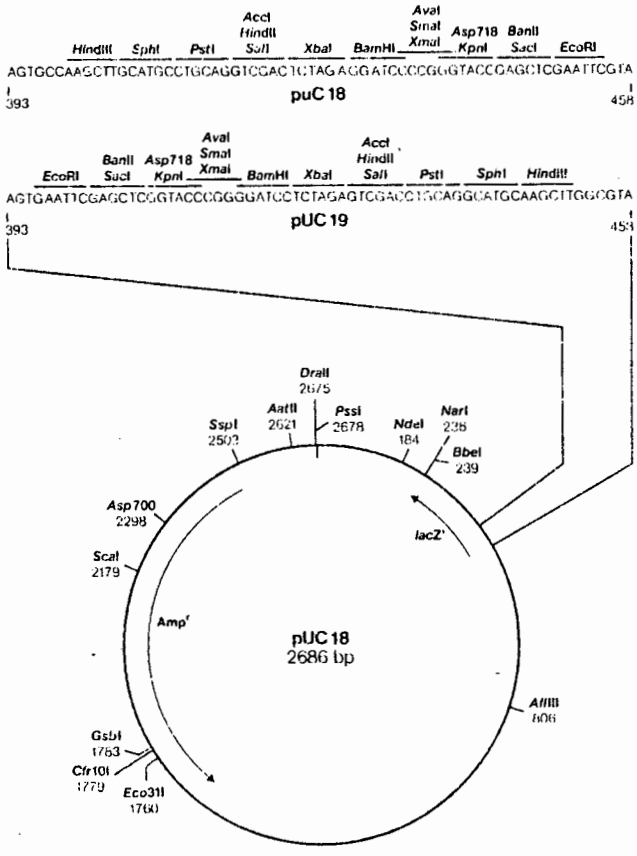
APPENDIX 1

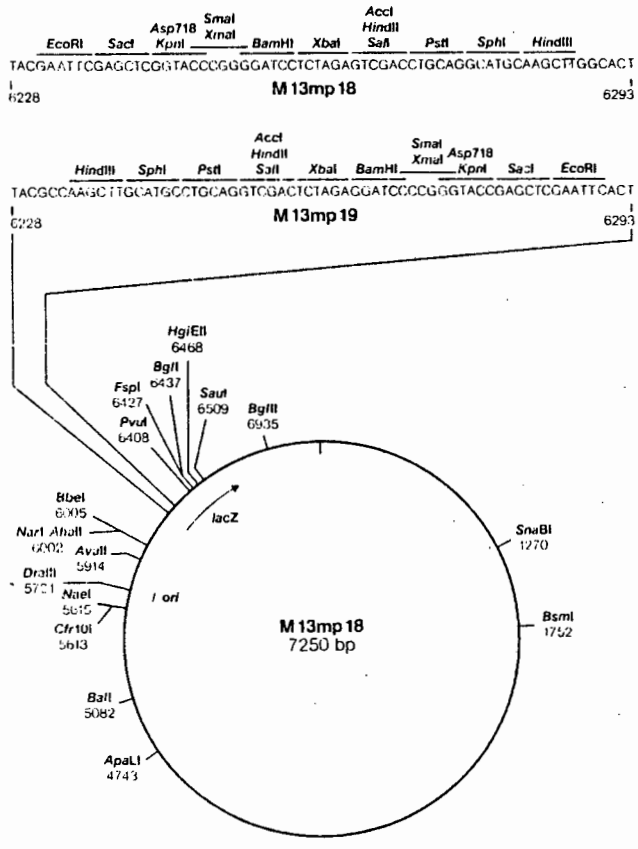
Cleavage maps of vectors used. pBR329 is almost identical to pBR328; pBR329 has a small deletion to the left of the HindIII site.





pAT153





DNA-synthesis ----->

17-mer "reverse sequencing primer"

5' CAGGAAACAGCTATGAC 3'

ATGTTGTGTGGAATTGTGAGCGGATAACAATTTCACACAGGAAACAGCTATGACCATGATTAC--

multiple cloning site

----------GGCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAACCCT

3' TGACCGGCAGCAAAATG 5'

17-mer "sequencing primer"

<----- DNA-synthesis

Sequence of M13/pUC19-primer binding sites

APPENDIX 2

Restriction fragment sizes (in kb) of markers used.

Lambda DNA digested with HindIII

23.130

9.416

6.682

4.361

2.322

2.027

.564

125

VarD digested with EcoRI

5.8

4.05

2.85

1.9

1.3

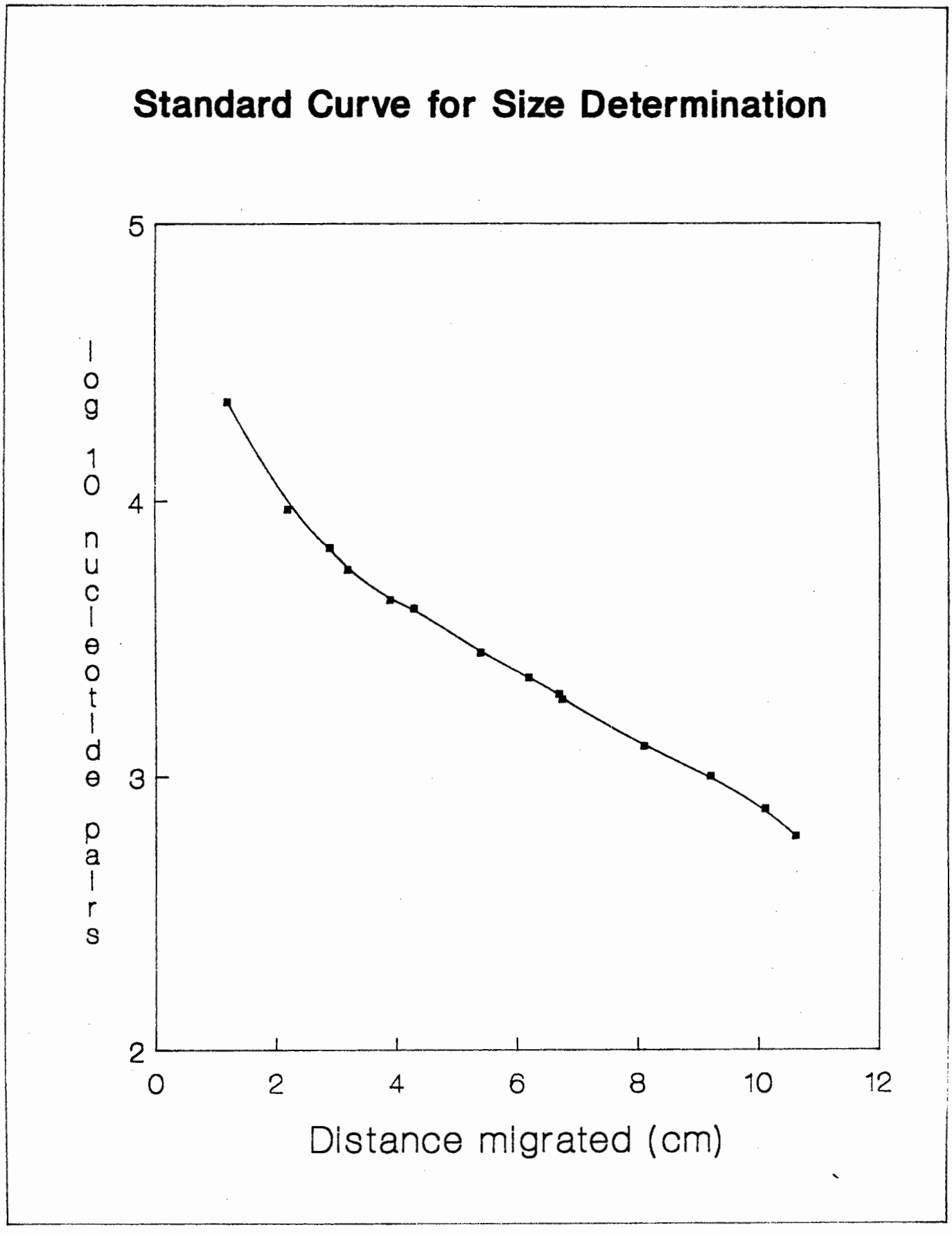
1.0 (X2)

.75

.2

APPENDIX 3

Standard curve of $\log_{10}$ nucleotide pairs versus distance migrated - based on figure 4.1.1.



REFERENCES

AHMED, A. (1985) A rapid procedure for DNA sequencing using transposon-promoted deletions in Escherichia coli. Gene. **39**: 305-310.

ARCHARD, L.C. and MACKETT, M. (1979) Restriction endonuclease analysis of red cowpox virus and its white pock variant. J. Gen. Virol. **45**(1): 51-63.

ARITA, I. and HENDERSEN, D.A. (1968) Smallpox and Monkeypox in non-human primates. Bull. WHO. **39**: 277-283.

ARITA, M. and TAGAYA, I. (1980) Virion polypeptides of poxviruses. Arch. Virol. **63**: 209-225.

ARITA, I., JEZEK, Z., KHODADEVICH, L. and RUTI, K. (1985) Human Monkeypox: A newly emerged Orthopoxvirus zoonoses in the tropical rain forests of Africa. Am. J. Med. Hyg., **34**(4): 781-789.

BAROUDY, B.M. and MOSS. B. (1982) Sequence homologies of diverse length tandem repetitions near ends of vaccinia virus genome suggest unequal crossing over. Nucleic Acids Res. **10**: 5673-5679.

BAROUDY, B.M., VANKATESAN, S. and MOSS, B. (1982) Incompletely base-paired flip-flop terminal loops link the two DNA strands of the vaccinia virus genome into one uninterrupted polynucleotide chain. Cell. **28**: 315-324.

BAXBY, D. (1975) Identification and interrelationships of the variola/vaccinia subgroup of poxviruses. Prog. Med. Virol. **19**: 215-246.

BAXBY, D. (1977) Poxviruses hosts and reservoirs. Arch. Virol. **55**: 169-179.

BERTHOLET, C., DRILLIEN, R. and WITTEK, R. (1985) One hundred base pairs of 5' flanking sequence of a vaccinia virus late gene are sufficient to temporally regulate late transcription. Proc. Natl. Acad. Sci. USA **82**(7): 2096-2100.

BERTHOLET, C., VAN MEIR, E., TEN HEGGELER-BORDIER, B. and WITTEK, R. (1987) Vaccinia Virus produces late mRNAs by discontinuous synthesis. Cell. **50**: 153-162.

BIGGIN, M.D., GIBSON, T.J. and HONG, G.F. (1983) Buffer gradient gels and ^{35}S label as an aid to rapid DNA sequence determination. Proc. Natl, Acad, Sci. USA. **80**: 3963-3965.

BINNS, M.M., STENZLER, L., TOMLEY, F.M., CAMPBELL, J. and BOURSNELL, M.E. (1987) Identification by a random sequencing strategy of the fowlpoxvirus DNA polmerase gene, its nucleotide sequence and comparison with other viral DNA polymerases. Nucleic Acids Res. **15**(16): 6563-6573.

BIRNBOIM, H.C. and DOLY, J. (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. **7**: 1513-1523.

BLACK, D.N., HAMMOND, J.M. and KITCHING, R.P. (1986) Genomic relationship between capripoxviruses. Virus Res. **5**(2-3): 277-292.

BOYLE, D.B., COUPAR, B.E., GIBBS, A.J., SEIGMAN, L.J. and BOTH, G.W. (1987) Fowlpox virus thymidine kinase: nucleotide sequence and relationships to other thymidine kinases. Virology. **156**(2): 355-365.

CABIRAC, G.F., STRAYER, D.S., SELL, S. and LEIBOWITZ, J.L. (1985) Characterization, molecular cloning, and physical mapping of the Shope fibroma virus genome. Virology. **243**(2): 663-670.

DALES, S. and POGO, B.G. (1981) Biology of poxviruses. Virol. Monogr. **18**: 1-109.

DEFILIPPES, F.M. (1982) Restriction enzyme mapping of vaccinia virus DNA. J. Virol. **43**(1): 136-149.

DOWNIE, A.W. and DUMBELL, K.R. (1947) The isolation and cultivation of variola virus on the chorio-allantois of chick embryos. J. Pathol. Bacteriol. **59**: 189-198.

DRETZEN, G., BELLARD, M. and SASSONE-CORSI, P. (1981) A reliable method for the recovery of DNA fragments from agarose and acrylamide gels. Anal. Biochem. **112**: 295-298.

DRILLIEN, R., SPEHNER, D., VILLEVAL, D. and LECOCQ, J.P. (1987) Similar genetic organization between a region of fowlpox virus DNA and the vaccinia virus HindIII J fragment despite divergent location of the thymidine kinase gene. Virology. **160**(1): 203-209.

DUMBELL, K.R. and ARCHARD, L.C. (1980) Comparison of white pock (h) mutants of monkeypox virus with parental monkeypox and with variola-like viruses isolated from animals. Nature. **286**: 29-32.

EARL, P.L., JONES, E.V. and MOSS, B. (1986) Homology between DNA polymerases of poxviruses, herpesviruses, and adenoviruses: nucleotide sequence of the vaccinia virus DNA polymerase gene. Proc. Natl. Acad. Sci. USA. **83**(11): 3659-3663.

ESPOSITO, J.J., OBIJESKI, J.F. and NAKANO, J.H. (1977) Serological relatedness of monkeypox, variola, and vaccinia viruses. J. Med. Virol. **1**: 35-47.

ESPOSITO, J.J., OBIJESKI, J.F. and NAKANO, J.H. (1977) The virion and soluble antigen proteins of variola, monkeypox, and vaccinia viruses. J. Med. Virol. **1**: 95-110.

ESPOSITO, J.J., OBIJESKI, J.F. and NAKANO, J.H. (1978) Orthopoxvirus DNA: strain differentiation by electrophoresis of restriction endonuclease fragmented virion DNA. Virology. **89**(1): 53-66.

ESPOSITO, J.J., CABRADILLA, C.D., NAKANO, J.H. and OBIJESKI, J.F. (1981) Intragenomic sequence transposition in monkeypox virus. Virology. **109**: 231-243.

ESPOSITO, J.J and KNIGHT, J.C. (1984) Nucleotide sequence of the thymidine kinase gene region of Monkeypox and Variola viruses. Virology. **135**: 561-567.

ESPOSITO, J.J and KNIGHT, J.C. (1985) Orthopoxvirus DNA: a comparison of restriction profiles and maps. Virology. **143**(1): 230-251.

ESPOSITO, J.J., NAKANO, J.H. and OBIJESKI, J.F. (1985) Can variola-like viruses be derived from monkeypox virus? An investigation based on DNA mapping. Bull. WHO. **63**(4): 695-703.

ESSANI, K. and DALES, S. (1979) Biogenesis of vaccinia: Evidence for more than 100 polypeptides. Virology. **95**: 385-394.

EVANS, E. and TRAKTMAN, P. (1987) Molecular genetic analysis of a vaccinia virus gene with an essential role in DNA replication. J. Virol. **61**(10): 3152-3162.

GANGEMI, J.D. and SHARP, G.D. (1976) Use of a restriction endonuclease in analyzing the genomes from two different strains of vaccinia virus. J. Virol. **20**(1): 319-323.

GESHELIN, P. and BERNS, K.I. (1974) Characterization and localization of the naturally occurring crosslinks in vaccinia virus DNA. J. Mol. Biol. **88**: 785-796.

GISPEN, R. and BRAND-SAATHOF, B. (1974) Three specific antigens produced in Vaccinia, Variola and Monkeypox infections. The Journal of Infectious Diseases **129**(3): 289-295.

GISPEN, R. (1975) Relevance of some poxvirus infections in monkeys to smallpox eradication. Transactions of the Royal Society of Tropical Medicine and Hygiene. **69**(2):

GREENAWAY,P.J. and DALE, J. (1983) Genetic Manipulation. A practical introduction to basic techniques.

HAMILTON, A., KINCHINGTON, D., GREENAWAY, P.J. and DUMBELL, K. (1985) Recombinant bacterial plasmids containing inserts of variola DNA. Lancet. **2**(8468): 1356-1357.

HARPER, L., BEDSON, H.S. and BUCHAN, A. (1979) Identification of orthopoxviruses by polyacrylamide gel electrophoresis of intracellular polypeptides. I. Four major groupings. Virology. **93**(2): 435-444.

HARLEY, E.H. (1986) A general DNA analysis program for the Hewlett-Packard Model 86/87 microcomputer. Nucleic Acids Res. **14**(1): 467-477.

HENIKOFF, S. (1984) Unidirectional digestion with Exonuclease III creates targeted breakpoints for DNA sequencing. Gene. **28**: 351-359.

HIRT, P., HILLER, G. and WITTEK, R. (1986) Localization and fine structure of a vaccinia virus gene encoding an envelope antigen. J. Virol. **58**(3): 757-764.

HOLMES, D.S. and QUIGLEY, M. (1981) A rapid boiling method for the preparation of bacterial plasmids. Anal. Biochem. **114**: 193-197.

HRUBY, D.E., LYNN, D.L. and KATES, J.R. (1979) Vaccinia virus replication requires active participation of the host cell transcriptional apparatus. Proc. Natl. Acad. Sci. USA. **76**: 1887-1890.

HUTCHINSON, H.D., ZIEGLER, D.W., WELLS, D.E. and NAKANO, J.H. (1977) Differentiation of variola, monkeypox, and vaccinia antisera by radioimmunoassay. Bull. WHO. **55**(5): 613-623.

ICHIHASHI, Y. and OIE, M. (1980) Adsorption and penetration of the trypsinized vaccinia virion. Virology. **101**: 50-60.

JANECZKO, R.A., RODRIGUEZ, J.F. and ESTEBAN, M. (1987) Studies on the mechanism of entry of Vaccinia virus in animal cells. Arch. Virol. **92**: 135-150.

JEZEK, Z., SZCZENIOWSKI, M. and PALUKU, K.M. (1987) Human monkeyox: clinical features of 282 patients. J. Infect. Dis. **156**(2): 293-298.

JOHNSON, D.A., GAUTSCH, J.W., SPORTSMAN, J.R. and ELDER, J.H. (1984) Improved technique utilizing nonfat dry milk for analysis of proteins and nucleic acids transferred to nitrocellose. Gene. Anal. Techn. **1**: 3-8.

JONES, E.V., PUCKETT, C. and MOSS,B. (1987) DNA-dependent RNA polymerase subunits encoded within the vaccinia virus genome. J. Virol. **61**(6): 1765-1771.

KAO, S-Y., RESSNER, E., KATES, J. and BAUER, W.R. (1981) Purification and characterization of a superhelix binding protein from vaccinia virus. Virology. **111**: 500-508.

KADO, C.I. and LIU, S.T. (1981) Rapid procedure for detection and isolation of large and small plasmids. J. Bacteriol. **145**(3): 1365-1373.

KATES, J.R. and McAUSLAN, B.R. (1967) Poxvirus DNA-dependent RNA polymerase. Proc. Nat. Acad. Sci. USA. **58**: 134-141.

KHODAKEVICH, L., SZCZENIOWSKI, M., MANBU-MA-DISU, JEZEK, Z., MARENNIKOVA, S., NAKANO, J. and MESSINGER, D. (1987) The role of squirrels in sustaining monkeypox virus transmission. Tropical and Geographical Medicine. **39**: 115-122.

KINCHINGTON, D., DOLLERY, A., GREENAWAY, P. and DUMBELL, K. (1984) The detection of subtle differences between different orthopoxvirus viruses by heteroduplex analysis. Virus Res. **1**(5): 351-363.

LADNYJ, I.D., ZIEGLER, P. and KIMA, E. (1972) A human infection caused by monkeypox virus in Basankusu Territory, Democratic Republic of the Congo. Bull. WHO. **46**: 593-597.

LAKE, J.R. and COOPER, P.D. (1980) Deletions of the terminal sequences in the genomes of the white pock (u) and host-restricted (p) mutants of rabbitpox virus. J. Gen. Virol. **48**: 135-147.

MACKETT, M. and ARCHARD, L.C. (1979) Conservation and variation in Orthopoxvirus genome structure. J. Gen. Virol. **45**(3): 683-701.

MAXAM, A.M. and GILBERT, W. (1977) A new method for sequencing DNA. Proc. Natl. Acad. Sci. **74**: 560.

MAXAM, A.M. and GILBERT, W. (1980) Sequencing end-labelled DNA with base-specific chemical cleavages. Methods Enzymol. **65**: 499.

MANIATIS, T., FRITSCH, E.F. and SAMBROOK, J. (1982) Molecular cloning - A laboratory manual. Cold Spring Harbor Laboratory.

MARENNIKOVA, S.S., SELUHINA, E.M., MAL'CEVA, N.N. and LADNYJ, I.D. (1972) Poxviruses isolated from clinically ill and asymtomatically infected monkeys and a chimpanzee. Bull. WHO. **46**: 613-620.

McCARRON, R.J., CABRERA, C.V, ESTERBAN, M., McALLISTER, W.T. and HOLOWCZAK, J.A. (1978) Structure of vaccinia DNA: Analysis of the viral genome by restriction endonucleases. Virology. **86**(1): 88-101.

MERCER, A.A., FRASER, K., BARNS, G. and ROBINSON, A.J. (1987) The structure and cloning of orf virus DNA. **157**(1): 1-12.

MERCHLINSKY, M. and MOSS, B. (1986) Resolution of linear minichromosomes with hairpin ends from circular plasmids containing vaccinia virus concatamer junctions. Cell. **27**: 391-401.

MESSING, J. (1983) New M13 vectors for cloning. Methods Enzymol. **101, part C**: 20-78.

MORGAN, J.R., COHEN, L.K. and ROBERTS, B.E. (1984) Identification of the DNA sequences encoding the large subunit of the mRNA-capping enzyme of vaccinia virus. J. Virol. **52**(1): 206-214.

MOSS, B. (1985) Replication of poxviruses. In: Fields, B.N. (ed.) Virology. Raven Press, New York. 685-703.

MOYER, R.W. and GRAVES, R.L. (1981) The mechanism of cytoplasmic Orthopoxvirus DNA replication. Cell. **27**: 391-401.

MULLER, H.K., WITTEK, R., SCHAFFNER, W., SCHUMPERLI, D., MENNA, A. and WYLER, R. (1977) Comparison of five poxvirus genomes by analysis with restriction endonucleases HindIII, BamHI and EcoRI. J. Gen. Virol. 38(1): 135-147.

MUNYON, W., PAOLETTI, E. and GRACE, J.T. (1967) RNA polymerase activity in purified vaccinia virus. Proc. Nat. Acad. Sci. USA. **58**: 2280-2287.

NILES, E.G., CONDIT, R.C., CARO, P., DAVIDSON, K., MATUSICK, L. and SETO, J. (1986) Nucleotide sequence and genetic map of the 16-kb vaccinia virus HindIII D fragment. Virology. **153**(1): 96-112.

NOBLE, J., Jr. (1970) A study of new and old world monkeys to determine the likelihood of a simian reservoir of smallpox. Bull. WHO **42**: 509-514.

OLGIATI, D., POGO, B.G.T. and DALES, S. (1976) Biogenesis of vaccinia: specific inhibition of rapidly labeled host DNA in vaccinia inoculated cells. Virology. **71**: 325-335.

PANICALI, D., DAVIS, S.W., MERCER, S.R. and PAOLETTI, E. (1981) Two major DNA variants present in serially propagated stocks of the WR strain of vaccinia virus. **37**(3): 1000-1010.

PAYNE, L.G. and KRISTENSSON, K.(1979) Mechanism of vaccinia virus release and its specific inhibition by N_1-Isonicotinoyl-N_2-3-Methyl-4-Chlorobenzoyl-hydrazine. J. Virol. **32**: 614-622.

PENG, Z. and WU, R. (1986) A simple and rapid nucleotide sequencing strategy and its application in analyzing a rice histone 3 gene. Gene. **45**: 247-252.

PICKUP, D.J., BASTIA, D., STONE, H.O. and JOKLIK, W.K. (1982) Sequence of terminal regions of cowpox virus DNA: arrangement of repeated and unique sequence elements. Proc. Natl. Acad. Sci. USA. **79**(23): 7112-7116.

PILASKI, J., ROSEN, A. and DARAI, G. (1986) Comparative analysis of the genomes of orthopoxviruses isolated from elephant, rhinoceros, and okapi by restriction enzymes. Brief report. Arch. Virol. **88**(1-2): 135-142.

POGO, B.G.T., O'SHEA, M.T. and FREIMUTH, P. (1981) Initiation and termination of vaccinia virus DNA replication. Virology. **108**: 241-248.

RAFII, F. and BURGER, D. (1985) Comparison of contagious ecthyma virus genomes by restriction endonucleases. Arch. Virol. **84**(3-4): 283-289.

RICE, A.P. and ROBERTS, B.E. (1983) Vaccinia virus induces cellular mRNA degradation. J. Virol. **47**(3): 529-539.

ROBINSON, A.J., ELLIS, G. and BALASSU, T. (1982) The genome of orf virus: restriction endonuclease analysis of viral DNA isolated from lesions of orf in sheep. Arch. Virol. **71**(1): 43-45.

ROBINSON, A.J., BARNS, G., FRASER, K., CARPENTER, E. and MERCER, A.A. (1987) Conservation and variation in orf virus genomes. Virology. **157**(1): 13-23.

RODRIGUEZ, J.F., JANECZKO, R.and ESTEBAN, M. (1985) Isolation and characterization of neutralizing monoclonal antibodies to vaccinia virus. J. Virol. **56**: 352-356.

RONDLE, C.J.M, and SAYEED, K.A.R. (1972) Studies on monkeypox virus. Bull. WHO. **46**: 577-583.

ROSEL, J.L., EARL, P.L., WEIR, J.P. and MOSS, B. (1986) Conserved TAAATG sequence at the transcriptional and translational initiation sites of vaccinia virus late genes deduced by structural and functional analysis of the HindIII H genome fragment. J. Virol. **60**(2): 436-449.

ROSEMAN, N.A. and HRUBY, D.E. (1987) Nucleotide sequence and transcript organizaation of a region of the vaccinia virus genome which encodes a constitutively expressed gene required for DNA replication. J. Virol. **61**(5): 1398-1406.

SANGER, F. and COULSON, A.R. (1975) A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. J. Mol. Biol. **94**: 441-448.

SANGER, F., NICKLEN, S. and COULSON, A.R. (1977) DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA. **74**: 5463-5467.

SCHIDA, H. (1986) Nucleotide sequence of the vaccinia virus hemagglutinin gene. Virology. **150**(2): 451-462.

SCHUMPERLI, D, MENNA, A., SCHWENDIMANN, F., WITTEK, R. and WYLER, R. (1980) Symmetrical arrangement of the heterologous regions of rabbit poxvirus and vaccinia virus DNA. J. Gen. Virol. **47**(2): 385-398.

SCHWER, B., VISCA, P., VOS, J.C., STUNNENBERG, H.G. (1987) Discontinuous transcription or RNA processing of vaccinia virus late messengers result in a 5′ poly(A) leader. Cell. **50**: 163-169.

SEALY, P. and SOUTHERN, E. (1982) Gel electrophoresis of DNA. In: Rickwood, D. and Hames, B.D. (ed) Gel electrophoresis of nucleic acids. IRL Press, Washington DC. 39-76.

SETH, A. (1984) A new method for linker ligation. Gene. Anal. Techn. **1**: 99-103.

SILVER, M., McFADDEN, G., WILTON, S. and DALES, S. (1979) Biogenesis of poxviruses: role for the DNA-dependant RNA polymerase II of the host during expression of late functions. Proc. Nat. Acad. Sci. USA. **76**: 4122-4125.

SOLOSKI, M.J. and HOLOWCZAK. J.A. (1981) Characterization of supercoiled nucleoprotein complexes released from detergent-treated vaccinia virions. J. Virol. **37**(2): 770-783.

SOUTHERN, E.M. (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**: 503-517.

TURNER, A. and BAXBY, D. (1979) Structural polypeptides of Orthopoxvirus: their distribution in various members and location within the virion. J. Gen. Virol. **45**: 537-545.

UPTON, C. and McFADDEN, G. (1986) Identification and nucleotide sequence of the thymidine kinase gene of Shope fibroma virus. J. Virol. **60**(3): 920-927.

UPTON, C., DELANGE, A.M. and McFADDEN, G. (1987) Tumorigenic poxviruses: genomic organization and DNA sequence of the telomeric region of the Shope fibroma virus genome. Virology. **160**(1): 20-30.

VENKATESAN, S., BAROUDY, B.M. and MOSS, B. (1981) Distinctive nucleotide sequences adjacent to multiple initiation and termination sites of an early vaccinia virus gene. Cell. **25**(3): 805-813.

VENKATESAN, S., GERSHOWITZ, A. and MOSS, B. (1982) Complete nucleotide sequences of two adjacent early vaccinia virus genes located within the inverted terminal repetition. J. Virol. **44**(2): 637-646.

VON MAGNUS, P., ANDERSEN, E.K., PETERSEN, K.B. and BIRCH-ANDERSEN, A. (1959) A pox-like disease in cynomolgus monkeys. Acta. Pathol. Microgiol. Scand. **46**: 156-176.

WEIR, J.P. and MOSS, B. (1983) Nucleotide sequence of the vaccinia virus thymidine kinase gene and the nature of spontanious frameshift mutations. J. Virol. **46**: 530-537.

WEIR, J.P. and MOSS, B. (1987) Determination of the promoter region of an early vaccinia virus gene encoding thymidine kinase. Virology. **158**(1): 206-210.

WILIMZIG, M. (1985) LiCl-boiling method for plasmid mini-preps. Trends Genet.

WITTEK, R., MENNA, A., SCHUMPERLI, D., STOFFEL, S., MULLER, H.K. and WYLER, R. (1977) HindIII and SstI restriction sites mapped on rabbit poxvirus and vaccinia virus DNA. J. Virol. **23**(3): 669-78.

WITTEK, R., MENNA, A., MULLER, H.K., SCHUMPERLI, D., BOSELEY, P.G. and WYLER, R. (1978) Inverted terminal repeats in rabbit poxvirus and vaccinia virus DNA. J. Virol. **28**(1): 171-181.

WITTEK, R., MULLER, H.R., MENNA, A. and WYLER, R. (1978) Length heterogeneity in the DNA of vaccinia virus is eliminated on printng the virus. FEBS lett. **90**(1): 41-46.

WITTEK, R., KUENZLE, C.C. and WYLER, R. (1979) High C + G content in parapoxvirus DNA. J. Gen. Virol. **43**(1): 231-234.

WITTEK, R. and MOSS, B. (1980) Tandem repeats within the inverted terminal repetition of vaccinia virus DNA. Cell. **21**: 277-284.

WITTEK. R. (1982) Organization and expression of the poxvirus genome. Experientia. **38**(3): 285-297.

YANISCH-PERRON, C., VIEIRA, J. and MESSING, J. (1985) Improved M13 phage cloning vectors and host strains; nucleotide sequences of the M13mp18 and pUC19 vectors. Gene. **33**: 103-119.

ZWARTOUW, H.T. (1964) The chemical composition of Vaccinia virus. J. Gen. Microbiol. **34**: 115-123.