ANALYSIS OF THE INTERNAL REPLICATION SEQUENCE OF MINUTE VIRUS OF MICE

by

JOHN BRUNSTEIN

B.Sc. (Hon). Simon Fraser University. 1992

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE STUDIES

Department of Biochemistry and Molecular Biology

THE UNIVERSITY OF BRITISH COLUMBIA

December 1997

Abstract

Minute Virus of Mice (MVM) is a member of the *Parvovirinae* genus of the *Parvoviridae* family of viruses. This family of small, single-stranded DNA viruses infect a wide range of eukaryotic hosts ranging from insects to humans. Due to their small size and limited coding capacity parvoviruses may serve as a suitable tool with which to examine elements of the host cell DNA replication machinery. Previous studies on MVM have localized a region of approximately 200 nucleotides inboard of the right viral genomic termini, known as the Internal Replication Sequence (IRS), which contributes in *cis* to viral replication competence.

A comprehensive library of linker scanning mutations across the IRS of MVM was constructed and assayed in the context of a minigenomic system for replication competence in an effort to identify short sequence elements contributing to viral replication. Three elements required for efficient replication were observed. Elements of the library were also examined for interactions with host cell nuclear factors, and such interactions were localized to four sites with evidence being obtained to suggest positive interactions between the sites. These sites were found to be directly adjacent to two of the elements required for efficient replication and overlapping the third, suggesting a possible correlation between these functions. Simultaneous deletion of two of these binding sites was observed to abrogate function of the IRS, further supporting a functional relationship between factor binding and origin activity.

The sequence element found to both bind host factors and be required for replication competence was employed as 'bait' in a yeast one-hybrid genetic screen of a murine cDNA library in an attempt to clone interacting factor(s). A clone recovered from this, while not considered likely to be directly relevant, leads to postulation of a known origin binding protein RIP60 as a prospective candidate for action at the IRS origin. Preliminary studies conducted to determine whether RIP60 binds at the viral origin were conducted but failed to provide evidence of RIP60 association with the viral genome.

A model is proposed whereby a leading-strand only origin of DNA replication within the region of the IRS affords a mechanism for the viral rearrangement of its 5' termini from an extended to a hairpin form during replication. Studies employing the viral IRS and right-hand terminus as an origin driving the replication of attached unrelated vector sequences are presented in support of this model.

Table of Contents

List of Tables

List of Figures

Abbreviations

ACS	ARS consensus sequence
ARS	autonomously replicating sequence
bp	base pairs
DEAE	diethyl aminoethyl
DHFR	Dihydrofolate reductase locus
DMEM	Dulbecco's Modified Eagle's Medium
DUE	DNA Unwinding Element
EP	early palindrome
EtBr	Ethidium Bromide
HEPES	N-2-Hydroxyethyl Piperazine N-2-Ethanesulfonic Acid
m.u.	map unit
mRF	monomer replicative form
NS	non-structural
nt	nucleotide
OBR	Origin of bidirectional replication
ORC	origin recognition complex
ori	origin of replication
PBS	Phosphate-buffered Saline
PCNA	proliferating cell nuclear antigen
RE	Replication Efficiency
RF-C	replication factor C
RPA	replication protein A
SR proteins	Serine-Arginine proteins
SRE	Scaled Replication Efficiency
SRRE	Scaled Relative Replication Efficiency
SSB	single-stranded binding protein
VP	viral protein (structural)

Acknowledgment

In the course of this study the support, guidance, and input of a number of people has been of great help and deserves recognition. First and foremost the patient supervision of Dr. Caroline Astell, without whose able direction this project would not have come to fruition, warrants my sincere thanks. Dr. Pat Tam should be recognized not only for his work in uncovering the existence of the IRS which laid the groundwork for my project, but also for his help and encouragement as I set out upon the project. Many late nights at the lab were livened by entertaining discussions with Dr. Wesley Hung; likewise, Glen Coburn earns thanks for his keen insight into many topics (even on occasion scientific in nature) over a pitcher of beer. Marcia MacDonald deserves special thanks for her unassailable good humor which did much to make the lab an entertaining workplace. Chris Houchens and Dr. N. Heintz from the University of Vermont were kind enough to supply both reagents for the detection and analysis of RIP60's potential interaction with the IRS, along with unpublished data on the topic. Finally, I would like to extend my appreciation to other members of the Astell lab past and present for assistance in many diverse aspects of the project: Qingquan, Colin, Jan, Carl, Darren, Rick, and Warren.

When I look back on graduate school as having been both a productive and enjoyable experience it is in no small measure due to the input of all of these people; I hope that while inadequate this acknowledgment serves to express my sincere appreciation to all of them.

Introduction

In a number of organisms examined to date, DNA replication is initiated by the specific interaction of protein factors with sequences at a *cis*-acting origin of replication (ori) to form a preinitiation complex. Action of the protein factors results in a localized kinking, bending, and/or unwinding of the associated sequences to create a suitable target for the initiation of DNA synthesis by primase and polymerase functions. Studies on prokaryotic and some simple eukaryotic systems have allowed for the identification of the basic structure and common components of such origins. Common components consist of an essential core sequence element containing binding sites for origin recognition proteins and sequence segments predisposed to easily unwind or bend, complemented by associated nonessential enhancer elements which frequently include transcription-factor binding sites[1,2,3,4]. Despite recent advances in these studies, ori nature and function in metazoans is as yet poorly understood.

Parvoviral systems may act as a model with which to examine aspects of eukaryotic replication. Minute Virus of Mice, fibrotropic strain (MVM(p), henceforth referred to as MVM) is a member of the genera *Parvovirus* whose members are characterized by unenveloped icosahedral capsids containing small single-stranded DNA genomes with long terminal palindromic sequences. Unlike their close relatives in the *Dependovirus* genus, *Parvovirus* are capable of replication autonomously of helper viruses. Current models of this replication[5,6,7] propose a strand-extension mechanism mediated by host-cell replication machinery, with a single viral protein NS-1 acting to resolve concatamer junctions. That this replication cycle is observed to be dependent on

host cell S phase[8] suggests a strong reliance on host cell replication machinery, and not simply a repair synthesis mechanism as might be surmised from the replication model. An elucidation of required viral *cis*-acting sequences and their associated host-cell proteins in this system may thus serve to better our understanding both of MVM and eukaryotic replication in general.

The Parvoviruses

The *Parvoviridae* comprise a large family of viruses related on the basis of structural and genomic similarity. The name comes from the Latin 'parvus' or small, and the virions are among the smallest known consisting of an unenveloped T=1 icosahedral capsid 18-24 nm across which contains a single-stranded DNA genome of approximately 5000 nucleotides. Generally, one strand is preferentially packaged over the other, with the negative-sense being the most common; there are however exceptions to this both for specific members of the family, and for specific members in differing host cell types. Long palindromic sequences at the genomic termini allow for the formation of double-stranded 'hairpin' telomeres by folding back of the sequence at its axis of symmetry. Aside from the provision by one of these hairpin structures of a free 3' hydroxyl as a primer for DNA synthesis in the replication of the genome, these telomeres may also serve to sequester the free ends of the genome from cellular surveillance mechanisms or degradation.

The genome encodes two major open reading frames: in the context of a negative-sense genome the first of these codes for a set of non-structural proteins known as NS or Rep, while the second codes for a set of structural proteins known as VP or Cap. In

addition, at least B19 (of the *Erythrovirus* subgenera) codes for several smaller expressed ORFs whose function(s) are currently unknown[9]. Viral ORFs are driven by either one (i.e., in B19), two (i.e., in MVM), or three promoters (i.e., AAV) and are expressed in a temporally ordered fashion[10]. NS or Rep proteins serve in the resolution of viral replicative intermediates and in packaging mature genomes into capsids, while the VP or Cap proteins form the capsid.

Due in large part to the limited coding capacity of the viral genome, parvoviruses are highly dependent on host cell functions for a productive infection. Replication only occurs in the nucleus of permissive cells during S-phase, suggesting the need for specific cellular replication machinery in the viral replication process. This need for actively dividing host cells appears as a tropism for tissues such as erythroid precursors in bone marrow (B19), epithelial cells of the gastrointestinal tract (CPV), or fetal tissues (MVM). The majority of pathogenic effects attributable to parvovirus infection reflect this tropism: in the case of the three examples given above, B19 infection can result in a transient aplastic crisis in humans, CPV in severe gastrointestinal distress in canids, and MVM can cause teratogenic effects in hamster.

Their simplicity and resulting reliance on the host cell have made parvoviruses an attractive model system with which to probe features of DNA replication and gene expression in the host-cell mileu.

Taxonomy

The family *Parvoviridae* is divided into two subfamilies on the basis of host range[11]. The *Parvovirinae* infect a range of vertebrate species, while the *Densovirinae* infect arthropods. Each is further subdivided, with the *Parvovirinae* containing the genera *Parvovirus* (whose replication does not require the coinfection of a helper virus such as Adenovirus or a Herpesvirus), *Dependovirus* (which normally do require helper-virus coinfection), and *Erythrovirus* (members of which infect a limited subset of erythroid precursor cells). The *Densovirinae* also consist of three genera (although there has been a recent suggestion to split this into four), the *Densovirus*, *Iteravirus*, and *Contravirus* on the basis of genome organization and host range. Table 1 lists some of the known parvoviruses and their host species.

MVM

MVM discovery

The prototypical fibrotropic strain of Minute Virus of Mice, MVM(p), was first isolated in 1966 by Crawford as a contaminant in a culture of mouse adenovirus and polyomavirus grown on monolayers of secondary mouse embryo cells[12]. A distinct band under isopycnic ultracentrifugation in CsCl at 1.43 g/cm^3 was observed to haemagglutinate murine erythrocytes but was found not to produce plaques on mouse embryo cell monolayers. Cultures of mouse embro cells innoculated with material from this band were found to produce more haemagglutinin, suggesting the band comprised a previously unrecognized virus. Further characterization of the virus demonstrated it to

Parvoviridae

Parvovirinae

Parvovirus	Host species
Minute Virus of Mice	mouse
Mouse Parvovirus	mouse
Kilham Rat Virus	rat
Aleutian Mink Disease Virus	mink
Mink Enteritis Virus	mink
Canine Parvovirus	dog
Minute Virus of Canines	dog
Feline Panleukopenia Virus	cat
Goose Parvovirus	goose
LuIII	(unknown)
Porcine Parvovirus	pig
H1 Virus	rodents
Bovine Parvovirus	cattle
Dependovirus	
Adeno-Associated Virus (5 serotypes)	most mammals
Bovine Adeno-Associated Virus	cattle
Avian Adeno-Associated Virus	birds
Erythrovirus	
B19	human
Simian Parvovirus	monkey
Pig-tailed Macaque Parvovirus	monkey

Table 1: Parvovirinae

Some of the known members of the *Parvovirinae* genera of the *Parvoviridae* family, along with their natural hosts (where known).

have an unenveloped icosahedral capsid approximately 19nm across containing a single-stranded DNA genome. Based on these properties, similarity to Kilham Rat, H1, H3, and X14 viruses was suggested; however antigenic properties proved the new isolate to be distinct and it was named as MVM.

Structural properties

MVM has an unenveloped icosahedral capsid displaying T=1 symmetry. The capsid consists of two proteins, a major component (VP2) and a minor component (VP1), with a proteolytic processing product of VP2 (VP3) present at low frequency. Structural determination by X-ray crystallography[13] has recently been successful at a resolution of less than 3Å on empty capsids, and at 3.5Å on both empty and full capsids of the allotropic variant MVM(i). The capsid consists of 60 asymmetric subunits of the structural proteins, each in the form of an eight-stranded antiparallel β-barrel. Each capsomer holds a loop of 20 nucleotides in an icosahedrally ordered structure. The capsid surface is irregular, with a distinct spike at each of the five-fold axes of symmetry surrounded by a circular cleft or 'canyon'; evidence from the structural differences between MVM(p) and MVM(i) suggests the viral receptor binding site lies within this canyon. In addition, in MVM there appears to be a distinct depression along the two-fold axes of symmetry, a region referred to as the 'dimple'. Current data also suggests the presence of an open channel through each of the five-fold spikes, the importance of which is unknown.

In culture, the virus is found to produce mainly (by a ratio ranging from 3:1 to 50:1)[14] empty capsids (ρ =1.35 gm/cm^3) , with full capsid banding on CsCl at ρ =1.41 to 1.43 g/cm^3.

Genome Organization

The MVM genome has been sequenced in its entirety and contains 5149 nucleotides in the mature virus[15]. The left-end (3') imperfect palindrome consists of 115 nucleotides of which 104 can arrange in a stable base-paired conformation, while the one at the right (5') end is 206 nucleotides of which 200 can base-pair. The genome is negative-sense, with the complementary strand encoding two large open reading frames[16]. While sequences in the left end are unique, the right end exists in two equally abundant complementary forms known as 'flip' and 'flop'[15]. The significance of this difference between the termini on the viral replication strategy will be dealt with in detail in the section on viral replication. In addition to this variation, the 5' end sequences in double-stranded monomeric replicative-form (mRF) DNA contain an additional 18 nucleotides not found in packaged mature genomes; these are complementary to genomic nucleotides 4858 through 4876 which lie immediately inboard of the 5' hairpin.

Transcripion

The two major open reading frames give rise to two overlapping transcriptional units which are expressed in a temporally ordered fashion[10]. A canonical eukaryotic Pol II promoter at 4 map units (P_4) drives the expression of two major families of transcripts, the R1 and R2 families. A second Pol II promoter at m.u. 39 (P_{39}) drives expression of R3, the third major transcript family. Each of these families undergoes three alternative

patterns of splicing at a small intron between m.u. 44 and 46, resulting in at least nine known transcripts[17]. All transcripts are complementary to the genomic DNA, and essentially all are polyadenylated at the last of four canonical AATAAA sequences at viral nucleotide 4908[18] (see Figure 1).

The common intron at m.u. 44-46 contains two potential splice-donor (D) and two potential splice-acceptor (A) sites, known as D1 (nt. 2280), D2 (nt. 2317), A1(nt. 2377), and A2 (nt. 2399), respectively. In all cases the major spliced product is D1-A1, with D2-A2 occuring at lower frequency and the D1-A2 products only being observed rarely. In no case is D2-A1 observed, apparently due to steric constraints as the donor and acceptor sites in this case are only 59 nt. apart. Observations by Haut *et. al.* support this hypothesis, as insertions between D2 and A1 render this splice viable[19]. The frequency of selection of splice donor-acceptor pairs appears to be influenced by a combination of spacing and sequence at the donor and acceptor sites, with the first donor site being used preferentially in all cases; this bias is modulated by the D2 donor being a closer match to the canonical splice-donor consensus sequence than is D1 to allow for its selection at a greater frequency than if both splice sites had identical sequences. In addition, evidence has been put forth by Gersappe *et. al.*[20] to suggest that sequences within either termini of the NS-2 specific exon (discussed below in connection with the R2 transcript) are required for accurate splicing, suggesting the possible interaction of SR proteins in the splicing of viral transcripts.

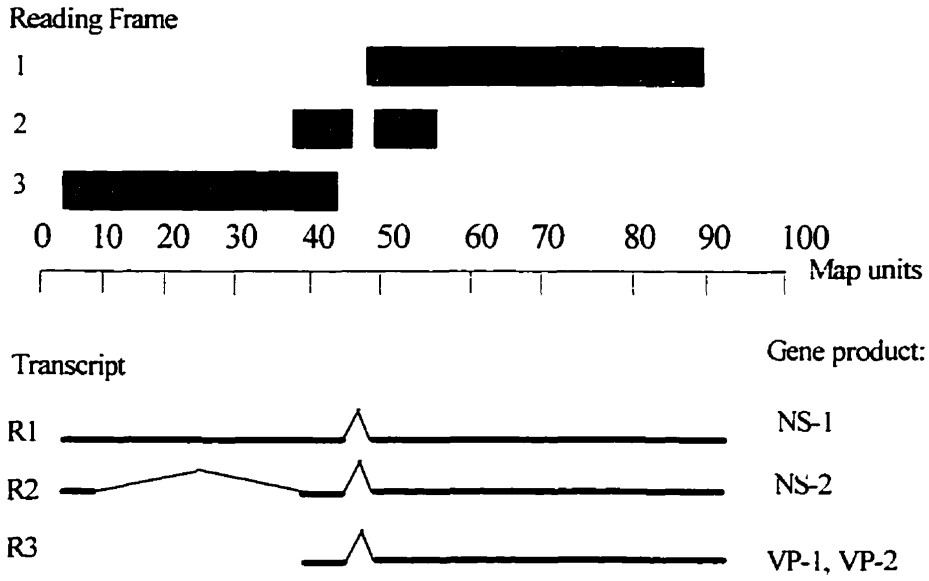


Figure 1:MVM Transcriptional Map

Transcriptional and translational map of MVM. Shaded boxes indicate relative locations of open reading frames. Heavy lines in transcript map indicate regions of mRNA present in mature transcripts, lighter lines indicate introns. Three alternative splicing patterns for the middle intron (m.u. 44 to 46) are not indicated (see text). Adapted from Reference 16.

The R1 transcript family is approximately 4.8 kb in length and makes up roughly 15% of the observed viral transcripts during late infection in murine cell culture. None of the three variant splicing patterns of the internal small intron affect the coding region of these transcripts, thus all are translated to a single non-structural protein NS-1. This protein carries out a variety of functions in viral replication and the resolution of genomic concatamers which will be discussed later.

The R2 transcript family is approximately 3.3 kb in length, extending from m.u. 4 through to the polyadenylation site but with m.u. 10 through m.u. 39 (viral nucleotides 514 through 1990) spliced out. This transcript family, which also makes up approximately 15% of the total viral transcripts, thus results in a protein-coding message whose 5' end corresponds to the 5' end of the NS-1 coding frame and whose 3' end exists in three unique forms depending on the splicing of the common small middle intron. The three closely related proteins arising from this message family are known collectively as NS-2; their function(s) are poorly understood as outlined in the section on viral polypeptides.

The R3 transcript makes up the majority (70%) of viral transcripts. Arising from the P_{39} promoter, this family of messages requires transactivation by the viral NS-1 product of the R1 transcript in order to be expressed,[21,22] leading to a temporal ordering in viral gene expression whereby the R1 and R2 transcripts are expressed immediately upon infection, with R3 not being detectable until several hours later[10]. The R3 transcript contains the right-hand viral ORF coding for the VP1 or VP2 structural proteins, which are differentially expressed based on the choice of donors and acceptors for the common

middle intron (outlined further in the following section on viral polypeptides). Use of a suboptimal first splice donor site (D1) helps to maintain a proper ratio between these two messages as the presence of a better splice donor in this preferred first position would greatly reduce the levels of VP1 message.

Viral Proteins

MVM encodes four polypeptides in its two ORFs (considering the variant forms of NS-2 mentioned above as a single species). The left-hand end of the genome codes for NS-1 and NS-2 nonstructural proteins, while the right-hand end codes for VP1 and VP2 capsid proteins.

NS-1 is an 83 kDa phosphoprotein which is localized to the nucleus of murine cells in the course of natural infections[23]. Coded for in reading frame 3 by the R1 transcript, it is 672 amino acids in length; as its stop codon lies upstream of the small middle intron D1 donor, all three of the R1 transcripts encode the same polypeptide. In contrast to other viral proteins, the NS-1 polypeptide is highly conserved across the rodent parvoviruses, with sera from animals infected with CPV, PPV, H-1, H-3, or KRV all displaying cross-reactivity to MVM NS-1[24]. The degree of highest DNA sequence homology lies between nt. 1428-1833, with values in the 90% range when compared to other members of the rodent parvoviruses, 51% to the Dependovirus AAV, and 41% to the Erythrovirus B19 [15,25,7]. Within this region a purine nucleoside triphosphate binding fold consensus sequence [$GX_4GK(T/S)X_{5-6}(I/L/V)$] is found[26].

While biochemical functions for this protein have been difficult to demonstrate in crude extracts of infected cells due to its low abundance, overexpressed and purified protein has been demonstrated to have ATPase, helicase, and site-specific single-stranded endonuclease activities[27,28]. The protein has been recently shown to bind directly to the DNA sequence $[ACCA]_2$[29], with there being evidence that a flanking (A/T) on both sides is preferred[30]. Helicase activity is thought to be dependent on ATP hydrolysis, as nonhydrolyzable analogs do not support this activity. Furthermore, mutations in the ATP binding pocket which uncouple these processes such that ATP binding and hydrolysis can still occur but helicase activity has been abolished suggest that the ATP hydrolysis serves as the energy source for helicase activity[31]. In contrast, DNA binding activity appears normal in the presence of nonhydrolyzable analogs[32]. These activities of NS-1 are differentially modulated by its phosphorylation state: relative to the phosphorylated state, the dephosphorylated protein shows 300% binding activity, <100% nicking activity, <3% helicase activity, and from 15-30% ATPase activity[33]. Based on these observed activities as well as by the observation that NS-1 is found covalently linked to the 5' ends of mature unit-length viral genomes[34] , it is predicted that NS-1 acts in a manner analogous to the gene-A protein of ϕX-174 in the resolution of unit-length single-stranded progeny genomes from double-stranded concatameric replicative intermediates.

NS-1 is also responsible for both trans-activation and trans-inhibition of the viral P_4 and P_{38} promoters[21]. It is capable of either up or down-regulating expression of the R1 transcript by interacting with the P_4 promoter (presumably in a dose-responsive manner) and acts to upregulate expression of the R3 message by approximately 1000-fold[22]. The

source of this transactivation is not as yet completely understood; while NS-1 itself contains a potent acidic activation domain (on the order of 30-50% as strong as that of VP16) located within the C-terminal 88 amino acids[35,36] it has also been shown to have a direct interaction with the transcription factor Sp-1 in transactivation of the viral P_{38} promoter[36]. Given that Sp-1 acts in a different fashion than the acidic activators to enhance transcription, there is a distinct possibility that the interaction with NS-1 is synergistic. This interaction may also conceivably result in an upregulation of host-cell genes with associated Sp-1 sites during the course of viral infection.

The NS-2 proteins are a set of three 25 kDa phosphoproteins coded for by the R2 transcript family. These species are labile, with half-lives under one hour[37], and are found primarily in the cytoplasm of infected cells (with phosphorylated forms being found only in the cytoplasm)[38]. The coding frame is identical to that of NS-1 for the first 85 amino acids (reading frame 3, viral nt. 260 to 514), however the splice from m.u. 10 to m.u. 39 shifts the remainder of the message into reading frame 2. In addition the three splicing patterns observed for the small common middle intron result in three isoforms of NS-2 which differ at their carboxy termini. While the functions of this protein are somewhat of a mystery, it has been shown to be required for efficient replication and for production of infectious virus in some cell types[39]. It has been suggested that NS-2 may play a specific role in enhancing the translation of viral mRNAs, in particular VP-2[40]. More recently, a role for NS-2 in capsid assembly in murine cells has been reported[41]. The significance, if any, of there being three isoforms of this protein is unknown.

Both NS-1 and NS-2 have cytotoxic effects, which are especially pronounced in transformed cell lines. Such effects appear to be mediated through the action of host oncogenes, including c-*ras* [42]. Studies have shown that the cytotoxic domains of NS-1 can be mapped to the same regions as the transactivation functions[43], suggesting that the roles may be related through the inappropriate regulation of such host genes by the NS-1 protein. In particular, NS-1 has been shown to cause an upregulation of the human c-*erbA* gene which encodes the thyroid hormone (T3) receptor α subunit; this in turn increases the host cell's sensitivity to parvoviral attack[44].

VP-1 is the larger of the two capsid proteins, at 83 kDa; it accounts for 15-18% of the capsid mass[14]. This low fraction is due to its being coded for by one form of the R3 transcript in which the small common middle intron is spliced using the D2-A2 (low-frequency) splice donor-acceptor pair, thus retaining an initiation codon at nt. 2286. VP-1 thus has a unique basic N-terminus while the rest of the protein is identical with VP-2. This N-terminal region has been found to be resistant to protease cleavage in either full or empty viral particles; furthermore, in the closely related H-1 virus chemical cross-linking between VP-1 molecules can be observed in empty capsids, but not in full capsids[45]. Taken together with evidence that aggregates of MVM's VP-1 bind the 3' genomic hairpin internally in the particles[46], this suggests that this protein's basic N-terminus is involved in sequestration of the viral genome within the capsid. While VP-1 is not essential for genome encapsidation, data has been presented which demonstrate a requirement for this protein in allowing viral progeny to be infectious[47].

VP-2 is smaller than VP-1, at 64 kDa. It is coded for by the most common subfamily of R3 transcript in which the small common middle intron is spliced out using the D1-A1 donor-acceptor pair, and also by the rare third splicing product D1-A2; in either case the AUG codon at nt. 2286 is removed, allowing a second AUG in the same reading frame at nt. 2795 to act as an initiation codon. VP-2 makes up the majority of the capsid, and in fact structurally (by electron microscopy) and antigenically 'normal' capsids can be made in systems only expressing VP-2 [47,48]. VP-2 is found to be resistant to protease cleavage in empty capsids, but in full capsids it is sensitive to a specific cleavage of its N-terminal domain to yield VP-3, suggesting some structural rearrangement occurs upon genome encapsidation.

VP-3 is a 61 kDa proteolytic product of VP-2. While VP-2 to VP-3 conversion accompanies the infection of internalized virus in a time-dependent fashion, particles are found to be equally infectious regardless of their VP-2/VP-3 content. It has been suggested that [illegible] a heterogenous population of differently phosphorylated species, and that only the most highly-phosphorylated forms are susceptible to this cleavage step in what may be a form of regulation of viral particle maturation[49].

DNA Replication

Features of Replicons

The crucial event in the replication of all cells is the duplication of its genetic material. As befits so critical a process, it has been extensively studied in order to gain insights into a number of aspects central to biology. DNA replication is thought to begin at discrete sites known as origins of replication or oris. An ori is postulated to be a DNA

sequence or set of sequences which act to bind cellular replication factors in an organized manner such as to signal for the initiation of synthesis of a new DNA strand. Techniques for the identification and mapping of such sequences have identified them successfully in prokaryotes and their viruses, as well as in lower eukaryotes (*Saccharomyces* sp.) and some viruses of eukaryotes. In the case of higher eukaryotes, while certain specific instances of discretely localized sites where replication reliably initiates every cell cycle have been identified, there is still debate as to whether the majority of genome replication occurs from defined loci or is more of a stochastic process, with replication initiating on average as frequently as once every 12 kbp of sequence[50].

All of the well-characterized origins to date appear to share certain similarities both in their basic organization and in their mode of action. The best studied of these is the OriC responsible for the establishment of replication forks on the *E. coli* chromosome. In this system[51], outlined here as a paradigm for replication initiation in general, multiple copies of a cellular sequence-specific DNA binding protein (dnaA) bind tightly to four repeats of a 9-base target sequence (the 'dnaA box') to yield an 'initial complex'. In this, the DNA is maintained in a double-stranded form tightly wound around a central core of from 20-40 copies of dnaA. Immediately adjacent to this looped complex are four repeats of a 13-base AT-rich sequence element, which are sequentially melted open by an ATP-driven activity of the dnaA protein to form an 'open complex' with approximately 45 base pairs of the AT-rich DNA element in a denatured single-stranded form. This open complex then recruits a dnaB-dnaC protein complex, apparently through recognition of the single-stranded structure of the 13-mer elements by

dnaB. Binding of this complex creates the 'prepriming complex', stabilizing the unwound DNA and localizing a DNA helicase activity (dnaB) at the site for subsequent activity by a primase function and the initiation of DNA synthesis.

Examination of a large number of other less well characterized prokaryotic replication origins has revealed the conservation of an AT-rich repeat motif in most cases. Extant footprinting data also suggests that there is some conservation in the protein-binding action of these origins. With this in mind, and based on the OriC system, a general model for the initiation of DNA replication was proposed[52] whereby an origin-specific initiator protein binds strongly to a short reiterated sequence motif in one section of the origin and induces the melting of an adjacent AT-rich sequence motif, to which the initiator protein may demonstrate a weaker binding. The AT-richness of this second sequence element presumably is important in its propensity to be easily unwound, leading to its being referred to as a DNA Unwinding Element or DUE by some investigators. This general model seems to fit well with the observed organization of not only the *E. coli* OriC, but also OriC from *B. subtilis*, the plasmid-born origins of pSC101, F, P1, R1, RK6, RK2, and P4, as well as the origins of lambdoid phages λ, ϕ80, and ϕ82. In each case, binding of a sequence-specific protein at or near an AT-rich element and subsequent melting of this element and recruitment of a DNA helicase function is observed as a prerequisite for priming of DNA synthesis.

Examination of the initiation of DNA replication in eukaryotes has lagged behind similar studies in prokaryotes and their viruses due to the increased complexity and size of eukaryotic systems. However in recent years advances in molecular biological

techniques have allowed for studies on a number of eukaryotic models, most notably the papovavirus SV40. The successful purification of all components required to observe *in vitro* replication of this virus in a reconstituted system has allowed for the elucidation of many of the components and physical processes which are likely common to all eukaryotic replication. While a number of aspects of replication must of necessity vary between a small circular sytem such as SV40 and the immensely larger linear chromosomes of the host, similarities between the SV40 system and the generalized model derived from prokaryotic systems and presented above suggests that the mechanism of initiation of replication may be conserved; a model for SV40 replication[53,54] is thus presented here as a paradigm for eukaryotic initiation.

SV40 has a single ori sequence which consists of an essential 65 bp core containing a 27 bp dyad symmetry element with four 5'-GAGGC-3' sequence repeats, flanked on one side (the 'Early' side) by an imperfect palindrome (EP) and on the other (the 'Late' side) by an AT-rich sequence motif (AT). A viral protein, the Large T-antigen, binds first to the 27 bp element in the form of monomers on each of the 5'-GAGGC-3' sequences; additional molecules of T-antigen then associate through protein-protein interaction to form two hexamers, each arranged as a ring around the DNA but interacting almost exclusively with the sugar-phosphate backbone on one strand. This results in structural deformations in both of the flanking regions, notably the melting of an 8 bp bubble in the EP which is expanded by an ATPase-dependent helicase activity of the T-antigen hexamers causing DNA unwinding to proceed bidirectionally outward from the central dyad element. The limitation of T-antigen contacts to a single strand allows

for the other strand to bind SSB proteins (in particular replication protein A (RPA), a trimeric SSB specifically associated with replication forks) in order to stabilize the nascent fork structures. These nascent forks then attract the Pol α-primase complex, which is responsible for laying down a short RNA primer followed by a longer DNA segment. The 3' end of these growing strands is then bound by replication factor C (RFC), which acts to recruit proliferating-cell nuclear antigen (PCNA) and pol δ in a replacement of the Pol α-primase complex.

Recent advances have also been made in the understanding of the initiation of replication in the budding yeast *S. cerevisiae*. Shotgun cloning of the genome into replication-incompetent vectors and selection for plasmid maintenance has successfully identified short conserved DNA sequences, known as Autonomous Replication Sequences (ARSs), which act in their native chromosomal loci as bi-directional origins of replication. These ARSs all contain an 11-base sequence 5'-(A/T)TTTA(T/C)(A/G)TTT(A/T)-3' known as the ARS consensus sequence (ACS) which is essential for origin function. In addition, ARSs contain a discrete second DNA element 3' to the ACS's T-rich strand which consists of an easily-unwound sequence; the exact sequence can vary between loci but often seems to contain two or three distinct elements, known as B1, B2, and B3. While these elements are individually dispensable, they act in concert with the ACS to achieve full origin function[55,56]. A multiprotein complex (the Origin Recognition Complex, ORC) has been found to bind specifically to the ACS, and all six subunits in the complex have been cloned. Although two of the subunits (Orc1p and Orc5p) appear to have nucleotide binding folds[57], and the binding of

the complex to the ACS has been shown to be ATP dependent, no other biochemical functions have been identified to date. In yeast, conditional mutants of Orc2, Orc 3, and Orc5 have all been demonstrated to have defects in DNA replication[58] while in *Xenopus*, Orc1 and Orc2 have been shown to be essential for the initiation of semi-conservative DNA replication[59,60].

Genomic footprinting experiments have demonstrated that the yeast ACS is bound by a protein complex (or complexes) throughout the cell cycle, which exists in at least two distinguishable forms. The post-replicative footprint observed closely matches that seen *in vitro* with purified ORC, suggesting that rather than acting as an activator of replication it may serve to suppress the origin after a single round of replication has occured in S phase. The pre-replicative footprint is larger than that of the post-replicative state, extending from the ACS well into the 'B' domain, and has been hypothesized to consist of ORC with a number of accessory factors associated[61]. A model is arising whereby ORC acts to mark the location of origins of replication, while the association and dissociation of other factors serve to restrict the initiation of DNA replication to a single occurrence in S phase.

In this model[62], loss of Cyclin B activity in anaphase allows for the association of RLF-B (Replication Licensing Factor B, a putative cell-cycle dependent activity which may be identical with Cdc6 in *S. cerevisiae*) with ORC on chromatin during G1, followed by the recruitment of Replication Licensing Factor M (RLF-M). RLF-M appears in turn to be composed of some or all of the six MCM (mini-chromosome maintenance) proteins. This entire complex constitutes the pre-replicative complex, and following transient

association with a postulated S-phase promoting factor, (SPF), DNA replication is initiated at the origin. This results in loss of both RLF factors and inactivation of RLF-B by cyclin-dependent kinases, leaving ORC in place as the post-replicative complex but incapable of re-initiating replication until the cell proceeds through M phase again. Evidence for this model has come from a number of studies indicating a requirement for ORC and MCM proteins for the initiation of DNA replication, as well as genetic and biochemical data suggesting an interaction between these complexes and their epistatic control by cyclin-dependent kinases (CDKs) (reviewed in Rowles and Blow[62]).

Among the origin sequences identified in higher eukarotes the best studied is that found downstream of the dihydrofolate reductase (DHFR) locus in Chinese Hampster Ovary cells, where an Origin of Bidirectional Replication (OBR) has been localized to a 450 bp segment of DNA[63]. This origin is composed of a section of bent DNA containing an A/T rich sequence motif including an (5'-ATT-3') repeat which is specifically bound by a 60 kDa protein known as RIP60 (for Replication Initiation Protein). While studies to date on this origin have not elucidated a complete model for initiation of DNA synthesis, several of the steps in the formation of a 'prepriming complex' have been observed. Studies indicate that RIP60 binds within this region as dimers and larger concatamers at two copies of the (5'-ATT-3') motif, one within the OBR and another in inverse orientation approximately 700 bp downstream. Protein-protein interactions between the RIP60 multimers results in a looping of the intervening DNA sequence and the imposition of torsional stress[64]. Affinity purification of RIP60 from nuclear extracts by oligonucleotide affinity chromatography indicate it associates with a 100 kDa

ATPase/helicase protein[65], perhaps in a manner analogous to the dnaA/dnaB-dnaC interaction in formation of a prepriming complex.

In addition to specific protein-binding and associated AT-rich sequence elements there is evidence for conservation of certain higher-order structural motifs at origins of replication. These motifs can be divided into two categories: in the first, sequences with repeated short oligo(dA) tracts spaced in synchrony with the helical phase of the DNA result in a predicted aberration of the local DNA from a linear to a bent conformation; in the second, short oligo(dA) tracts are again found, but in this case arranged such as to avoid synchrony with the helical phase, thereby maintaining the DNA in a macroscopically unperturbed or 'anti-bent' form[66]. In a number of examples from both prokaryotes (such as OriC of *B. subtilis* or the ori of pSC101) and eukaryotes (ARS1 of *S. cerevisiae*) both motifs are present, while in other cases (such as the SV40 ori) only the bent motif is observed. It is proposed that these structural motifs have a functional significance, with the bent region acting to localize the first origin-binding proteins and the unbent region acting as an easily unwound element (a DUE) or acting as a rigid region which avoids binding of duplex-stabilizing proteins such as histones[66]. In some cases the structure of these elements may be of equal or greater importance than their sequence, as in the case of the bent region from ARS1; while deletion of this element reduces replication activity, replacement of the original sequence with an unrelated bent element allows activity to be maintained[67].

MVM replication

The problem of replication of the termini of a linear genome without sequential and eventually fatal loss of genomic size has been solved in a number of different ways by different organisms. Among the *Parvoviridae* the first step in addressing this problem is the inclusion of palindromic sequences at both termini, allowing for the formation of the double-stranded viral 'hairpin' telomeres characteristic of the family. A method for the replication of telomeres of this form was first put forward as the 'Hairpin Transfer' model[68], whereby the 3' end of the single-stranded region folds back on itself and primes extension back towards the 5' end of the parental strand. A site-specific single-stranded endonucleolytic cleavage event on the original parental strand is then postulated to occur just upstream of the palindromic sequences, thereby simultaneously 'transferring' the hairpin structure to the progeny strand as a template, and providing a new 3' hydroxyl opposite it to prime what is essentially repair synthesis to regenerate the terminal palindrome on the parental strand. A necessary consequence of this model is that any imperfections in the terminal palindrome sequence will not be copied exactly from progeny to parent, but will be regenerated on the parent as the inverted complement of the original sequence (see Figure 2).

This model can be applied directly to members of the *Dependovirus* subfamily whose termini contain both direct and inverted repeat sequences, and are identical at both ends. Extension from the genomic 3' hydroxyl allows for copying back all the way to the 5' terminus of the molecule by displacement of the 5' hairpin to make a double-stranded monomer replicative form (mRF) in which the 3' end is in a loop and the 5' end is

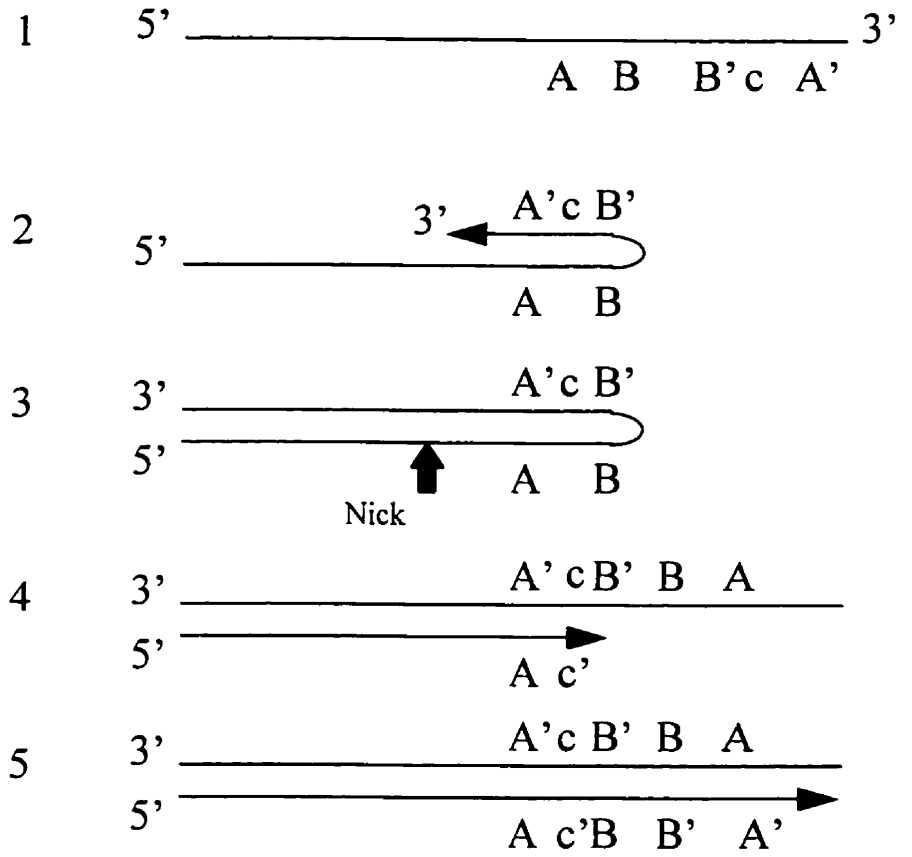


Figure 2: Mechanism of Hairpin Transfer

Presence of palindromic sequences (A and A', B and B') at the 3' terminus [1] allows for hairpin formation and replication back along parental strand [2]. Site-specific nicking (shaded arrow,[3]) transfers terminal hairpin onto new strand and provides a template for repair synthesis [4]. End product has completely replicated the terminus, but imperfection within the palindrome {c} is regenerated on parental strand as its complement {c'} [5].

extended. The process of hairpin transfer is then invoked, leading to a linear double-stranded form of the incoming genome with both termini extended. It is then postulated that the termini dissociate from their linear *trans* double-stranded conformation and reform in the *cis* double-stranded (hairpin) conformers; each strand is thus presenting a suitable 3' primer from which strand extension can remake the mRF form while simultaneously displacing the other strand (see Figure 3). It is suggested this process is concurrent with viral packaging, with one of the two strands not extending but rather being sequestered by the capsid (possibly through binding of the genome's 3' end by aggregates of VP1 noted above) from further rounds of replication. This model makes two predictions: firstly, as hairpin transfer occurs equally at both termini, any imperfections in terminal palindromes will result in two populations for each terminus, each being the complement of the other; and secondly, positive and negative stranded progeny are produced in an equal ratio. Support for this model comes from the observations that in AAV the terminal sequences do in fact exist as an equal ratio of two complementary forms, and the virus encapsidates both positive and negative strands equally. The packaging-driven generation of single-stranded genomes is also supported by evidence from temperature-sensitive mutants of H1, where capsid instability at elevated temperatures does not interfere with the generation of double-stranded intermediates but does limit the production of single-stranded forms[69], and from the studies done on MVM minigenomes in this lab where double-stranded but not single-stranded viral DNAs are observed in the absence of capsid proteins[70].

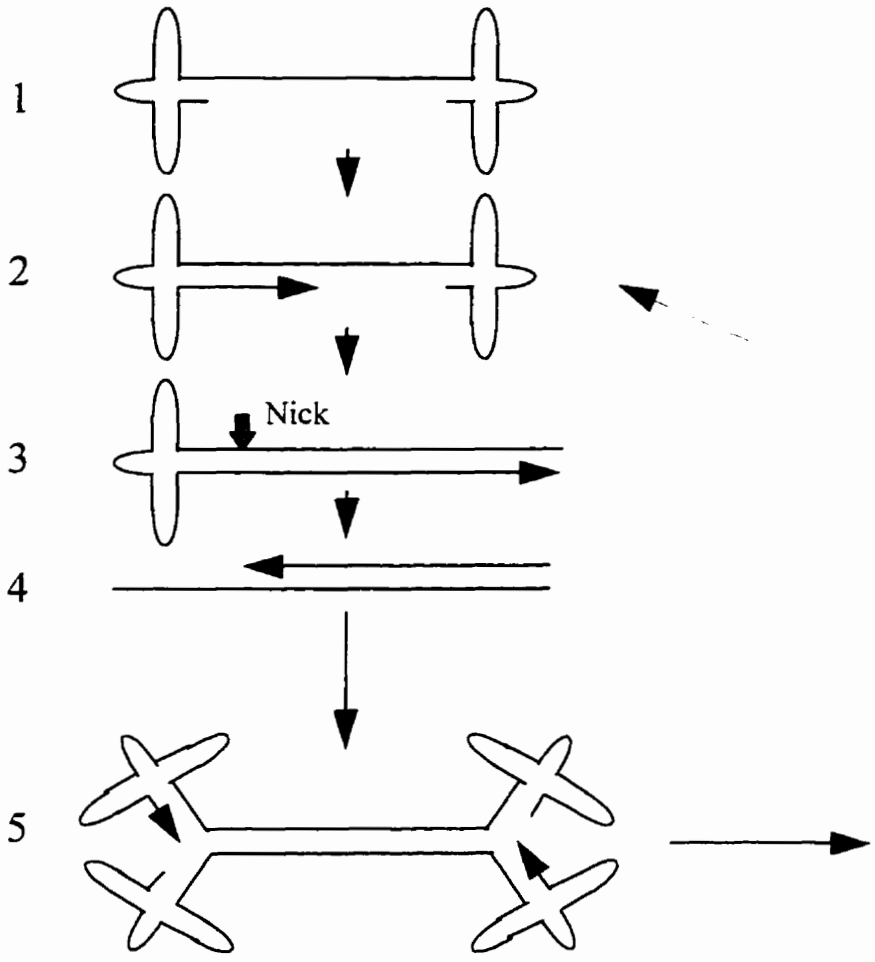


Figure 3: Dependovirus replication

Incoming genome [1] is extended from free 3' hydroxyl [2] to generate mRF [3]. Hairpin transfer allows generation of dimeric form with extended termini [4]. Rearrangement of terminal palindromes into hairpins regenerates [2], which may either be encapsidated as progeny virus or be used to amplify replication intermediates [5].

Studies have been conducted in an effort to determine the minimal *cis*-acting sequences required for origin function in the *Dependovirus* AAV-2 and indicate a requirement for sequences derived from the extreme viral termini only[71,72]. Three regions, referred to as 'A', 'B', and 'C', are all within the terminal palindromes and are essential for genome replication *in vitro*[73]. A fourth region, 'D', lies immediately inboard of the terminal palindrome sequences and contains the terminal resolution site (TRS, the site at which the viral Rep protein nicks the hairpin to commence the process of hairpin transfer); and is therefore necessary for viral replication. That no other functionality lies within the 'D' region is strongly suggested by the further observation that linear duplex vector sequences ending in AAV terminal sequences up to but not including the 'D' region are capable of replication[73]. In light of this, it would seem that there is little or no similarity between the roles of the AAV 'D' region and the MVM IRS which is the subject of the current study.

This relatively simple replication model is not however consistent with the data observed for the replication of the autonomous parvoviruses including MVM. As the *Parvovirinae* are far from being unanimous in the following traits, they will be described for MVM with major exceptions being noted. Firstly, only negative-strand genomes are packaged to an appreciable extent (BPV packages roughly 20-30% positive-sense genomes, and LuIII packages equal numbers of positive and negative-strand genomes in some hosts)[74,75]. Secondly, while both hairpin palindromes are imperfect, only that at the 5' end of mature genomes is found to exist in two complementary forms (known as 'flip' and 'flop'), while the 3' termini exists in a single form[15,76]. (Again, BPV is unusual in

this regard, having up to 10% of the 3' ends existing in the complementary form). Thirdly, analysis of viral DNAs in infected cells reveals the presence of large numbers of viral concatamers at dimeric and higher multiples of the genomic length. Any replication model for the autonomous parvoviruses must take these observations into account.

Although likely incorrect in some aspects, the best supported model for MVM replication at present is the Modified Rolling Hairpin (MRH) model[5,6] (see Figure 4). This model varies from that for AAV by requiring the mRF form to have the 5' hairpin dissociate from its extended form and re-form as two hairpins without the nicking at the hairpinned 3' termini, thus allowing for the polymerase to extend from a 3' hydroxyl using first the strand it just created and then the original parental strand as templates. The net result is a dimeric Replicative Form (dRF) which is a concatamer of two viral genomes (parental and progeny) each across from its complement; there are two 5' terminal palindrome sequences, one in an extended form and one in a hairpinned form, and two 3' terminal palindrome sequences, both in an extended form in the middle of the molecule. This molecule may have its extended 5' hairpin undergo another cycle of dissociation, hairpinning, and extension to form longer concatamers, or may undergo resolution as outlined below. As resolution of the longer concatamers is exactly analogous to that for the dRF species it will not be discussed further.

Resolution occurs through the asymmetric hairpin transfer of the central extended 3' palindromic sequences (the 'dimer bridge'). While the exact mechanism of this process is currently unclear, staggered site-specific nicking by NS-1 occurs on both strands with one end being selectively ligated across to reform a closed 3' hairpin. The

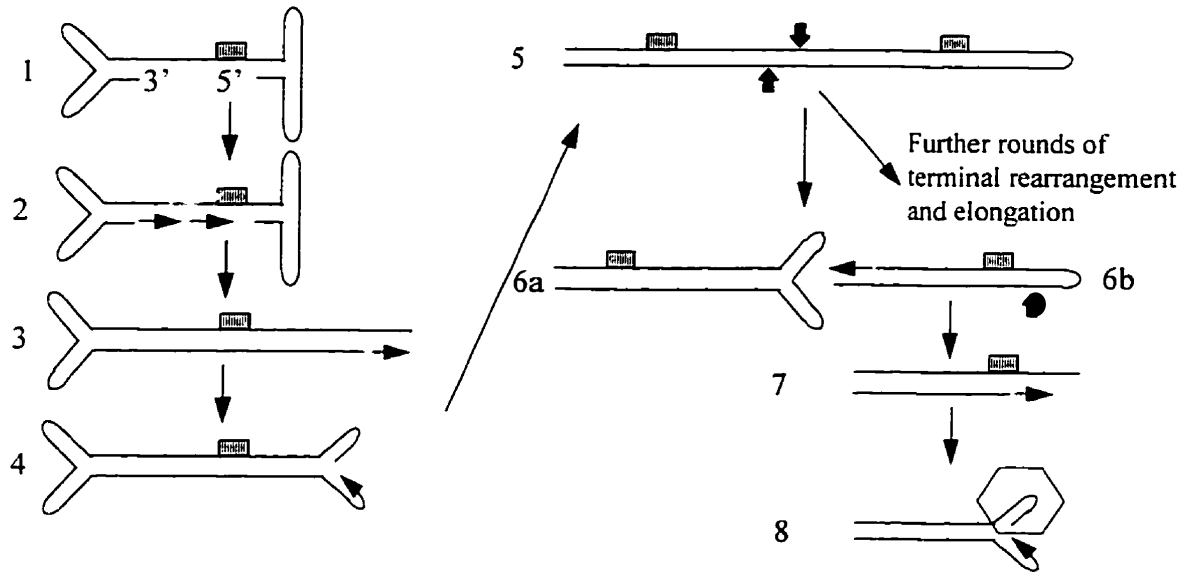


Figure 4: MVM DNA replication cycle

MVM replication cycle. Monomeric genome [1] is extended from base-paired 3' terminus [2], displacing and extending the 5' hairpin structure (mRF, [3]). The extended hairpin terminus undergoes rearrangement to self-associate [4], affording a primer for further extension to a dimer-length molecule (dRF, [5]. This may then undergo further rounds of terminal rearrangement as in [4] and extension to higher-order concatamers, or asymmetric resolution by site-specific nicking by NS-1 (shaded arrows) and ligation to generate [6a] and [6b]. [6a] is identical to [3] and may be used to continue the cycle, while [6b] serves as a template to generate viral progeny by site-specific nicking by NS-1 (shaded arrow) and hairpin transfer [7]. A terminal rearrangement [8] identical to that in [4] allows for the lower strand to regenerate [6b], simultaneously displacing the upper strand for encapsidation as a mature genome [1].

Approximate location of the IRS is indicated by a shaded box above one strand of forms [1] through [7].

other end is filled in to generate an extended 3' end. Hairpin transfer is simultaneously invoked for the 5' termini in the hairpin configuration. The net yield of these molecular gymnastics is production of one mRF molecule identical to that arising from the original strand extension of the infecting genome which may thus go back and repeat this cycle, and a linear double-stranded genome with both termini extended. Generation of single-stranded progeny is achieved from this linear form, by having the right (genomic 5' sequence) end of the molecule reform in independent hairpins, followed by extension from the 3' hydroxyl of the complementary strand and hairpin transfer; this re-creates the linear molecule while simultaneously displacing a mature single-stranded genome (again, concurrent packaging is suggested). The linear molecule is capable of going through this process multiple times, each time generating a mature genome with the same 3' terminal sequence but with alternating complementary 5' termini. It should be noted that this model requires there to be imperfections within the 3' palindromic sequences such that the resolution of the central dimer bridge can be asymmetric, and similarly requires the two ends of the genome to be nonidentical.

One theoretical concern with both of these models is the requirement for the extended termini to spontaneously melt from stable double-stranded conformations in order to reform as hairpins that act as primers for further extension, a transition which must have a very appreciable activation energy barrier to overcome[77]. While models for an enzymatically assisted transition can be put forth, a certain amount of credulity is required for their serious adoption. A model put forth by Rhode *et. al.*[78] overcomes this by postulating an origin of replication somewhat inboard of the 5' hairpin. This model is

essentially identical to the Modified Rolling Hairpin model except that the origin is postulated to function both on the mRF and on the linear double-stranded molecule with both ends extended arising from resolution of the dimer bridge. Extension of nascent strands outward from this origin displaces the template's complementary strand, thereby eventually releasing one strand of the palindromic terminal sequences into a single-stranded form from which hairpin formation is thermodynamically plausable as a spontaneous event. Similarly to the MRH model, this requires imperfections within the palindromic sequences and non-identical termini. It also predicts that Okazaki fragments should be detectable at the viral origin of DNA replication.

In both cases the models call for the existence of a site-specific endonuclease to create the nicks required for hairpin transfer; the fact that this endonuclease must recognize viral sequences, coupled with the observation that the sequences of the parvoviruses do not show conservation at their termini, strongly suggests this enzyme is virally encoded. That this is in fact the case, and that the enzyme functions are provided by the NS-1 polypeptide, is now clear[27].

Previous studies and goals of the project

Previous work from this lab had demonstrated that an artificial defective interfering (DI) genome of MVM, referred to as a 'minigenome', was replication competent when supplied with the viral NS-1 protein *in trans*. Deletional analysis of this minigenome indicated the presence of *cis*-acting sequences inboard of the 5' palindrome which were required for DNA replication in COS-7 cells (the Internal Replication Sequence or IRS)[70]. Bandshift studies employing a partially purified HeLa cell extract

and restriction fragments from the region of the IRS indicated the formation of several specific complexes in the region, and further fractionation of the nuclear extract allowed for the footprinting of one of these complexes. It was shown to cover a bipartite region between viral nucleotides 4589-4646[79].

The goals of this project were to further refine on these studies by examining the IRS sequences in finer detail to determine if specific short sequence elements were important in replication competence, and to attempt to localize all host-factor binding sites within the region. Furthermore, it was hoped that the identity of these host factor(s) and their role(s), if any, in viral replication could be elucidated.

Materials and Methods

General methods

General protocols for routine cloning procedures, growth of bacteria, composition of bacterial media, and common buffers and solutions have been previously described[80].

Sources of enzymes etc.

Restriction and modification enzymes were purchased from New England Biolabs or GIBCO/BRL and used according to the supplier's protocols.

Bacterial culture

Cell lines

Plasmids were propagated and cloning procedures performed exclusively in SURE cells (Stratagene), which were chosen for their recombination-deficient phenotype. This strain was previously found in our lab to be capable of stably maintaining plasmids containing the extended palindromic sequences present in clones containing the MVM terminal sequences.

Growth

Bacteria were routinely grown on TY solid media (10g tryptone, 5g yeast extract, 5g NaCl per litre, 2% agar) or 2xTY liquid media (16g tryptone, 10g yeast extract, 5g NaCl per liter). Media was supplemented with ampicillin (Penbritin, Wyeth-Ayerst) to 200 μg/ml when required for the selection of plasmids.

Transformation and electroporation

Routine transformations were done with cells made chemically competent by a modified calcium chloride treatment. Briefly, the bacteria were grown to mid log phase in 2xTY media containing 10mM $MgSO_4$, chilled, pelleted by centrifugation, resuspended in sterile Transformation Buffer I (30 mM KOAc, 50mM $MnCl_2$, 100 mM KCl, 10 mM $CaCl_2$, 15% glycerol), recentrifuged, resuspended in Transformation Buffer II (10mM Na-MOPS pH 7.0, 75 mM $CaCl_2$, 10 mM KCl, 15% glycerol), aliquoted, and flash-frozen on dry ice before storage at -70°C. Transformations were done by mixing plasmid or ligation mixture with 50-100 µl thawed competent cells, incubating on ice 30 minutes, heat shocking at 37°C for one minute with gentle agitation, chilling on ice for one minute, followed by addition to 2 ml 2xTY media for growth at 37°C for one to two hours prior to plating on selective media.

For transformation of extremely small quantities of plasmid DNA such as that obtained from yeast minipreps, electroporation was employed using a Bio-Rad Gene Pulser. Electrocompetent cells were prepared and used in accordance with the manufacturer's supplied protocols.

DNA recovery and Hirt extraction

Small scale 'miniprep' isolation of plasmid DNAs was performed with a rapid modified alkaline lysis protocol[81]. Mid- and large-scale plasmid isolations were performed by alkaline lysis followed by polyethylene glycol precipitation.

Hirt extraction[82] was performed on 10cm dishes of adherent cells by removal of culture media and scraping of cells into 1 ml of Hirt lysis buffer (100 mM NaCl, 10 mM

EDTA, 1% SDS). NaCl was added to 1.1 M final concentration from a 5 M stock, and lysates incubated at 4°C for 12-16 hours prior to centrifugation at 12 000 × *g*, 4°C, for 30 minutes. The resulting supernatant was digested with Proteinase K (50 μg/ml) at 37°C for 2 hours, phenol-chloroform extracted, and DNAs precipitated from the aqueous phase by the addition of an equal volume of isopropyl alcohol and centrifugation. Recovered DNAs were resuspended in 60 μl of sterile water. For 60mm tissue culture dishes, all reagent volumes used were 40% that employed for 10cm dishes.

Cell culture

Cell lines

COS-7 cells[83] were grown in DMEM supplemented with 10% fetal calf serum and 20μg/ml Gentamicin. A9 oub^{r}11 variant mouse L cells (LA9 cells)[84] were grown in DMEM supplemented with 5% fetal calf serum and 20μg/ml Gentamicin.

Maintenance

Cells were maintained in culture by growth in complete media in 10 cm tissue culture dishes at 37°C in 5% CO_2. When cells reached ~90% confluency they were passaged by trypsinization and dilution 1/10 in fresh media (approximately once every 3 days). Cells were maintained in fresh culture for the minimal amount of time needed, with fresh stocks from frozen culture being employed if passaging of more than a few weeks was anticipated between experiments.

Transfection

Confluent cells were passaged at 1/5 dilution 24 hours prior to transfection by the DEAE-Dextran method. Cell monolayers were incubated with 5 μg of each plasmid to be transfected in 2.5 ml of DMEM (without serum) containing 400 μg DEAE-Dextran for 8 hours. Dishes were gently rocked every few hours to ensure even distribution of the transfection media. Transfection media was then aspirated and cells shocked with 10% DMSO in PBS for 5 minutes (COS-7 cells) or 2 minutes (LA9 cells). The DMSO shock was aspirated, cells were rinsed twice in PBS, and recovered in complete media for growth. (Volumes given are for 10 cm culture dishes; for 60 mm dishes 40% of these volumes were used).

Viral Isolation

Virus was grown and isolated in a modification of the method of Tattersall *et. al.*[14]; briefly, confluent monolayers of LA9 cells were split 1/5 in complete media to 10 cm dishes and allowed to grow for 24 hours prior to low-efficiency transfection with 1 μg SmaI linearized infectious clone and growth for 72 hours following DMSO shock. Monolayers were then scraped and collected, centrifuged at 300 × *g* for 5 minutes, pellet washed with 5 ml TNE (0.15 M NaCl, 50 mM Tris-HCl, 0.5 mM EDTA, pH 7.5), centrifuged at 300 × *g* for 5 minutes, pellet resuspended in TE (50 mM Tris-HCl, 0.5 mM EDTA, pH 8.7) containing 1 mM PMSF and cells lysed with a Tenbroek homogenizer (15 strokes) followed by centrifugation at 300 × *g* for 5 minutes at room temperature. The supernatant was then centrifuged at 12 000 × *g* for 30 minutes, and the supernatant made 25 mM in $CaCl_2$. Virus was allowed to precipitate on ice for 30 minutes, then

centrifuged at 12 000 × *g* for 10 minutes. The pellet was resuspended by vortexing in 100 µl TE (pH 8.7) containing 20 mM EDTA and centrifuged 12 000 × *g* for 10 minutes. The supernatant, equivalent to Fraction 3 of Tattersall *et. al.*[14], was used as viral stock for plaque assays.

Plaque Assay

LA9 cells were trypsinized, counted in a haemocytometer, and resuspended at 1.25 $\times 10^4$ cells/ml in complete media. Four-ml aliquots (5×10^4 cells) were seeded per 60 mm dish and were allowed to grow for 24 hours prior to having the media removed and viral dilutions in 300 µl of DMEM/1% FCS/10 mM HEPES pH 7.3 applied. Virus were allowed to adsorb at 37°C for one hour, with the dish being gently rocked every 30 minutes to ensure even distribution. Viral dilutions were removed, and the monolayers were recovered in DMEM/ 5% FCS/ 20 µg/ml Gentamicin/ 1% tryptose phosphate/ 0.75% agarose equilibrated to 45°C. Cells were left to grow for 5 days before being fixed in 10% formaldehyde and stained with 0.3% Methylene Blue.

Cloning techniques

Digestions and Ligations

Digestions and ligations were performed in the supplier's buffer system. Routine digestions were performed at the recommended temperatures for 1-3 hours; ligations were carried out at 16°C for 4 hours (common cohesive-end simple ligations) to overnight (for ligations with more than two fragments or involving blunt ends).

Oligonucleotides

Synthetic oligonucleotides employed in this study are presented in Table 2. All oligonucleotides were prepared on ABI 371 or 374 DNA synthesizers.

Plasmid constructs

Details of the construction of pCA4.0, pPTLR, and pCMVNS-1 have been previously published[70]. Construct pJB2.0 was made by cloning a PCR product corresponding to viral nt. 4484 through 4777 between EcoRI and BamHI tags (obtained using primers MVMA5 and MVMB3 on pCA4.0 template) between the EcoRI and BamHI sites of pUC19. Oligonucleotide MVMA5 simultaneously introduced a single T→C silent nucleotide substitution at viral nt. 4486 to introduce a unique BstEII site. Minigenome pJBLR4 was constructed from pPTLR by introduction of a single T→C silent nucleotide substitution at viral nt. 4486 to introduce a unique BstEII site inboard of the IRS. The minigenome used in replication studies, pJBLR4.3, was derived from pJBLR4 by digesting, filling in, and religating the BamHI linkers flanking the viral minigenome to convert them to ClaI sites and substitution of the viral sequence TAGGTTAAT at nt. 4780 with CCCTAGGC, thereby introducing a unique BamHI site (see Figure 5). Clone pJB3.0 was constructed by deleting the SspI to SmaI fragment of pUC19 followed by inserting the XbaI to SphI fragment of pCA4.0 into the same sites in the polylinker of the deleted vector. Its derivative pJB3.0S1 was constructed by replacing the BstEII to SspI fragment of pJB3.0 with the same fragment from pJB2.0S1 (see below), and pCA4.0S1 was constructed by replacing the XbaI to SphI fragment of pCA4.0 with that from pJB3.0S1. Clone pJBR1

Oligo name	Oligonucleotide sequence
MVMA5	GGAATTCCGGTTACCAACTGCTACTGGAA
MVMB3	CGGGATCCCGAACCACCCTTCCACCCTTTTA
S20E	CAGATCTGTTATAACAAGACC
S21B	CAGATGTGTAAGTACCATATTA
S22B	CAGATCTGATATGAAGTACAG
S23B	CAGATGTGAAAGAAAAAGCATG
S25-2	ATTATATTTCTCAGATCTGTCTTTATTAGTCTTAATAATATATG
S26C	TATGTTGTATCTTTATTAGTCTTAATAATATATG
S26D	CAGATCGATCTGTACATATAGATTTAAGAAATAG
S822C	AAGCATGGTTAGTTAG
S822D	pTTTCTTTCTGTACTTC
T1C	CCTTTATTAGTCTTAATAATATATG
C2A	TATTTATTAGTCTTAATAATATATG
T3G	TCGTTATTAGTCTTAATAATATATG
T4A	TCTATATTAGTCTTAATAATATATG
A6C	TCTTTCTTAGTCTTAATAATATATG
T8G	TCTTTATGAGTCTTAATAATATATG
S1P	pTACAACATAGAAATATAATATTAC
pGADforward	TACCACTACAATGGATG
Vecpro	CGCGTTTCGGTGATGAC
JAVA	TTTTGGTCCTTAACATCAAG
Forward seq.	GTAAAACGACGGCCAGT
Reverse seq.	AACAGCTATGACCATG
S1BE	AATTCCTAATAAAGATACTAATAAAGATACTAATAAAGATACTAATAAAGATA
S1BEC	TATCTTTATTAG TATCTTTATTAG TATCTTTATTAG TATCTTTATTAGG

Table 2: Oligonucleotides Used

A listing of oligonucleotides referred to in text. Sequences are reported 5' to 3'; a leading 'p' indicates a 5' phosphorylation.

was constructed from pJBLR4.3 by digestion with NheI and XbaI to remove sequences derived from the viral 3' genomic region, gel isolation of the large fragment retaining vector and viral genomic 5' sequences, and religation. Linker-scanning mutant derivatives of this were made from the corresponding pJBLR4.3 clone in the same fashion.

Construction of the Linker-scanning library

The linker-scanning library constructed and employed in this study is shown in Table 3. Digestion of pJB2.0 with either EcoRI or BamHI followed by limited digestion with Bal-31 nuclease, blunt-ending with Klenow fragment in the presence of dNTPs, and religation in the presence of excess synthetic double-stranded BglII linker (pCAGATCTG) created pools of clones with partial deletions of the cloned viral sequences, known as ΔEco or ΔBam clones respectively. Clones were sequenced (Sequenase 2.0, USB) to determine deletion endpoints, and ΔEco and ΔBam clones which matched up to re-create the original cloned viral sequence with a central gap of 8 nt filled by the BglII linker were joined together at the linker directly. Clones which were found to together reconstitute the original cloned viral sequence with a gap of exactly 12 base pairs were treated by digestion with BglII, the recessed 3' end filled in by Klenow fragment in the presence of all four dNTPs, and ligated together to recreate the original viral sequence with an internal 12 bases replaced with a ClaI linker (CAGATCGATCTG) (see Figure 6). Clones produced in this fashion are referred to as

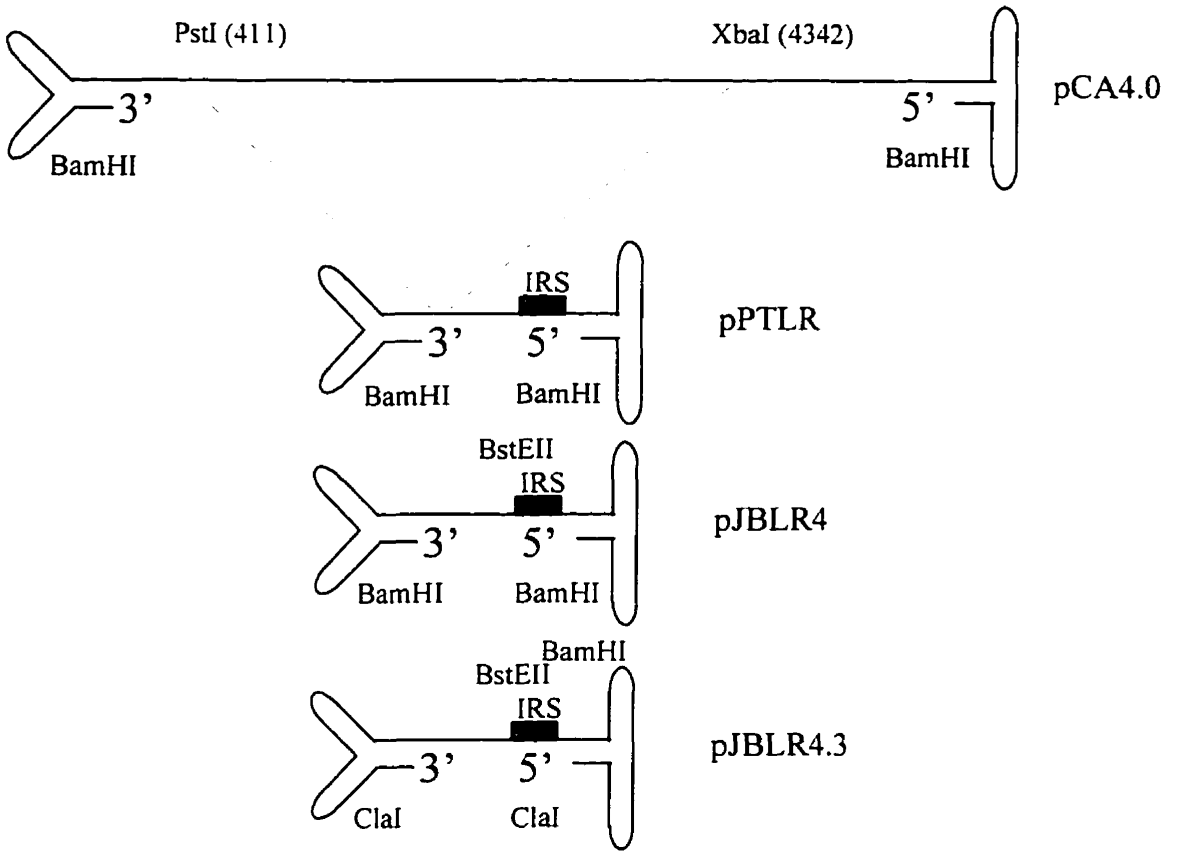


Figure 5: Schematic Representation of Minigenomes

Schematic diagram of the relationship between full-length viral clone pCA4.0 and the minigenomes discussed in this study. Relevant restriction sites are noted, as is the location of the IRS. pPTLR was the minigenome employed in previous studies; pJBLR4.3 was the primary minigenome employed in this study; and pJBLR4 was a construction intermediate between the two employed in one experiment (see Figure 7).

Note that all four of these constructs are in fact double-stranded plasmid clones; for simplicity only one strand of the virally-derived sequence element from each clone is represented here.

pJB2.0S(x) where (x) is a number specific to the particular linker-scanning mutation. In some cases where a matching set of ΔEco and ΔBam clones could not be found, PCR was used to create a suitable match to an existing deletion; specifically clones S20, S21, S22, and S23 were created this way with the matching clone being made by PCR between primer pairs S20E/MVMB3, S21B/MVMA5, S22B/MVMA5, and S23B/MVMA5 respectively, using pJB2.0 as template. These products were cloned into pUC19 vector linearized with SmaI, and then digested with BglII (followed by fill-in with Klenow fragment for S23) and either EcoRI (S21, S22, S23) or BamHI (S20) and used as were the native ΔEco and ΔBam clones.

Clones S25 and S26 were made totally by PCR. Mutation S25 was made by PCR using oligos S25-2 (phosphorylated with T4 polynucleotide kinase) and MVMA5 on pJB2.0 template; the product was ligated to the small SspI fragment of pJB2.0, and the resulting mixture of products used as a template for PCR with primers MVMA5 and MVMB3. The resulting product was digested with EcoRI and Bam HI and cloned into pJBLR4.3 vector digested with the same enzymes, directly yielding pJBLR4.3S25. Mutation S26 was constructed by PCR using primer pairs S26C/MVMA5 and S26D/MVMB3 on pJB2.0 template, gel-isolating the products, ligating them together, and then using the ligation mix as a template for a PCR with primers MVMA5 and MVMB3. The resulting product was digested with EcoRI and Bam HI and cloned into pJBLR4.3 vector digested with the same enzymes, directly yielding pJBLR4.3S26.

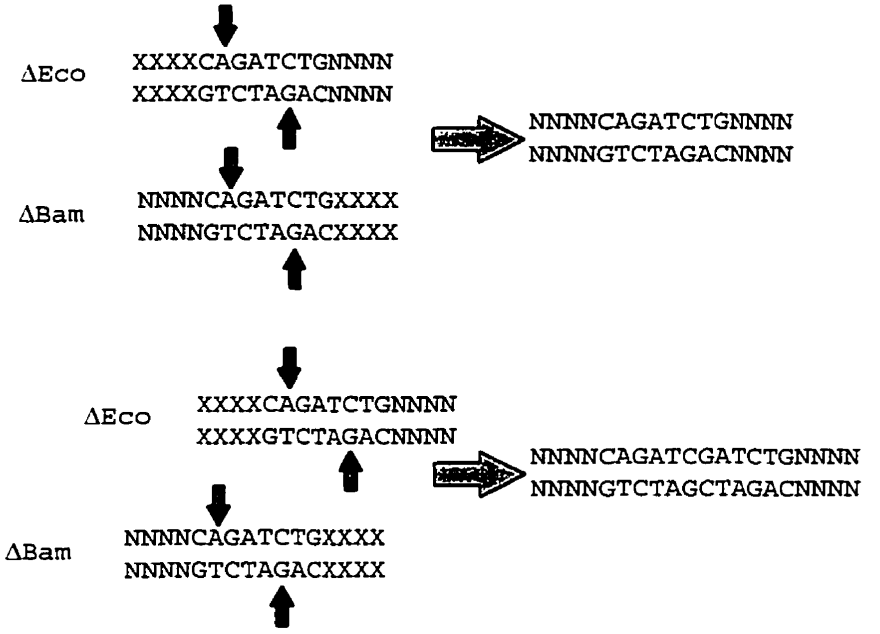


Figure 6: Linker-scanner mutant construction

Construction of linker scanning mutants was carried out by digesting pJB2.0 clones deleted from BamHI or EcoRI sites and having BglII linkers either in register (directly ligated to generate 8 bp. scanner mutations; upper part of figure), or four bases out of register (digested, filled in with Klenow fragment in the presence of all four dNTPs, and ligated to generate a 12 bp. scanner mutation; lower part of figure). 'N' represents a nucleotide of cloned viral sequence, and 'X' represents a nucleotide of vector; shaded arrows indicate the cut site within the BglII linker.

Dual-mutant S8/22 was made by using pBLR4.3S8 as a template for PCR with primers MVMA5 and S822C, pJBLR4.3S22 as a template for PCR with primers S822D and MVMB3 and ligation of the two products together. The ligation product was used as template for PCR of with MVMA5 and MVMB3, and the resulting product digested with BamHI and BstEII prior to cloning into pJBLR4.3 vector digested with the same enzymes. Scanner mutant clones in the context of pJB2.0 were introduced into the pJBLR4.3 context by replacing the equivalent BstEII to BamHI fragment of pJBLR4.3 with that of pJB2.0S(x) to yield pJBLR4.3S(x). All clones were sequenced to ensure proper construction.

Construction of point mutants

Mutations T1C, C2A, T3G, T4A, A6C, and T8G were made by PCR with the respective oligonucleotides and MVMA5 on pJB2.0 template. Gel-isolated products were ligated to isolated PCR product obtained with primers S1P and MVMB3 on pJB2.0 template. The resulting ligation products were used as template for PCR with primers MVMA5 and MVMB3, and the product gel-isolated, digested with BstEII and BamHI, and cloned into pJBLR4.3 digested with the same enzymes to yield the respective point-mutant minigenomes.

Gel isolations

DNA fragments were routinely isolated from agarose gels after electrophoresis by visualization of EtBr stained bands under longwave ultraviolet light, excision of desired band from the gel, and centrifugation of the gel slice through a mesh of silanized glass wool in a microfuge tube for 2 minutes at 12 000 × *g*. The aqueous phase collected

below the wool was used directly in cloning or precipitated as needed to concentrate the fragment.

DNA End-Labelling

Oligonucleotides and double-stranded DNA fragments were routinely 5' end-labelled with ^{32}P by incubating 50-200 ng of DNA in a 20 µl reaction containing 1× T4 Polynucleotide Kinase buffer, 5 µl γ-^{32}P rATP (3000 Ci/mMol), and 10 units T4 Polynucleotide Kinase at 37°C for 30 minutes. For oligonucleotides to be used for hybridization probes, the reaction was then stopped by incubation at 65°C for 10 minutes and the probe used directly; for double-stranded fragments to be employed as bandshift probes, the reaction was precipitated by addition of 2 µl 3.5 M NaOAc pH 5.4 and 60 µl EtOH, incubation on ice for 20 minutes, centrifugation, removal of supernatant, and washing of the pellet with 200 µl ice-cold 70% EtOH three times prior to being resuspended at 1 ng/µl final concentration for use.

Sequencing

DNA sequencing was carried out by a Sanger dideoxy protocol, using the Sequenase 2.0 kit (United States Biochemicals/Amersham) and α-^{32}P dATP as label. Templates, generally consisting of 10 µl of miniprep DNA or 2 µg of large-scale purified DNA, were denatured in 0.2M NaOH, 0.2mM EDTA for 30 minutes prior to precipitation and annealing in sequencing buffer with 1 pmol of the desired primer.

Polymerase Chain Reaction

PCR was routinely carried out using Vent polymerase (New England Biolabs) in its supplied buffer system, supplemented with BSA to 100μg/ml. Standard reactions were done in 50 μl volumes and contained 500 μM dNTPs and 25 pmol of each primer. Thermocycling was done with an MJ Research Minicycler. Conditions were individually optimized for each reaction.

DNA quantitation

Routine DNA quantitations of plasmids and oligonucleotides were performed by spectrophotometry.

Replication Assay

Sets of 10cm dishes of confluent COS-7 monolayers were rinsed in phosphate-buffered saline, trypsinized, and evenly resuspended in a pooled volume of complete media at 1/5 dilution. Four-ml aliquots of the resuspended cells were seeded onto 60mm tissue culture dishes and allowed to grow for 24 hours prior to transfection. Cells were transfected by a standard DEAE-Dextran protocol using 2 μg of each plasmid; minigenomic clones were linearized by digestion with EcoRI prior to transfection. Complete media was then added to the cells and they were allowed to grow for 48 hours. Low molecular weight DNAs were recovered by Hirt extraction[82] and used for analysis by slot-blotting and hybridization.

DNAs transfected in each assay consisted of linearized pJBLR4.3 alone (two samples), linearized pJBLR4.3 with pCMVNS-1 (two samples), and each of the linearized mutants being assayed with pCMVNS-1. Replication was quantitated by slot-

blotting 5 μl of Hirt extract from each sample onto Hybond-N+ along with a standard range (concentration determined by spectrophotometry) of both linearized pJBLR4.3 and pCMVNS-1 (the latter included as a control for non-specific hybridization). Blots were prehybridized in 6×SSPE, 1%SDS, 5× Denhardt's Reagent, and 10 μg/ml fragmented, denatured salmon sperm DNA at 42°C for 4 hours, followed by hybridization in 6×SSPE, 1% SDS with 5 ng of 5' ^{32}P end-labeled Vecpro oligonucleotide at 39°C for 10-15 hours. Blots were washed twice at room temperature in 6×SSPE/1% SDS for 3-5 minutes and once in 2×SSPE at 39°C for 3 minutes prior to visualization by phosphorimaging (Molecular Dynamics PhosphorImager SI; IP Lab Gel Version H (Signal Analytics) was used in image analysis). Results of this exposure provide data on input DNA (*input*) present in each sample as Vecpro hybridizes uniquely to the vector backbone. Blots were then stripped (washed in 6×SSPE, 50% formamide, 65°C, 30 minutes followed by 2×SSPE, 65°C, 30 minutes) and again visualized to verify removal of probe. Prehybridization, hybridization, washing, and exposure were then carried out exactly as before, but with 5' end-labeled oligonucleotide probe JAVA which allows for determination of the amount of viral sequence present in each sample (*viral*) by hybridizing uniquely to sequences present in the minigenome.

Analysis of replication data consisted of preparing a standard curve from the amount of signal obtained from the range of linearized pJBLR4.3 standard applied to each blot, and using this standard curve to determine the quantity of DNA present in each of the samples corresponding to each of the transfections. Replication efficiency (RE) was calculated as $RE = (viral - input) / (input)$. Samples consisting of viral minigenome

alone were taken as having no replication and a Scaled Replication Efficiency (SRE) was calculated for each sample (n) as $SRE(n) = RE(n) - \overline{RE(pJBLR4.3)}$. Samples consisting of wild-type viral minigenome and pCMVNS-1 were taken as having 100% replication, and a Scaled Relative Replication Efficiency (SRRE) was calculated for each sample (n) as $SRRE(n) = SRE(n) / \overline{SRE(pJBLR4.3 + pCMVNS - 1)}$. Individual data sets thus normalized for 0 and 100% replication were pooled for statistical analysis.

Competitive Replication Assay

Competitive replication assays were carried out in the same fashion as were regular replication assays, with the exception that 1 μg of EcoRI-linearized mutant minigenome was cotransfected with an equal amount of EcoRI-linearized pJBLR4.3 and 2 μg of pCMVNS-1.

Assays of each mutant were examined by slot-blotting 5 μl of Hirt extract on Hybond-N+ along with a set of standard dilutions of EcoRI-linearized mutant minigenome and one control sample each of pJBLR4.3 and pCMVNS-1. Blots were prehybridized as were regular replication assays, and hybridized in 6× SSPE/ 1% SDS with 5ng of a 5` ^{32}P end-labelled oligonucleotide capable of specific hybridization to the mutant being assayed (generally, such oligonucleotides were designed to include the scanner mutation present in the mutant minigenome and its flanking sequences). Washing was performed as for regular replication assay slot-blots, and phosphorimaging used to quantitate the amount of mutant minigenome present in each sample (*mutant*). Following this, blots were stripped by the same protocol used in regular replication

assays and examined by phosphorimaging to verify removal of probe. Prehybridization and hybridization of the blot was then again carried out, using 5' ^{32}P labelled JAVA as probe (which hybridizes equally to mutant and wild-type minigenomes present in the sample). After washing, the blot was quantitated by phosphorimaging, with this value corresponding to the sum of wild-type and mutant minigenome present in each sample (*total*).

Replication efficiency of the mutant relative to the wild-type minigenome was calculated directly from this data as (*mutant*)/(*total*). Specificity of the probes employed was verified by a lack of hybridization of either probe to the pCMVNS-1 control sample. hybridization to the pJBLR4.3 control sample by the JAVA probe only, and hybridization to the mutant minigenome standards by both probes.

Competition Bandshift assay

Gel retardation assays were carried out on gel-purified PCR products or restriction fragments prepared from the wild-type or scanner mutant template or clone. Scanner mutants S4, S5, S20, S2, and S7 were analyzed with the BstEII-SspI restriction fragment (viral nt. 4484-4626; BS fragment); mutants S8, S14, S3, S23, S22, S1, and S15 were analyzed with a PCR product obtained with primers S20E and S21B (viral nt. 4524-4668, flanked on either side by CAGATCTG; 2021 fragment); and mutants S16, S17, S21, S12, S13, and S18 were analyzed with a PCR product obtained with primers S1P and MVMB3 (viral nt. 4610-4781; SB fragment). All purified fragments were quantitated by binding of Hoescht 33258 dye and fluorometry on a Hoeffer TKO-100 fluorometer. Assays were performed by incubating 1 ng of end-labeled wild-type fragment in 50mM NaCl, 10mM

Tris, 1 mM DTT, 5% glycerol pH 7.1 containing approximately 3.75 μg LA9 nuclear extract[7] for 30 minutes. Competition by a given mutant was assayed by including 40 ng of unlabelled mutant-derived fragment in the binding reaction and allowing a 10-minute preincubation with the nuclear extract before addition of the labeled wild-type probe. Reactions were run on 4% polyacrylamide/1% glycerol vertical gels in 1/2× TBE/1% glycerol buffer. Gels were dried and the results visualized by phosphorimaging.

Southern Blotting

Routine Southern blotting was performed by standard protocols. DNA species were separated by agarose gel electrophoresis in TAE buffer and transferred to nylon membrane by vacuum transfer, using a Pharmacia Vacu-Gene apparatus in accordance with its supplied protocol. Membranes were dried to fix the DNAs prior to prehybridization for a minimum of one hour in 6 × SSPE/ 1% SDS/ 50% formamide/ 50 μg/ml sheared, denatured salmon sperm DNA at 42°C. Hybridizations were carried out by adding a denatured random-primed probe directly to the membrane and prehybridization buffer and allowed to proceed at 42°C for 12-18 hours. Blots were washed twice in 50 ml of 6 × SSPE/ 1% SDS at 25°C for 15 minutes and then twice more in 50 ml of 2 × SSPE/ 1% SDS at 37°C for 15 minutes. Visualization was by phosphorimaging.

Two-Dimensional Neutral-Alkaline Agarose Electrophoresis

Two-dimensional neutral-alkaline agarose electrophoresis was carried out in a modification of the method of Nawotka and Huberman[85,86]. Hirt extracts were resolved first on the basis of native size in a 1% TAE agarose gel (approximately 8 V/cm for 2-3

hours); an adjacent lane contained a size standard for comparison. The lane containing Hirt extract was then excised from the gel and soaked in 50mM NaOH/1 mM EDTA for 45 minutes to denature the DNA species. This denatured gel slice was placed at the top of a support tray and a 1% alkaline agarose gel (50mM NaOH/1 mM EDTA) was poured below the slice and allowed to set. The composite gel was then run (in alkaline running buffer (50mM NaOH/1 mM EDTA) at 2 V/cm for 7 hours) at right angles to the first electrophoresis, separating individual DNA strands on the basis of size. Following this, DNA species in the gel were transferred to Hybond-N nylon membrane by vacuum transfer and treated as a Southern blot.

Yeast

Cell lines

Plasmid pLacZi/S1 was made by annealing oligonucleotides S1BE and S1BEC, and ligation of this product into pLacZi vector digested with HindIII and SmaI. Plasmid pHis-i/S1 was prepared in an analogous fashion from vector pHIS-i.

Reporter strain YMLHS1 was prepared by sequentially integrating plasmids pHISi-S1 and pLacZi-S1 into the HIS3 and URA3 loci of strain YM4271 (his3-200 ura3-52 trp1-903 tyr1-501 leu2-3,112 ade2-101 lys2-801 MATa ade5:hisG gal4Δ gal80Δ) supplied with the Clontech One-Hybrid kit in accordance with their supplied protocols. Strain YM53Blue was prepared by integrating p53Blue (Clontech) into the Ura3 locus of YM4271 by the supplier's protocol, thereby placing a tandem repeat of 3 copies of the p53 consensus binding site upstream of a β-galactosidase reporter. Both strains were tested for leaky expression, and YM53Blue for proper reporting by transformation with

pGAD53m (Clontech). A range of 3-aminotriazole concentrations was tested for effective inhibition of uninduced His3 expression by the supplied protocol, with 30mM being found optimal. β-galactosidase assays were performed by filter lift of colonies according to the supplied protocol.

One-hybrid screen

One-hybrid genetic screens were performed with the Matchmaker One-Hybrid System (Clontech) using the strains and integrants described above. A Matchmaker cDNA library (Catalog # ML4005AB, Lot# 49011) made from whole 11-day Swiss Webster/ NIH mouse embryos was used in screening. cDNAs for the library were obtained by oligo(dT)/random priming, ligation of EcoRI adapters, and cloning into the EcoRI site of pGAD10. These clones thus use the ADH1 promoter to express a fusion of the SV40 Large-T antigen start codon and nuclear localization signal, followed by the transcriptional activation domain of GAL4, followed by the cDNA. Information supplied with the library indicated it contained 3.0×10^6 independent clones, with an insert size range of 0.5-3.5 kb and an average size of 1.3 kb. The frequency of clone hybridization to a human β-actin probe was 0.16%.

Strain YMLHS1 was used to screen this library by transformation and selection on SD/-Leu/-His/30 mM 3-aminotriazole media according to the supplied protocol. Plasmid isolation from positive clones was by yeast miniprep as described in the supplied protocol, and recovered plasmids were transformed into *E. coli* SURE cells by electroporation. Strain YMLHS1 was used to assay potential positive clones for β-galactosidase expression by the filter-lift protocol as supplied by Clontech.

Results

Minigenome Replication studies

Previous studies on MVM indicated the presence of *cis*-acting sequences at either genomic termini with characteristics of replication origins. Spontaneously occuring defective interfering (DI) particles are only found to be replication competent when the genomic terminal regions are maintained[87] and studies on full-length infectious clones demonstrated that even small deletions within the right-hand hairpin are sufficient to block viral replication[88]. In order to attempt to localize any sequence regions inboard of the terminal palindromes required for viral DNA replication, an artificial DI genome or 'minigenome' was employed for deletional analysis studies. Observations with this system identified the Internal Replication Sequence (IRS) as a region between viral nt. 4489 and 4695, the presence of which in full or in part was required for efficient minigenome replication[70].

In order to extend these observations the current study aimed to examine the IRS with a linker-scanner library in the context of the viral minigenome system. As the minigenome construct employed in previous studies (pPTLR) contained no convenient restriction sites for replacement of the IRS element with homologs carrying linker scanning mutations, a new minigenomic construct was made to simplify cloning. As outlined in more detail in the Materials and Methods section, minigenome pJBLR4.3 contains the same regions as its progenitor pPTLR with two new unique restriction sites introduced to flank the IRS: a BstEII site at viral nt. 4486 and a Bam HI site at viral nt. 4780 (see Figure 5). The mutation to introduce the BstEII site was a single-nucleotide

silent substitution already known to exist in some MVM isolates and was thus expected to have no effect on replication levels. Introduction of the BamHI site required more extensive modification of the viral sequence, with the substitution of 8 nucleotides directly at the junction between two 65 nt. direct-repeat elements present in the viral genome. Conservation of this repeat element in a variable number of copies across several parvoviral species together with evidence indicating an importance of these elements in viral replication[88] suggested that mutations in this region might have an adverse effect on minigenome replication.

In order to assess whether the modifications used to create pJBLR4.3 interfered with minigenome replication competence, the replication efficiency of this minigenome was compared to that of pJBLR4 by the replication assay used in prior studies[70]. As the only difference between pPTLR and pJBLR4 is the presence of the BstEII site in pJBLR4, this was used as the 'wild-type' minigenome control. Each minigenome (linearized with EcoRI) was cotransfected with pCMVNS-1 into COS-7 cells for 72 hours followed by Hirt extraction. The extract was examined for replicated species by DpnI digestion and Southern blotting for minigenomic sequences. Input plasmid species carry a characteristic bacterial methylation pattern $G^{(N6Me)}ATC$ which is efficiently cleaved by DpnI, while replicated DNAs are either hemimethylated or unmethylated and thus resistant to cleavage. As is shown in Figure 7, the presence of DpnI resistant species at mRF and higher concatameric species at approximately equal levels in samples with

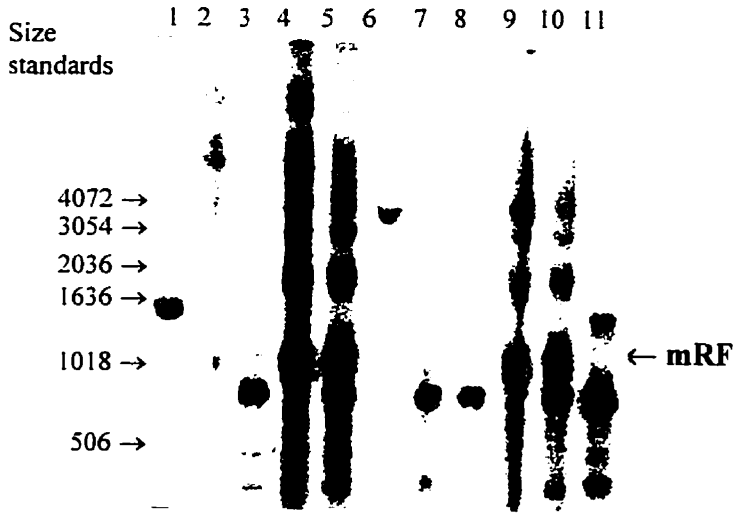


Figure 7: Replication of pJBLR4 vs. pJBLR4.3

EcoRI linearized pJBLR4 or pJBLR4.3 plasmids were transfected into COS-7 cells with or without pCMVNS-1 for 72 hours prior to Hirt extraction. 5 µl samples of Hirt extract from each sample were treated as indicated below and run on a 1% agarose gel at 8 V/cm for 3 hours, followed by vacuum tranfer to Hybond-N membrane. Southern blotting was performed with a random-primed probe made with viral sequences (BamHI insert of pJBLR4) as template and blot visualized by phosphorimaging. Details of protocols are as per Methods and Materials.

Lane #:Plasmids transfected, treatment of Hirt extract
(Lane 1: Marker (BRL 1kb ladder))*
Lane 2: pJBLR4, undigested.
Lane 3: pJBLR4, DpnI digested.
Lane 4: pJBLR4 + pCMVNS-1, undigested.
Lane 5: pJBLR4 + pCMVNS-1, DpnI digested.
Lane 6: pJBLR4.3, undigested.
Lane 7: pJBLR4.3, DpnI digested.
Lane 8: pJBLR4.3, DpnI and BamHI digested.
Lane 9: pJBLR4.3 + pCMVNS-1, undigested.
Lane 10: pJBLR4.3 + pCMVNS-1, DpnI digested.
Lane 11: pJBLR4.3 + pCMVNS-1, DpnI and BamHI digested.

Similar levels of mRF in lanes 5 and 10 indicate that the changes introduced in making pJBLR4.3 from pJBLR4 do not interfere with replication competence.

*Hybridization of the viral minigenomic probe to the 1636 bp DNA size standard used in this and subsequent figures is due to a short region of homology arising from a remnant of pUC19 polylinker present in minigenome clones and thus in the template for random-primed probe.

either pJBLR4 or pJBLR4.3 (compare lanes 5 and 10) indicates that the modifications introduced to create pJBLR4.3 do not have any apparent effect on its replication competence.

Having verified that pJBLR4.3 was a suitable construct for testing the linker scanning mutations, a library of minigenomic constucts based on this vector and containing the linker-scanning mutations shown in Table 3 and Figure 9 was constructed. As the DpnI resistance replication assay used in previous studies[70] was found to be unsuitable for obtaining quantitative data on replication, a new form of assay (see Figure 8, and as outlined in Materials and Methods) was carried out in triplicate on these constructs. The raw data from these assays and the preliminary data analysis (preparation of standard curves, determination of quantity of viral and input sequences for each sample, and calculation of RE, SRE, and SRRE values for all samples in each assay) is presented in Appendix A. The pooled analyzed data from these assays is given in Table 4, and shown graphically in Figure 10.

Results of this set of experiments clearly indicate three distinct regions of the IRS which appear to contribute to replication competence. Mutations S2 and S7, S23, and S1 all cause marked reductions in minigenome replication efficiency by replacing short sequence elements with unrelated sequences. Mutation S7 causes a loss of more than 50% of replication competence, while mutations S2, S23, and S1 all reduce minigenome replication efficiency by more than 75%. In all four cases the result is highly significant statistically, with $p<0.01$ of the observed result being a random

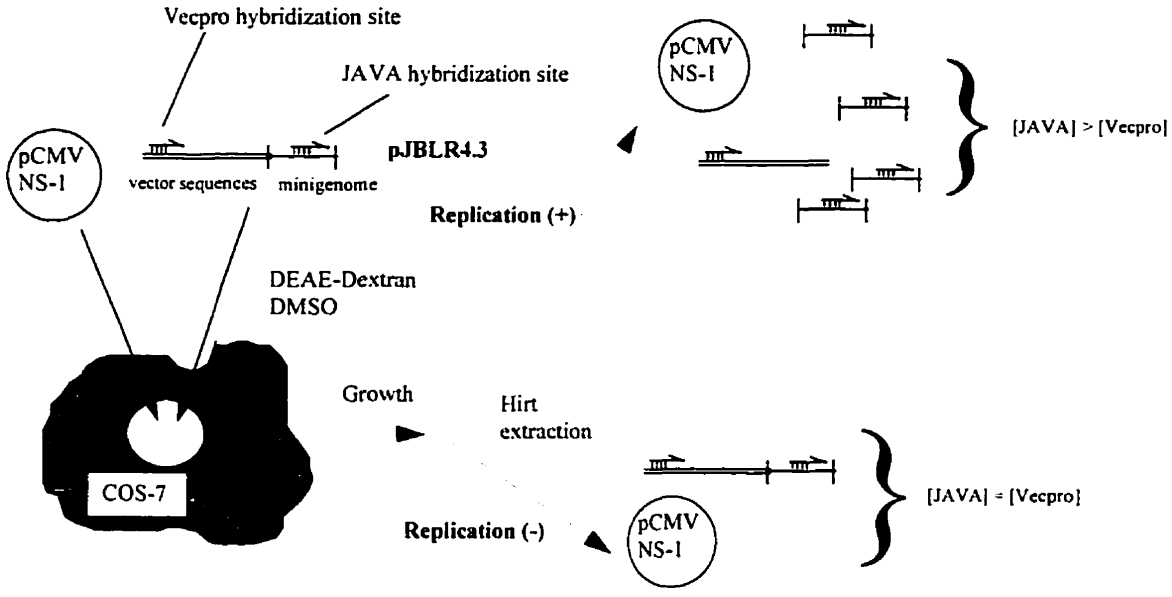


Figure 8: Replication Assay

Simplified schematic representation of the replication assay. EcoRI-linearized minigenome attached to vector sequences is cotransfected into permissive cell line with NS-1 expression vector pCMVNS-1. Following growth (routinely for 48 hours) low molecular weight DNAs are isolated by Hirt extraction. If replication has occurred, nicking by NS-1 will have released minigenome sequences from associated vector sequences and minigenomes will outnumber associated pJBLR4.3 vector molecules; thus more sites for hybridization by JAVA probe will exist than for Vecpro probe. If replication has not occurred, minigenome and associated vector sequences will still be attached and will be in a 1:1 ratio; therefore JAVA and Vecpro probes will have an equal number of hybridization sites. Differential quantitation of the hybridization of these two probes thereby allows for a measurement of replication ability of a transfected minigenome.

For a detailed description of assay technique and mathematical analysis of data derived from this assay, see Materials and Methods section.

Clone	Mutation with flanking sequences
pJBLR4.3S4*	(4484) ggttaccaactgctactgga (4503) ggttaccCAGATCTG--gga
pJBLR4.3S5	(4504) aacatgcagtct (4515) CAGATCGATCTG
pJBLR4.3S20	(4514) tgtgccgc (4522) CAGATCTG
pJBLR4.3S2*	(4523) ttataac- (4529) CAGATCTG
pJBLR4.3S7	(4530) aagacctgttgct (4542) aCAGATCTGtgct
pJBLR4.3S8	(4543) agaaatacttact (4555) CAGATCTGttact
pJBLR4.3S14	(4556) aactaaccatgct (4568) CAGATCGATCTGt
pJBLR4.3S3	(4569) ttttcttt (4576) CAGATCTG
pJBLR4.3S23	(4577) ctgtacttcatat (4589) CAGATCGATCTGt
pJBLR4.3S22	(4590) attattaagact (4601) CAGATCTGgact
pJBLR4.3S1	(4602) aataaaga (4609) CAGATCTG
pJBLR4.3S25	(4610) tacaacata (4618) CAGATCTGa
pJBLR4.3S26	(4619) gaaatataatatt (4631) CAGATCGATCTGt
pJBLR4.3S15	(4632) acatatagatttaag (4646) CAGATCGATCTGaag
pJBLR4.3S16	(4647) aaatagaataat (4658) CAGATCGATCTG
pJBLR4.3S17*	(4656) aatatggtacttag (4669) -CAGATCGATCTG-
pJBLR4.3S21	(4669) gtaactgtta (4678) CAGATCTGta
pJBLR4.3S12	(4679) aaaataatagaacc (4692) CAGATCGAagaacc
pJBLR4.3S13	(4693) tttggaat (4700) CAGATCTG
pJBLR4.3S18	(4699) ataacaagatag (4710) CAGATCGATCTG

Table 3: Scanner Mutation Sequences

Twenty linker-scanning mutations are presented in order of sequence position. For each clone, the associated wild-type sequence is presented with the nt positions relative to full-length MVM of the first and last nt shown in brackets (top line). The same sequence region from the mutant is then presented in alignment, with the mutation indicated in upper-case type (bottom line).

Three mutations (*) do not maintain exact spacing; dashes indicate placeholders in sequence to maintain wild-type spacing. The gap indicated in the wild-type sequence associated with mutation S2 indicates that this inserts a single nucleotide relative to the wild-type sequence.

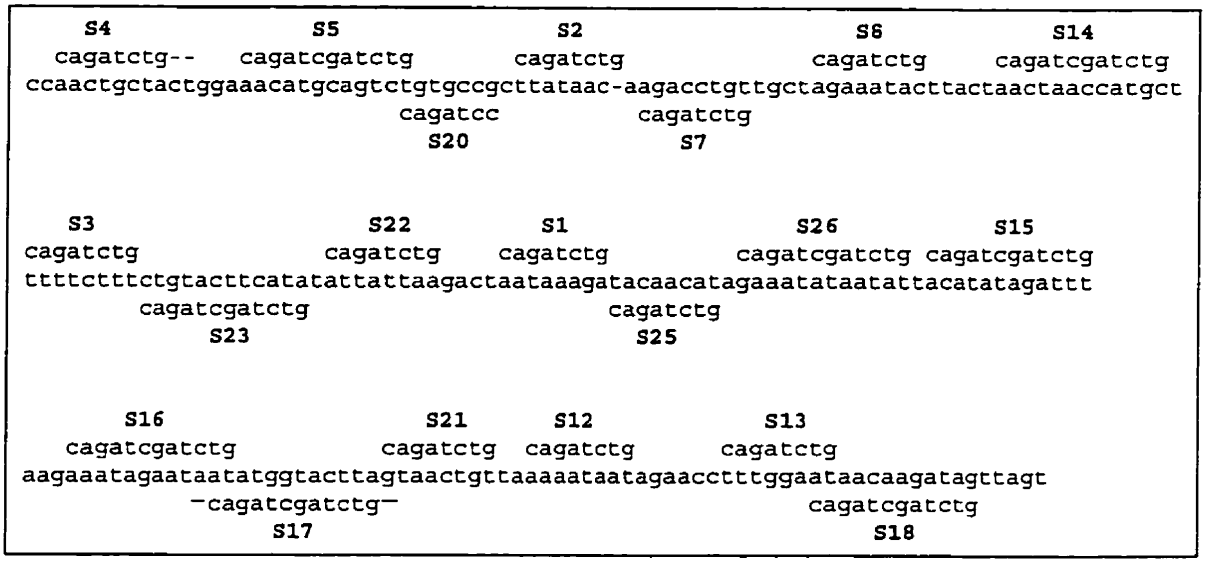


Figure 9: Overview of Scanner Mutants

Sequence of the native and linker-scanner substitution mutants of the IRS. Viral nt 4484 through 4715 are represented by the continuous central line of text, with the sequences modified in each of the linker scanning mutations being noted above or below their corresponding location in the native sequence. Notations in bold (i.e. **S4**) indicate the designation of the associated scanner mutation. Dashes (-) represent single nt gaps; that indicated in the wild-type sequence under mutant **S2** indicates that this particular mutant introduces an extra base pair.

Clone	Assay 1 SRRE	Assay 2 SRRE	Assay 3 SRRE	Average	Std. Dev.	Confidence [a]
pJBLR4.3S4	115%	138%	182%	145%	34%	39%
pJBLR4.3S5	132%	99%	161%	131%	31%	36%
pJBLR4.3S20	92%	113%	108%	104%	11%	13%
pJBLR4.3S2	11%	12%	26%	16%	9%	10%
pJBLR4.3S7	22%	71%	37%	43%	25%	28%
pJBLR4.3S8	109%	107%	111%	109%	2%	3%
pJBLR4.3S14	60%	137%	112%	103%	39%	44%
pJBLR4.3S3	140%	132%	195%	156%	34%	39%
pJBLR4.3S23	14%	26%	29%	23%	8%	9%
pJBLR4.3S22	77%	146%	173%	132%	49%	56%
pJBLR4.3S1	7%	28%	22%	19%	10%	12%
pJBLR4.3S25	101%	120%	98%	107%	12%	13%
pJBLR4.3S26	123%	104%	115%	114%	10%	11%
pJBLR4.3S15	63%	152%	108%	108%	45%	51%
pJBLR4.3S16	79%	126%	115%	107%	25%	28%
pJBLR4.3S17	84%	130%	111%	108%	23%	26%
pJBLR4.3S21	73%	139%	180%	131%	54%	62%
pJBLR4.3S12	121%	172%	131%	141%	27%	31%
pJBLR4.3S13	129%	112%	120%	121%	9%	10%
pJBLR4.3S18	106%	198%	118%	141%	50%	57%

Table 4: Scanner Mutation SRRE Values

Pooled results of replication assays on linker-scanning mutations in Table 3. Mutations are ordered by location of the mutation within the IRS. Scaled Relative Replication Efficiencies (SRRE) values from three independant assays are given, their average, the standard deviation of the assays, and ([a]) the 95% confidence interval on the reported average. For individual assay results see Appendix A.

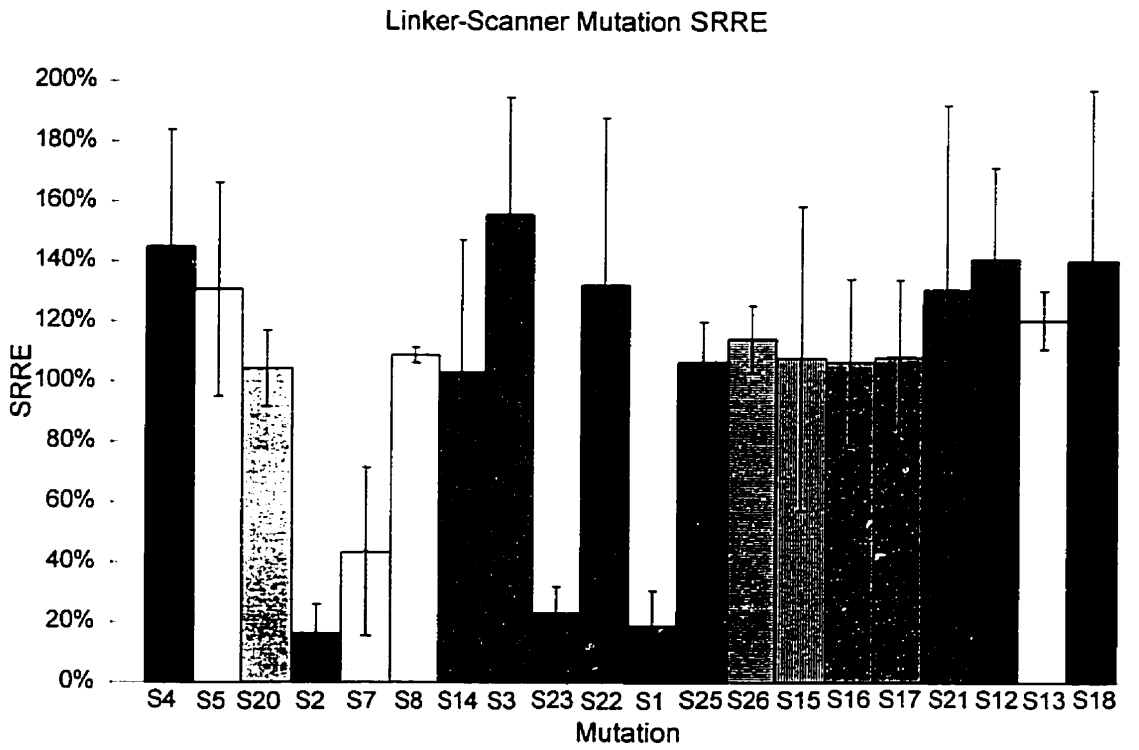


Figure 10: Scanner Mutation SRRE Values

Scaled Relative Replication Efficiency (SRRE) values for each of the linker-scanner mutations in the minigenome context. Mutations are arranged in order of location, and each is designated by 'S(x)' to indicate the corresponding linker scanning mutation as listed in Figure 2. Values graphed are the average of triplicate assays, with error bars indicating a 95% confidence interval.

observational effect. The fact that these observed replication deficits sum to greater than 100% has possible implications which are examined in the Discussion.

Several of the mutations demonstrate a weakly hypercompetent phenotype which is also statistically significant. The magnitude of this was appreciably smaller than that of the replication-defective mutations, and was not studied further. This hypercompetence (and that observed with some point mutations, described below) is considered in the Discussion.

Competitive Replication Assays

In an effort to compare more directly the replication competence of the linker-scanner mutation containing minigenomes with the wild-type, competitive replication assays as described in the Materials and Methods section were used. It was hoped that by simultaneously cotransfecting two minigenomic constructs, small differences in replication competence might be magnified by the faster-replicating construct sequestering required replication machinery from the less efficient construct.

While attempts were made to examine several of the linker-scanning mutants (including both ones found to be replication defective and replication competent by the standard assay) with this technique, unacceptable levels of cross-hybridization by the mutant-specific probe were encountered in most cases leading to the discontinuation of this technique. Results with no detectable interference by nonspecific hybridization were however obtained for two constructs (pJBLR4.3S20 and pJBLR4.3S21) and are presented here for comparison with results from the standard assay. Raw data from these assays

and preliminary data analysis is presented in Appendix B; pooled analyzed data is presented in Table 5.

While the values obtained by this technique for the replication competence of the mutant minigenomes examined relative to that of wild-type are not identical with values obtained for the same mutants with the standard assay, the differences observed are not beyond what might be expected by experimental error and are thus not considered significant. The relatively better replication of pJBLR4.3S21 than pJBLR4.3S20 noted in the regular assay is also apparent in these assays, suggesting that replication trends noted in the standard assay are mirrored in the competition assay. Given the technical difficulties encountered in the competition assays and the fact that only two mutants were sucessfully examined by this technique, no further conclusions can be based on these assays.

Bandshift studies

Previous studies to analyze interactions of host-cell nuclear factors with sequences from the region of the IRS employed two restriction fragments, known as Rsa 'A' (viral nt 4431-4579) and Rsa 'B' (viral nt. 4580-4662) as probes for gel-retardation assays. Results with an LA9 cell nuclear extract indicated the formation of several specific complexes on each probe, and cross-competition experiments indicated that at least some of the complexes formed were identical between the two probes. One of these common complexes was partially purified and footprinting performed, revealing a bipartite area of protection on both strands[79].

Sample Quantitation, S20 Competition Replication Assay

Sample	*ng S20*	*ng WT*	*Ratio S20/WT*
1	44.38	59.48	75%
2	30.90	48.48	64%
3	45.10	59.76	75%
		Average:	71%
		Std. Dev:	7%
		Confidence:	7%

Sample Quantitation, S20 Competition Replication Assay

Sample	*ng S21*	*ng WT*	*Ratio S20/WT*
1	73.58	42.29	174%
2	49.42	24.22	204%
3	51.48	27.20	189%
		Average:	189%
		Std. Dev:	15%
		Confidence:	17%

Table 5: Results of Competitive Replication Assays

Pooled analyzed data from competition replication assays on clones pJBLR4.3S20 and pJBLR4.3S21. Total ng of mutant and wild-type plasmid DNA detected are reported, the ratio of them as the replication efficiency of the mutant as compared to wild-type minigenome, along with the average, standard deviation, and 95% confidence interval resulting from 3 assays.

In order to localize further sites of protein-DNA interaction within the IRS the elements of the linker-scanning library were examined by competition bandshift assay as detailed in the Materials and Methods section. Bandshift experiments were originally carried out with a probe corresponding to the entire region of interest as a BstEII - BamHI fragment from each of the respective minigenomic clones. Use of such a large probe resulted in the formation of extremely large complexes which failed to enter polyacrylamide gels. Other gel matrices were tested but found to give poor resolution. Tentative results obtained with agarose/glycerol gels indicated the formation of a single extremely large complex, and indicated that this complex was only interacting with sequences near the middle of the IRS (data not shown). Based on these experiments the region was divided into three shorter overlapping fragments of approximately equal size (BS, 2021, and SB probes, see Materials and Methods and Figure 11) whose endpoints were chosen so as to be well away from any suspected interactions. The smaller size of these probes allowed for clear resolution of the complexes formed on polyacrylamide gels.

Each linker scanning mutation was analyzed in the context of only one of these probes such that none of the mutations were tested near the end of a probe fragment (see Materials and Methods). Scanner mutants interfering with a host factor binding site are observed by a loss of ability to compete with a radiolabeled wild-type probe for binding by factors in a nuclear extract, and thus show a retention of shifted bands present in the control sample without competitor.

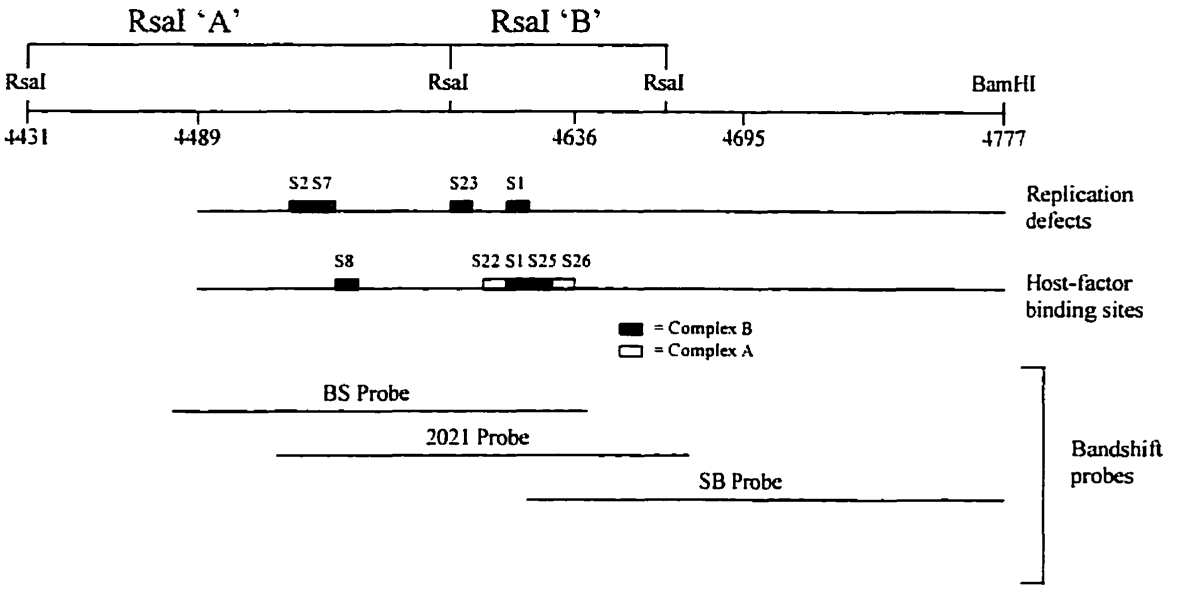


Figure 11: Relative Location of Bandshift Probes

Schematic representation of the IRS region and location of sequence elements observed in this study. Upper line represents sequence of IRS (Element A is nt. 4489-4636, Element B is nt. 4637-4695) plus nearby flanking regions. Location of RsaI sites delimiting RsaI"A" and "B" fragments are indicated, with the fragments being marked by name. Linker scanning mutations causing loss of minigenome replication competence are indicated by number and shown as shaded boxes at their relative positions along the next line. The third line indicates the location of linker scanning mutations interfering with the binding of host-cell factors; linkers are marked by number and shaded according to whether they interfere with the binding of complex 2021-A ("Complex A") or 2021-B ("Complex B"). Bottom section of the diagram indicates the fragments used as bandshift probes (see text).

Examination of the leftward side of the IRS using the BS fragment (see Figure 12) reveals the formation of two specific complexes (BS-A and BS-B; other smaller complexes are not competed by any of the competitors and appear to be non-specific interactions). Both of these complexes are competed equally well by a 40-fold excess of unlabelled wild-type or mutant BS fragment, indicating that none of the bands observed arise from interactions in the sequences mutated in scanners S4, S5, S20, S2, or S7.

The central region was examined with the 2021 probe, which also was found to form two distinct specific complexes (2021-A and 2021-B, at approximately the same mobilities as BS-A and BS-B; see Figure 13). Both complexes were effectively competed for by a 40-fold excess of several of the mutant competitors examined with this probe. However competitor 2021-S8 failed to compete for the binding of complex 2021-B; competitors 2021-S22 and 2021-S26 lost ability to compete for the binding of complex 2021-A, with the S22 mutation having the least competition activity; and adjacent-mutation competitors 2021-S1 and 2021-S25 had a reduced ability to compete for binding of complex 2021-B. The appearance of a number of faint, smaller shifted forms in the presence of competitors S8, S26, and S15 is most likely due to partial formation of the disrupted species.

Of note is that in each case where there is a loss of competition for one of the complexes there appears to be a weaker loss of competition for the other of the two complexes as well; the implications of this observation will be considered in the Discussion.

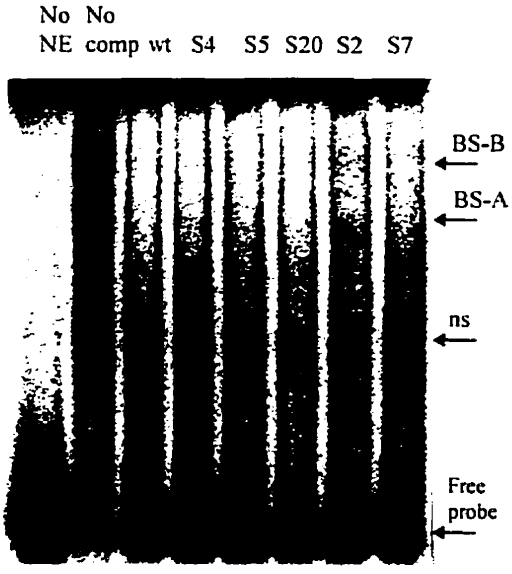


Figure 12: BS Bandshifts

Competition bandshift assays performed on BS probes of linker scanner mutations S4, S5, S20, S2, and S7. Lane marked 'No NE' contains no nuclear extract, lane marked 'No comp' contains no competitor fragment, lane marked 'wt' contains a 40-fold excess of unlabelled wild-type competitior, and lanes marked 'S(x)' contain a 40-fold excess of the respective linker-scanning mutation derived competitor fragment. 'BS-B' and 'BS-A' indicate specific complexes 'A' and 'B' obtained with this probe. 'ns' indicates a non-specific interaction.

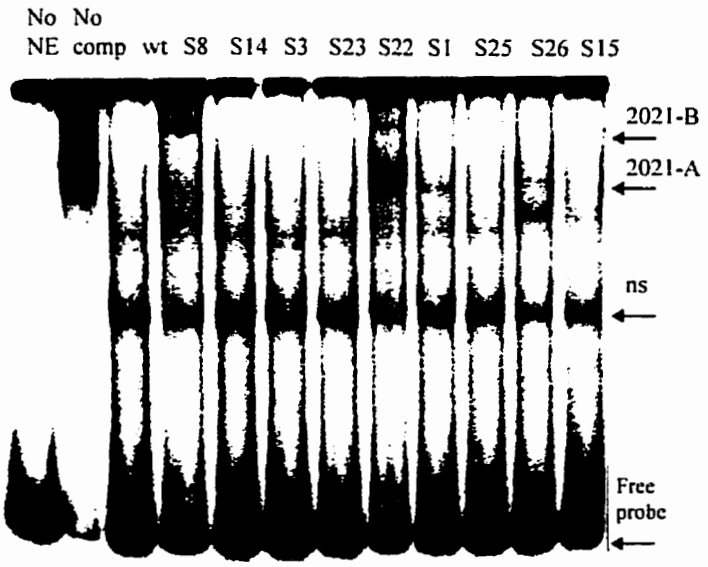


Figure 13: 2021 Bandshifts

Competition bandshift assays performed on 2021 probes of linker scanner mutations S8, S14, S3, S23, S22, S1, S25, S26, and S15. Lane marked 'No NE' contains no nuclear extract, lane marked 'No comp' contains no competitor fragment, lane marked 'wt' contains a 40-fold excess of unlabelled wild-type competitior, and lanes marked 'S(x)' contain a 40-fold excess of the respective linker-scanning mutation derived competitor fragment. '2021-B' and '2021-A' indicate specific complexes 'A' and 'B' obtained with this probe. 'ns' indicates a non-specific complex.

Examination of the region towards the 5' viral termini with the SB probe (see Figure 14) reveals the formation of two specific complexes SB-A and SB-B. The SB-B complex is by far the stronger of the two complexes formed, with only a small fraction of probe existing in the SB-A complex. Both complexes are competed for equally by an excess of unlabelled wild-type or mutant probe, indicating that none of the scanner mutants from S16 to S18 interfere with host factor binding sites.

In the case of all three probes, a distinct band (indicated as 'ns' in Figure 12, Figure 13, and Figure 14) becomes apparent only in the presence of competitor. Based on the observations that this band was not present without competitor, and that none of the mutations examined as competitors resulted in any apparent variation in this band, it was tentatively assigned as being due to a non-specific interaction. Presumably, displacement of labelled probe from the A and B specific complexes allowed for its lower-affinity interaction with this species. At least one other interpretation of the observed data was plausible, however; the 'ns' species might be a partially-formed subunit of either the A or B complex. In this model, addition of competitor would effectively dilute out the proteins contributing to the formation of the A and B complexes, resulting in a population of probes each bearing monomers or small subsets of the larger complexes. The lack of variation in the 'ns' band between competitor samples could be accounted for by its component(s) having a distributed binding over the IRS, such that no one mutation greatly interfered with its capacity to form.

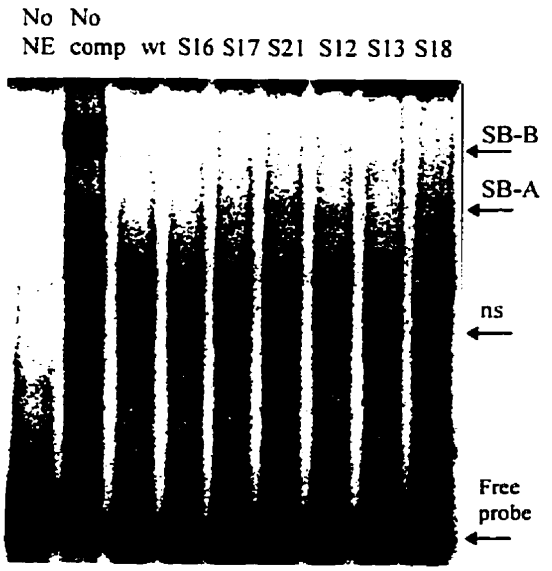


Figure 14: SB Bandshifts

Competition bandshift assays performed on SB probes of linker scanner mutations S16, S17, S21, S12, S13, and S18. Lane marked 'No NE' contains no nuclear extract, lane marked 'No comp' contains no competitor fragment, lane marked 'wt' contains a 40-fold excess of unlabelled wild-type competitor, and lanes marked 'S(x)' contain a 40-fold excess of the respective linker-scanning mutation derived competitor fragment. 'SB-B' and 'SB-A' indicate specific complexes 'A' and 'B' obtained with this probe. 'ns' indicates a non-specific interaction.

In order to discriminate between these possibilities an experiment was performed in which competitors (both specific and non-specific) were titrated for their ability to compete with the formation of the A, B, and ns bands across a wide concentration range relative to probe. Shifted species arising from specific interactions should be competed to a greater extent by specific competitor than by non-specific; use of a wide concentration range ensures that effective levels of competitor are achieved. Specific competitor was the 2021 probe; as non-specific competitor a PCR product of 93 bp was generated using forward and reverse sequencing primers and pUC19 as template. For this particular experiment, competitor fragments were purified by ethanol precipitation in the presence of ammonium acetate rather than gel isolation in order to minimize loss of product. Quality of the purified competitors was verified by agarose gel electrophoresis and ethidium bromide staining.

Results of this experiment (see Figure 15) indicate that the original assignment of 'A' and 'B' as specific complexes, and 'ns' as a nonspecific interaction, are valid. The 'A' and 'B' complexes are competed to a much greater extent by specific probe than by non-specific probe, whereas there is no apparent difference in the ability of the two probes to compete for formation of the 'ns' species.

Examination of the specific shifted species observed with each of the three probes shows that there is an apparent comigration of the BS-A, 2021-A, and SB-A complexes, and of complexes BS-B, 2021-B, and SB-B. If as is suggested by the data from the 2021 probe the major binding sites lie under S8 and S22 with weaker sites under S1, S25, and S26, consideration of the endpoints of the probes used in this assay lends credence to the

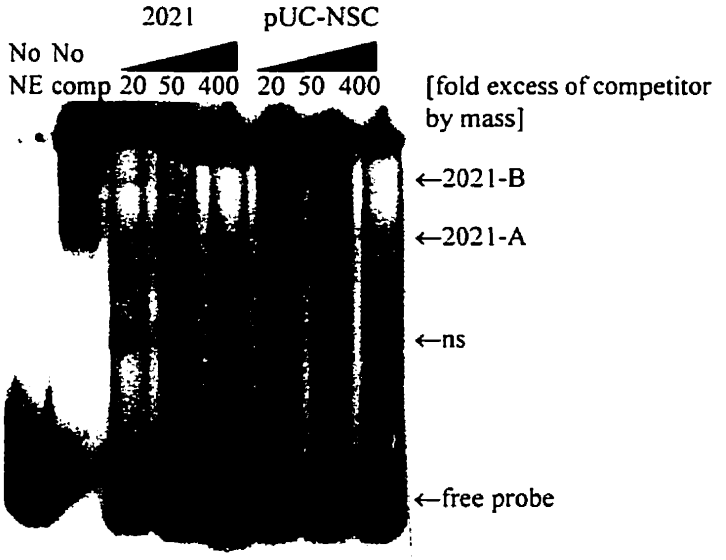


Figure 15:Competitor Titration of Bandshifts

Titration of shifted species observed with LA9 nuclear extract and 2021 bandshift probe with varying excess of specific competitor (2021) or non-specific competitor (pUC-NSC). 'A' and 'B' species are effectively competed for by a 20-fold excess of specific competitor but not non-specific competitor: 'ns' species is competed for equally by both competitors at all concentrations.

supposition that this comigration is not coincidental. Both the BS and 2021 probes contain all of these sites, and the SB probe's left end is coterminal with the observed weak site under S25 and S26 (see Figure 11, bottom section). Given that all three probes employed are of approximately the same size one would predict that identical complexes should migrate approximately equally within the resolution of the employed assay.

While the proximity of these two major binding sites (under S8 and S22) to sequences determined to be important for replication suggested a possible correlation between binding of host-cell factors and replication competence, the lack of any apparent replication deficit in either of these mutations appears to contradict this. As the bandshift data however suggests a co-operative interaction between the factors binding at these sites, it was hypothesized that each of the single mutants might still be competent to localize the entire set of factors necessary in replication and thus not appear replication defective with the assay employed.

Studies on a dual-mutant minigenome

Results of the bandshift experiments indicated the presence of several distinct binding sites for host factors, and in all cases these binding sites were directly adjacent to or overlapping with sequences indicated to be important in minigenome replication. This suggested a possible correlation between these activities, despite the fact that neither of the mutations which caused a major loss in binding of host-factor activity (mutation S8 for complex 2021-B, mutation S22 for complex 2021-A) apparently compromised minigenome replication. In order to try to address whether the coincidence of sequences important in minigenome replication (S2 and S7, S23, and S1) with those binding host-

cell factors (especially S8 and S22) is biologically relevant, a dual-site scanner mutant was constructed in which the major binding sites for complexes 2021-A (under scanner mutation S22) and 2021-B (under scanner mutation S8) were simultaneously substituted.

Replication assays carried out on this mutant are summarized in Table 6. Strikingly, while either of these linker-scanner mutations individually has no detectable effect on minigenome replication efficiency, the dual-site mutant can only replicate at approximately 20% of the level of the unmutated minigenome, approximately the same basal level as observed with the S2, S23, and S1 mutations. As will be further considered in the Discussion, this result strongly suggests a direct role of the bound host factors in stimulating replication functions of the IRS, and furthermore supports a co-operativity of host factor binding previously suggested by the bandshift results.

This dual-site mutation was also examined by competition bandshift assay with the 2021 probe. As each of the single mutations alone had demonstrated a distinct loss in binding activity it was expected that the dual mutation should fail to compete for either the 2021-A or 2021-B complexes; however results with this competitor (see Figure 16) indicate that it is capable of effectively competing for formation of the 'B' complex and retains at least some competition activity towards formation of the 'A' complex. While there is no obvious explanation for this unexpected observation it is suggestive of a high degree of complexity in the interactions controlling factor binding.

Clone	Assay 1 SRRE	Assay 2 SRRE	Assay 3 SRRE	Average	Std. Dev.	Confidence[a]
pJBLR4.3S8/22	18%	26%	18%	20%	4%	5%

Table 6: Dual-mutant SRRE Values

Pooled results of replication assays on dual-mutant minigenome pJBLR4.3S8/22. Scaled Relative Replication Efficiencies (SRRE) values from 3 independant assays are reported, their average, standard deviation, and ([a]) the 95% confidence interval on the average. For individual assay results see Appendix A.

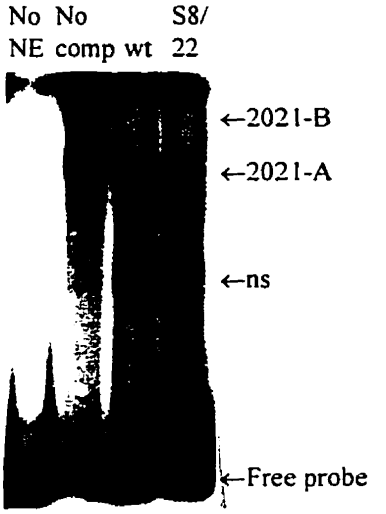


Figure 16:Dual-Site mutant competition bandshift

Competition bandshift assays performed on 2021 probe of dual linker scanner S8/22. Lane marked 'No NE' contains no nuclear extract, lane marked 'No comp' contains no competitor fragment, lane marked 'wt' contain a 40-fold excess of unlabelled wild-type competitior, and lane marked 'S8/22' contain a 40-fold excess of the 2021 S8/22 competitor fragment. '2021-A' and '2021-B' indicate specific complexes 'A' and 'B' obtained with this probe. 'ns' indicates a non-specific interaction.

Studies on point mutants

As scanner mutation S1 was shown to have an effect on minigenome replication competence as well as being the site of binding of a host cell factor, it was desirable to identify which nucleotides within this 8 bp segment contribute to these properties. To examine this, single nucleotide point mutations were made at each of the six bases altered in the scanner mutation, converting the wild-type base to that present in the scanner mutant individually (see Table 7). These mutations were then assayed both for replication competence in a minigenome-based assay (summarized in Table 8), and for ability to interact with host-cell nuclear factors by bandshift assay (see Figure 17). While two of the point mutations (A6C and T8G) appear to lose some competition ability towards the formation of both the 'A' and 'B' complexes and thereby indicate that these nucleotides likely make direct contributions to binding of the complex, none of the point mutations display a detectable replication defect. As with the linker scanning mutations, some of the point mutations (notably C2A and T8G) demonstrate a replication hypercompetence; this observation was not further studied.

Studies on Full-length Virus

In order to address the question of whether the replication-defective phenotype exhibited by some of the linker-scanner mutants in the context of the minigenomic assay system was relevant in the context of whole virus, the S1 mutation was introduced into the full-length infectious clone of MVM (pCA4.0) to make pCA4.0S1. This mutation was chosen for its observed effect on both replication competence and factor binding. To examine the replication competence of this construct relative to the wild-type virus, both

Clone	*Mutation with flanking sequences*
pJBLR4.3T1C	(4602) aataaaga (4609) aataaagG
pJBLR4.3C2A	(4602) aataaaga (4609) aataaaTa
pJBLR4.3T3G	(4602) aataaaga (4609) aataaCga
pJBLR4.3T4A	(4602) aataaaga (4609) aataTaga
pJBLR4.3A6C	(4602) aataaaga (4609) aaGaaaga
pJBLR4.3T8G	(4602) aataaaga (4609) Cataaaga

Table 7: Point mutant sequences

Six point mutations are presented in order of sequence position. For each clone, the associated wild-type sequence is presented with the nt positions relative to full-length MVM of the first and last nt shown in brackets (top line). The same sequence region from the mutant is then presented in alignment, with the mutation indicated in upper-case type (bottom line).

Clone	*Assay 1 SRRE*	*Assay 2 SRRE*	*Assay 3 SRRE*	*Average*	*Std. Dev.*	*Confidence*[a]
pJBLR4.3T1C	100%	92%	94%	95%	4%	5%
pJBLR4.3C2A	166%	123%	135%	141%	22%	25%
pJBLR4.3T3G	105%	125%	129%	120%	13%	14%
pJBLR4.3T4A	97%	109%	109%	105%	7%	8%
pJBLR4.3A6C	112%	87%	166%	122%	41%	46%
pJBLR4.3T8G	136%	137%	131%	135%	3%	3%

Table 8:Point Mutant SRRE Values

Pooled results of replication assays on linker-scanning mutations in Table 7. Mutations are ordered by location of the mutation within the IRS. Scaled Relative Replication Efficiencies (SRRE) values from three independant assays are given, their average, the standard deviation of the assays, and ([a]) the 95% confidence interval on the reported average. For individual assay results see Appendix A.

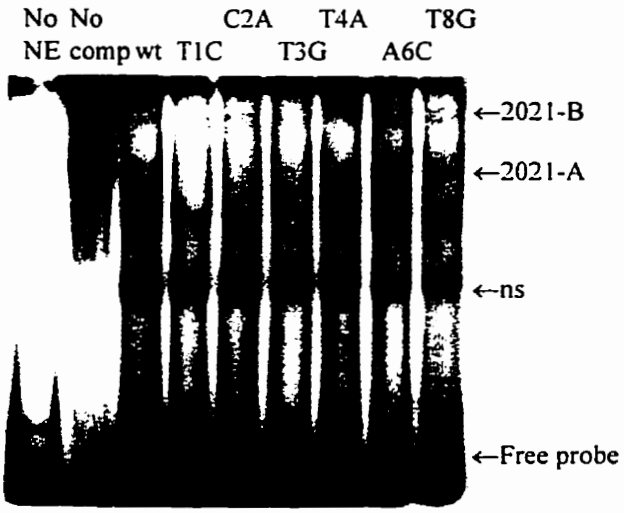


Figure 17: Point mutant competition bandshifts

Competition bandshift assays performed on 2021 probes of point mutations T1C, C2A, T3G, T4A, A6C, and T8G. Lane marked 'No NE' contains no nuclear extract, lane marked 'No comp' contains no competitor fragment, lane marked 'wt' contains a 40-fold excess of unlabelled wild-type competitior, and lanes marked with a point-mutant designation contain a 40-fold excess of the respective point mutation derived competitor fragment. '2021-A' and '2021-B' indicate specific complexes 'A' and 'B' obtained with this probe. 'ns' indicates a non-specific interaction.

clones were transfected into LA9 cells and viral lysates were prepared (see Materials and Methods). Lysates were titred by plaque assay on LA9 cells. No reproducible difference in titre between the wild-type and mutant virus was observed in these studies (data not shown).

During these experiments a definite difference in cytopathogenicity of the wild-type as compared to the mutant virus was noted. Examination 24 hours post-shock revealed dishes transfected with wild-type virus to have fewer live cells than those transfected with S1 mutant virus. In addition, the morphology of the two cell populations appeared different, with many cells in the wild-type samples having an abnormally elongated or sickled shape as compared to cells in the S1 mutant virus samples (see Figure 18). While these observations are only qualitative, this effect was observed reproducibly and suggests the S1 mutation has an effect on viral pathogenesis when introduced into full-length virus.

Evidence for the IRS as a leading-strand only origin

The presence of elements contributing to viral replication in the location of the IRS on the viral genome lead to a consideration of what their possible role in viral replication might be. Considering the simplified replication model shown in Figure 4, one possibility was that the elements in the IRS contribute to the 5' terminal rearrangements from extended to paired hairpins shown in steps 4 and 8. Given that the length of the viral 5' palindrome is 206 base pairs of which 200 can participate in a stably base-paired hairpin, there exists a prohibitively high thermodynamic barrier to its spontaneous interconversion between the forms under

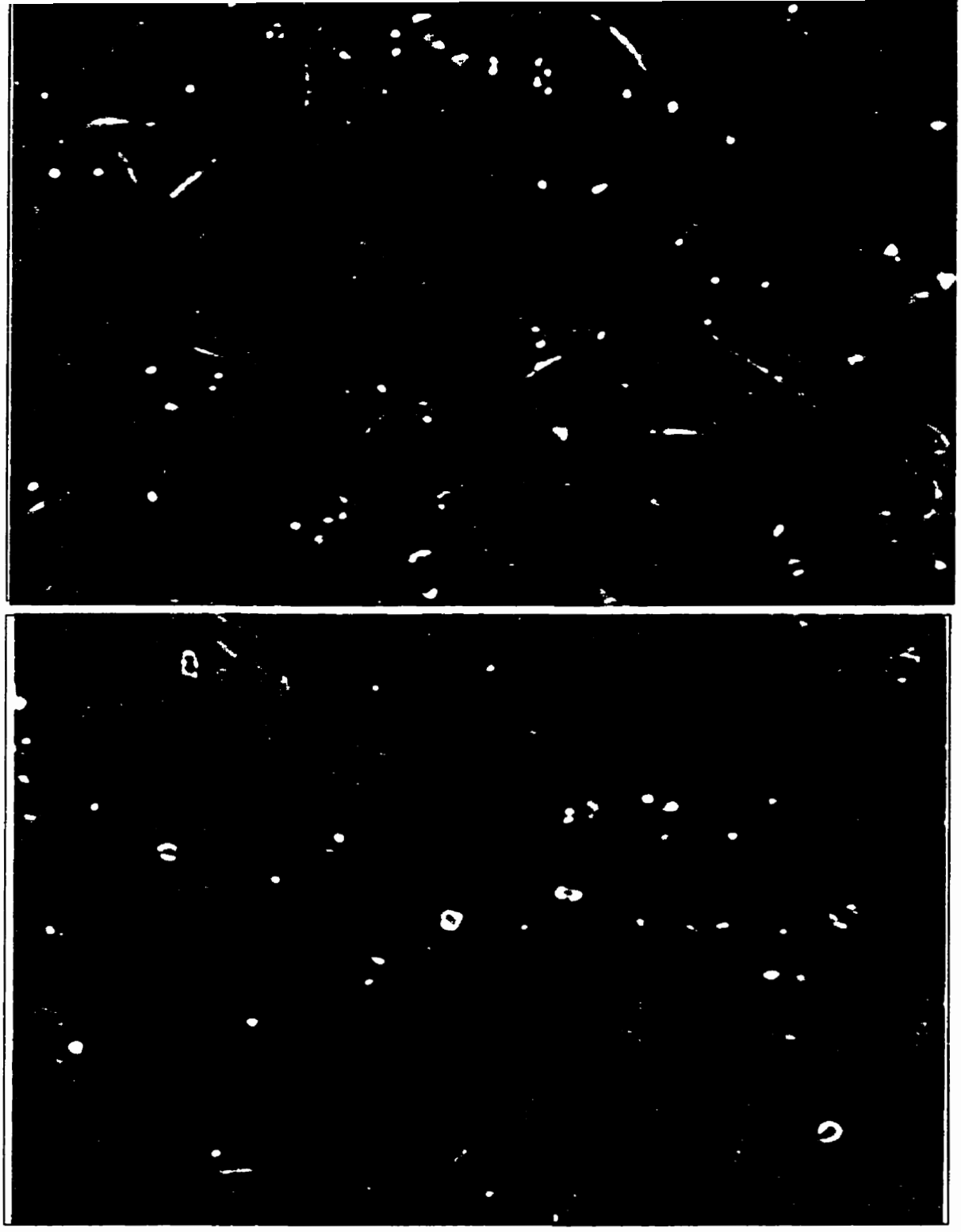


Figure 18: Differential CPE of Wild-Type and S1 Virus

Photomicrographs (microscope magnification 200 ×, photographic enlargement ~3 ×) of dishes of LA9 cells 24 hours post DMSO-shock after being transfected with either wild type infectious clone pCA4.0 (Panel A) or S1 mutant infectious clone (Panel B). Dishes were seeded identically and cell density and morphology were indistinguishable between dishes prior to transfection.

physiological conditions. No satisfactory model for this has yet been proposed. A model which incorporates the observed findings in the IRS and simultaneously provides a mechanism for this terminal rearrangement is presented in Figure 19(a). An origin of DNA replication inboard of the 5' palindromic sequences generates two leading strands of nascent DNA and thereby displaces one strand of the extended palindrome, thus allowing self-annealing into a hairpin conformation to occur spontaneously. As mentioned in the Introduction, a similar model was proposed by Rhode and Klaassen in 1982[78]; however a critical difference between the models is that Rhode and Klaassen proposed both leading and lagging strand synthesis outward from the origin. The importance of this difference becomes apparent when one considers the effect that Okazaki fragments produced between the origin and the extended 5' palindrome termini might have on the proposed terminal rearrangement step. Any such Okazaki fragments extending into the palindromic region would effectively block the complete formation of the terminal hairpin and thus not allow the viral mRF's free 3' hydroxyl to act as a primer for synthesis to dRF (see Figure 19(b)).

In order to test this model it was considered advantageous to examine only the right-hand viral telomeric sequences as an isolated system. As studies conducted previously[70], and up to this point in the current study, all examined the IRS in the context of either a minigenome or full-length virus they have been complicated by the unknown effect of the presence of what appears to be another (albeit weak) origin within the left-hand viral telomere. Furthermore, all evidence for the function of the IRS as an origin of

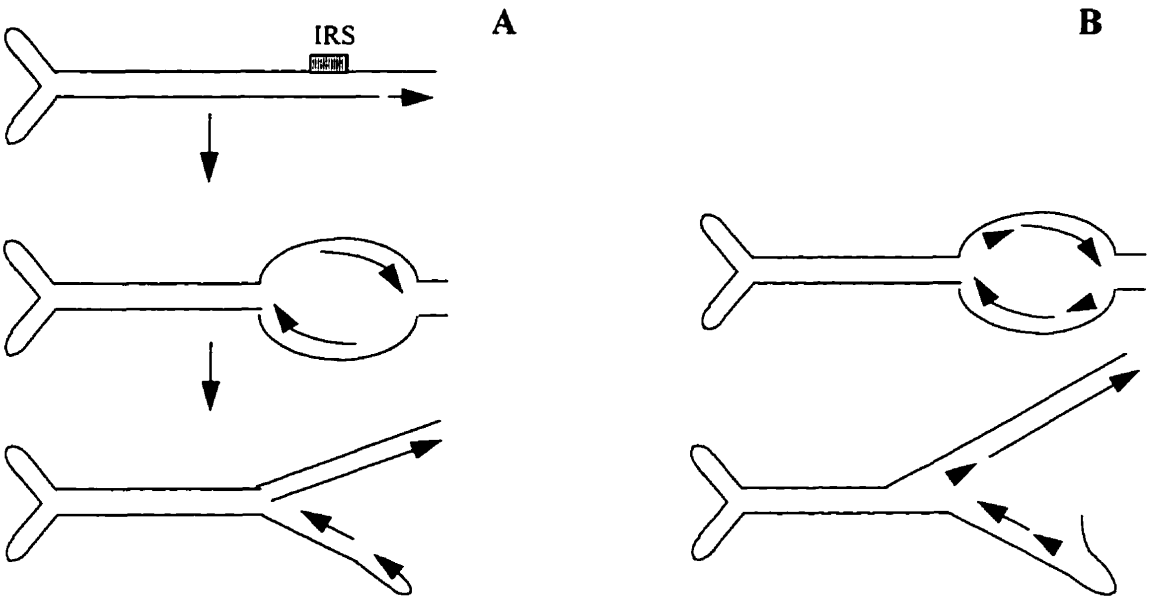


Figure 19: Leading-Strand Only Model for Hairpin Rearrangement

Unidirectional initiation model for action of the IRS to allow for rearrangement of viral 5' termini from extended to hairpin conformation. Panel (a) represents the model as described in text; panel (b) the Rhode and Klaassen model[72] where presence of a lagging strand would interfere with hairpin formation and lead to incomplete replication of template. Solid lines with arrows indicate nascent leading strands; dashed lines with arrows indicate nascent lagging strands.

replication has come from studies whereby the replication competence of the system is compromised by alterations to sequences within the IRS.

If the IRS is in fact a canonical origin of replication it should be possible to isolate it from the putative left-hand viral ori sequences and use it to drive replication of unrelated DNA sequences. Such a system would allow both for an examination of the IRS elements uncomplicated by effects from the left-hand viral ori, and allow for a positive demonstration of the ori function of the IRS. That such a system should be feasible was suggested by prior results with a chimeric minigenome consisting of two right-hand viral telomeres joined at their inboard ends; when supplied with NS-1 in *trans*, this construct (pPTRR) was replication competent[70]. Interestingly, however, other preliminary studies from our lab with the isolation of the IRS on a circular plasmid lacking a eukaryotic origin of replication failed to detect any evidence for plasmid maintenance in LA9 cells[89].

Construct pJBR1 (see Figure 20) was used to examine if the isolated IRS is able to act as an origin of replication, and if so whether it acts in a canonical eukaryotic leading and lagging strand fashion or in the hypothesized leading-strand only manner. Sequences downstream of the IRS as far as the genome 5' end were maintained on the construct for two reasons: firstly it was considered possible that as-yet uncharacterized sequence elements within this region might be required for the postulated origin function, and secondly the hypothesized leading-strand only model requires a terminal palindrome; retention of the existing viral sequence was simpler than replacing it with an unrelated palindrome of similar size.

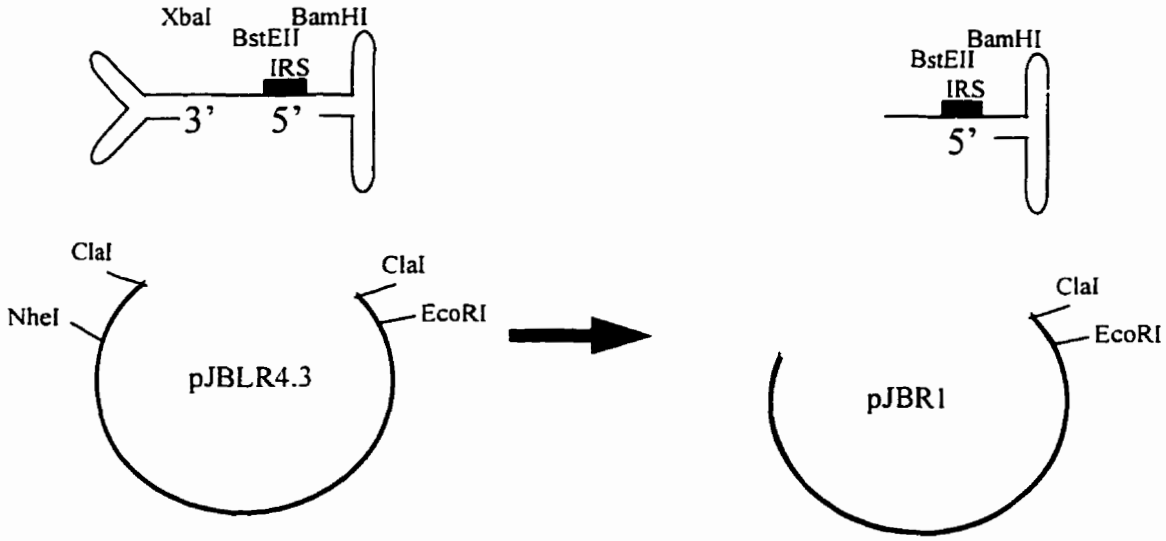


Figure 20: pJBR1

Deletion of the XbaI-NheI fragment of pJBLR4.3 creates pJBR1. Both constructs are shown as 'inserts' in the context of associated vector plasmid; both are linearized at the EcoRI site prior to transfection.

Models of NS-1 function in MVM replication up to this point have only pointed to its having roles in the resolution of dimer bridges between viral concatamers, and in the process of hairpin transfer; it was not believed to play a direct role in the action of either viral origin of replication. It was thus not predicted to be absolutely required for replication of pJBR1, although the products of the Leading-Strand Only (LSO) model without NS-1 present should be concatameric. Figure 21 diagrams what the predicted replication products of either the LSO or a leading- and lagging-strand origin, in the presence of NS-1, should be. Notably, for a linear replicon such as EcoRI digested pJBRI, a bidirectional origin with both leading and lagging strands should produce a uniform set of products starting at the size of the input plasmid and getting progessively smaller with each successive round of replication by an inability of the plasmid termini to replicate fully while a bidirectional LSO origin should lead to the production of three distinct species ([1], [4a], and [4b] of Figure 21).

The replication competence of pJBR1 in both the presence and absence of NS-1 was analyzed by transfection of EcoRI linearized plasmid into COS-7 cells with and without the cotransfection of pCMVNS-1. The standard DEAE-Dextran protocol was employed, and transfected cells were allowed to grow for 48 hours post DMSO shock before Hirt extraction. The extracted DNAs were analyzed by DpnI digestion and Southern blotting with a random-primed probe for virally-derived sequences (the XbaI-EcoRI fragment of pJBR1 was used as template). Results of this study are shown in Figure 22.

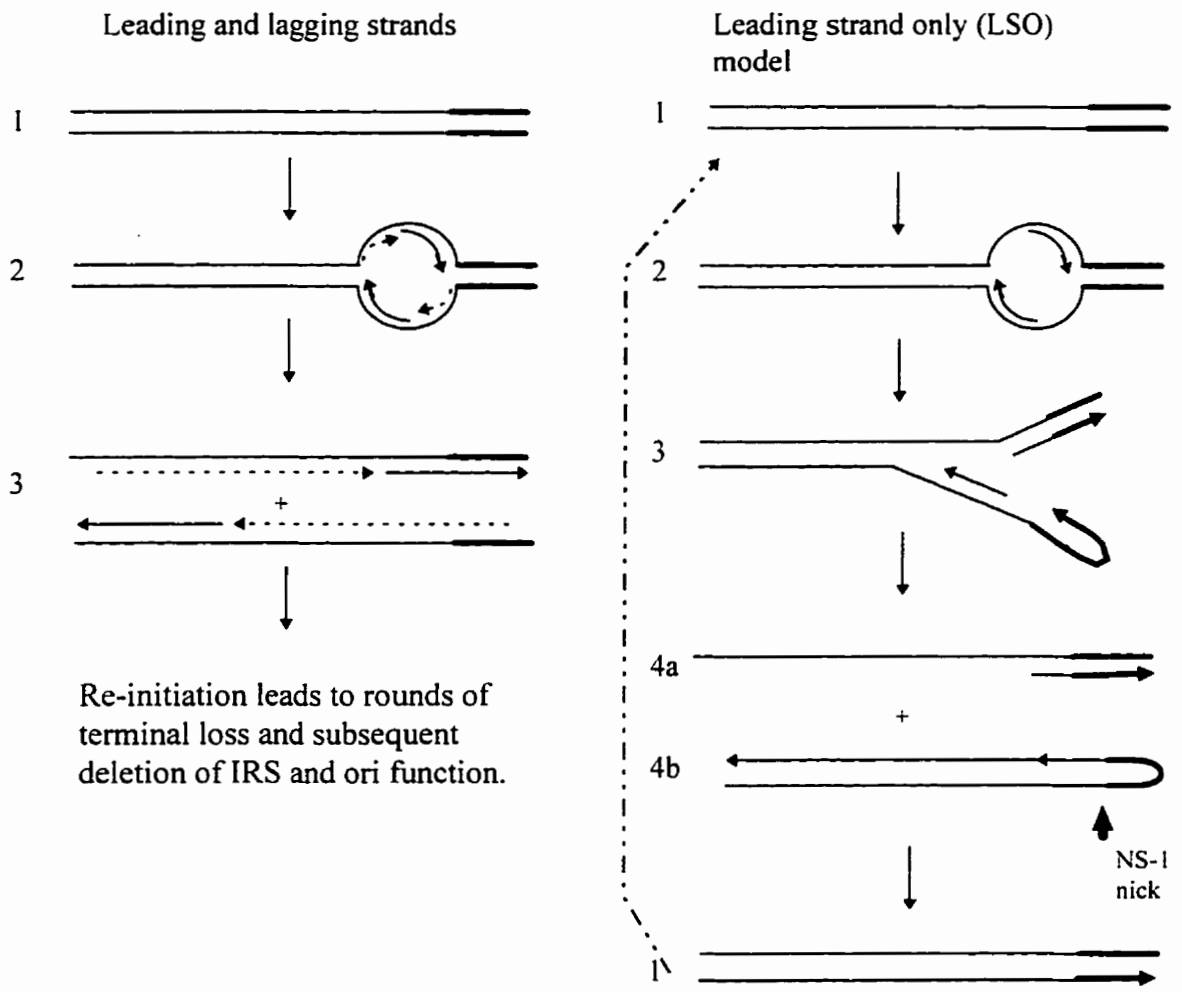


Figure 21:Predicted replication products of pJBR1

Proposed replication products of pJBR1 construct for an origin generating both leading and lagging strands (left) or leading strands only (right). Solid lines represent pre-existing or nascent leading-strand DNA, dashed lines represent nascent Okazaki fragments. Regions indicated in bold are the palindromic terminal sequences.

Leading and lagging strand synthesis will block hairpin formation and result in a population of double-stranded linear molecules with incompletely replicated termini ([3], left side). Synthesis of a leading strand only will allow hairpin formation ([3], right) and result in three distinct populations of molecules ([4a], [4b], and [1]) in the presence of NS-1. In the absence of NS-1, [4b] can re-initiate replication to form a dimeric intermediate.

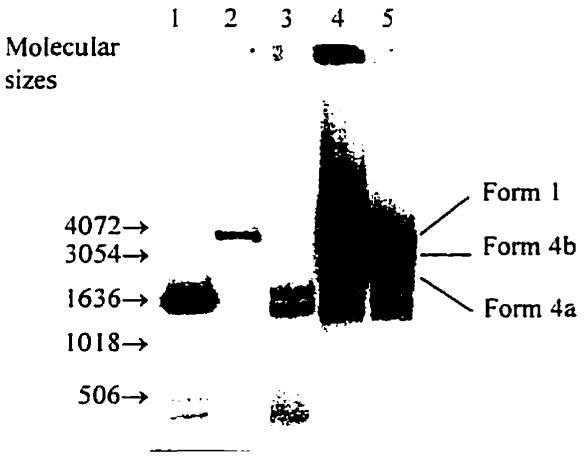


Figure 22: pJBR1 Replication Products

COS-7 cells were transfected with EcoRI linearized pJBRI with or without pCMVNS-1 and grown for 48 hours prior to Hirt extraction. 5 µl aliquots of Hirt extract were treated as listed below and run on a 1% agarose gel, followed by vacuum transfer to Hybond-N and Southern blotting with a random-primed probe directed against virally-derived sequences (XbaI to BamHI fragment of pJBLR4 corresponding to viral 5' terminus).

Lane #:Plasmids transfected, treatment of Hirt extract
(Lane 1: BRL 1kb ladder.)
Lane 2: pJBR1, undigested.
Lane 3: pJBR1, DpnI digested.
Lane 4: pJBR1 + pCMVNS-1, undigested.
Lane 5: pJBR1 + pCMVNS-1, DpnI digested.

In Lane 5, three distinct forms are tentatively identified as Form 1, Form 4b, and Form 4a (see text and Figure 21) based on mobilites relative to size markers (Lane 1).

In the absence of NS-1 (lanes 2 and 3) no replication to full-length species is observed. While lane 2 contains a single band corresponding to linear input plasmid, its disappearance in lane 3 demonstrates its DpnI sensitivity and thus lack of replication. (Faint higher molecular weight bands present in Lane 2 are trace amounts of incompletely digested plasmid). Of the two bands present in lane 3, the upper is attributable to a DpnI digestion product of full-length input plasmid; the other however is not readily identifiable and most likely arises from an abortive replication initiation event, the probable cause of which will be considered in the Discussion. Although not successful in fully replicating the construct, this results in hemimethylation and thus DpnI resistance[90] of an additional region of the plasmid.

In contrast, the presence of NS-1 (lanes 4 and 5) allows for effective replication leading to the formation of three additional DpnI resistant bands (lane 5) than are not present without NS-1. The largest of these is tentatively assignable as full-length plasmid (Form 1 from Figure 21) based on its equal mobility to input plasmid (lane 2). The next two smaller species, observed only when NS-1 is present, are tentatively assigned as Forms 4b and 4a (Figure 21) based on mobility. (While DNA species with extended single-stranded regions may under some circumstances demonstrate decreased electrophoretic mobility relative to an equal length double-stranded species, previous results with MVM indicate that single-stranded minigenomes migrate more quickly than do the duplex mRF forms[70]; thus by analogy form 4a is predicted to exhibit greater mobility than does form 4b). The two smallest forms present appear to be analogous to those in lane 3, with one being a fragment of unreplicated input plasmid and the other

being a failed replicative intermediate. There is an apparent loss of size on the larger of these two species, which may be attributable to NS-1 catalyzed nicking of the band.

Several observations, some of them unexpected, can be made from the results of this experiment. Firstly, the right-hand end of the MVM genome from the IRS through the terminal palindrome appears to be a true origin as under the correct conditions it is capable of driving the replication of the attached plasmid sequences. Secondly, it appears that the viral NS-1 protein is required in some direct manner at this origin; complete plasmid replication is only observed if NS-1 is available in *trans*. As will be considered in the Discussion however, this particular observation may be an artifact of the test system. Thirdly, comparison of the replicated (pJBR1 + pCMVNS-1) samples versus the unreplicated samples (pJBR1 alone) show the formation of three distinct DNA species, as would be predicted exclusively by the LSO origin model presented here. Under no conditions tested did the construct replicate to form a set of near full-length species with serially decreasing terminal lengths as predicted by a leading and lagging strand model.

The LSO model makes predictions regarding the form of each of these three DNA species. In order to examine whether the observed products had the predicted forms (1, 4b, and 4a from Figure 21) the Hirt extract from COS-7 cells cotransfected with EcoRI-linearized pJBR1 and pCMVNS-1 was examined by two-dimensional neutral/alkaline agarose gel electrophoresis, transfer to hybridization membrane, and probing with viral sequences as detailed in the Materials and Methods. Results of this experiment are presented in Figure 23.

In this form of analysis, linear duplex species are resolved as a curving diagonal band; species above the diagonal in the second dimension arise from folded DNA molecules in which the single strands are longer than the duplex form, while species below the diagonal in the second dimension arise from duplex molecules in which the single strands are of less then unit length. Form 1 DNA would thus be predicted to run on the diagonal, Form 4b DNA should run above the diagonal, and Form 4a should resolve into two spots, one above the diagonal and one below the diagonal (this assumes that the Form 4a species will be retarded under neutral conditions by the single-stranded extension to some point between the duplex mobility of either of its two constituent strands). The Form 2 intermediate would only be expected to be present in small amounts, and should form a continuous arc above the duplex diagonal.

Examination of Figure 23 demonstrates the presence of all of these forms with the exception of the small strand of Form 4a. Form 1 can be distinguished as lying on the duplex diagonal. The Form 4b band, predicted to be a single large palindrome and running at something slightly less than full-length plasmid under neutral conditions, is resolved as a continuous set of species ranging from dimer-length single strands down to sub unit length species; most of the signal is of sub unit length. Consideration of the LSO model suggests that if the ligation step was slow and a variable degree of displacement synthesis occurs during Form 4b synthesis (see lower strand of Panel 3, Figure 21) all of these forms could be expected. The Form 4a band, predicted to be mostly single-stranded unit length with a short double-stranded region at one termini, resolves into a single spot above the duplex diagonal arising from the longer of its two strands. While the shorter of

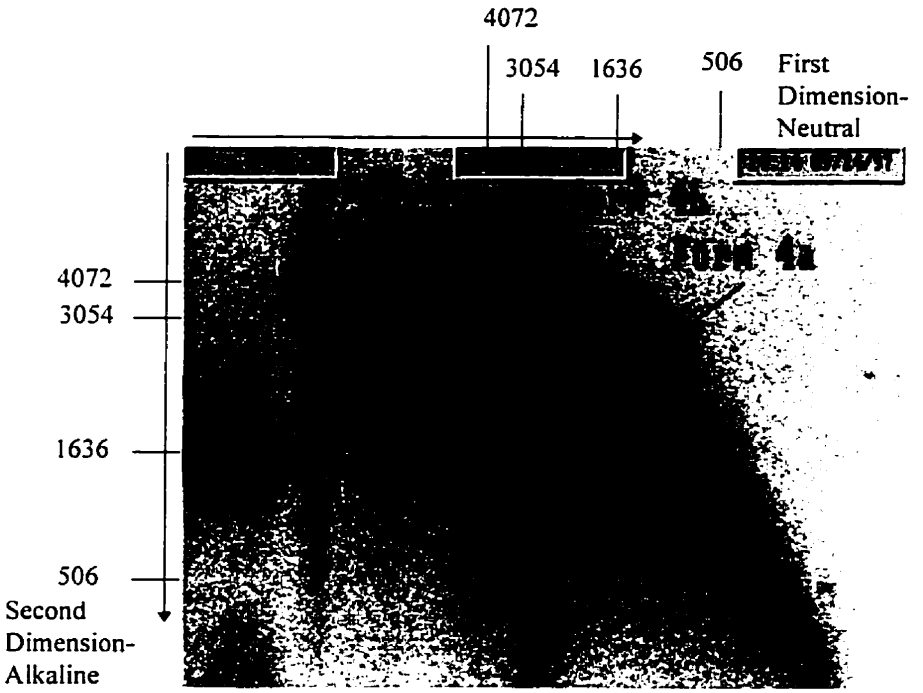


Figure 23: 2D Neutral-Alkaline Analysis of pJBR1 Replication Products

Two-dimensional neutral-alkaline gel analysis of pJBR1 + pCMVNS-1 Hirt extract. Hirt extract was run in first dimension in 1% agarose/TAE gel matrix next to a size standard (approximate mobilites as marked) prior to lane excision, soaking in 50mM NaOH/1mM EDTA for 45 minutes, and pouring of a 1% alkaline gel below the excised lane. This gel was then run in 50mM NaOH/1mM EDTA at 2 V/cm for 7 hours, vacuum transfered to Hybond-N membrane, and Southern blotted for viral sequences. As a standard, BRL 1kb ladder was denatured in alkaline gel loading buffer and run from a well even with the lower edge of the sample lane; weak cross-hybridization to these sequences allows for size estimation in the second dimension (as marked).

Species tentatively identified as Form 1, Form 4b, and the longer strand from Form 4a are indicated (see text).

the two bands is not visible here, the low abundance of the longer strand suggests that the signal from the shorter strand may merely be too weak for clear detection above background.

Signals arising from discrete species which run at larger than unit length in the first dimension probably arise from a heterogeneous population of replicative intermediates (Form 3 from Figure 21) which due to their forked nature run anomalously large under nondenaturing conditions.

Having demonstrated that the IRS and its associated terminal sequences were capable of driving plasmid replication in a system independent of the viral left-hand origin, it was decided to analyze the linker-scanning mutations which compromised minigenome replication efficiency in the context of this isolated system. Accordingly, constructs pJBR1S1, pJBR1S2, pJBR1S7, pJBR1S23, and pJBR1S8/22 were all made (from the respective linker-scanner mutant pJBLR4.3 derivatives in a manner analogous to that for pJBR1, see Materials and Methods) and analyzed for replication competence by DpnI resistance assay. Following transfection into COS-7 cells either in the presence or absence of pCMVNS-1 for 48 hours, Hirt extracts were prepared, digested with DpnI, and analyzed by Southern blotting for virally-derived sequences. The results (see Figure 24) indicate that all of the mutations render the pJBR1 construct replication incompetent. As further considered in the Discussion, this has implications on the function of the sequences replaced by replication-defective mutants.

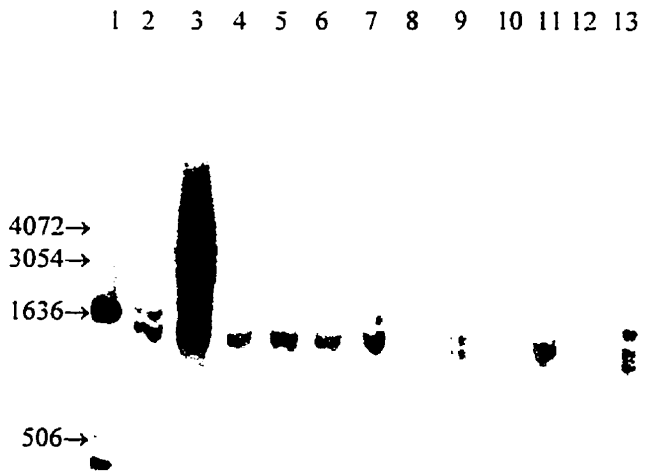


Figure 24: Replication of pJBR1S(x) Scanner Mutations

COS-7 cells were transfected with pJBR1 and its linker scanner derivatives with or without pCMVNS-1 and allowed to grow for 48 hours prior to Hirt extraction. 5 µl aliquots of Hirt extract from each transfection were digested with DpnI and run on a 1% agarose gel, followed by vacuum transfer to Hybond-N membrane and Southern blotting with a random-primed probe directed against virally-derived sequences (pJBLR4 XbaI-BamHI fragment, viral 5' terminus).

Lane #:Plasmids transfected
(Lane 1: BRL 1 kb marker)
Lane 2: pJBR1
Lane 3: pJBR1 + pCMVNS-1
Lane 4: pJBR1S1
Lane 5: pJBR1S1+ pCMVNS-1
Lane 6: pJBR1S2
Lane 7: pJBR1S2+ pCMVNS-1
Lane 8: pJBR1S7
Lane 9: pJBR1S7+ pCMVNS-1
Lane 10:pJBR1S23
Lane 11:pJBR1S23+ pCMVNS-1
Lane 12:pJBR1S8/22
Lane 13:pJBR1S8/22+ pCMVNS-1

Presence of DpnI resistant forms in Lane 3 (pJBR1 + pCMVNS-1 sample) are indicative of replication.

One-hybrid screen

In an effort to identify some of the host-cell nuclear factor(s) interacting with the IRS a yeast genetic one-hybrid screen was employed. This technique, as first employed by Wang and Reed[91], takes advantage of the fact that for many transcriptional activators DNA-binding and transactivation activities are separable. A reporter gene (generally an auxotrophic marker) is placed downstream of a minimal promoter such that without transactivation of the promoter little or no transcription of the reporter is observed. A 'bait' sequence consisting of DNA which is a recognition site for the unknown cellular protein is cloned directly upstream of this minimal promoter, usually in the form of a tandem multimer to improve binding affinity. A library consisting of cDNAs fused to a nuclear localization signal and a transactivation domain is then transformed into the cell line harboring the reporter, and transformants screened for expression of the reporter gene. In theory, transactivation and gene expression should arise when a cDNA encoding the bait sequence's cognate DNA binding domain is fused in-frame to the library vector's transactivation domain. The DNA binding activity of the cDNA-derived portion of the fusion polypeptide acts to localize the transactivation domain of the fusion in close proximity to the reporter's minimal promoter and enhances transcription, resulting in detectable gene expression. As employed here, the system uses both auxotrophic selection (HIS3) and a secondary reporter (β-galactosidase) to minimize the number of false positive transformants observed.

As linker-scanning mutation S1 was the only one shown to simultaneously affect both replication competence and host-cell factor binding capacity, the sequences replaced

in this mutation plus the two flanking nucleotides from each side were chosen as a suitable bait sequence. An oligonucleotide duplex containing a tetramer of this bait sequence was cloned upstream of a His3 gene (creating pHis-i/S1) and also upstream of a β-galactosidase gene (creating pLacZi/S1). These clones were integrated into yeast strain YM4271 to create YMLHS1 as a dual-reporter construct. (See Methods and Materials for construction details).

A mouse 11-day whole embryo cDNA fusion library was screened by transformation into YMLHS1 and selection for histidine prototrophy. Out of 2.49×10^6 clones, a total of 11 primary transformants were isolated as being His+ and were re-streaked on selective media (SD/-Leu/-His/30 mM 3-aminotriazole). All grew following reselection hence attempts were made to isolate plasmid from all. Plasmid was successfully obtained from all but two of the eleven clones. These nine plasmids were re-transformed into YMLHS1 and selected again on SD/-Leu/-His/30 mM 3-aminotriazole media. All nine yielded viable colonies, and all were tested for β-galactosidase expression by filter lift (see Methods and Materials). Only one clone, designated 2B1, demonstrated β-galactosidase expression. This clone exhibited both rapid growth in the absence of histidine and strong β- galactosidase expression (equal to or greater than that of the positive control system supplied with the one-hybrid system). This clone was selected for further analysis.

Two colonies obtained from the original plasmid recovery into SURE cells were miniprepped and the resulting DNA analyzed by EcoRI digestion. This should release the cDNA insert and leave a 6.7 kb pGAD10 vector band. Both isolates, 2B1-1 and 2B1-

2, were identical and contain a 1.4 kb cDNA insert. Isolate 2B1-1 was used for sequencing from the 5' end with primer pGADforward.

Initial sequencing efforts yielded 152 bases of reliable sequence starting from the EcoRI site. Analysis of this recovered sequence with BLAST[92] against a non-redundant database revealed the clone to have high homology to human and murine Hoxb-13, with 100% identity between the last 62 bases of sequence obtained and the first 62 bases of coding sequence reported for murine Hoxb-13. Homologies upstream of nucleotide 91 of the obtained sequence were only noted to the putative 5' untranslated region of the human Hoxb-13 mRNA (the 5'UTR for the murine clone was not available for comparison). Translation of the obtained sequence in the proper reading frame for expression as a fusion with the GAL4 and SV40 T-Antigen fragments upstream revealed no stop codons within the obtained sequence. Comparison of this translated sequence against a non-redundant database by BLASTP revealed no significant homologies within amino acids derived from the first 91 nucleotides, and a 100% identity between the 20 amino acids derived from the last 61 nucleotides and the first 20 amino acids of both human and murine Hoxb-13 protein (see Figure 25). The probability of this homology occuring randomly was less than 1×10^{-8}, strongly suggesting that clone 2B1 encoded the gene for Hoxb-13. As the reported total mRNA size for human Hoxb-13 message is only 1026 nt, the 1.4 kb insert in clone 2B1 likely contains a full-length cDNA.

Hoxb-13 was originally discovered by Zeltser et. al.[93] by screening of an expression library with an oligonucleotide probe carrying a (5'-TAA-3') repeat sequence. Given that the bait sequences employed here include four repeats of the sequence 5'-

Results of BLAST search with 2B1 sequence

DNA alignment

```
Mus musculus Hoxb-13 mRNA, complete cds

Plus Strand HSPs:

 Score = 310 (85.7 bits), Expect = 4.8e-18, P = 4.8e-18
 Identities = 62/62 (100%), Positives = 62/62 (100%), Strand = Plus / Plus


Query:91 ATGGAGCCCGGCAATTATGCCACCTTGGACGGGGCCAAGGATATCGAAGGCTTGCTGGGAGC 152
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 1 ATGGAGCCCGGCAATTATGCCACCTTGGACGGGGCCAAGGATATCGAAGGCTTGCTGGGAGC 62
```

Amino acid alignment

```
Protein alignment
(U57051) Hoxb-13 [Mus musculus]
           Length = 286

 Score = 105 (47.5 bits), Expect = 8.3e-08, P = 8.3e-08
 Identities = 20/20 (100%), Positives = 20/20 (100%)


Query:      1 MEPGNYATLDGAKDIEGLLG 20
              ||||||||||||||||||||
Sbjct:      1 MEPGNYATLDGAKDIEGLLG 20
```

Figure 25: Alignment of Clone 2B1 with Hoxb-13

Results of BLAST searches (both BLASTN for nucleotide sequence, and BLASTP for putative translation product of the nucleotide sequence) against a non-redundant set of sequence databases returned a single match of very high significance to Hoxb-13. The reported alignments from both searches are represented here.

TAATAA-3', the observed interaction is very likely real in the context of the one-hybrid assay. However, Hoxb-13 is a member of a transcription factor family which is conserved from nematodes through vertebrates. Members of the Hox gene family are important in anterior/posterior patterning in embryogenesis and development and as such show highly restricted expression both spatially and temporally[94]. This has been shown to be specifically true of murine Hoxb-13. Given this restricted expression pattern relative to the observed wide permissiveness of fetal murine tissue for MVM replication, it seems unlikely that this observed interaction bears a functional relationship to MVM replication.

Examination of RIP60 Interaction with the IRS

The original screen which identified Hoxb-13[93] was intended to clone RIP60 (Replication Initiation Protein 60), the DNA-binding protein associated with initiation of the DHFR OBR. Given that RIP60 is a known origin-binding protein from a murine system and its binding affinity is for a 5'-TAA-3' repeat motif[95] which occurs twice within the MVM IRS region, the possibility that RIP60 was a component of the observed complexes binding this region suggested itself.

In order to examine this possibility, I obtained two polyclonal rabbit antisera from the Heintz lab (Dept. of Pathology, University of Vermont). The first of these, α-Z1, was raised against the bacterially expressed N-terminal zinc-finger domain of RIP60 fused with Glutathione-S-Transferase (GST); the second, α-Z2, was also raised against a fusion between the second zinc-finger domain of RIP60 and GST. Both antisera have been found to function in supershifting of RIP60-DNA complexes and in the

immunoprecipitation of these complexes when using nuclear extracts prepared from cells containing an RIP60 expression vector[96].

A preliminary experiment was conducted in which end-labelled 2021 bandshift probe was incubated with LA9 nuclear extract under conditions identical to those used for bandshifting, followed by a 30 minute incubation on ice with 2 μl of antisera, addition of prepared SAC I immunoprecipitin and incubation on ice for a further 30 minutes. The immunoprecipitate was washed three times in RIPA buffer and precipitated DNA quantitated by scintillation counting. Antisera used in this experiment were α-Z1 and α-Z2. Control prebleed sera from each of the rabbits used to generate α-Z1 and α-Z2 and a rabbit α-GST polyclonal antisera (Sadowski lab, Dept. of Biochemistry, UBC) were also tested. As shown in Table 9, neither of the α-RIP60 antisera allowed for the immunoprecipitation of the DNA-nuclear extract complex beyond control levels.

As a more sensitive assay for the presence of RIP60 in the complexes observed formed on the IRS by an LA9 nuclear extract, the ability of the α-Z1 and α-Z2 antisera to induce a supershift in either of the complexes observed with the 2021 fragment was tested. A series of bandshift reactions containing a radiolabelled 2021 probe were set up as for a normal bandshift reaction (see Materials and Methods), but prior to loading reactions on the gel, 2 μl of antisera were added to samples which were allowed to incubate on ice for a further 30 minutes. Antisera examined in this experiment were the same as in the previous immunoprecipitation experiment: α-Z1, α-Z2, prebleed sera on

Antibody	*Counts Immunoprecipitated*
αZ1	115
αZ1-prebleed	85
αZ2	60
αZ2-prebleed	64
αGST	190

Table 9: Anti-RIP60 Immunoprecipitation Results

Results of immunoprecipitation experiment with anti-RIP60 polyclonal antisera and controls. Neither of the RIP60 specific antisera (αZ1 or αZ2) demonstrates a capacity to immunoprecipitate more of a radiolabelled 2021 probe following incubation with LA9 nuclear extract, relative to prebleed or anti-GST controls.

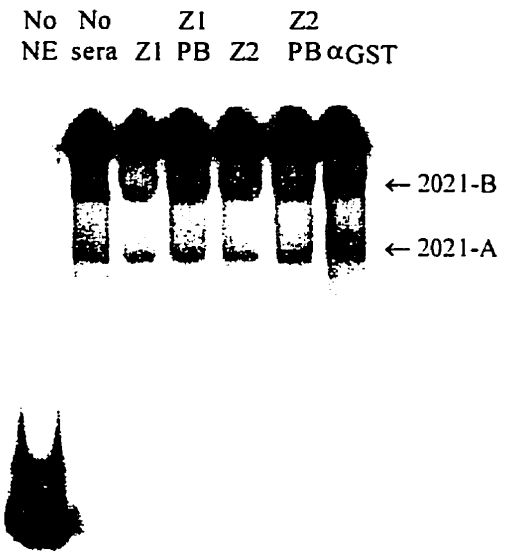


Figure 26: Anti-RIP60 supershift of 2021 probe

Results of supershifting experiment with anti-RIP60 polyclonal antisera and controls. Lane marked 'No NE' contains no nuclear extract; lane marked 'No sera' contains probe, nuclear extract, but no antisera; lanes marked 'Z1' and 'Z2' contain each of the indicated specific antisera; lanes marked 'Z1 PB' and 'Z2 PB' contain prebleed antisera for Z1 and Z2 repectively; and lane marked 'αGST' contains anti-GST antisera. No supershift or disruption of either the 'A' or 'B' specific complex is noted with any of the sera.

both of these, and rabbit α-GST polyclonal. The results (see Figure 26) indicate no evidence for either a supershifting or disruption of either the 2021-A or 2021-B complexes.

Taken together, these two experiments indicate that RIP60 is probably not a component of the complexes observed in the bandshift studies presented here. Whether this indicates that RIP60 does not in fact have a role in MVM replication is considered more fully in the Discussion.

Discussion

Previous evidence indicated the existence of an origin of DNA replication within a region of the MVM genome known as the IRS. Data presented in the current study supports this and expands on prior data by localizing specific sequence elements involved in origin activity and defining binding sites for nuclear factors within the region.

Construction of a comprehensive library of linker-scanning mutations across the region of interest allowed for its analysis in a minigenomic replicon system for sequences directly contributing to replication competence. As the method used in earlier studies to analyze replication of MVM minigenomes[70] was not found to be suitable for accurate quantitative analysis, a novel assay method was devised and implemented. Using this technique, a set of mutant viral minigenomes were examined for their capacity to replicate in the presence of the viral NS-1 protein.

Results of this analysis (see Figure 10) indicate the presence of three distinct regions within the IRS contributing to replication competence. The level of replication deficit observed by the substitution of each of these three elements with unrelated sequences (S2 and S7, S23, and S1) was approximately equal, suggesting that the elements contribute equally to the function of the viral right-end origin. Summation of the replication deficits observed in each case yields a value greater than 100% replication, an observation which is most easily explained by a functional interaction between the elements; that is, the net replication competence imparted by all elements together is greater than the sum of their individual contributions.

While the S1 mutation was able to reduce minigenome replication efficiency to 23% that of wild-type levels, individual substitution of each of the six nucleotides in the S1 mutation from wild-type sequence did not result in any detectable replication defect (see Table 8). This observation suggests that no single nucleotide within this element is individually responsible for the replication activity associated with the element. The small but definite replication hypercompetence exhibited by two of these point mutations (C2A and T8G) is puzzling, as is the weak hypercompetence observed with several of the scanner mutations (e.g. S3, Figure 10). While no wholly satisfactory explanation for these observations can be put forward at this time, it is interesting to note that similar observations have been made in studies of other viral origins[97].

Elements of the linker-scanning library were also analyzed by gel-retardation assay for their capacity to specifically interact with factors present in a nuclear extract. All observed interactions were localized to a central region of the IRS, between nucleotides 4524 and 4668. A probe spanning this region was observed to form two major specific complexes, 2021-A and 2021-B. A probe bearing nucleotide substitutions in the location of S8 was unable to effectively compete for the formation of complex 2021-B and thereby defines a sequence element important for the formation of this complex. Similarly, a probe with substitutions in the location of S22 lost capacity to compete for the formation of complex 2021-A. Weaker interactions were observed with probes bearing mutations S1 and S25 (which lost some competition ability for complex 2021-B) and S6 (which lost some competition ability for complex 2021-A). In all cases, an observed loss of competition ability towards one of the complexes was paralleled by a

weaker but apparent loss of competition ability towards the other complex. These data suggest that the binding of complexes responsible for the shifted species is a cooperative interaction, a hypothesis which is supported by other observations from both this and previous studies.

One line of evidence comes from the previous report that the IRS bound several distinct complexes as analyzed by gel retardation assay and that some of these complexes were the same on the RsaI 'A' and 'B' probes as evidenced by cross-competition[79]. The apparent discrepancy between these findings and the data presented here is clearly explained if one assumes that the binding of the 2021-A and 2021-B complexes is cooperative; as the boundary between the RsaI 'A' and 'B' fragments lies at viral nt. 4579 the major binding site for complex 2021-B is on one fragment (S8, on RsaI 'A') while the binding sites for complex 2021-A are on the other (see Figure 11). Studies with the RsaI fragments individually would then be expected to detect some shared complexes through protein-protein interactions. The lack of the major site for the binding of the other complex on each probe could readily be expected to result in a reduction of complex stability and its potential partial dissociation into constituent subunits resulting in a number of smaller shifted forms.

Another line of evidence for cooperative interaction comes from consideration of the dual-site mutation. In the present study, the factor binding sites observed are adjacent to (or in the case of sequences replaced in the S1 mutation, overlying) sequence elements required for efficient minigenome replication. This suggests a possible functional relationship between factor binding and viral origin function, which is strongly supported

by the results obtained with the S8/22 dual mutation. Neither of these mutations individually resulted in an apparent loss of minigenome replication competence, however the dual mutation reduces minigenome replication to the same basal level observed with the S2 and S1 mutations. The simplest interpretation of these observations would be that the factors responsible for formation of the 2021-A and 2021-B complexes are required for activity of the origin within the IRS. The already postulated cooperative interaction between the elements of these two complexes could account for the lack of a replication deficit in either of the individual mutations; binding of either complex and subsequent localization of the second complex through protein-protein interaction may be sufficient to activate replication to wild-type levels observed in the minigenome assay employed.

At present, the identity of the constituents of these complexes can not be stated. Evidence from other origins suggests that transcription factors, helicases, and single-stranded binding proteins are all potential candidates. While the interacting gene found through a one-hybrid screen (Hoxb-13) can most likely be discounted as being meaningful in viral replication due to limitations in its expression pattern relative to observed viral tropism, the fact that this gene was also identified by other workers looking for a known murine origin-binding protein (RIP60) may not be coincidental. The recovery of Hoxb-13 by Heintz *et. al.* by interaction with a DNA sequence known to bind RIP60[93] suggests that these two proteins happen to share recognition sequence specificity; the preferential identification of the Hox gene in both studies may merely indicate that this gene product is better able to work out of its native context than is RIP60.

The reported RIP60 binding element is an $(ATT)_n$ repeat sequence, and while this does not seem to be associated with either replication-active or factor-binding sites reported in this study, the sequence (TAATAA) does occur in both the elements replaced by scanner mutations S22 and S1 (in opposite orientations at the two sites). Given that this is simply the inverse complement of the reported RIP60 binding site, shifted by one nucleotide in phase, RIP60 would seem to be a likely candidate for a potential IRS-binding factor. The observed capacity of this protein to associate into large aggregates on the DHFR OBR in association with several other proteins, if representative of RIP60's normal mode of action, could explain both the large apparent size of the complexes observed by gel-retardation assay and the apparent cooperativity between binding of the complexes noted in this report. The fact that both immunoprecipitation and supershift experiments reported here failed to indicate RIP60 as a component of the complexes observed in bandshift experiments strongly suggests that this protein is not involved; however the circumstantial evidence seems strong enough to warrant further *in vivo* studies before RIP60 involvement in MVM replication is dismissed. The recent cloning of full-length RIP60 in a mammalian expression vector and the demonstration that cotransfection of this construct with polyoma virus results in increased viral replication[96] suggests one technique through which this might be examined.

The apparent lack of a replication deficit in full-length virus following introduction of the S1 mutation is surprising, but perhaps not innately contradictory to data from the minigenome system. Replication capacity in the full-length viral system was analyzed by measurement of final viral titres, as opposed to levels of viral DNA in

the minigenome-based assay. Given that NS-1 is known to be cytotoxic, it is conceivable that in native infections replication capacity of an infected cell is limited by NS-1 level: that is, replication proceeds until such time as NS-1 levels reach a threshold resulting in cell death and release of encapsidated viral progeny. A slower-replicating viral genome would result in fewer copies of the NS-1 gene available for transcription and might thereby allow for a longer period of effective replication ending at the same NS-1 threshold, and thus at approximately the same number of viral particles as the wild-type virus. If this is indeed the case, careful timecourse analysis of replicative intermediates in infected cells should demonstrate an initial lag in mutant viral replication relative to wild-type. While a preliminary attempt at such an analysis was made, inconclusive results were obtained and have not been presented in this study.

Further insights into the activity of the origin of DNA replication within the IRS are provided by studies reported here in which this origin, in conjunction with one viral telomere, were used sucessfully to drive replication of a linear construct containing unrelated sequences. Replication of pJBR1 constructs clearly and unequivocally demonstrate that the IRS alone functions as an origin of replication irrespective of any contributions to replication capacity the viral genome gains from a postulated left-terminal origin. While prior models of MVM replication have not included any direct role for the viral NS-1 protein at the replication origin, the apparent lack of replication of these constructs in the absence of NS-1 suggests that such an interaction could be functionally important. Upon consideration however it seems more likely that the reason for this requirement could lie in the nature of the pJBR1 construct itself. While the viral

5' palindromic sequences end at the ClaI site (see Figure 20) the construct is linearized with EcoRI prior to transfection leading to the presence of a 23 nt. 'tail' which is not capable of forming a stable duplex when the terminus rearranges from a linear to hairpin conformation. As this extension is not base-paired across from single-stranded template, it cannot serve as a primer for strand extension back to a double-stranded form (Figure 21, panel 3 of LSO model). While cellular DNA repair mechanisms might be expected to remove this unpaired segment, it is possible that this does not occur to a large enough extent to yield detectable product in this assay and that NS-1 is needed to nick the terminus and release the extension. Alternatively, NS-1 may be required in some more cryptic role at the IRS origin; and while the purpose or mechanism of such interaction is not suggested at this time, the known helicase function of NS-1[28] lends itself to at least one plausible hypothesis in which it serves a role analogous to that of dnaB at the *E. coli* OriC origin[52].

While evidence existed to suggest the presence of a second origin in the viral left-end terminus, no data was previously available to accurately measure its contribution to viral replication. Data presented here addresses this by allowing for an upper limit to its activity relative to the origin within the IRS to be estimated at 20%. This conclusion comes from the assumption that if any of the linker-scanning mutations examined in the minigenome based assay completely abolished activity of the right-hand origin, any residual replication activity must arise from the left-hand origin. All of the replication defective mutants (except S7, which may very well only touch on the edge of a replication element) reduce minigenome replication efficiency to approximately the same

level (20%) suggesting the possibility that each of these mutations (S2, S23, S1, and the S8/22 dual mutant) completely abrogate function of the right-hand origin. This possibility is supported by observations from the pJBR1 based analysis of these same mutations in which all of the mutants lose replication capacity. If this is indeed the case, the simultaneous requirement for all of these discrete sequence elements for function of the right-hand origin would indicate there is a functional interaction between the elements.

A model is presented in this study to address the question of how parvoviral genomes effect the rearrangement of their terminal palindrome(s) (either a specific one in the case of MVM and other autonomous parvoviruses, or both for the dependoviruses) from an extended double-stranded linear form to an intramolecularly-hybridized hairpin form. At present, no other models exist to address this problem. Experimental results with the pJBR1 construct lend support to the proposed mechanism whereby a leading-strand only (LSO) origin displaces one strand of the terminal palindrome, allowing the other to spontaneously form into the hairpin structure required by viral replication models. In particular, the LSO model predicts that pJBR1 should produce three classes of replicative intermediates, in contrast to a canonical bidirectional leading and lagging strand origin which would predict a single class of replicative intermediates. The strongest evidence presented here in support of this model is the analysis of replication products of pJBR1 by one-dimensional neutral agarose gel electrophoresis (Figure 22). This reveals three DNA species produced uniquely under conditions in which full-length input plasmid is replicated (supply of NS-1 in *trans* being a required condition).

Examination of these species by two-dimensional neutral-alkaline electrophoresis (Figure 23), while unfortunately complex due to the heterogeneous nature of the many replicative intermediate forms involved, suggests that the observed products contain forms corresponding to the predicted Forms 1, 4a, and 4b. These results, while not uniquely interpretable with the LSO model, suggest that it should be considered as a possible model for IRS activity. With further refinements this experimental technique may prove useful in firmly elucidating the IRS's mechanism of action.

There are several attractive features of this model. First, presence of an origin would explain the virus' requirement for a host-cell S phase to replicate, a requirement which otherwise is puzzling when one considers that all of the intermediates shown in Figure 4 would seem to be suitable templates for repair synthesis and should thus be replicated regardless of host cell cycle. Evidence in support of the leading-strand only nature of the proposed origin model comes from studies by Goulian *et. al.*[98], which reports that Okazaki fragments cannot be detected in parvoviral replication. The idea of an exclusively leading-stranded origin brings to mind mitochondrial DNA replication, which proceeds in a Pol γ dependent leading-strand only fashion. Intriguingly, studies by Kollek *et. al.*[99,100] employing polymerase inhibitors in the course of native H1 infections led to an identification of an early Pol α/Primase dependent step (initiation) followed by a longer Pol γ dependent process.

Two recent reports in the literature have bearing on this model. Cossons *et. al.*[101] observed rearrangement of the 5' extended palindrome of MVM from a linear extended form to a hairpin-containing structure *in vitro* and concluded on the basis of selective

sensitivity of this process to polymerase inhibitors that pol δ was required for the rearrangement. Of particular interest is their observation that while hairpin formation was observed, extension synthesis to form the dRF did not occur. Current models for the action of pol δ indicate it acts to synthesize leading and lagging strands in a coordinate manner. The replication model presented here would suggest that an *in vitro* system in which pol δ is acting at this viral origin would be able to form an incomplete hairpin structure at the 5' terminus but not be able to prime synthesis to the dRF replicative intermediate. It is conceivable that the apparent involvement of pol δ in *in vitro* studies by Cossons *et. al.* may not reflect the *in vivo* mechanism.

Other *in vitro* studies by Baldauf *et. al.*[102] demonstrated that an LA9 cellular extract was capable of supporting the conversion of viral genomes into a covalently closed (cRF) monomer form. Only when the extract was supplemented with purified NS-1 was terminal resolution of the cRF, extension to the mRF, and the appearance of dRF species observed. The ability of the supplemented extract to support replication to the level of dRF was suggested by the authors to result entirely from site-specific nicking by NS-1 as a requirement to open the covalently closed cRF species and allow for subsequent generation of the linear 5' termini, a prerequisite for terminal rearrangement and elongation to the dimeric species. This explanation is somewhat unsatisfactory given that the system without NS-1 in the Cossons *et. al.* study[101] achieved both the linear mRF and some form of terminal rearrangement. A possibility not previously considered is that NS-1 may play an obligate role in suppressing lagging-strand synthesis from the IRS

origin, either by influencing the choice of polymerase used or through a more direct mechanism.

Comparison of the results reported in this thesis with earlier results from our lab[70,79] clarifies some observations but raises new questions. Previously it was observed that while deletions in the right side of minigenomic constructs as far as nt. 4489 had no effect, deletions to nt. 4636 abolished replication competence in LA9 cells and resulted in roughly a 70% loss in replication competence in COS-7 cells (thereby defining Element A as nt. 4489 through 4636), and extension of the deletion to nt. 4695 cause a complete loss of replication competence in COS-7 cells (thereby defining Element B as nt. 4637 to 4695)[70]. In contrast, all of the elements contributing to replication competence noted in this study lie within Element A, despite the fact that the studies were done in COS-7 cells. Given that the apparent sensitivity of the replication assay employed in the current study is on the order of 32%, it is possible that the functions within Element B were too weak to be detected.

Data presented here agrees well with previously reported results regarding the rescue of replication deficient minigenomes by the RsaI "A" and "B" fragments. Given that RsaI"A" contains the site under mutants S2 and S7 while RsaI"B" contains the sequences under mutants S23 and S1, it is not surprising that both fragments were capable of restoring replication competence. Similarly, previous results using the RsaI"A" (containing the S8 site) and "B" (containing the S22, S1, and S25 sites) fragments as bandshift probes had suggested that at least one common complex was formed by both probes, and that it was in turn composed of at least two other

complexes[79]. This is born out by observations here which indicate a cooperative interaction between the two major complexes formed in this region whereby binding of one of the two will result in at least some binding of the other. It is tempting to postulate that the observed greater rescue of replication capacity by the RsaI"A" fragment than the "B" fragment is related to the results shown here which indicate that the site on RsaI"A" for complex 2021-B is a higher-affinity site than that on the "B" fragment, however at the moment this must remain purely speculative.

Previous footprinting of host factors binding in the IRS indicated that the activity referred to as MRFB5[79] protected a bipartite region, with 'Site I' overlying scanner mutations S22 and S1 and 'Site II' overlying mutations S26 and S15 of this study. The observations presented here localize the majority of binding activity under 'Site I' along with some sequences important for replication, while only weak binding activity and no replication activity is associated with sequences under 'Site II'. The apparent cooperativity of binding between the complexes observed would suggest that the protection of 'Site II' may arise from factors being strongly bound at 'Site I'. Considering the small gap between Sites I and II (6 bp) a simple model to account for this would have the viral genome bend sharply between the regions of protection, such that 'Site II' arises from the DNA coming into close proximity with the backside of the complex localized to 'Site I' (see Figure 27). Such dramatic kinking or bending of DNA has been a common feature of many origins characterized to date. Notably, RIP60 has been shown to induce DNA distortions in the DHFR OBR[64] of a nature similar to that proposed here; however, in view of the lack of any direct evidence to support either

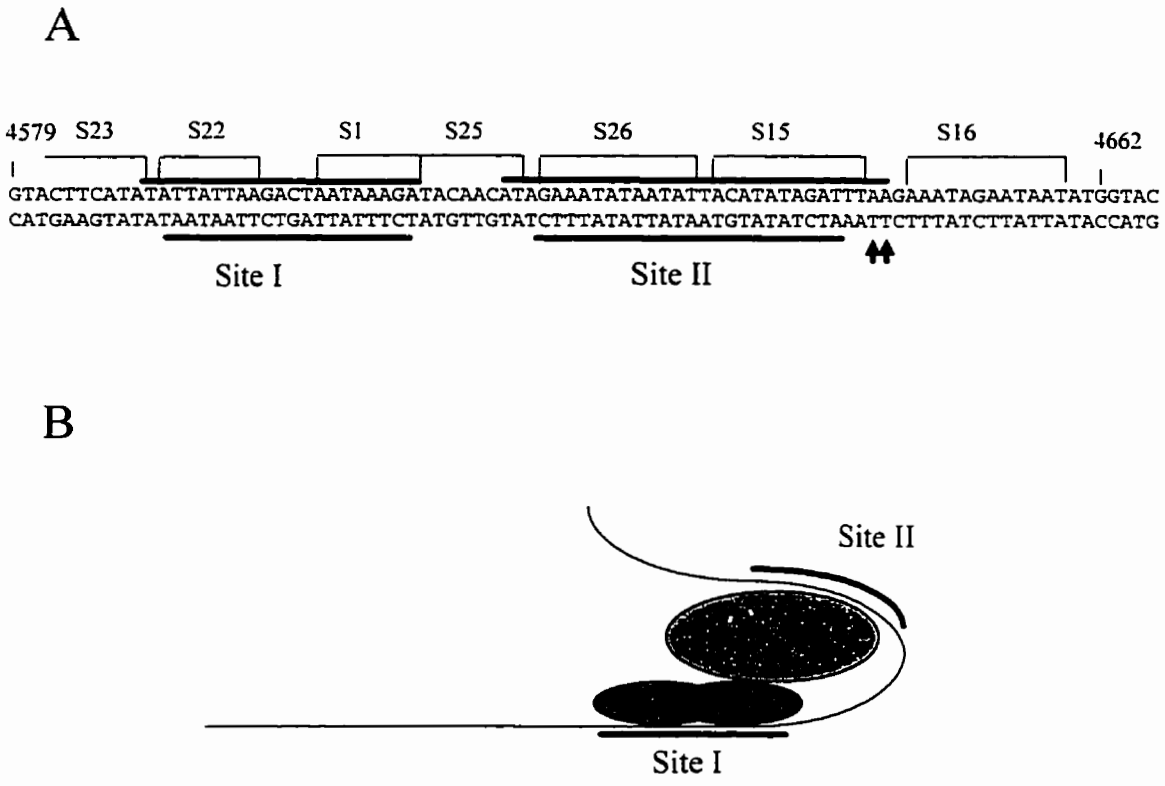


Figure 27: Hypothetical Bent Model for Protection of Sites I and II

Diagram of hypothetical bent-DNA model for origin of Site I and Site II footprints from a single protein complex with sequence specificity conferred by Site I. Panel A diagrams the RsaI 'B' fragment and associated MRFB5 footprint of Reference 79; heavy overbars and two arrows represent regions of DNAse I protection. The locations of linker scanning mutations within this region are indicated above the sequence along with their designations. Panel B is a representation of a protein complex localized to Site I and inducing a DNA bend such that a second region (Site II) contacts the back side of the complex and is protected from DNAse I cleavage.

RIP60 binding to the IRS or the postulated DNA bending in this region this observation must be considered purely conjectural.

Taken together, evidence presented here indicates the IRS may function as a multipartite origin of DNA replication consisting of a number of short sequence elements which interact in a cooperative fashion. Some of these elements act through the binding of nuclear factors and some appear to act only as required sequence elements, independent of any observable protein-binding function. Such organization is reminiscent of other origins of replication characterized to date, both from viral systems (human papillomavirus type 11, SV40) and simple eukaryotic systems (ARS1 of *S.cerevisiae*), in which several short discreet sequence elements contribute to the initiation of replication either through the localization of protein factors or through intrinsic properties of the sequence element.

Appendix A

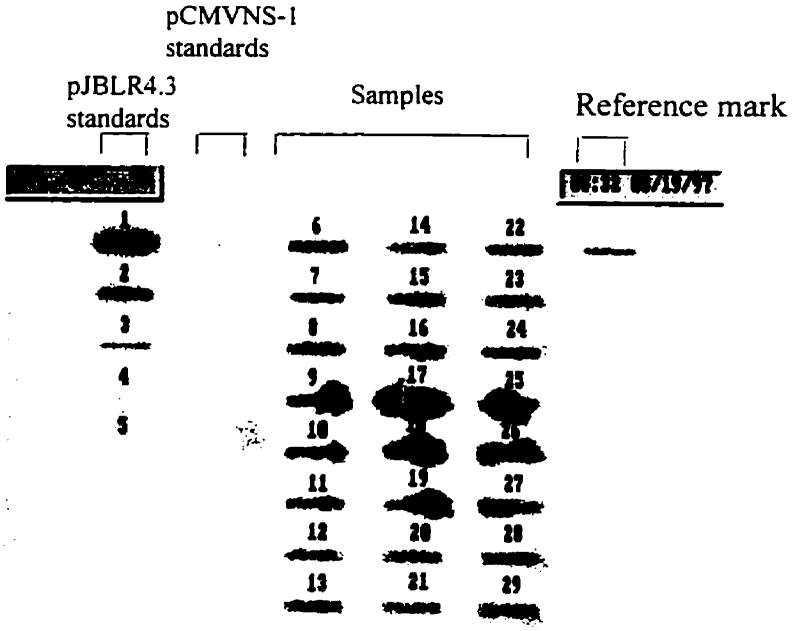
This appendix contains the original data from the replication assays referred to in the main body of the thesis, along with the quantitation and preliminary data analyses which are summarised in Table 4, Table 6, and Table 8.

Data is presented in the following manner: a description of the assay and what clones were examined is followed on the next page by images of the associated slot-blot following hybridization with Vecpro and JAVA oligonucleotides. An image of the blot after stripping between hybridizations is not presented. Numbers above each of the slots correspond to the sample numbers for the quantitation data, which follows each set of blot images. The next page presents quantitation data for the pJBLR4.3 standard curve associated with each blot image (samples 1 through 5), with the amount of standard in each sample and the associated quantitation value presented in a table. These values are plotted graphically with a least-squares linear regression curvefit displayed. The equation of the calculated curvefit is presented in each case. Data on this page correspond to the blots on the preceeding page, i.e. the upper data table and graph refer to the upper blot image preceeding. Finally, the fourth page contains the quantitation values for each of the experimental samples from the two blots, and the RE, SRE, and SRRE values calculated from this data.

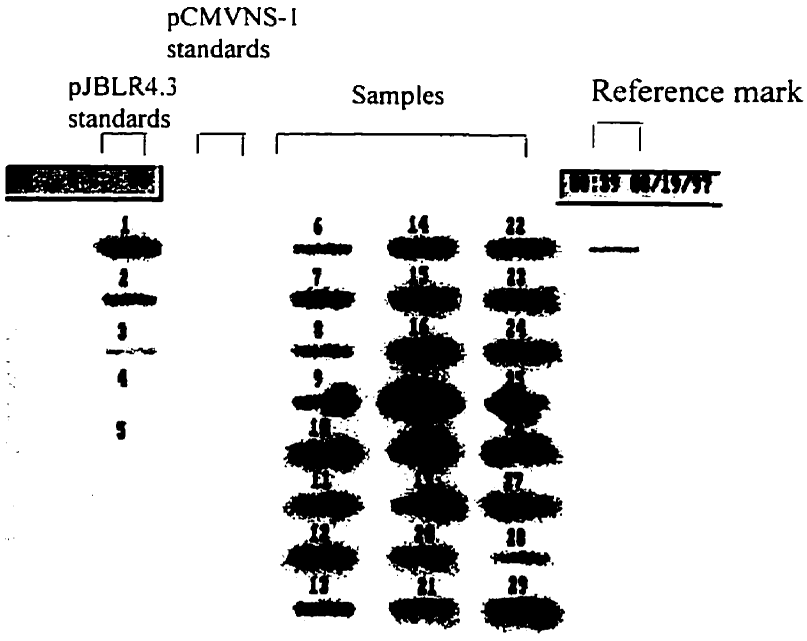
Note that in Assays 1, 2, and 3, there are two samples on each blot (segments 26 and 27) which are not analyzed. Analysis of these clones after the assay indicated the presence of extraneous mutations in these clones and data pertaining to them was discarded.

Replication Assay 1:

Sample #	Contains
1	pJBLR4.3 104 ng
2	pJBLR4.3 21 ng
3	pJBLR4.3 4.3 ng
4	pJBLR4.3 833 pg
5	pJBLR4.3 167 pg
6	pJBLR4.3
7	pJBLR4.3 + pCMVNS-1
8	pJBLR4.3S1 + pCMVNS-1
9	pJBLR4.3S2 + pCMVNS-1
10	pJBLR4.3S3 + pCMVNS-1
11	pJBLR4.3S4 + pCMVNS-1
12	pJBLR4.3S5 + pCMVNS-1
13	pJBLR4.3S7 + pCMVNS-1
14	pJBLR4.3S8 + pCMVNS-1
15	pJBLR4.3S12 + pCMVNS-1
16	pJBLR4.3S13 + pCMVNS-1
17	pJBLR4.3S14 + pCMVNS-1
18	pJBLR4.3S15 + pCMVNS-1
19	pJBLR4.3S16 + pCMVNS-1
20	pJBLR4.3S17 + pCMVNS-1
21	pJBLR4.3S18 + pCMVNS-1
22	pJBLR4.3S20 + pCMVNS-1
23	pJBLR4.3S21 + pCMVNS-1
24	pJBLR4.3S22 + pCMVNS-1
25	pJBLR4.3S23 + pCMVNS-1
26	Sample rejected due to mutation in clone
27	Sample rejected due to mutation in clone
28	pJBLR4.3
29	pJBLR4.3 + pCMVNS-1



Assay 1: Vecpro probe

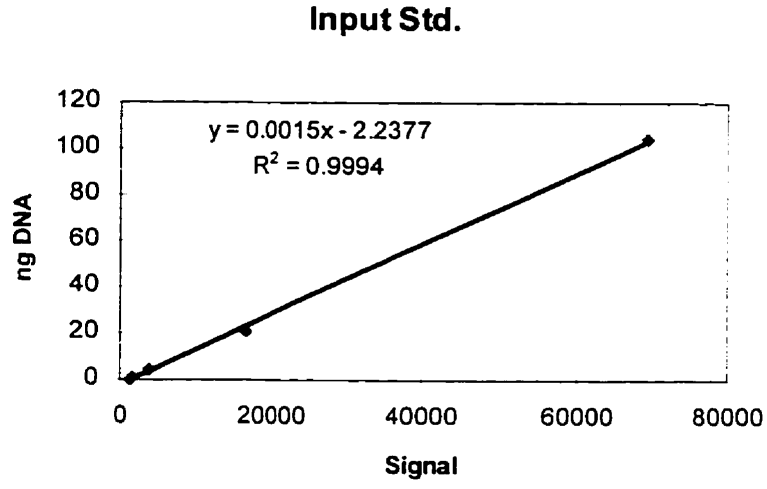


Assay 1: JAVA probe

Quantitation data and standard curves, Assay 1

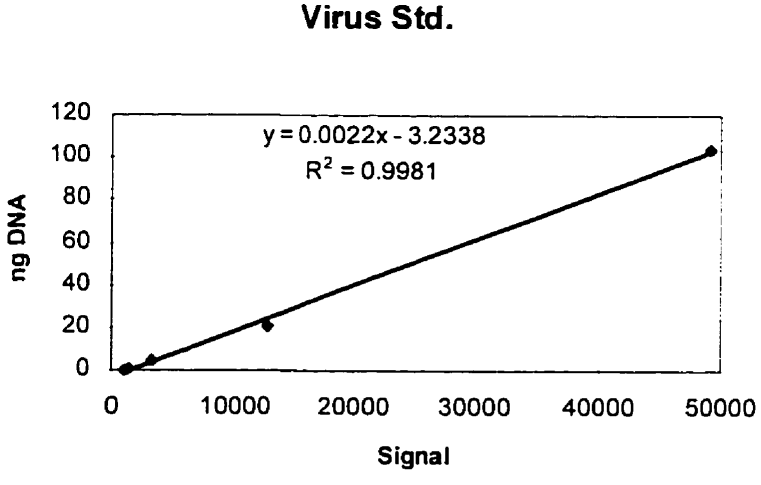
pJBLR4.3 Standards, Vecpro probe:

Segment #	*Volume*	*ng Std.*
1	69275.3	104
2	16473.21	21
3	3810.85	4.2
4	1593.14	0.833
5	1396.27	0.167



pJBLR4.3 standards, JAVA probe:

Segment #	*Volume*	*ng Std.*
1	49109.33	104
2	12776.66	21
3	3237.12	4.3
4	1362.37	0.833
5	1103.31	0.167

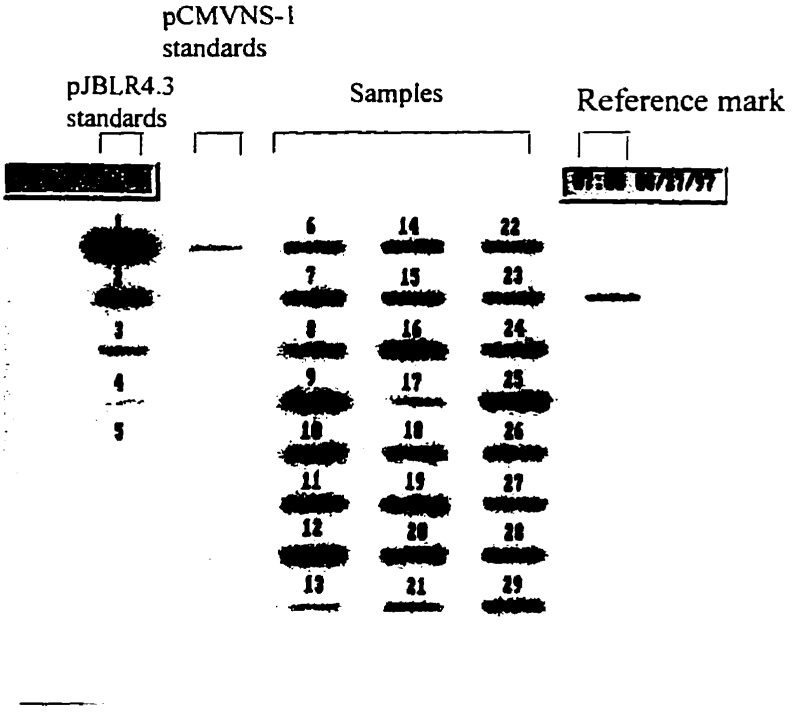


Sample quantitation, Assay 1

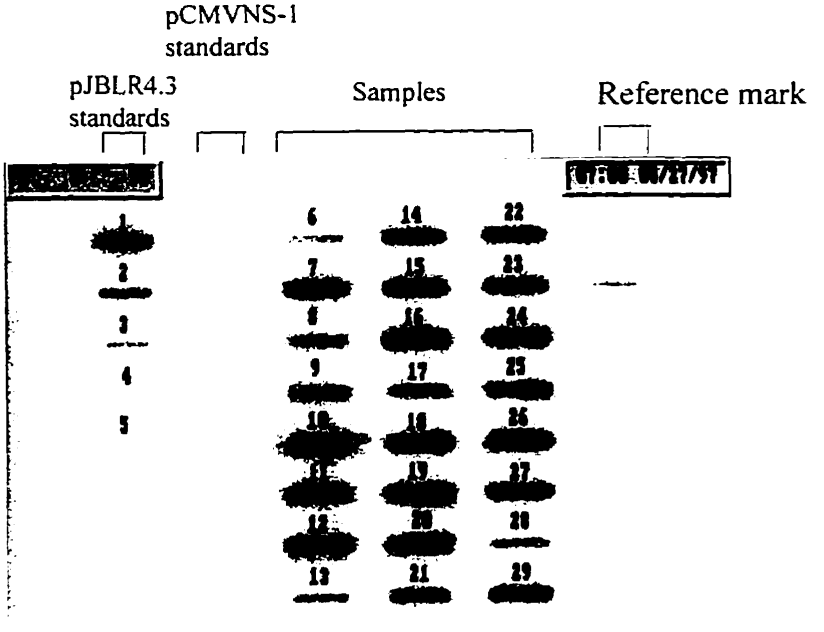
Clone	Vecpro signal	ng input	JAVA signal	ng out	RE	Scaled RE	Scaled RRE
4.3	12196.75	16.06	9305.84	17.24	0.07	-0.04	-1%
4.3 + pCMVNS1	8043.04	9.83	42665.96	90.63	8.22	8.11	117%
S1	10252.15	13.14	11126.45	21.24	0.62	0.50	7%
S2	25301.19	35.71	31617.64	66.33	0.86	0.74	11%
S3	11239.26	14.62	73379.47	158.20	9.82	9.70	140%
S4	9571.79	12.12	51424.86	109.90	8.07	7.95	115%
S5	7675.82	9.28	44738.58	95.19	9.26	9.15	132%
S7	10017.46	12.79	16878.50	33.90	1.65	1.54	22%
S8	9314.96	11.73	47557.20	101.39	7.64	7.53	109%
S12	10993.86	14.25	62847.88	135.03	8.47	8.36	121%
S13	11913.98	15.63	73074.39	157.53	9.08	8.96	129%
S14	35788.61	51.45	125128.90	272.05	4.29	4.17	60%
S15	30666.83	43.76	110258.68	239.34	4.47	4.35	63%
S16	24587.7	34.64	105197.26	228.20	5.59	5.47	79%
S17	8947.37	11.18	36920.45	77.99	5.97	5.86	84%
S18	8233.37	10.11	40294.40	85.41	7.45	7.33	106%
S20	9420.97	11.89	41864.16	88.87	6.47	6.36	92%
S21	10347.32	13.28	38579.58	81.64	5.15	5.03	73%
S22	10498.32	13.51	41249.48	87.52	5.48	5.36	77%
S23	26365.71	37.31	37001.07	78.17	1.10	0.98	14%
4.3	10658.85	13.75	8693.15	15.89	0.16	0.04	1%
4.3+PCMVNS1	12421.34	16.39	52703.15	112.71	5.88	5.76	83%

Replication Assay 2:

Sample #	Contains
1	pJBLR4.3 104 ng
2	pJBLR4.3 21 ng
3	pJBLR4.3 4.3 ng
4	pJBLR4.3 833 pg
5	pJBLR4.3 167 pg
6	pJBLR4.3
7	pJBLR4.3 + pCMVNS-1
8	pJBLR4.3S1 + pCMVNS-1
9	pJBLR4.3S2 + pCMVNS-1
10	pJBLR4.3S3 + pCMVNS-1
11	pJBLR4.3S4 + pCMVNS-1
12	pJBLR4.3S5 + pCMVNS-1
13	pJBLR4.3S7 + pCMVNS-1
14	pJBLR4.3S8 + pCMVNS-1
15	pJBLR4.3S12 + pCMVNS-1
16	pJBLR4.3S13 + pCMVNS-1
17	pJBLR4.3S14 + pCMVNS-1
18	pJBLR4.3S15 + pCMVNS-1
19	pJBLR4.3S16 + pCMVNS-1
20	pJBLR4.3S17 + pCMVNS-1
21	pJBLR4.3S18 + pCMVNS-1
22	pJBLR4.3S20 + pCMVNS-1
23	pJBLR4.3S21 + pCMVNS-1
24	pJBLR4.3S22 + pCMVNS-1
25	pJBLR4.3S23 + pCMVNS-1
26	Sample rejected due to mutation in clone
27	Sample rejected due to mutation in clone
28	pJBLR4.3
29	pJBLR4.3 + pCMVNS-1



Assay 2: Vecpro probe

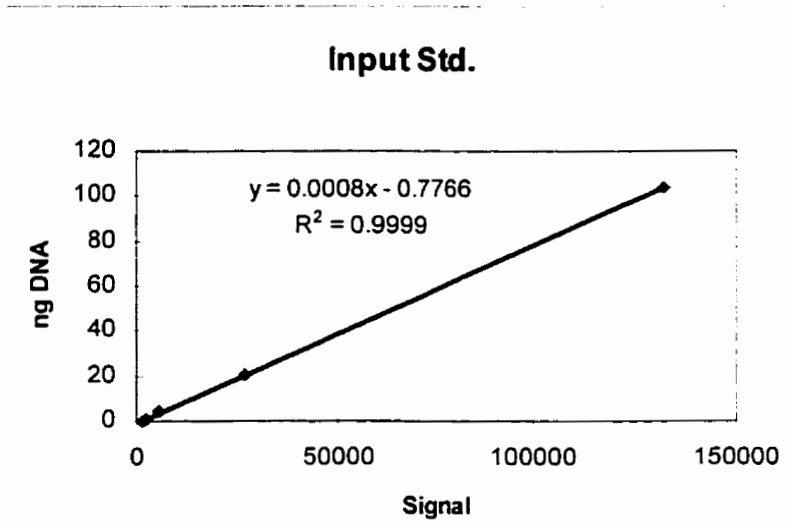


Assay 2: JAVA probe

Quantitation data and standard curves, Assay 2

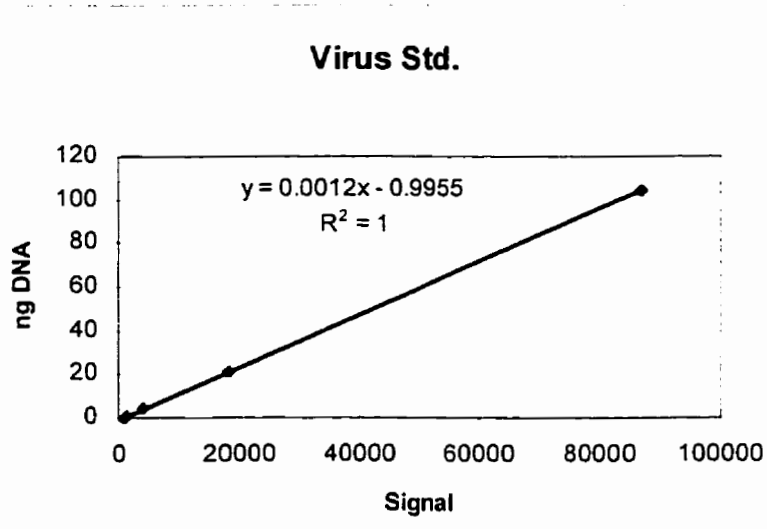
pJBLR4.3 Standards, Vecpro probe:

Segment #	*Volume*	*ng Std.*
1	132075.09	104
2	26784.39	21
3	5842.47	4.2
4	2677.78	0.833
5	1468.4	0.167



pJBLR4.3 Standards, JAVA probe

Segment #	*Volume*	*ng Std.*
1	87083.16	104
2	18173.6	21
3	4255.07	4.3
4	1542.38	0.833
5	1121.71	0.167

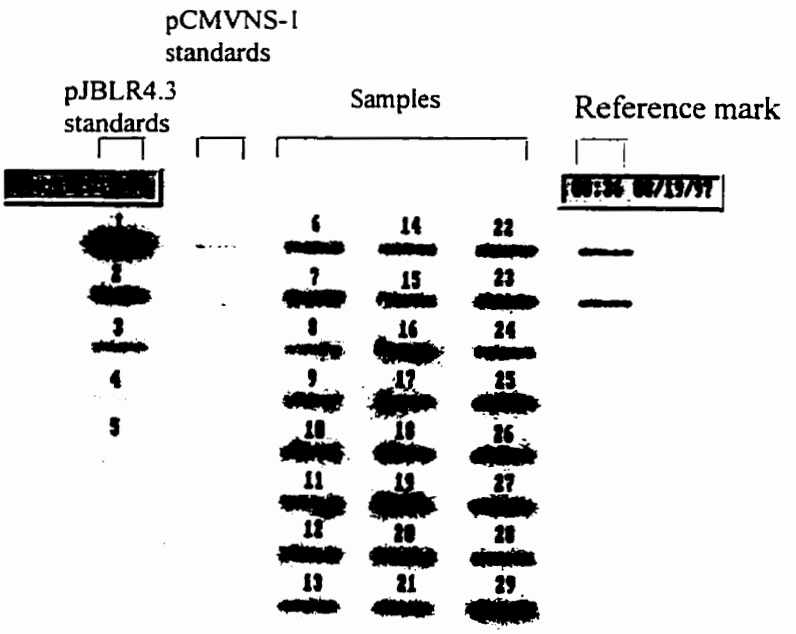


Sample quantitation, assay 2:

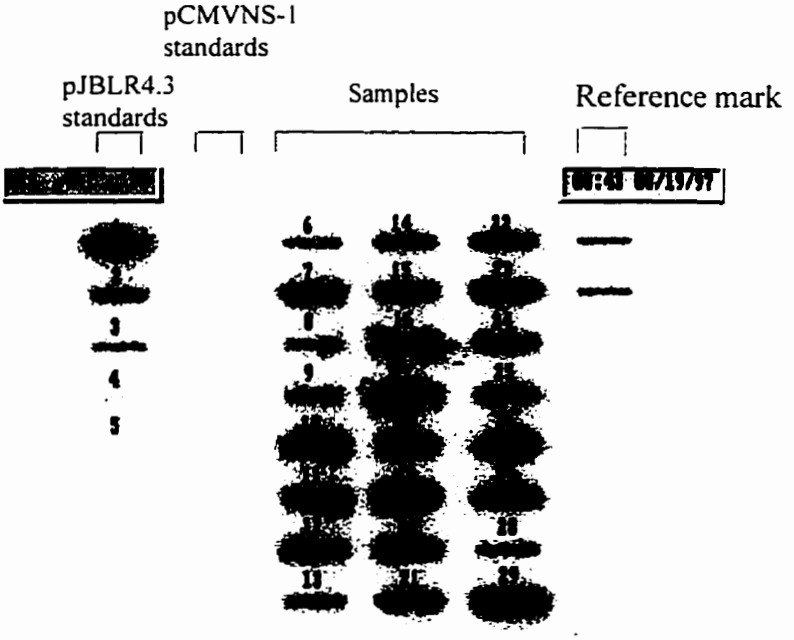
Clone	Vecpro signal	ng input	JAVA signal	ng out	RE	Scaled RE	Scaled RRE
4.3	11091.62	14.40	8145.05	8.78	-0.39	-0.01	0%
4.3 + pCMVNS1	17848.18	24.53	86226.37	102.48	3.18	3.56	101%
S1	11914.22	15.63	21435.18	24.73	0.58	0.97	28%
S2	19835.21	27.52	24652.60	28.59	0.04	0.42	12%
S3	14669.8	19.77	87262.68	103.72	4.25	4.63	132%
S4	14979.18	20.23	92693.44	110.24	4.45	4.83	138%
S5	20870.48	29.07	99572.12	118.49	3.08	3.46	99%
S7	4408.23	4.37	12093.18	13.52	2.09	2.47	71%
S8	12284.32	16.19	59644.40	70.58	3.36	3.74	107%
S12	12076.08	15.88	88700.34	105.44	5.64	6.03	172%
S13	21772.28	30.42	116229.23	138.48	3.55	3.94	112%
S14	5521.13	6.04	28067.61	32.69	4.41	4.79	137%
S15	11771.59	15.42	77429.59	91.92	4.96	5.35	152%
S16	17449.49	23.94	101138.92	120.37	4.03	4.41	126%
S17	16678.09	22.78	98894.58	117.68	4.17	4.55	130%
S18	7265.87	8.66	55445.55	65.54	6.57	6.95	198%
S20	12149.75	15.99	61855.64	73.23	3.58	3.97	113%
S21	11087.72	14.39	66846.18	79.22	4.50	4.89	139%
S22	15394.35	20.85	100636.55	119.77	4.74	5.13	146%
S23	20302.2	28.22	36947.09	43.34	0.54	0.92	26%
4.3	14960.72	20.20	11184.63	12.43	-0.38	0.00	0%
4.3+PCMVNS1	13858.72	18.55	63773.15	75.53	3.07	3.46	98%

Replication Assay 3:

Sample #	Contains
1	pJBLR4.3 104 ng
2	pJBLR4.3 21 ng
3	pJBLR4.3 4.3 ng
4	pJBLR4.3 833 pg
5	pJBLR4.3 167 pg
6	pJBLR4.3
7	pJBLR4.3 + pCMVNS-1
8	pJBLR4.3S1 + pCMVNS-1
9	pJBLR4.3S2 + pCMVNS-1
10	pJBLR4.3S3 + pCMVNS-1
11	pJBLR4.3S4 + pCMVNS-1
12	pJBLR4.3S5 + pCMVNS-1
13	pJBLR4.3S7 + pCMVNS-1
14	pJBLR4.3S8 + pCMVNS-1
15	pJBLR4.3S12 + pCMVNS-1
16	pJBLR4.3S13 + pCMVNS-1
17	pJBLR4.3S14 + pCMVNS-1
18	pJBLR4.3S15 + pCMVNS-1
19	pJBLR4.3S16 + pCMVNS-1
20	pJBLR4.3S17 + pCMVNS-1
21	pJBLR4.3S18 + pCMVNS-1
22	pJBLR4.3S20 + pCMVNS-1
23	pJBLR4.3S21 + pCMVNS-1
24	pJBLR4.3S22 + pCMVNS-1
25	pJBLR4.3S23 + pCMVNS-1
26	Sample rejected due to mutation in clone
27	Sample rejected due to mutation in clone
28	pJBLR4.3
29	pJBLR4.3 + pCMVNS-1



Assay 3: Vecpro probe

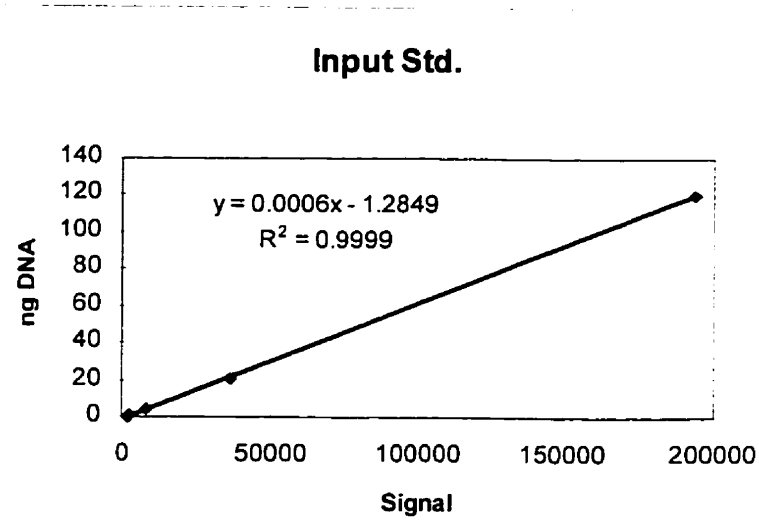


Assay 3: JAVA probe

Quantitation data and standard curves, Assay 3

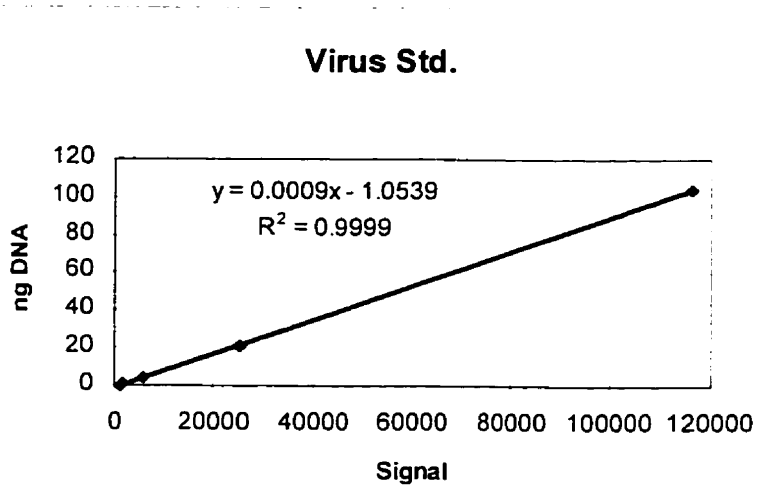
pJBLR4.3 Standards, Vecpro probe:

Segment #	*Volume*	*ng Std.*
1	193766.25	120.8
2	36756.64	21
3	8484.9	4.2
4	2490.2	0.833
5	2302.47	0.167



pJBLR4.3 Standards, JAVA probe:

Segment #	*Volume*	*ng Std.*
1	116349.04	104
2	25384.07	21
3	6002.69	4.3
4	1586.22	0.833
5	1074.01	0.167

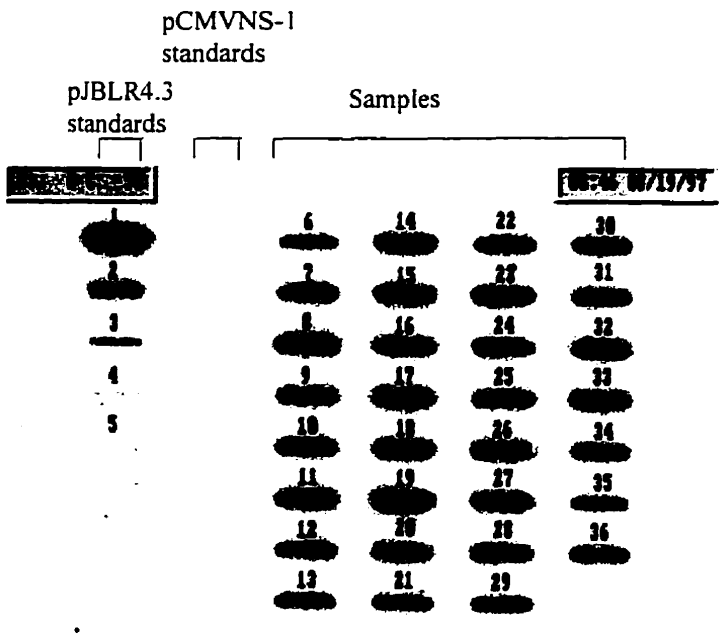


Sample quantitation, assay 3:

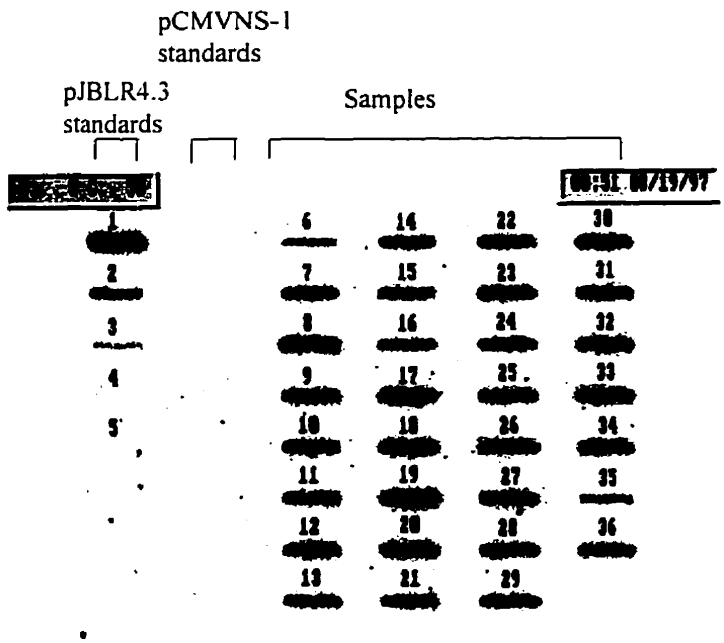
Clone	Vecpro signal	ng input	JAVA signal	ng out	RE	Scaled RE	Scaled RRE
4.3	13898.08	18.61	7997.44	6.14	-0.67	-0.01	-1%
4.3 + pCMVNS1	24390.4	34.35	58702.22	51.78	0.51	1.16	99%
S1	11357.45	14.80	11000.29	8.85	-0.40	0.25	22%
S2	13556.7	18.10	14282.34	11.80	-0.35	0.31	26%
S3	18488.11	25.49	75933.70	67.29	1.64	2.29	195%
S4	18059.07	24.85	69871.40	61.83	1.49	2.14	182%
S5	18559.86	25.60	64943.57	57.40	1.24	1.90	161%
S7	14802.05	19.97	18576.77	15.67	-0.22	0.44	37%
S8	13741.32	18.37	34892.54	30.35	0.65	1.31	111%
S12	18268.38	25.16	53778.67	47.35	0.88	1.54	131%
S13	23074.46	32.37	64474.59	56.97	0.76	1.41	120%
S14	21021.66	29.29	55393.27	48.80	0.67	1.32	112%
S15	21499.48	30.01	55173.67	48.60	0.62	1.27	108%
S16	27479.13	38.98	74602.29	66.09	0.70	1.35	115%
S17	23792.48	33.45	62398.59	55.10	0.65	1.30	111%
S18	19289.21	26.70	52464.09	46.16	0.73	1.38	118%
S20	16419.62	22.39	41410.89	36.22	0.62	1.27	108%
S21	18240.24	25.12	69960.63	61.91	1.46	2.12	180%
S22	14386.79	19.34	52263.24	45.98	1.38	2.03	173%
S23	22191.96	31.05	24748.40	21.22	-0.32	0.34	29%
4.3	19821.4	27.49	12136.66	9.87	-0.64	0.01	1%
4.3+PCMVNS1	35610.5	51.18	88375.52	78.48	0.53	1.19	101%

Replication Assay 4:

Sample #	Contains
1	pJBLR4.3 104 ng
2	pJBLR4.3 21 ng
3	pJBLR4.3 4.3 ng
4	pJBLR4.3 833 pg
5	pJBLR4.3 167 pg
6	pJBLR4.3
7	pJBLR4.3 + pCMVNS-1
8	pJBLR4.3S25 + pCMVNS-1
9	pJBLR4.3S25 + pCMVNS-1
10	pJBLR4.3S25 + pCMVNS-1
11	pJBLR4.3S26 + pCMVNS-1
12	pJBLR4.3S26 + pCMVNS-1
13	pJBLR4.3S26 + pCMVNS-1
14	pJBLR4.3S8/22 + pCMVNS-1
15	pJBLR4.3S8/22 + pCMVNS-1
16	pJBLR4.3S8/22 + pCMVNS-1
17	pJBLR4.3T1C + pCMVNS-1
18	pJBLR4.3T1C + pCMVNS-1
19	pJBLR4.3T1C + pCMVNS-1
20	pJBLR4.3C2A + pCMVNS-1
21	pJBLR4.3C2A + pCMVNS-1
22	pJBLR4.3C2A + pCMVNS-1
23	pJBLR4.3T3G + pCMVNS-1
24	pJBLR4.3T3G + pCMVNS-1
25	pJBLR4.3T3G + pCMVNS-1
26	pJBLR4.3T4A + pCMVNS-1
27	pJBLR4.3T4A + pCMVNS-1
28	pJBLR4.3T4A + pCMVNS-1
29	pJBLR4.3A6C + pCMVNS-1
30	pJBLR4.3A6C + pCMVNS-1
31	pJBLR4.3A6C + pCMVNS-1
32	pJBLR4.3T8G + pCMVNS-1
33	pJBLR4.3T8G + pCMVNS-1
34	pJBLR4.3T8G + pCMVNS-1
35	pJBLR4.3
36	pJBLR4.3 + pCMVNS-1



Assay 4: Vecpro probe

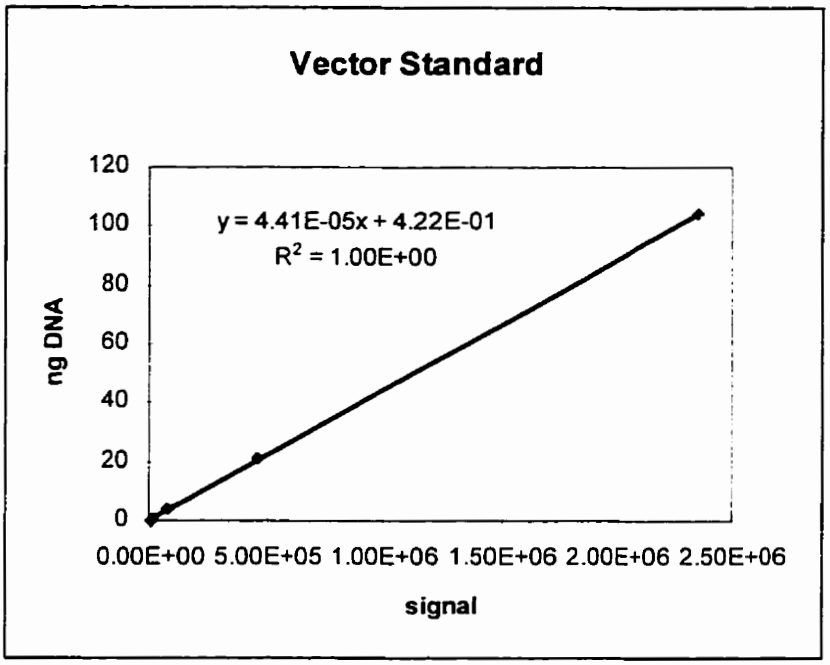


Assay 4: JAVA probe

Quantitation data and standard curves, Assay 4

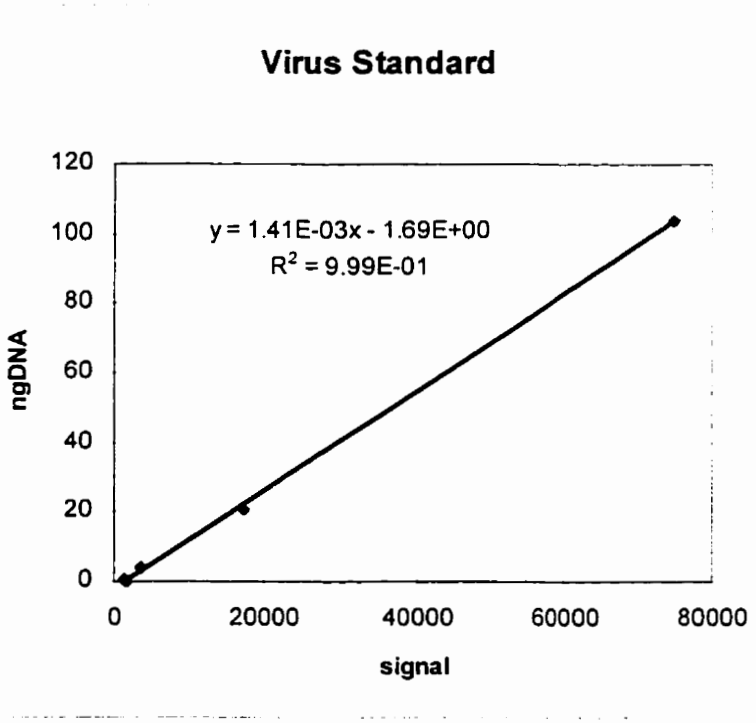
pJBLR4.3 Standards, Vecpro probe:

Segment #	*Volume*	*ng Std.*
1	2.35E+06	104
2	453834.44	21
3	79097.05	4.3
4	16984.6	0.833
5	5201.7	0.167



pJBLR4.3 Standards, JAVA probe:

Segment #	*Volume*	*ng Std.*
1	74744.21	104
2	17239.48	21
3	3550.33	4.3
4	1345.93	0.833
5	1533.75	0.167



Sample Quantitation, Assay 4:

Clone	*Vecpro signal*	*ng input*	*JAVA signal*	*ng out*	*RE*	*Scaled RE*	*Scaled RRE*
pJBLR4.3	325127.78	14.76	10317.36	12.86	-0.13	-0.01	-1%
4.3 + pCMVNS-1	545875	24.50	29315.48	39.64	0.62	0.74	93%
pJBLR4.3S25	742010.56	33.14	40692.82	55.69	0.68	0.80	101%
pJBLR4.3S25	498765.81	22.42	30283.07	41.01	0.83	0.95	120%
pJBLR4.3S25	527322.81	23.68	29017.1	39.22	0.66	0.78	98%
pJBLR4.3S26	427711.25	19.28	26570.29	35.77	0.86	0.98	123%
pJBLR4.3S26	432726.31	19.51	24729.4	33.18	0.70	0.82	104%
pJBLR4.3S26	424222.25	19.13	25505.9	34.27	0.79	0.91	115%
pJBLR4.3S8/22	754731.88	33.71	25594.54	34.40	0.02	0.14	18%
pJBLR4.3S8/22	713315.88	31.88	25660.43	34.49	0.08	0.20	26%
pJBLR4.3S8/22	616017.69	27.59	21182.63	28.18	0.02	0.14	18%
pJBLR4.3T1C	781041.25	34.87	42430.24	58.14	0.67	0.79	100%
pJBLR4.3T1C	682699.19	30.53	35975.45	49.04	0.61	0.73	92%
pJBLR4.3T1C	724139.38	32.36	38425.52	52.49	0.62	0.74	94%
pJBLR4.3C2A	370675.5	16.77	27290.89	36.79	1.19	1.31	166%
pJBLR4.3C2A	459295.88	20.68	28341.2	38.27	0.85	0.97	123%
pJBLR4.3C2A	523021.84	23.49	33589.07	45.67	0.94	1.06	135%
pJBLR4.3T3G	623429.19	27.92	35041.34	47.72	0.71	0.83	105%
pJBLR4.3T3G	451482.03	20.33	28134.71	37.98	0.87	0.99	125%
pJBLR4.3T3G	485523.75	21.83	30565.02	41.41	0.90	1.02	129%
pJBLR4.3T4A	672608.06	30.08	36308.58	49.51	0.65	0.77	97%
pJBLR4.3T4A	636161.75	28.48	36370.31	49.59	0.74	0.86	109%
pJBLR4.3T4A	535892.19	24.05	30907.01	41.89	0.74	0.86	109%
pJBLR4.3A6C	485079.84	21.81	28516.16	38.52	0.77	0.89	112%
pJBLR4.3A6C	574192.62	25.74	29841.53	40.39	0.57	0.69	87%
pJBLR4.3A6C	482930.88	21.72	35014.49	47.68	1.20	1.32	166%
pJBLR4.3T8G	638611.69	28.58	40785.86	55.82	0.95	1.07	136%
pJBLR4.3T8G	675706.62	30.22	43196.25	59.22	0.96	1.08	137%
pJBLR4.3T8G	515624.19	23.16	32708.62	44.43	0.92	1.04	131%
pJBLR4.3	334859.5	15.19	10864.7	13.63	-0.10	0.02	2%
4.3 + pCMVNS-1	532072.88	23.89	30383.78	41.15	0.72	0.84	107%

Appendix B

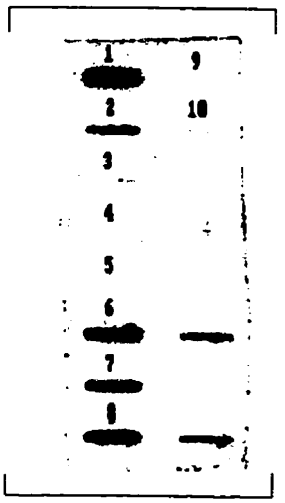
This appendix contains the raw data and analysis from two competition replication assays as summarised in the Results section, Table 5.

Data from each assay is presented on a single page, with each slot-blot hybridization presented next to a table reporting the contents and quantitation values for each of segments 1 through 8 and a graph displaying the least-squares curvefit for the standards (segments 1 through 5 of the blot). The equation of this curvefit was used to determine the amount of DNA present in each of the experimental samples (segments 6 through 8). Segments 9 and 10 of each blot are pCMVNS-1 and pJBLR4.3 hybridization controls, respectively, and are not quantitated.

For each mutant, the hybridization and analysis with mutant-specific probe is presented first and that with JAVA probe (for both mutant and wild-type minigenome) second. Mutant-specific probes were S20E oligonucleotide for pJBLR4.3S20, and S21B oligonucleotide for pJBLR4.3S21.

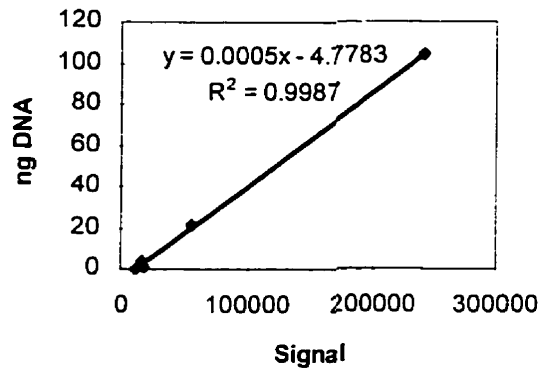
Following the results from the individual assays, the pooled data used to generate Table 5 is presented.

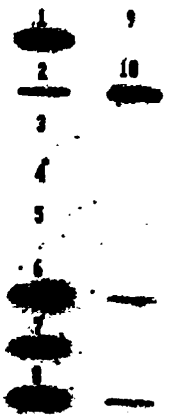
pJBLR4.3S20 Competition Replication Assays



S20-specific probe

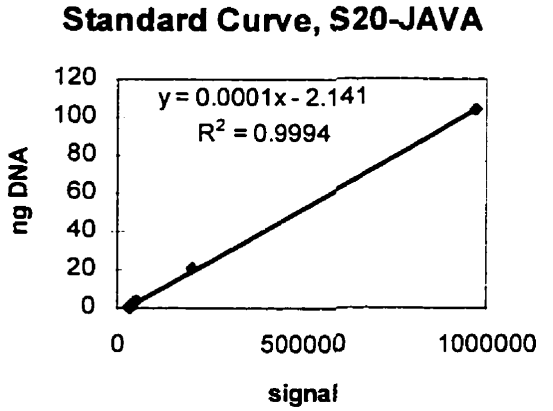
Segment #	*Contains*	*Volume*	*ng DNA*
1	pJBLR4.3S20	240317.44	104
2	pJBLR4.3S20	55696.29	21
3	pJBLR4.3S20	15678.42	4.3
4	pJBLR4.3S20	17823.71	0.833
5	pJBLR4.3S20	10897.33	0.167
6	Assay 1	98308.93	44.38
7	Assay 2	71361.42	30.90
8	Assay 3	99753.62	45.10



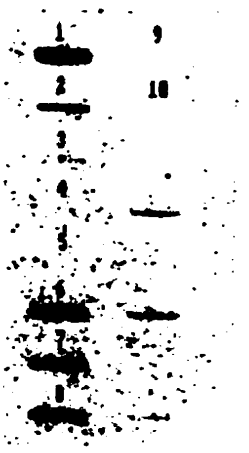


JAVA probe

Segment #	*Contains*	*Volume*	*ng DNA*
1	pJBLR4.3S20	973893.56	104
2	pJBLR4.3S20	202643.25	21
3	pJBLR4.3S20	48362.76	4.3
4	pJBLR4.3S20	32162.07	0.833
5	pJBLR4.3S20	34287.34	0.167
6	Assay 1	1.06E+06	103.86
7	Assay 2	815222.81	79.38
8	Assay 3	1.07E+06	104.86



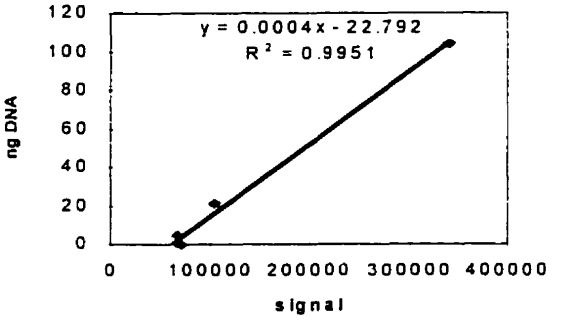
pJBLR4.3S21 Competition Replication Assays

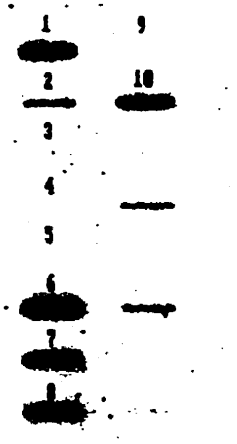


S21-specific probe

Segment #	*Contains*	*Volume*	*ng DNA*
1	pJBLR4.3S21	341062.06	104
2	pJBLR4.3S21	105871.18	21
3	pJBLR4.3S21	68071.81	4.3
4	pJBLR4.3S21	67478.59	0.833
5	pJBLR4.3S21	71762.8	0.167
6	Assay 1	240925.11	73.58
7	Assay 2	180537.38	49.42
8	Assay 3	185691.91	51.48

Standard Curve, S21-Specific

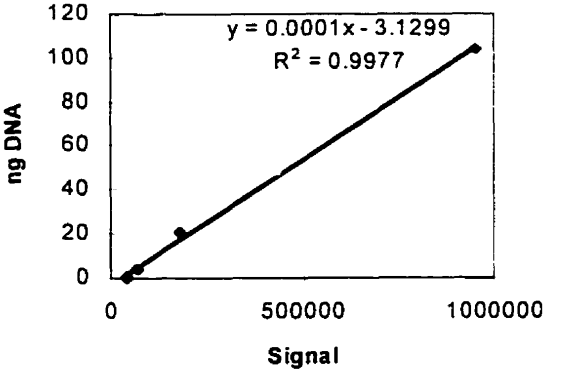




JAVA probe

Segment #	*Contains*	*Volume*	*ng DNA*
1	pJBLR4.3S21	953157.38	104
2	pJBLR4.3S21	180199.53	21
3	pJBLR4.3S21	70820.98	4.3
4	pJBLR4.3S21	45432.23	0.833
5	pJBLR4.3S21	42311.44	0.167
6	Assay 1	1.19E+06	115.87
7	Assay 2	767760.12	73.65
8	Assay 3	818158.19	78.69

Standard Curve, S21-JAVA



References

[1] Diffley,J., J. Cocker, S. Dowell, and A. Rowley. 1994. Two steps in the assembly of complexes at yeast replication origins in vivo *Cell* **78**:303-316

[2] Diffley,J. 1994. Eukaryotic DNA replication. *Current Opinion in Cell Biology* **6**:368-372

[3] Bell, S. 1995. Eukaryotic replicators and associated protein complexes. *Current Opinion in Genetics and Development* **5**:162-167

[4] Eckdahl, T., and J. Anderson. 1990. Conserved DNA structures in origins of replication. *Nucl. Acids Res.* **18**(6):1609-1612.

[5] Astell, C., M.B. Chow, and D.C. Ward. 1985. Sequence analysis of the termini of virion and replicative forms of minute virus of mice DNA suggests a modified rolling hairpin model for autonomous parvovirus dna replication. *J.Virol.* **54**(1):171-177

[6] Astell, C., Q. Liu, C. Harris, J. Brunstein, H. Jindal, and P. Tam. 1996. Minute virus of mice *cis*-acting sequences required for genome replication and the role of the *trans*-acting viral protein, NS-1. *Progress in Nucleic Acids Research and Molecular Biology* **55**:245-285.

[7] Cotmore, S., and P. Tattersall. 1987. The autonomously replicating parvoviruses of vertibrates. *Adv. Virus Res.* **33**:91-174.

[8] Tennant, R., and R. Hand. 1970. Requirement for cellular synthesis for killham rat virus replication. *Virology* **42**:1054

[9] Astell, C., W. Luo, J. Brunstein, and J. St.-Amand. 1997. B19 parvovirus: biochemical and molecular features. In: *Monographs in Virology*, Volume 20. L.J. Anderson and N. Young, Editors. Karger AG, Switzerland. Pages 16-41.

[10] Clemens, K, and D. Pintel. 1988. Two transcriptional units of the autonomous parvovirus minute virus of mice are transcribed in a temporal order. *J. Virol.* **62**(4):1448-1451

[11] Pringle C. 1993. *Arch. Virol.*494

[12] Crawford, L.V. 1966. A minute virus of mice. *Virology* **29**:605-612

[13] Angbandje-McKenna, M.m R. McKenna, I. Portman, E. Hernando, J. Almendral, A. Llamas-Saiz, C. Foces-Foces, F. Wang, P. Tattersall, and M. Rossman. 1997. The 3D structure of Minute Virus of Mice: implications for parvoviral tissue tropism and encapsidation of parvoviral genomic DNA. Abstract of presentation at the VII International Parvovirus Workshop, Heidelberg Germany.

[14] Tattersall, P., P. Cawte, A. Shatkin, and D. Ward. 1976. Three structural polypeptides coded for by minute virus of mice, a parvovirus. *J. Virol* **20:**(1) 273-289.

[15] Astell, C., M. Thomson, M. Merchlinsky, and D. Ward. 1983. The complete DNA sequence of minute virus of mice, an autonomous parvovirus. *NAR* **11**(4):999-1018.

[16] Pintel, D., D. Dadachanji, C. Astell, and D. Ward. 1983. The genome of minute virus of mice, an autonomous parvovirus, encodes two overlapping transcription units. *NAR* **11**(4):1019-1038.

[17] Morgan, W., and D. Ward. 1986. Three splicing patterns are used to excise the small intron common to all minute virus of mice RNAs. *J. Virol.* **60**(3):1170-1174.

[18] Clemens, K., and D. Pintel. 1987. Minute virus of mice (MVM) mRNAs predominantly polyadenylate at a single site. *Virology* **160:**511-514.

[19]Haut, D., and D. Pintel. 1996. Determinants which influence alternative splicing of parvovirus pre-mRNAs. Presentation at the annual meeting of the American Society for Virology, London Ontario.

[20] Gersappe, A., L. Burger, and D. Pintel. 1996. A small bipartite determinant within the MVM NS2-specific exon governs its inclusion into mature spliced mRNA. Presentation at the annual meeting of the American Society for Virology, London Ontario.

[21] Lorson, C., L. Burger, M. Mouw, and D. Pintel. 1996. Efficient transactivation of the minute virus of mice P38 promoter requires upstream binding of NS-1. *J. Virol.* **70**(2):834-842.

[22] Ahn, J., Z. Pitluk, and D. Ward. 1992. The GC box and TATA transcription control elements of the P38 promoter of the minute virus of mice are necessary and sufficient for transactivation by the nonstructural protein NS-1. *J. Virol.* **66**(6):3776-3783.

[23] Yeung, D., G. Brown, P. Tam, R. Russnak, G. Wilson, I. Clark-Lewis, and C. Astell. 1991. Monoclonal antibodies to the major non-structural nuclear protein of minute virus of mice. *Virology* **181**:35-45.

[24] Cotmore, S., L. Sturzenbecker, and P. Tattersall. 1983. The autonomous parvovirus MVM encodes two nonstructural proteins in addition to its capsid polypeptides. *Virology* **129**:333-343.

[25] Shade, R., M. Blundell, S. Cotmore, P. Tattersall, and C. Astell. 1986. Nucleotide sequence and genome organization of human parvovirus B19 isolated from the serum of a child during aplastic crisis. *J. Virol.* **58**(3):921-936.

[26] Astell, C, C. Mol, and W. Anderson. 1987. Structural and functional homology of parvovirus and papovavirus polypeptides. *J. Gen. Virol.* **68**:885-893.

[27] Liu, Q., C.B. Yong, and Caroline R. Astell. 1994. In vitro resolution of the dimer bridge of minute virus of mice (MVM) genome supports the modified rolling hairpin model for DNA replication. *Virology* **201**: 251-262.

[28] Wilson, Gary M., Hitesh K. Jindal, Douglas E. Yeung, Wei Chen, and Caroline R. Astell. 1991. Expression of minute virus of mice major nonstructural protein in insect cells: purification and identification of ATPase and helicase activities. *Virology* **185**: 90-98

[29] Christensen, J., S. Cotmore, and P. Tattersall. 1995. Minute virus of mice transcriptional activation protein NS-1 binds directly to the transactivation region of the viral P38 promoter in a strictly ATP-dependent manner. *J. Virol.* **69**(9):5422-5430.

[30] Gottleib, R., J. Christensen, S. Cotmore, and P. Tattersall. 1996. Characterization of the interaction between MVM NS-1 and its cognate DNA binding site. Presentation at the annual meeting of the American Society for Virology, London Ontario.

[31] Jindal, Hitesh K., Carl B. Yong, Gary M. Wilson, Patrick Tam, and Caroline R. Astell. 1993. Mutations in the NTP-binding motif of minute virus of mice (MVM) NS-1 protein uncouple ATPase and helicase functions. *J. Biol. Chem.* **269**(5):3283-3289.

[32] Mouw, M.B., J.L. Pearson and D.J. Pintel. 1996. Characterization of the biochemical properties of an MVM GST-NS-1 fusion protein. Presentation at the 15th Annual Meeting of the ASV,London Ontario.

[33] Nüesch,J., Corbau, R., Tattersall, P., and Rommelaere,J. 1996. Replication activites of MVM NS-1 are regulated through phosphorylation. Presentation at the 15th Annual Meeting of the ASV, London Ontario.

[34] Cotmore, Susan F., and Peter Tattersall. 1988. The NS-1 polypeptide of minute virus of mice is covalently attached to the 5' termini of duplex replicative-form DNA and progeny single strands. *J. Virol.* **62** (3):851-860.

[35] Colin Harris, unpublished data

[36] Krady, J.K., and D.C. Ward. 1995. Transcriptional activation by the parvoviral nonstructural protein NS-1 is mediated via a direct interaction with SP-1. *Mol. Cell. Biol.* **15**(1):524-533.

[37] Schoborg, R., and D. Pintel. 1991. Accumulation of MVM gene products is differentially regulated by transcription initiation, RNA processing and protein stability. *Virology* **181**:22-34.

[38] Cotmore, S., and P. Tattersall. 1990. Alternate splicing in a parvoviral nonstructural gene links a common amino-terminal sequence to downstream domains which confer radically different localization and turnover characteristics. *Virology* **177**:477-487.

[39] Naeger,Lisa, Jean Cater, and David J. Pintel. 1990. The small nonstructural protein (NS2) of the parvovirus minute virus of mice is required for efficient DNA replication and infectious virus production in a cell-type-specific manner. *J. Virol.* **64**(12):6166-6175.

[40] Naeger,Lisa, Nathalie Salomé, and David J. Pintel. 1993. NS2 is required for efficient translation of viral mRNA in minute virus of mice-infected murine cells. *J. Virol.* **67**(2):1034-1043.

[41] Cotmore, S., A. D'Abramo Jr., L. Carbonell, J. Bratton, and P. Tattersall. 1997. The NS2 polypeptide of parvovirus MVM is required for capsid assembly in murine cells. *Virology* **231**(2):267-280.

[42] Mousset, Suzanne, Youssef Ouadrhiri, Perrine Caillet-Fauquet, and Jean Rommelaere. 1994. The cytotoxicity of the autonomous parvovirus minute virus of mice nonstructural proteins in FR3T3 rat cells depends on oncogene expression. *J. Virol.* **68**(10):6446-6453.

[43] Legendre, D., and J. Rommelaere. 1992. Terminal regions of the NS-1 protein of the parvovirus minute virus of mice are involved in cytotoxicity and promoter *trans*-inhibition. *J. Virol.* **66**(3):5705-5713.

[44] Vanacker, Jean-Marc, V. Laudet, G. Adelmant, D. Stéhelin, and J. Rommelaere. 1993. interconnection between thyroid hormone signalling pathways and parvovirus cytotoxic functions. *J. Virol.* **67**(12):7668-7672.

[45] Paradiso, Peter R. 1983. Analysis of the protein-protein interactions in the parvovirus H-1 capsid. *J. Virol.* **46**(1):94-102.

[46] Willwand, Kurt, and Bernard Hirt. 1991. The minute virus of mice capsid specifically recognizes the 3' hairpin structure of the viral replicative-form DNA: mapping of the binding site by hydroxyl radical footprinting. *J. Virol.* **65**(9):4629-4635.

[47] Tullis, G. L. Burger, and D. Pintel. 1993. The minor capsid protein VP1 of the autonomous parvovirus minute virus of mice is dispensable for encapsidation of progeny single-stranded DNA but is required for infectivity. *J. Virol.* **67**(1):131-141.

[48] Willwand, K., and B. Hirt. 1993. The major capsid protein of mniute virus of mice (MVM) can form particles which bind to the 3' terminal hairpin of MVM replicative form DNA and package single-stranded viral progeny DNA. *J. Virol.* **67**(9):5660-5663.

[49] Santarén, Juan F., Juan C. Ramírez, and Jose M. Almendral. 1993. Protein species of the parvovirus minute virus of mice strain MVMp: involvement of phosphorylated VP-2 subtypes in viral morphogenesis. *J. Virol.* **67**(9):5126-5138.

[50] Krysan, P., J. Smith, and M. Calos. 1993. Autonomous replication in human cells of multimers of specific human and bacterial DNA sequences. *Mol. Cell. Biol.* **13**(5):2688-2696.

[51] Bramhill, D., A. Kornberg. 1988. Duplex opening by dnaA protein at novel sequences in intitiation of replication at the origin of the E. coli chromosome. *Cell* **52**(5):743-755

[52] Bramhill, D., A. Kornberg. 1988. A model for initiation at origins of DNA replication. *Cell* **54**(7):915-918

[53] Waga, S., and B. Stillman. 1994. Anatomy of a DNA replication fork revealed by reconstitution of SV40 replication *in vitro*. *Nature* **369**:207-212.

[54] SenGupta, Dhruba, and James A. Borowiec. 1994. Strand and face: the topography of interactions between the SV40 origin of replication and T-antigen during the initiation of replication. *EMBO J.* **13**(4): 982-992.

[55] Toyn, J., W. Toone, B. Morgan, and L. Johnston. 1995. The activation of DNA replication in yeast. *TIBS* **20**:70-73.

[56] Almouzni, G. 1994. The origin recognition complex (ORC):the stone that kills two birds. *BioEssays* **16**(4):233-235.

[57] Wang, T., and J. Li. 1995. Eukaryotic DNA replication. *Curr. Opin. Cell Biol.* 7:414-420.

[58] Donovan, S., and Diffley J. 1996. Replication origins in eukaryotes. *Curr. Opin. Genet. Dev.* **6**:203-207.

[59] Carpenter, P., P. Mueller, and W. Dunphy. 1996. Role for a *Xenopus* Orc-2 related protein in controlling DNA replication. *Nature* **379**:357-360.

[60] Rowles, A., J. Chong, L. Brown, M. Howell, G. Evan, and J. Blow. 1996. Interaction between the origin recognition complex and the replication licensing system in *Xenopus*. *Cell* **87**:287-296.

[61] Diffley, J. J. Cocker, S. Dowell, and A. Rowley. 1994. Two steps in the assembly of complexes at yeast replication origins *in vivo*. *Cell* **78**:303-316.

[62] Rowles, A, and J. Blow. 1997. Chromatin proteins involved in the initiation of DNA replication. *Curr. Opin. Genet. Dev.*7:152-157.

[63] Burhans, W., L. Vassilev, M. Caddle, N. Heintz, and M. DePamphilis. 1990. Identification of an origin of bidirectional DNA replication in mammalian chromosomes. *Cell* **62**:955-965.

[64] Mastrangelo, I, P. Held, L. Dailey, J. Wall, P. Hough, N. Heintz, and N. H. Heintz. 1993. RIP60 dimers and multiples of dimers assemble link structures at an origin of bidirectional replication in the dihydrofolate reductase amplicon of chinese hamster ovary cells. *J. Mol. Biol.* **232**:766-778.

[65] Dailey, L., M. Caddle, N. Heintz, and N.H. Heintz. 1990. Purification of RIP60 and RIP100, mammalian proteins with origin-specific DNA-dinding and ATP-dependent DNA helicase activities. *Mol. Cell. Biol.* **10**(12):6225-6235.

[66] Eckdahl, T., and J. Anderson. 1990. Conserved DNA structures in origins of replication. *Nucl. Acids Res* **18**(6):1609-1612.

[67] Williams, J.S., T. Eckdahl, and J. Anderson. 1988. Bent DNA functions as a replication enhancer in *Saccheromyces cerevisiae*. *Mol. Cell. Biol.* **8**:2763-2769

[68] Cavalier-Smith, T. 1974. Palindromic base sequences and replication of eukaryotic chromosome ends. *Nature* **250**:467-470.

[69] Rhode, S.L. 1976. *J. Virol.* **17**:659-667.

[70] Tam, P., and C. Astell. 1993. Replication of minute virus of mice minigenomes: novel replication elements required for MVM DNA replication. *Virology* **193**:812-824.

[71] Samulski, R., A. Srivastava, K. Berns, and N. Muzyczka. 1983. Rescue of adeno-associated virus from recombinant plasmids:gene correction within the terminal repeats of AAV. *Cell* **33**:135-143.

[72] Senapathy, P, J. Tratschin, and B. Carter. 1984. Replication of adeno-associated virus DNA. Complementation of naturally occurring *rep-* mutants by a wild-type genome or an *ori-* mutant and correction of terminal palindrome deletions. *J. Mol. Biol.***179**(1):1-20.

[73] Ward, P., and K. Berns. 1995. Minimum origin requirements for linear duplex AAV DNA replication *in vitro*. *Virology* **209**:692-695.

[74] Bates, R., C. Snyder, P. Banerjee, and M. Sankar. 1984. Autonomous parvovirus LuIII encapsidates equal amounts of plus and minus DNA strands. *J. Virol.* **49**:319-

[75] Berns, K. 1991. Parvoviridae and their replication. *Fundamental Virology*. Chapter 32, Second edition. Raven Press Ltd. New York.

[76] Astell C.R., M. Thomsom, M.B. Chow, and D.C. Ward. 1983. Structure and replication of minute virus of mice DNA. *Cold Spring Harbour Symposia on Quantitative Biology*, Vol. XLVII: 751-761.

[77] Maniloff, J. 1975. Thermodynamic considerations of the deoxyribonucleic acid helix-cruciform transition. *J. Theor. Biol.* **53**:497-501.

[78] Rhode, S.L. and Belia Klaassen. 1982. DNA sequence of the 5' terminus containing the replication origin of parvovirus replicative form DNA. *J. Virol.* **41**(3):990-999.

[79] Tam, P., and C. Astell. 1994. Multiple cellular factors bind to cis-regulatory elements found inboard of the 5' palindrome of minute virus of mice *J. Virol.***68**(5):2840-2848.

[80] Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. Molecular cloning. *Cold Springs Laboratory Press*. Second Edition.

[81] Xiang, C., H. Wang, P. Shiel, P. Berger, and D.J. Guerra. 1994. A modified alkaline lysis miniprep protocol using a single microcentrifuge tube. *BioTechniques* **17**(1):30-32.

[82] Hirt, B. 1967. Selective extraction of polyoma DNA from infected mouse cell cultures. *J. Mol. Biol.* **26**:365-369.

[83] Gluzman, Y. 1981. SV-40 transformed simian cells support the replication of early SV40 mutants. *Cell* **23**:175-182.

[84] Littlefield, J. 1964. Three degrees of guanylic acid-inosinic acid pyrophosphorylase deficiency in mouse fibroblasts. *Nature* **203**:1142-1144.

[85] Nawotka, K., and J. Huberman. 1988. Two-dimensional gel electrophoretic method for mapping DNA replicons. *Mol. Cell. Biol.* **8**(4):1408-1413.

[86] Huberman, J., L. Spotilla, K. Nowatka, S. El-Assouli, and L. Davis. 1987. The *in vivo* replication origin of the yeast 2 μm plasmid. *Cell* **51**:473-481.

[87] Faust, E., and D. Ward. 1979. Incomplete genomes of the parvovirus minute virus of mice: selective conservation of genome termini, including the origin for DNA replication. *J. Virol.* **32**:276-292.

[88] Salvino, M., M. Skiadopoulos, E. Faust, P. Tam, R. Shade, and C. Astell. 1991. Two spatially distinct genetic elements constitute a bipartite DNA replication origin in the minute virus of mice genome. *J. Virol.* **65**(3):1352-1363.

[89] Yong, C. Unpublished data.

[90] Vovis, G. F. and Lacks, S. 1977. Complimentary action of restriction enzymes endo R-DpnI and endo R-DpnII on bacteriophage f1 DNA. *J. Mol. Biol.* **115**:525-538

[91] Wang, M., and R. Reed. 1993. Molecular cloning of the olfactory neuronal transcription factor Olf-1 by genetic selection in yeast. *Nature* **364**:121-126.

[92] Atschul, S., W. Gish, W. Miller, E. Myers, and D. Lipman. 1990. Basic local alignment search tool. *J. Mol. Biol.* **215**:403-410.

[93] Zeltser, L., C. Desplan, and N. Heintz. 1996. *Hoxb-13*: a new Hox gene in a distant region of the HOXB cluster maintains colinearity. *Development* **122**:2475-2484.

[94] Colleta, P. L., S. Shimeld, and P. Sharpe. 1994. The molecular anatomy of HOX gene expression. *J. Anat.***184**:15-22.

[95] Heintz, N. Personal communication. Dept. of Pathology, University of Vermont.

[96] Houtchens, C. Personal communication. Dept. of Pathology, University of Vermont.

[97] Russel, J. and M. Botchan. 1995. *Cis*-acting components of human papillomavirus (HPV) DNA replication: linker substitution analysis of the HPV type 11 origin. *J. Virol.* **69**:651-660.

[98] Goulian, M., R. Kollek, D. Revie, W. Burhans, C. Carton, and B. Tseng. 1983. Parvovirus DNA replication. In: *Mechanisms of DNA replication and recombination*, UCLA Symposium on Molecular and Cellular Biology, New Series. **10**:367-379.

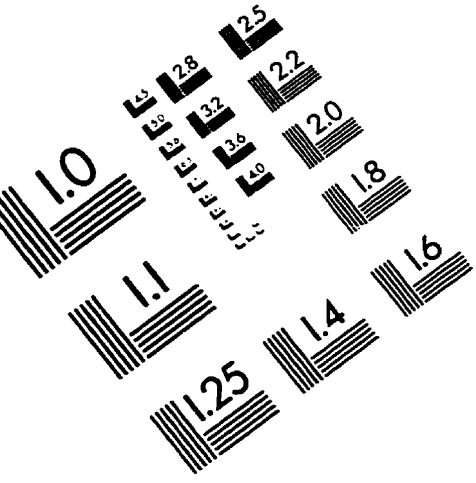
[99] Kollek, R., B. Tseng, and M. Goulian. 1982. DNA polymerase requirements for parvovirus H-1 DNA replication in vitro. *J. Virol.* **41**:982-989.

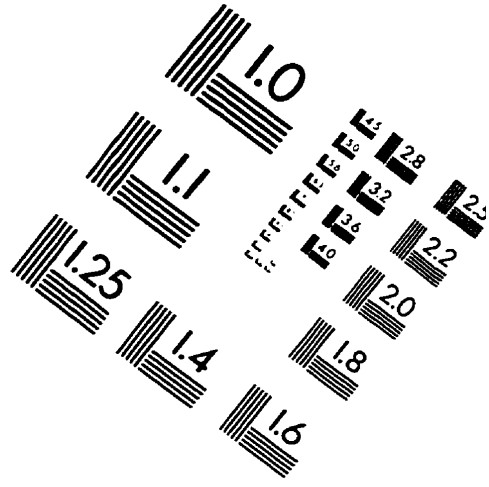
[100] Kollek, R., and M. Goulian. 1981. Synthesis of parvovirus H-1 replicative form from viral DNA by DNA polymerase γ. *Proc. Natl. Acad. Sci. USA.* **78**:6206-6210.

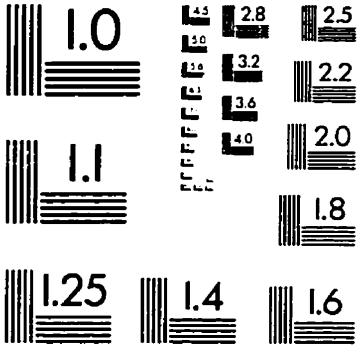
[101] Cossons, N., E. Faust, and M. Zannis-Hadjopoulos. 1996. DNA polymerase δ-dependent formation of a hairpin structure at the 5' terminal palindrome of the minute virus of mice genome. *Virology* **216**:258-264.

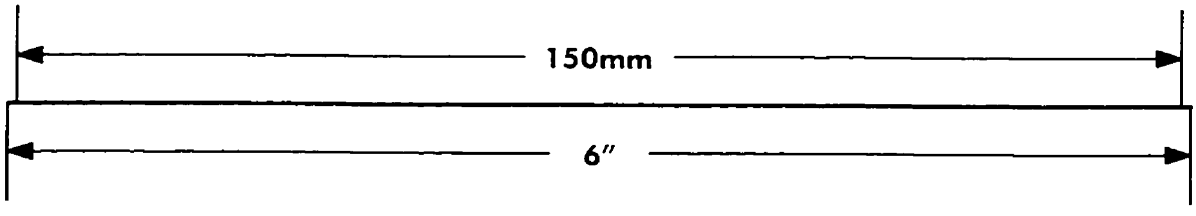
[102] Baldauf, A., K. Willwand, E. Mumtsidu, J. Nüesch, and J. Rommelaere. 1997. Specific initiation of replication at the right-end telomere of the closed species of minute virus of mice replicative-form DNA. *J. Virol.* **71**:971-980.

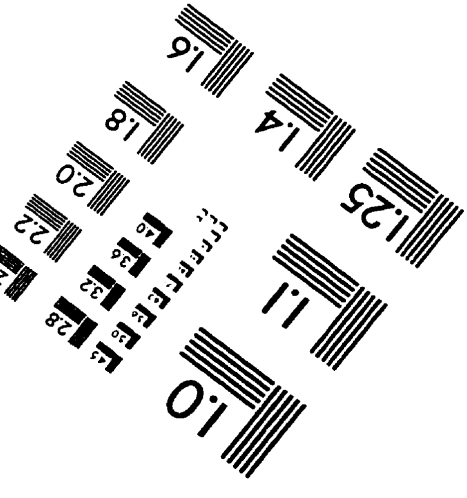
IMAGE EVALUATION TEST TARGET (QA-3)

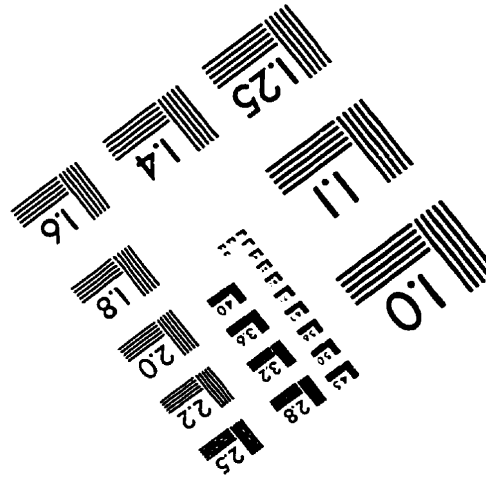












APPLIED IMAGE, Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989