The Molecular Determinants and Consequences of Recombination in the Evolution of Human Enteroviruses

Kym Sheree Lowry

A thesis submitted for the degree of Doctor of Philosophy

University of Warwick

School of Life Sciences

March 2011

Table of Contents

List of Tables

List of Figures

List of Appendices

Acknowledgements

This work was funded by a University of Warwick Studentship, to which I am extremely grateful for making this research project possible. I have appreciated the opportunity to gain valuable experience in such a rich research environment and have enjoyed all that Warwick University (including an outstanding virology section) has to offer.

I would like to thank my supervisor Professor David J. Evans for awarding me a scholarship – thank you for taking that leap of faith having never met the potential student on the other side of the world. I appreciate the endless time, help, and support throughout my PhD and realise how lucky I was to have such an approachable supervisor.

I need to thank the members of the Evan's lab for making it such an enjoyable work environment. I'll miss the cheeky sports banter, the constant laughs, and the wealth of knowledge and experience that made it so enriching.

Thanks must go to my family and friends who have offered constant support throughout and always provided help or a sympathetic ear. I can't wait to spend more time with you all.

I would like to dedicate this thesis to my mother, Judy, who overcame serious illness during the course of my study and taught me what it was to 'knuckle down' and keep going when the times were tough. You are brave and always look to the future.

Finally, I would like to acknowledge the unwavering support, direction, patience, understanding, occasional prodding, computer expertise, and shoulder to lean on, of my husband Mark. What more can I say? Thank you so much.

Declaration

This work was completed at the University of Warwick between July 2007 and November 2010 and has not been submitted for another degree. The work is original and unless otherwise stated in the text, has been completed by the author.

Signed

Date

Summary

Recombination is an important biological process in a diverse range of viruses, particularly those with single-stranded ribonucleic acid (RNA) genomes including the enteroviruses. Mutations caused by the error-prone RNA-dependent RNA polymerase (RdRp) and the vast population size of these virus populations are evolutionary mechanisms that generate genetic diversity – this allows viruses to survive under changing environmental pressures (e.g. adaptive host immunity). Ribonucleic acid recombination has been identified as another contributing mechanism involved in diversification, by removing interfering or lethal mutations from a virus genome, and by establishing new viruses.

Virus RNA recombination is well documented and is identified in several virus families including picornaviruses. However, recombination is a rare event and the study of the molecular mechanisms behind virus recombination is complicated by our ability to isolate and analyse recombinants from a mixed virus population including the parental viruses.

The objectives of this study were to firstly devise a method for generating populations of natural recombinant viruses, and secondly, to study the molecular processes that determine where and when recombination occurred in the enteroviral genome.

During this project, an *in vitro* system was developed to allow the recovery of recombinant enteroviruses in the absence of their parental viruses. Two virus RNA molecules containing “lesions” rendering them unable to generate viable virus on their own were co-transfected into mouse L929 cells. The method required two parental virus RNA molecules to be present in a single cell to produce a viable recombinant virus. Reverse Transcription-Polymerase Chain Reaction (RT-PCR) and sequencing analysis confirmed the recombinant nature of progeny virus genomes. Experimental data confirmed the effectiveness of the method and provided evidence that recombination occurs in at least two phases. Initial template switching, referring to the transfer of RdRp from one RNA template to another mid-replication, occurred apparently indiscriminately and with the addition of extra virus and non-virus sequence at the junction sites. In the second phase, any additional sequence was lost during subsequent rounds of replication and selection. The approach was expanded to incorporate viruses from different enterovirus species to investigate intra- and interspecies recombination events.

Abbreviations

ANOVA	Analysis of variance
aVDPV	Ambiguous vaccine-derived poliovirus
BLAST	basic local alignment search tool
bp(s)	base pair(s)
CD	cluster of differentiation
cDNA	complementary DNA
CI	confidence interval
CMV	cytomegalovirus
CO_2	carbon dioxide
CPE	cytopathic effect
CRE	*cis*-acting replication element
cVDPV	circulating vaccine-derived poliovirus
D	redundant nucleotide for adenine, guanine, or thymine
DAPI	4',6-diamidino-2-phenylindole
DEPC	diethylpyrocarbonate
dH_2O	distilled water
DI	defective interfering
DMEM	Dulbecco's modified Eagle's medium
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
ds	double stranded
DTT	dithiothreitol
dNTP	deoxyribonucleotide triphosphate
E. coli	*Escherichia coli*
EDTA	ethylenediaminetetraacetic acid
EMEM	minimum essential medium with Earle's salts
FBS	foetal bovine serum
FITC	fluoroscein isothiocyanate
FLC	full-length clone
g	gram
GFP	green fluorescent protein
GI	gastrointestinal
GMEM	Glasgow minimum essential medium
GuHCl	guanidine hydrochloride
HeLa	human cervical cancer cell line
HI	heat inactivated
hr	hour
ICAM-1	intracellular adhesion molecule type - 1
IF	immunofluorescence
IPV	inactivated poliovirus vaccine
IRES	internal ribosome entry site
iVDPV	immunodeficiency-associated vaccine-derived poliovirus
K	redundant nucleotide for guanine or thymine
kbp	kilobase pair
LB	Luria Bertani
M	redundant nucleotide for adenine or cytosine

MCS	multiple cloning site
MFED	minimal free energy differences
mg	milligram
min	minute
miRNA	microRNA
ml	millilitre
mM	millimolar
MOI	multiplicity of infection
mRNA	messenger RNA
NCBI	National Center for Biotechnology Information
NCR	non-coding region
ng	nanogram
nm	nanometre
nt	nucleotide
OPV	oral poliovirus vaccine
ORF	open reading frame
PABP	poly(A) binding protein
PBS	phosphate buffered saline
PCBP2	poly(rC) binding protein 2
PCR	polymerase chain reaction
PFU	plaque forming unit
pH	power of hydrogen
pmol	picomole
PS	Penicillin and streptomycin
PTB	polypyrimidine tract binding protein
PTGS	post-transcriptional gene silencing
PVR	poliovirus type 3 receptor (CD155)
R	redundant nucleotide for adenine or guanine
RD	rhabdomyosarcoma cells
RD-ICAM	rhabdomyosarcoma cells expressing ICAM-1
RdRp	RNA-dependent RNA polymerase
RF	replicative form
RI	replication intermediate
RISC	RNA-induced silencing complex
RNA	ribonucleic acid
RNAi	RNA-mediated interference
RNase	ribonuclease
rpm	revolutions per minute
RT-PCR	reverse transcription polymerase chain reaction
SDS	sodium dodecyl sulphate
s	second
SG	stress granule
shRNA	short hairpin RNA
siRNA	short interfering RNA
$TCID_{50}$	50% tissue culture infective dose
U	unit
µg	microgram
µl	microlitre

μM	micromolar
UV	ultraviolet
VAPP	vaccine-associated paralytic poliovirus
VDPV	vaccine-derived poliovirus
VPg	virus protein genome linked
v/v	volume per volume total
W	redundant nucleotide for adenine or thymine
w/v	weight per volume total
WHO	World Health Organisation
Y	redundant nucleotide for cytosine or thymine
°C	degrees Celsius

Virus abbreviations (enterovirus species where appropriate)

BVDV	bovine viral diarrhoea virus
CVA	coxsackievirus A serotypes
CVA21	coxsackievirus A21 (species C)
CVA16	coxsackievirus A16 (species A)
CVB	coxsackievirus B serotypes
CVB3	coxsackievirus B3 (species B)
CVB4	coxsackievirus B4 (species B)
E30	echovirus 30 (species B)
E7	echovirus 7 (species B)
EV70	enterovirus 70 (species D)
EV71	enterovirus 71 (species A)
FMDV	foot-and-mouth disease virus
HAV	hepatitis A virus
HCV	hepatitis C virus
HEV	human enterovirus
HEVC	human enterovirus C serotypes
HRV	human rhinovirus
HRV2	human rhinovirus type 2
HRV14	human rhinovirus type 14
PV	poliovirus
PV1	poliovirus type 1 (species C)
PV2	poliovirus type 2 (species C)
PV3	poliovirus type 3 (species C)
TBSV	tomato bushy stunt virus

Protein abbreviations

VP0	precursor protein before cleavage to make VP2 and VP4
VP1	capsid protein
VP2	capsid protein
VP3	capsid protein
VP4	capsid protein
$2A^{pro}$	virus protease
2BC	precursor protein, interacts with cellular membranes
2B	interacts with cellular proteins, induces vesicles in cells
2C	interacts with cellular proteins, induces vesicles in cells
3AB	precursor protein, binds to one or more CREs
3A	component of virus replication complex
3B	protein VPg
$3CD^{pro}$	precursor protein with protease activity, binds to one or more CREs
$3C^{pro}$	virus protease
$3D^{pol}$	RNA-dependent RNA polymerase (RdRp)
VPg	virion protein genome linked covalently to 5' NCR of virus genome
CRE	cis-acting replication element in 2C encoding region
p200	large subunit of eIF-4 degraded by $2A^{pro}$
G3BP	regulated effecter of stress granule assembly
p53	tumour suppressor protein that regulates cell cycle
PCBP2	cellular protein, promotes PV replication by binding to RNA secondary structure elements

CHAPTER ONE: Introduction

Evolution requires genetic variation. Positive-sense single stranded RNA viruses including important human pathogens like poliovirus (PV) and rhinovirus have error prone polymerases, short replication cycles, and high yields that together contribute to genetic diversity. Far greater genetic variation is achieved by chance recombination events than is possible in a single round of genome replication, and this is an important biological process that accelerates virus evolution further.

Although the result of many separate recombination events have been observed and characterised in nature, other than an understanding of the underlying principle, little is known about the mechanisms (viral and host) that establish the location of recombination breakpoints across the virus genome. Recombination is a rare event (possibly only one per million progeny), so this complicates isolation and analysis of recombinants from a mixed population. To this end, a method has been developed to recover recombinant viruses *in vitro* in the absence of infectious parental viruses.

This review will briefly consider the mechanisms of enteroviral evolution, methods of RNA recombination and evidence of recombination in enteroviruses both naturally and in the laboratory. Poliovirus, being the prototype strain of picornaviruses, is one of the best studied viruses and is used extensively during this project. As such, the majority of this literature review will focus on this virus.

1.1 Classification

Virus classification and the species concept

Classifying viruses into a taxonomic system is primarily performed by the International Committee on Taxonomy of Viruses (ICTV). The classification of a new virus into a family is based on morphology of the virion, the type of genomic nucleic acid it contains (e.g. single- or double stranded deoxyribonucleic acid [DNA] or RNA), the method of replication, and in some cases, the host organism (Hyypiä et al., 1997). A collection of genera form a family and these are generally defined on the basis of virion density, acid sensitivity, and host organisms. A virus "species" is a group of viruses that share similar morphological characteristics, can recombine amongst themselves to produce viable progeny, and can infect the same range of host

species. A group of species, in turn, comprise a genus. The term virus species implies a single genotype whereas RNA viruses, due to their large population sizes and high mutation rates, are in fact a diverse population of genotype variants or quasispecies (Lauring & Andino, 2010). Virus species are further subdivided into serotypes, which are classically defined as antigenically distinct viruses that are incapable of being neutralised to a significant extent by other serotype antibodies from the same species. Genetic variation within serotypes leads to the identification of virus strains, which are the limit to virus classification.

The picornavirus family

The picornaviruses are small (pico is a very small unit of measure), single stranded positive-sense RNA viruses belonging to the family *Picornaviridae*. Poliovirus is the type species (Flint *et al.*, 2004) and the family includes 12 genera that are responsible for morbidity in both humans and animals (summarised in table 1.1). After poliomyelitis, probably the most notable disease-causing picornavirus is foot-and-mouth disease virus (FMDV) (genus aphthovirus), which infects cloven-hoofed animals such as cattle and swine (Bachrach, 1968). This highly infectious virus has devastating financial effects on the agricultural industry through the necessary elimination of millions of animals and containment/sanitation measures. Other picornaviruses that greatly impact public health comprise enteroviruses (including PV and echoviruses); rhinoviruses (causing the majority of cases of the common cold); and hepatoviruses (hepatitis A virus [HAV]).

The enterovirus genus

Human enteroviruses are currently subgrouped into seven species, *Human Enterovirus* (HEV) A-D and *Human Rhinovirus* (HRV) A-C (Arden & Mackay, 2009). Over 100 distinct human enterovirus serotypes exist at present and these were mainly grouped according to pathogenicity in experimental animals before the introduction of more modern molecular methods for classification (Hyypiä et al., 1997). Recently, over 70 identified serotypes were assigned or reassigned to various

Table 1.1 Classification of Picornavirus family

Genus	Species	Serotype
Aphthovirus	*Equine rhinitis A virus*	ERAV
	Bovine rhinitis B virus	BRV-2
	Foot-and-mouth disease virus	FMDV-A, FMDV-Asia1, FMDV-C, FMDV-O, FMDV-SAT1-3
Avihepatovirus	*Duck hepatitis A virus*	DHV-1
Cardiovirus	*Encephalomyocarditis virus*	EMCV
	Theilovirus	TMEV, VHEV, TLV, RTV-1
Enterovirus	*Bovine enterovirus*	BEV-1, BEV-2
	Human enterovirus A	CVA2-8, 10, 12, 14, 16, EV71, 76, 89-92, SV46
	Human enterovirus B	CVA9, CVB1-6, E-1-9, 11-21, 24-27, 29-33, EV69, 73-75, 77-88, 95, 97, 100, 101, SVDV, unclass HVB
	Human enterovirus C	CVA1, 11, 13, 15, 17-22, 24, EV96, PV1-3, unclass EVC
	Human enterovirus D	EV68, 70, 94
	Human rhinovirus A	HRV-1, 2, 7-13, 15, 16, 18-25, 28-34, 36, 38-41, 43-47, 49-51, 53-68, 71, 73-78, 80-82, 85, 88-90, 94-96, 98
	Human rhinovirus B	HRV-3-6, 14, 17, 26, 27, 35, 37,42, 48, 52, 69, 70, 72, 79, 83, 84, 86, 91-93, 97, 99
	Human rhinovirus C	HRV-1-3, HRV-4-6, 8, 10, 12, 13, 17-20,
	(not defined)	22-28, 30-35, 37, 38, 40-48, 51-57, 59-61, 63, 64, 66-71, 73-84, 86-88, 90-100
	Porcine enterovirus B	PEV-9, 10, J10
	Simian enterovirus A	A-2PV, SVSA4, SVSV4, SVSV28, SPV7
Erbovirus	*Equine rhinitis B virus*	ERV-2
Hepatovirus	*Hepatitis A virus*	HAV, SAV
Kobuvirus	*Aichi virus*	AiV, unclass KV
	Bovine kobuvirus	BKV
Parechovirus	*Human parechovirus*	E1-8, 10, 11
	Ljungan virus	LV87-012, LV145SL, LV174F
Sapelovirus	*Avian sapelovirus*	ASV
	Porcine sapelovirus	PV8, PVJ4, PVJ6
Senecavirus	*Seneca Valley virus*	SVV
Teschovirus	*Porcine teschovirus*	PTV1-11, PV11, PVJ1-3, 5, 7, 9
Tremovirus	*Avian encephalomyelitis virus*	AEMV

As of 21 Jan 2011

Virus abbreviations are outlined on page 10

enterovirus species through partial or complete genome sequencing and phylogenetic analysis (Lukashev, 2005; Oberste *et al.*, 2004a; Oberste *et al.*, 2004b). Molecular typing has been adopted for the identification of enterovirus isolates, and this is based on comparisons of a portion of the VP1 capsid gene (Oberste et al., 1999). Enterovirus serotypes within a species typically share > 75% nucleotide sequence identity (> 85% amino acid identity) (Oberste *et al.*, 2007).

Poliovirus was recently reclassified as a member of HEV-C as a result of phylogenetic analysis which showed that members of this species were closely related in the non-capsid coding region (Benschop *et al.*, 2008; Brown *et al.*, 2003). Previous separation of PVs from HEV-C was based on disease manifestations, specifically the ability of PV to cause poliomyelitis, and to bind to the specific poliovirus receptor (PVR; CD[cluster differentiation]155) (Lukashev, 2005; Minor, 2004). There is generally little correlation between the virus serotype and the resulting disease in the host, as in virus serotypes belonging to different enterovirus species can cause similar symptoms.

Transmission

Enteroviruses infect the host by two main routes namely; (1) faecal-oral transmission (such as PV) and (2) respiratory routes (most predominantly with rhinoviruses). Faecal-orally transmitted viruses such as PV are acid-stable (pH 3.0 to 5.0) for 1 to 3 hours (hrs), and replicate at 37°C which makes them suited to enteric infection. These faecal-orally transmitted enteroviruses are also resistant to inactivation by many common disinfectants and soaps (Sutter *et al.*, 2008). In contrast, rhinoviruses are acid-labile and replicate more efficiently at 33°C (hence their preference for the upper respiratory tract) (Kistler *et al.*, 2007; Pallansch & Roos, 2001).

Diseases associated with enterovirus infection

Human enterovirus infections are usually asymptomatic or cause mild illness (such as a rash or a “common cold”); however some enteroviruses can cause a range of more severe illness in the host including respiratory diseases, acute hemorrhagic conjunctivitis, myocarditis, and undifferentiated rash. Rarer conditions such as aseptic meningitis and irreversible paralytic poliomyelitis can lead to death (Huang *et al.*, 1999; Khetsuriani *et al.*, 2006; Melnick, 1984; Pallansch & Roos, 2001).

Table 1.2 outlines a number of enteroviruses and their clinical symptoms (adapted from Flint et al., 2004). The patient age also influences the severity of disease experienced after infection by enterovirus serotypes. For example, children tend to have mild or asymptomatic symptoms following PV infection while older children and adults experience more severe symptoms. Neonates are at higher risk of disease caused by non-poliovirus enteroviruses (44% of total infections by age group) (Khetsuriani *et al.*, 2006) presumably due to underdeveloped immunity, while all ages are at risk of rhinovirus infection (causing predominantly upper and lower respiratory illness).

Overview of PV pathogenesis and eradication strategy

Poliovirus was discovered by Landsteiner and Popper in 1909 and was the cause of extensive poliomyelitis disease outbreaks up to the 1950s (Wimmer et al., 1993). Infection with PV can lead to one of the following outcomes: (1) asymptomatic infection which is the most common outcome (72%); (2) minor illness such as fever, vomiting, constipation, or sore throat; (3) non-paralytic poliomyelitis (aseptic meningitis); or (4) paralytic poliomyelitis which is a rare outcome (usually <1% of cases) (reviewed in Sutter et al., 2008). Poliovirus progressively penetrates other tissues after escaping the tonsils and gut following initial infection and can cause symptoms elsewhere, e.g. myalgia, vomiting, and pain in the limbs. Studies in the 1940s and 1950s determined that this form of disease lasted from 5 to 25 days, and resulted in complete recovery (reviewed in Nathanson & Kew, 2010). Progression to severe symptoms occurs if the virus penetrates the blood brain barrier of the central nervous system (CNS) by infecting motor neurons (Minor, 2004), as the name suggests – (*polio* = grey, *myelitis*/myelo = marrow). Virus replicates and destroys motor neurons in the grey matter of the spinal cord, brain stem, or motor cortex, and can ultimately lead to paralysis or even death (reviewed in Sutter et al., 2008).

There are three antigenically distinct serotypes of PV (serotypes 1, 2, and 3) and infection with one serotype does not confer protection against the other serotypes

Table 1.2 Clinical symptoms associated with certain enteroviruses

Virus	Paralytic Disease	Encephalitis, Meningitis	Carditis	Neonatal disease	Hand-foot-and-mouth disease	Rash	Acute hemorrhagic conjunctivitis	Respiratory Tract infections	Fever	Diarrhea, GI disease	Diabetes, pancreatitis	Orchitis	Congenital anomalies
Poliovirus 1-3	+	+						+	+				
Coxsackieviruses A1-24	+	+	+		+	+	+	+	+				+
Coxsackieviruses B1-6	+	+	+	+		+			+		+	+	+
Echovirus 1-33	+	+	+	+		+		+	+	+			
Enterovirus 70	+						+						
Enterovirus 71	+	+			+								
Parechoviruses 1-3	+	+						+	+	+			
Rhinoviruses 1-100								+					

(Adapted from Flint et al., 2004)

(Bodian, 1951; Toyoda *et al.*, 1984). The introduction of trivalent, inactivated polio vaccine (IPV) (Salk) and trivalent, live attenuated (Sabin) oral poliovirus vaccine (OPV) (Wimmer et al., 1993) reduced the incidence of poliomyelitis between 1950 and 1970. The former was a formalin trivalent IPV developed by Dr. Jonas Salk and licensed in 1955 (Salk, 1953). In 1963, an attenuated oral vaccine containing all three serotypes was developed and licensed by Dr. Albert Sabin (Sabin & Boulger, 1973). Oral polio vaccine contained attenuated strains of PV and was developed by rapidly passaging neurovirulent strains at suboptimal temperatures in simian cells and tissues. Numerous studies have since identified the multiple genetic determinants of Sabin vaccine strain attenuation, and Kew *et al* have summarised the critical mutations responsible in all three strains (Kew *et al.*, 2005). Briefly, Sabin-1 vaccine strain was distinguishable from its neurovirulent progenitor by 55 nucleotide (nt) substitutions which were located throughout the genome (McGoldrick *et al.*, 1995). Through reverse genetics, it was determined that the most important attenuating substitution occurred in the internal ribosome entry site (IRES) at position 480 (Kawamura *et al.*, 1989) while four other substitutions in the capsid region contributed to the attenuating phenotype. Sabin-2 contained only two substitutions (in the IRES and VP1 coding region) from the original P712 strain that contributed to the attenuated phenotype (Macadam *et al.*, 1993). Lastly, the attenuated Sabin-3 strain differed from the neurovirulent Leon strain by 10 nt substitutions, with the principal attenuating nts located in the IRES, VP3, and VP1 coding regions (Stanway *et al.*, 1984). The attenuation of the neurovirulent phenotype in PV vaccine strains provides one of the best model for rapid virus evolution and how we exploited it for the generation of highly effective vaccines.

Poliovirus vaccines are included in the national paediatric vaccine schedules of most developed countries. In 1988, the World Health Organisation (WHO) implemented the global eradication target to improve vaccination efforts in developing countries where PV infections resulted in annual cases of poliomyelitis in sub-optimally vaccinated or other vulnerable populations (WHO, 1988). To date, only four countries (Afghanistan, India, Nigeria, and Pakistan) still have wild PV transmission (CDC, 2006). However, the rapid evolution of PV during replication (with particular reference to mutation rates of the viral polymerase mentioned later) means that

vaccine strains are inherently unstable and occasionally mutate to cause PV infections in vaccinated individuals or close contacts.

Cases of vaccine-associated paralytic poliomyelitis (VAPP) are rare independent events that occur following administration of OPV (Kew *et al.*, 2005). The risk is low - approximately one case per 2.4 million doses delivered - and is a direct consequence of genetic mutation of OPV strains (CDC, 1997; Kew *et al.*, 2005). During virus replication in a vaccinee, Sabin vaccine strains can spontaneously revert to enhanced virulence by key nt substitutions that conferred attenuation in the first place (outlined above). Sabin-3 contributes to the highest rates of VAPP post-vaccination, with low genetic stability of critical substitutions the main cause (Chumakov *et al.*, 1992; Kew *et al.*, 2005). Such vaccine-derived outbreaks of VAPP arise as a direct result of OPV replication in vaccinees, and highlights the role of rapid genetic evolution within these virus populations.

Several outbreaks caused by vaccine-derived polioviruses (VDPVs) have been identified in recent years due to more modern sequencing methods. Vaccine-derived PVs differ from 1% to 15% of VP1 sequence from the original Sabin vaccine strains indicating longer circulation post-vaccination (Kew *et al.*, 2005). These VDPVs may circulate within a population, especially in areas where wild PV no longer occurs and where there is low vaccine coverage (CDC, 2006; Rousset *et al.*, 2003). There are several categories of VDPVs including circulating-VDPVs (cVDPVs) that arise from recombination events between VDPVs and species C enteroviruses (Kew *et al.*, 2002; Rousset *et al.*, 2003; Shimizu *et al.*, 2004; Yang *et al.*, 2003), immunodeficiency-associated VDPVs (iVDPVs) which are excreted from immunodeficient recipients usually over the long term, and ambiguous-VDPVs (aVDPVs) which do not fall into either of the previous categories and have no defined origin (CDC, 2006).

1.2 Virus Structure and Life Cycle

Virus morphology

Enteroviruses consist of a ~7.5 kilobase (kb) RNA genome encapsulated in a non-enveloped icosahedral capsid with a diameter of approximately 30 nanometres (nm) (Putnak & Phillips, 1981). The capsid shell is formed from 60 copies of each of the

viral capsid proteins (VP4, VP2, VP3, and VP1) and forms a virion with a single virus protein genome (VPg)-linked positive-sense RNA (Hogle *et al.*, 1985; Lee *et al.*, 1977; Rossmann *et al.*, 1985). Each heteromeric structural unit comprises one copy of VP1, VP2, and VP3, while one copy of VP4 lies on the inner surface of the viral capsid, purportedly to support the capsid interior and interact with the RNA genome (Guttman & Baltimore, 1977; Lee *et al.*, 1993). Each of the 12 five-fold axes of the capsid (with prominent peaks) is surrounded by a surface depression, or "canyon", where the host receptor CD155 docks (Bernhardt *et al.*, 1994; Colston & Racaniello, 1994).

Tissue tropism

Pathogenesis of enteroviruses is influenced by extracellular factors including the presence of cellular receptors required for virus attachment and cell infection, and this largely determines tissue tropism. As well as requirements for cellular entry, viruses must also be able to replicate within the cell by utilising the cell's own machinary.

Experimental studies suggest that many enteroviruses exploit more than one cell surface protein in order to infect host cells (Evans & Almond, 1998). However, PV and the major group of HRV are exceptions in that they only require one receptor, CD155 and the intracellular adhesion molecule type-1 (ICAM-1) respectively (Greve *et al.*, 1989; Mendelsohn *et al.*, 1989). Other enteroviruses use different immunoglobulin-like molecules and adhesion proteins for attachment, and a current list of identified receptors used by a selection of enteroviruses is outlined in table 1.3.

Tissue tropism of PV is predominantly determined by expression of CD155, also known as poliovirus receptor (PVR), therefore natural PV infection is only supported by human and simian cells in culture (McLaren *et al.*, 1959). These observations suggest that PV did not emerge from an animal reservoir, nor do they circulate in such hosts. In 1989, Mendelsohn and colleagues cloned the human PVR gene and as a result, were able to express PVR on the surface of mouse cells

Table 1.3 Some enterovirus serotypes and their receptors

Serotype	Receptor	Reference
Poliovirus 1-3	CD155 (PVR)	(Mendelsohn *et al.*, 1989)
Coxsackievirus A21	ICAM-1, DAF	(Shafren *et al.*, 1997)
Rhinovirus (major group)	ICAM-1	(Greve *et al.*, 1989)
Coxsackievirus B1, B3, B5	DAF	(Shafren *et al.*, 1995) (Bergelson *et al.*, 1994)
Echovirus 3, 6, 7, 11-13, 20, 21, 24, 29, 33	DAF	(Powell *et al.*, 1998) (Ward *et al.*, 1994)
Echovirus 22	$\alpha_v\beta_3$ vitronectin	(Roivainen *et al.*, 1994)
Enterovirus 71	SCARB2, PSGL-1	(Nishimura *et al.*, 2009) (Yamayoshi *et al.*, 2009)

CAR, coxsackievirus and adenovirus receptor; DAF, decay accelerating factor; ICAM-1, intracellular adhesion molecule 1; PVR, poliovirus receptor; SCARB2, scavenger receptor class B member 2; PSGL-1, human P-selectin glycoprotein ligand-1

(L cells) which ordinarily do not support PV infection (Mendelsohn *et al.*, 1989). As a result, L cells also became susceptible to PV infection, and not just permissive to PV replication. In further experiments, transgenic mice containing the human PVR gene in the germ line were established for PV studies (Koike *et al.*, 1991; Ren *et al.*, 1990). Both groups observed that transgenic mice were susceptible to either infection by PV1 (Mahoney) or all PV serotypes, and presented with the same clinical symptoms identified in humans and monkeys. PVR transcripts were shown to be expressed in a wide range of transgenic mouse tissues, and this was also observed in humans (e.g. CNS) (Koike *et al.*, 1991; Mendelsohn *et al.*, 1989).

Genome organisation

The enterovirus genome consists of a single stranded, non-segmented positive-sense RNA molecule approximately 7500 nts long. The genome encodes a long single open reading frame preceded by a highly structured 5' non-coding region (NCR) and followed by a short 3' NCR with a poly(A)-tail. The polyprotein is cleaved co- and post-translationally by the virus's own encoded proteases ($2A^{pro}$ and $3C^{pro}/3CD^{pro}$). The enterovirus genome encodes four structural (capsid) proteins and seven non-structural proteins (figure 1.1A). The P1 region encodes the proteins necessary for the virion capsid, proteins from the P2 region are involved in host cell membrane rearrangement, protein processing and RNA replication, and P3 proteins are required for immune response interference and virus RNA replication. Figure 1.1B demonstrates how the cleavage products from P1, P2, and P3 regions are processed from the polyprotein.

A series of fast and slow proteolytic cleavages occur throughout the polyprotein in order to release virus proteins necessary for disrupting host cell processes and for fulfilling virus replication. Cleavage of the polyprotein initially occurs at the P1/P2 junction *in cis* by $2A^{pro}$ at a Tyr/Gly pair, while the second processing step also occurs in *cis* to release P3 from P2 by $3C^{pro}/3CD^{pro}$ at a Gln/Gly pair (Lawson & Semler, 1992; Toyoda *et al.*, 1986). The remaining cleavage events (except one) occur *in trans* by $3C^{pro}/3CD^{pro}$ at Gln/Gly pairs to generate mature proteins (Kitamura *et al.*, 1981; Semler *et al.*, 1981). It remains to be determined the mechanism that cleaves VP0 to yield VP4 and VP2

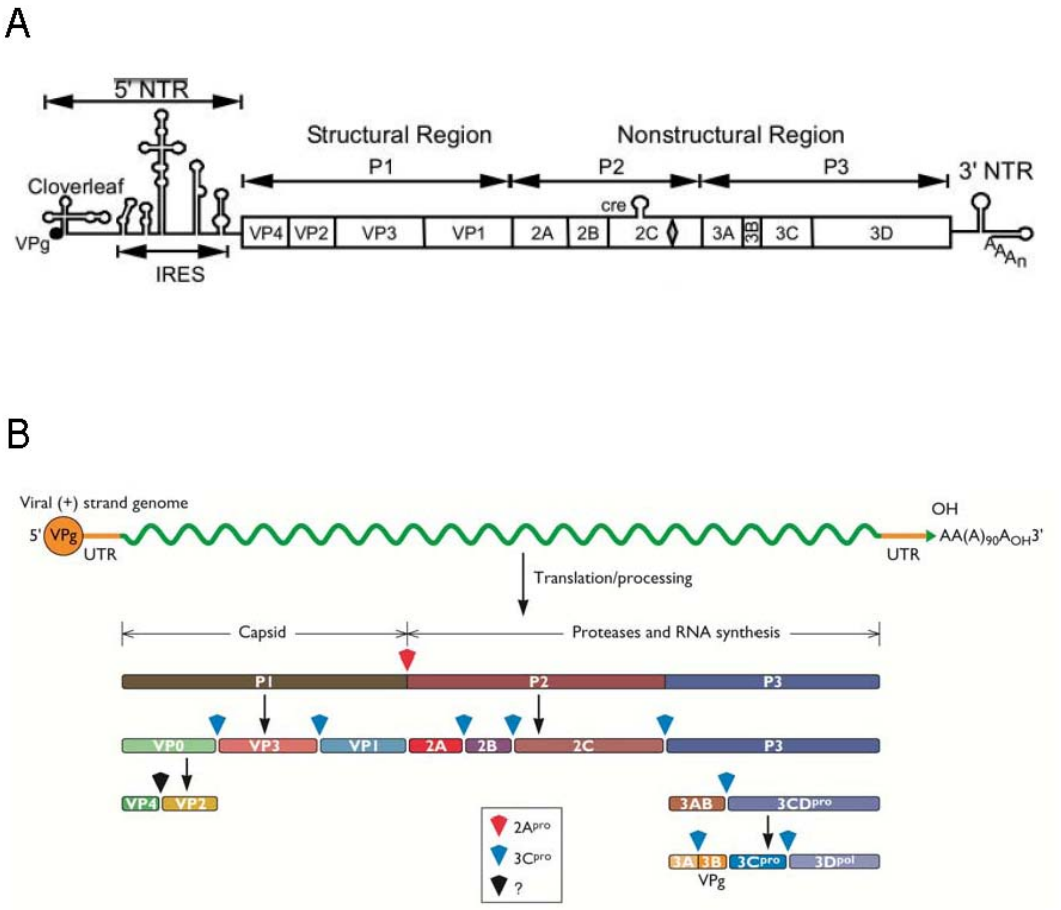


Figure 1.1 Genome structure of enteroviruses. (A) The virus RNA genome contains the essential signals for viral translation, replication, virion assembly and release of viral progeny from the host cell and is divided into 3 regions, P1-P3 (Source: Jiang *et al*., 2007). (B) Genome structure and cleavage products of picornavirus coding sequence. The red triangle indicates cleavage performed by 2A; the blue triangles are cleaved by 3C; the black triangle is cleaved by an unknown protease (Source: Flint *et al.*, 2004).

(Basavappa *et al.*, 1994). Not only do virus proteases cleave the polyprotein, they are also involved in cleaving host cell proteins in order to inactivate cellular targets and inhibit host defences. Virus proteinase $2A^{pro}$ plays an essential role in disrupting cap-dependent translation in the cell, by cleaving subunit p220 of the eukaryotic translation initiation factor 4 (eIF-4) (Kräusslich *et al.*, 1987). Likewise, PV $3C^{pro}$ has been implicated in many host cell processes. In one instance, $3C^{pro}$ cleaves a cellular factor (G3BP), which is involved in the formation of cellular messenger (m)RNA stress granules (SGs) (White *et al.*, 2007). Stress granules are thought to regulate mRNA metabolism during stress and favour the translation of stress response proteins. Additionally, $3C^{pro}$ mediates cleavage of transcriptional activator p53 with the help of cellular factors, thereby regulating the transcription of cellular genes (Weidman *et al.*, 2001).

The enterovirus genome also consists of *cis*-acting features required for translation and replication, and are discussed in further detail in the following section.

Regulatory elements in enterovirus/picornavirus genomes

The picornavirus replication cycle relies not only on virally encoded proteins, but is also mediated by *cis*-acting elements within the genome that drive translation and replication.

The 5' NCR contains *cis*-acting features required for translation and replication including the IRES, responsible for recruiting ribosomal subunits to translate the single polyprotein from which mature virus cleavage products are derived – these are essential for subsequent RNA replication and virion packaging (Pelletier & Sonenberg, 1988). An 88 nt cloverleaf structure is also located in the 5' NCR upstream of the IRES. This is crucial for RNA replication and plays a role in forming a ternary complex with virus protein 3CD, cellular poly(rC)-binding protein (PCBP), and poly(A)-binding protein (PABP), which circularises the genome and initiates replication from the 3' poly(A)-tail (Andino *et al.*, 1990; Herold & Andino, 2001; Parsley *et al.*, 1997).

The 3' NCR and poly(A)-tail are involved in RNA synthesis, as this is the site for initiation of negative-strand synthesis mentioned above. A pseudoknot structure has

been identified in this region (Melchers *et al.*, 1997; Mirmomeni *et al.*, 1997; Rohll *et al.*, 1995), and although the role of the 3' NCR in general is not clear, it may provide a role in stabilising spatial interactions during replication (Pilipenko *et al.*, 1996). Rohll and colleagues found that extensive mutagenesis of the single stem-loop structure formed by the HRV-14 3' NCR and replaced in a PV3 replicon, severely debilitated virus replication, indicating that this structure was essential for replication. Herold and Andino suggest that the 3' NCR plays a regulatory role, but is not an origin of replication, as PVs are still viable if this region is interchanged with another enterovirus (HRV14) or if the region is deleted (Herold & Andino, 2001; Rohll *et al.*, 1995; Todd *et al.*, 1997). Similarly, the poly(A)-tail is an important *cis*-acting element for RNA replication, since removal or significant shortening results in defects in replication (Herold & Andino, 2001; Spector *et al.*, 1975).

Within the coding region itself, a further *cis*-acting replication element (CRE) is located in the 2C coding region of PV and all other enteroviruses, and is a conserved 61 nt stem-loop structure (Goodfellow *et al.*, 2000). The CRE acts as a template for VPg uridylylation and VPg-pUpU functions as a primer for RNA synthesis by $3D^{pol}$. Uridylylation is the process of covalently linking two uridine nts to a tyrosine residue (Crawford & Baltimore, 1983; Kuhn *et al.*, 1988) and this process relies on an $A^1A^2A^3CA$ motif in the terminal loop of the CRE structure (Goodfellow et al., 2003) . It has been demonstrated using a cell-free *in vitro* translation and replication reaction with PV that $2C^{CRE}$ is not required for negative-strand synthesis, even though both positive- and negative-strand synthesis are primed by VPg (Goodfellow et al., 2003). In this case it is possible that the poly(A)-tail functions to template VPg uridylylation prior to negative-strand synthesis, as shown in *in vitro* assays (Paul *et al.*, 1998).

Similar CRE structures have been identified in various regions of other non-enterovirus genomes, including the VP1 encoding region of HRV14 (McKnight & Lemon, 1998), the 5' NCR of FMDV (Mason *et al.*, 2002), and 2A of HRV2 (Gerber *et al.*, 2001), indicating that the CRE location is not conserved. Goodfellow *et al* later demonstrated that the CRE could be relocated to the 5' NCR (from its original 2C position in PV) without this impacting on activity (Goodfellow et al., 2003).

Likewise, the FMDV CRE can be moved to the 3' end of the genome without affecting replication (Mason *et al.*, 2002).

In a competitive intracellular environment, virus genomes contain *cis*-acting elements, including the cloverleaf, CRE, and componants of the 3' NCR, to allow 3D polymerase ($3D^{pol}$) to differentiate between viral and cellular RNAs (reviewed in Bedard & Semler, 2004).

Overview of the enterovirus life cycle

Enteroviruses share similar genome structure and morphology, and as a result, have a similar life cycle. The following overview describes the PV life cycle as this is well studied and the subject of research presented in this thesis. Many aspects of PV infection have been mentioned, therefore will be shortened for brevity. A diagrammatic summary of the life cycle is shown in figure 1.2.

Poliovirus infection is initiated by the attachment of virus particles to PVR, and this triggers conformational changes in the virus particle. Upon cell attachment, both VP4 and the amino terminus of VP1 are externalised and are subsequently inserted into the host cell membrane (Levy *et al.*, 2010). This results in the formation of pores/channels that allow for the movement of virus RNA into the cell's cytoplasm. Viral RNA is translated under the control of the IRES, by recruitment of ribosomes to the site, into a single polyprotein by a cap-independent mechanism (Pelletier & Sonenberg, 1988). The long polyprotein is cleaved co- and post-translationally by the virus's encoded proteases to yield 11 virus proteins, some of which are necessary for virus replication (Kräusslich *et al.*, 1988; Toyoda *et al.*, 1986). The switch from translation to transcription occurs once sufficient viral proteins (such as $3CD^{pro}$ and $3D^{pol}$) have been synthesised, and these associate with cellular proteins to the 5' NCR to begin this process. Host cell endoplasmic reticulum membranes are rearranged and associate with non-structural virus proteins and cellular proteins to form replication complexes where viral RNA synthesis takes place (Bienz et al., 1983, Bienz et al., 1990). Newly synthesised negative-sense strands of PV serve as template for positive-sense strand synthesis, of which a proportion are further translated to increase the number of virus RNA capsid proteins in the cytoplasm. Virus capsid proteins are required to encapsidate single copies of VPg-linked virus

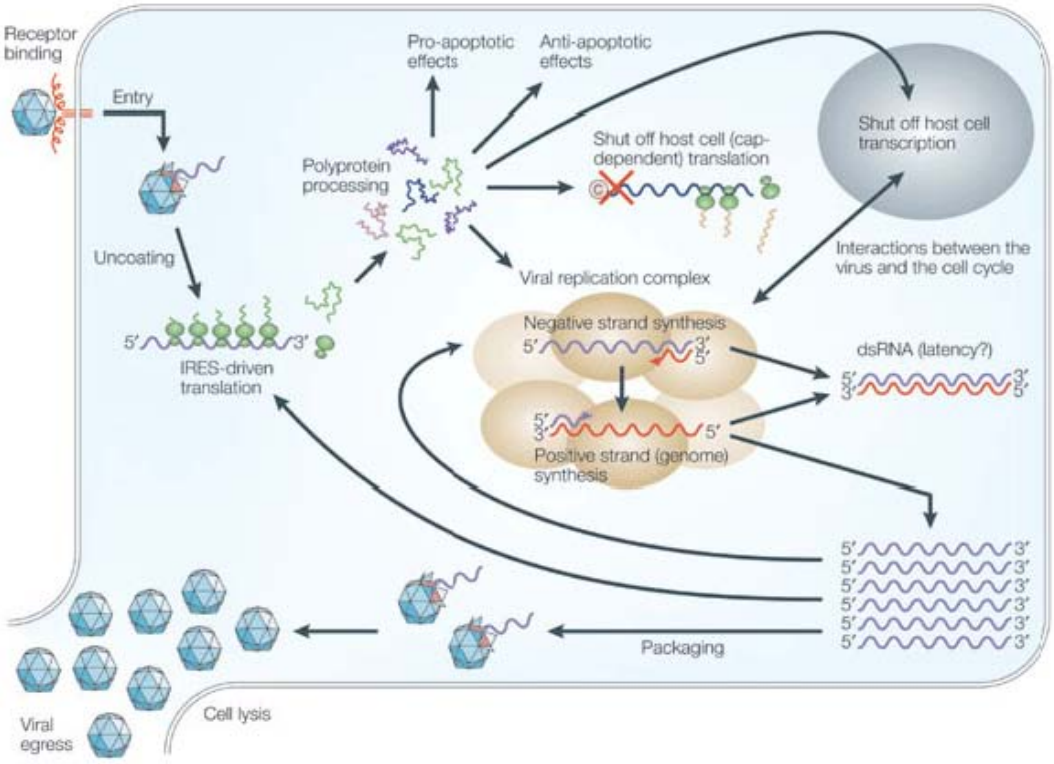


Figure 1.2 Summary of the PV life cycle. A brief outline is described in the text. © represents the m^7G cap which is present on most host mRNAs, but not on PV. Negative-strand RNA is purple, while positive-strand RNA is red. (Source: Whitton *et al.*, 2005).

RNA genomes before mature virions are released from the cell following lysis (Novak & Kirkegaard, 1991). In all, PVs have a short growth cycle which takes approximately 8 hrs from cell attachment to virus release (Mueller *et al.*, 2005).

Poliovirus RNA replication

Although PV RNA synthesis is an intricate process, it has been examined extensively to provide a clear understanding of the viral and cellular proteins involved. Once the virus genome is released into the cytoplasm, VPg is cleaved by a cellular enzyme (Ambros et al., 1978). The genome is translated and virus proteins necessary for RNA replication and for inhibiting host cell processes are produced. Virus RNA replication proceeds in a replication complex, formed by the rearrangement of cellular membranes, and involves virus and cellular proteins that initiate negative-strand RNA synthesis by genome circularisation (Ansardi *et al.*, 1996; Herold & Andino, 2001).

Replication complexes have been studied extensively, including their role in virus replication (Bienz *et al.*, 1980; Bienz *et al.*, 1990). The P2 virus proteins are found to be associated with the endoplasmic reticulum from 3 to 3.5 hrs post-infection, and RNA synthesis begins to decline approximately 1 hr later (Ansardi et al., 1996, Bienz et al., 1987). These virus-induced vesicles are described as closed 'entities' which limit or prevent the exchange of viral protein, RNA, or membranes, in order to concentrate necessary components (Egger *et al.*, 2000; Miller & Krijnse-Locker, 2008). They may also protect newly synthesised virus RNA from cellular factors and competing processes (Ahlquist, 2002).

Virus proteins 2C and 3AB are responsible for binding virus RNA to membranes during replication, as $3D^{pol}$ cannot associate with membranes (Plotch & Palant, 1995; Semler *et al.*, 1982; Teterina *et al.*, 1997). Figure 1.3 outlines PV negative-strand RNA synthesis. Virus precursor protein $3CD^{pro}$, together with poly(rC)-binding protein 2 (PCBP2) facilitates the formation of a ribonucleprotein complex as both bind to the 5' terminal cloverleaf structure on the positive-strand RNA (Andino *et al.*, 1990; Andino *et al.*, 1993; Gamarnik & Andino, 1997). Poly(A)-binding protein 1 (PABP1), which is bound to the 3' poly(A)-tail also interacts with the ribonucleoprotein complex to circularise the genome (Wang *et al.*, 1999). The CRE

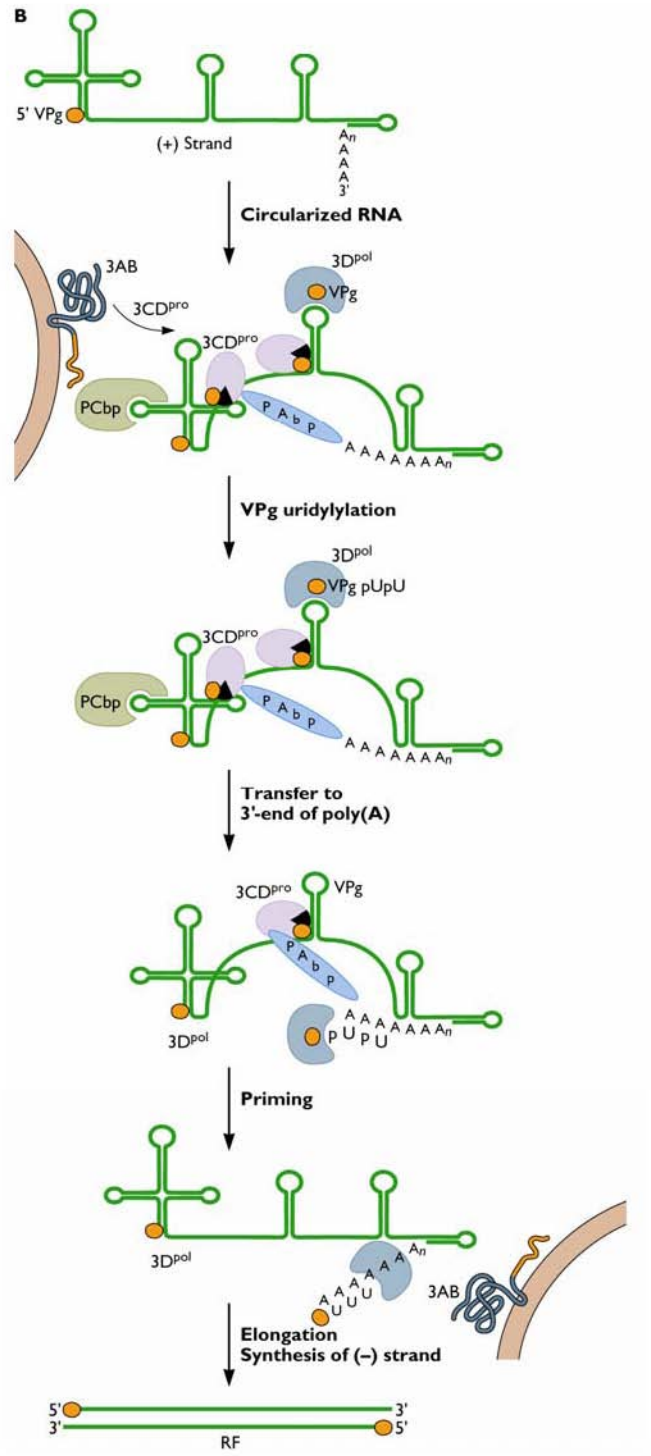


Figure 1.3 Poliovirus negative-strand synthesis. Once the virus genome is released into the cytoplasm, VPg, a virus protein covalently linked to the 5' terminal of RNA, is cleaved by a cellular enzyme. The genome is translated and virus proteins necessary for RNA replication are produced. Virus RNA replication proceeds in a replication complex, which incorporates cellular membranes associated with initial translation, and virus and cellular proteins that initiate negative-strand RNA synthesis by genome circularisation. (Source: Flint *et al*., 2004).

structure also binds $3D^{pol}$, $3CD^{pro}$, and VPg in which the latter serves as a protein primer for RNA synthesis. Negative-strand RNA synthesis proceeds by $3D^{pol}$ elongation once the complex is transferred to the 3' end of the genome. As a result, a double stranded replicative form (RF) is produced, containing a positive- and negative-strand virus genome.

Following synthesis of negative RNA strands, membrane-associated virus protein 2C anchors this RNA strand to replication membranes (Banerjee & Dasgupta, 2001) (figure 1.4). It has been demonstrated that guanidine inhibits initiation of negative-strand synthesis and mutants resistant to guanidine contain mutations mapped to the 2C region (Barton & Flanegan, 1997). Binding to protein 2C might aid recruitment of uridylylated VPg-containing complexes to the 3' end of the negative-strand to initiate positive-strand RNA synthesis. VPg-pUpU (bound to the 3' end of negative-strand) is uridylylated by $3D^{pol}$ using 3' terminal A residues, and serves as the primer for positive-sense RNA synthesis. VPg-pUpU is elongated by $3D^{pol}$ to synthesise multiple copies of positive-sense strands (Barton & Flanegan, 1997; Richards *et al.*, 1984) and allows for suitable genome packaging with structural proteins prior to virion release.

The use of reverse genetics (infectious virus clones and subgenomic replicons)

The study of viruses, such as their genome structure, functions, and interactions (both internally and with the host cell) has benefited greatly from the ability to make mutations in cDNA and then determine the phenotype by recovery of virus in cell culture. Engineered PV genomes, for example, have been a powerful tool for investigating the critical substitutions that attenuate neurovirulent vaccine strains during development of the OPV (reviewed in Almond, 1987). The sequencing and characterisation of virus RNA genomes was originally difficult due to instability of the nucleic acid as well as its susceptibility to ribonucleases, and this limited virus studies. The molecular study of RNA viruses was only improved in the last 30 years when groups were able to recover infectious virus from cloned cDNA copies. Briefly, these clones comprise partial- or full-length virus cDNA genomes inserted into self-replicating vectors that allow for the generation of multiple copies through culture in transformed bacteria. Ribonucleic acid genomes that are transcribed from cDNA clones *in vitro* are shown to generate infectious virus upon transfection into

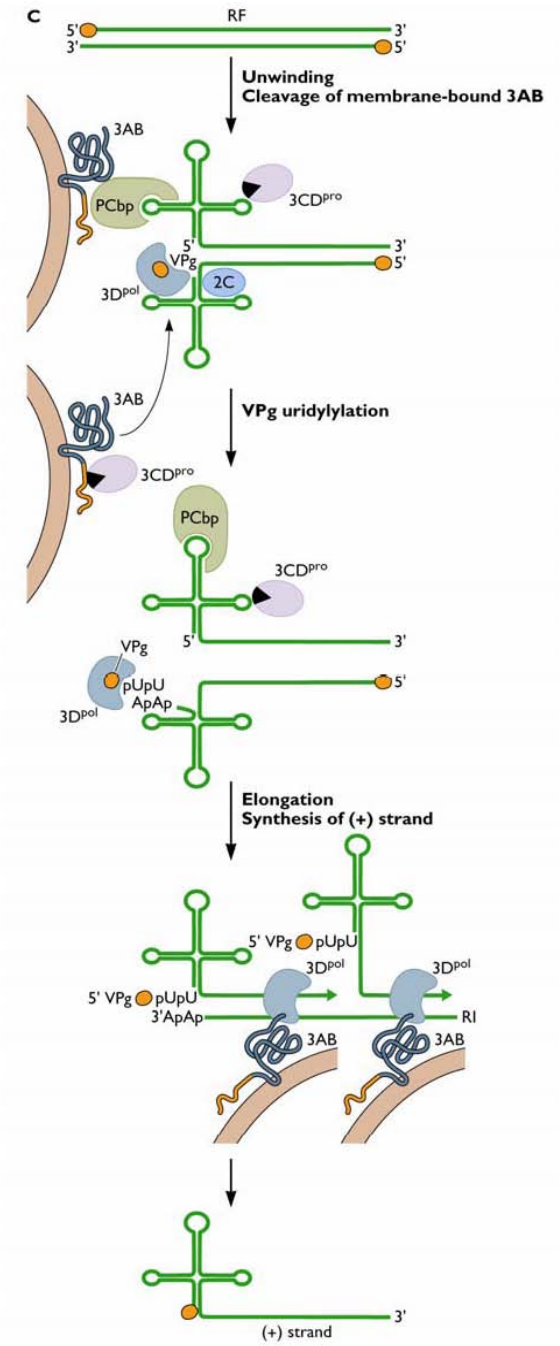


Figure 1.4 Poliovirus positive-strand RNA synthesis. Following synthesis of negative RNA strands, it is hypothesised that membrane-associated virus protein 2C anchors this RNA strand to replication membranes. This might aid recruitment of uridylated VPg-containing complexes to the 3' end of the negative-strand to initiate positive-strand RNA synthesis. VPg-pUpU (attached to the 3' end of negative-strand) is synthesised by $3D^{pol}$ using 3' terminal A residues, and serves as the primer for positive-sense RNA synthesis. VPg-pUpU is elongated by $3D^{pol}$ to synthesise multiple copies of positive-sense strands and allows for suitable genome packaging with structural proteins prior to virion release. (Source: Flint *et al*., 2004).

permissive cells that are able to support infection. This was first achieved by Taniguchi and colleagues with the recovery of an RNA bacteriophage, Qβ, from cloned cDNA (Taniguchi et al., 1978). The cloning of the first animal RNA virus, PV, occurred only a few years later. Racaniello and Baltimore initially cloned three fragments spanning the PV1 genome into plasmids before combining these into one final clone to make the first infectious PV clone (Racaniello & Baltimore, 1981a; b).

Likewise, the HRV14 genome was cloned in its entirety and infectious virus was recovered from RNA transcripts transfected in HeLa cells (Mizutani & Colonno, 1985). Subsequently, the coxsackievirus B3(CVB3) genome was cloned directly into pBR322, and the entire coxsackievirus B4 (CVB4) (strain JVB) genome was cloned also into pBR322, both by the cDNA:RNA hybrid method to construct a first full-length infectious clone for sequence comparison (Jenkins *et al.*, 1987; Tracy *et al.*, 1985). Since then, molecular techniques have improved to construct complete enterovirus cDNA plasmids, especially using amplification by polymerase chain reaction (PCR) and cloning of entire genomes (Lindberg *et al.*, 1997).

Infectious cDNA in plasmid form allows for genetic manipulations that cannot be performed on an RNA genome alone. Mutations or modifications can be introduced into the cDNA in order to characterise the properties of subsequently recovered virus *in vitro*. One such example was the disruption of the CRE structure by site-directed mutagenesis, which led to the inhibition of positive-strand RNA synthesis and ultimately a lack of recovered virus in cell culture (Goodfellow *et al.*, 2000).

Virus cDNA plasmids are also useful in order to map spontaneous mutations that arise when wild-type viruses are able to grow under restrictive conditions (selective pressures). Mutations that allow viruses to escape such restrictive conditions can therefore be identified by sequencing and used for recombination studies. One such example was the mapping of PV mutants that were resistant to 2 millimolar (mM) guanidine, which usually inhibits replication (Emini *et al.*, 1984). Resistance to guanidine was mapped to mutations in the 2C region, and similar studies have identified genetic markers linked to selection by temperature, escape from monoclonal antibody neutralisation, and actinomycin D (Dewalt *et al.*, 1990; Kean *et al.*, 1989; Reynolds *et al.*, 1992).

Subgenomic replicons emerged as a powerful tool in the mid 1980s when researchers became interested in the mechanism of PV replication and the role of defective interfering (DI) particles during virus infection. Defective interfering particles are shown to replicate in cells but due to missing portions of their genome, are not able to propagate virus unless in the presence of helper virus (Kuge *et al.*, 1986). Minimal genome elements required for replication and signals necessary for virion packaging can be determined as a result of such particles. Poliovirus DI particles have been isolated from virus populations propagated in cell culture and results indicate that the locations of deletions were limited to the capsid coding region (Hagino-Yamagishi & Nomoto, 1989; Kajigaya *et al.*, 1985). This suggests that such incomplete genomes retain the ability to replicate and package due to *cis*-acting non-structural proteins, in the presence of infectious virus. This knowledge was applied to the construction of subgenomic replicons in which large fragments of the P1 region were deleted from infectious PV cDNA clones (Kaplan & Racaniello, 1988). Enterovirus cDNA genomes can be altered by the replacement of the capsid region with a reporter gene (e.g. luciferase) to create subgenomic replicons. The capsid can be removed without any consequences to RNA synthesis and allows for accurate quantitative detection of RNA, replication, and translation by *in vitro* assay. A subgenomic replicon for PV3 was constructed by the removal of the capsid region and replacement by a reporter gene encoding chloramphenicol acetyltransferase (CAT) (Percy *et al.*, 1992). A luciferase-expressing Mahoney strain of PV1 replicon was also engineered to study the ribonucleoprotein complex during virus replication (Andino *et al.*, 1993). These are a few of the several virus clones that exist for picornaviruses.

1.3 Enterovirus Evolution

Virus evolution

Evolution requires genetic variation. Positive-sense single-stranded RNA viruses including important human pathogens like PV and hepatitis C virus (HCV) have error prone polymerases, short replication cycles, and high virus yields that together contribute to genetic diversity. The small genome of PV ultimately allows for it to replicate quickly so that more progeny RNA can be generated, and this has advantages when adaptation to host conditions is required.

Virus genome adaptability in a host depends on three broad factors; (1) the genetic heterogeneity of the population (or quasispecies) or the number of stable mutations in individual genomes compared to the consensus sequence; (2) the population size; and (3) the complexity of the viral genome (Domingo & Holland, 1997). Consequently, a viral population can contain genomes with one or several mutations that make mutants dissimilar from the type species or reference genomes (Bouslama *et al.*, 2007). This high level of variation allows RNA viruses to exist as a heterogeneous population and is the reason why viruses are constantly changing, with the continuous emergence of new strains from a quasispecies population. Mutations within viral populations lead to variations in virulence and possibly modifications in host cell tropism as a result of selective pressures imposed by the host.

Mechanisms of evolution in enteroviruses

Most genetic diversity in positive-sense RNA virus populations arise from mutations caused by the error-prone nature of the virus's own RdRp (Domingo & Holland, 1997, Steinhauer et al., 1992). As a result, enteroviruses (and indeed all positive-sense RNA viruses), exist as quasispecies; a collection of non-identical but related genomes (Lauring and Andino, 2010; Domingo & Holland, 1997). This is significant given the size of such quasispecies genome populations. Virally encoded RdRps lack exonucleolytic editing ability (Ishihama *et al.*, 1986), and it has been shown that PV polymerases have an estimated mis-insertion error rate ranging from 1.2×10^{-4} to 1×10^{-6} for transition mutations and 3.2×10^{-5} to 4.3×10^{-7} for transversion mutations (Freistadt *et al.*, 2007). As a result, repeated passages of large populations of virus over time enrich those mutant genomes that replicate more efficiently, and can also alter virulence by modifying host cell tropism (Domingo & Holland, 1997). The best example of this is the attenuation of the three serotypes of PV to create the OPV. Strains of PV were repeatedly passaged through monkey cells (*in vivo* and *ex vivo*) to eventually select viruses with multiple spontaneous point mutations that led to the inability of these strains to replicate as efficiently in human neurons while retaining the ability to replicate in the human gut cells where they were capable of inducing the desired protective immune response (Kew et al., 2005, Sabin & Boulger, 1973). Conversely, reversion to a virulent phenotype by point mutation is a risk associated with oral polio vaccination (Rezapkin *et al.*, 1995; Taffs *et al.*, 1995)

and was demonstrated by Westrop *et al*., when they showed that attenuation of PV3 Sabin vaccine strain from its neurovirulent wild-type parent only required two point mutations at positions 472 and 2034 of the genome (Westrop et al., 1989).

The natural replication of all three PV serotypes can be followed easily due to rapid genome evolution, and this can be calculated with high precision. Nucleotide substitutions (of which > 80% are in the coding region) accumulate at an average rate of ~ 1% per year at each position, and this rate increases to ~ 3% per year when calculated for synonymous sites only (Kew *et al.*, 1995; Martín *et al.*, 2000). Rapid evolution of vaccine viruses, such as PV, also accounts for the adverse reactions associated with the use of live-attenuated vaccines. The level by which VDPVs differ from the vaccine Sabin strains can help determine when such viruses were originally transmitted. Isolated VDPVs can differ from OPV strains at 1% to 15% of VP1 nts when resequenced, which provides useful timelines for tracing origins of outbreaks (Shimizu *et al.*, 2004).

Over a decade ago, extensive sequence and phylogenetic analysis of enterovirus genomes revealed that evolutionary rates differed depending on the genomic region studied (Hyypiä et al., 1997). Different selection pressures are most likely applied to different regions; the capsid region is directly linked to host immune responses and over a significant amount of time, has diversified and evolved due to genetic drift, and this is clear from the wide range of receptors recognised for cell attachment. The 5' NCR shows the slowest rate of evolution (Hyypiä *et al.*, 1997; McWilliam Leitch *et al.*, 2009). Only two distinct phylogenetic groups exists for 5' NCR in human enteroviruses, those belonging to species A and B, and the remaining serotypes in species C and D (Santti *et al.*, 1999).

Even though the RdRp of RNA viruses are usually cited as the reason for high sequence variability, the evolutionary rates observed in PV sequences are in fact higher than those attributed to RdRp fidelity alone (Freistadt *et al.*, 2007). Genetic recombination may well provide the opportunity for evolutionary advancement that may be greater than that seen by the steady accumulation of point mutations. Coupled with the paucity of information of the molecular mechanisms and influences

on the process, the causes and outcomes of recombination were evaluated as part of this project.

Evidence of virus recombination

Recombination is a way for viruses to achieve dramatic alterations of a significant part of the genome. Evolution therefore occurs quicker than by the gradual acquisition and accumulation of natural mutations during replication (genetic drift). As a general rule, recombinants arise and evolve from either initial dual infections in the same cell, or as a result of superinfections. Superinfection is the process by which a cell that has previously been infected with a virus is subsequently co-infected by a different virus strain or type altogether.

The phenomenon of recombination in non-segmented RNA viruses has been demonstrated previously and is an important evolutionary mechanism. Interspecies recombination in alphaviruses between an eastern equine encephalitis-like virus (EEEV) and a Sindbis-like virus gave rise to a related new virus – western equine encephalitis virus (WEEV) (Hahn *et al.*, 1988). This virus combined the encephalogenic properties of EEEV but maintained the antigenic properties of Sindbis virus.

In a further example, a recombination event between virus and host RNA enhanced the cytopathic properties of the virus, bovine viral diarrhoea virus (BVDV). Cytopathic and non-cytopathic forms of BVDV are classified by culture through cells, and it was determined that the cytopathic form contained host-cellular RNA encoding ubiquitin (Meyers et al., 1989). It was suggested that in persistently infected cattle, non-cytopathic BVDV evolved to the cytopathic form, triggered by recombination events that inserted cellular RNA into the virus genome. Recombination events (not involving host-cellular RNA) have enhanced virus pathogenicity in other animal viruses infecting cats, including feline infectious peritonitis viruses (FIPVs) (Vennema *et al.*, 1998) and feline coronavirus (FCoV) type II (Herrewegh *et al.*, 1995).

Recombination in enteroviruses

Recombination in RNA viruses was first described in PV in the early 1960s (Hirst, 1962; Ledinko, 1963) and it has been identified since then in many enteroviruses, and studied extensively in PV (Agol *et al.*, 1985; Minor *et al.*, 1986; Romanova *et al.*, 1980; Savolainen-Kopra *et al.*, 2009b). In the early PV experiments, both groups were able to characterise recombinants generated *in vitro* by mapping different selection markers passed onto recombinant progeny virus. Since then, recombination between PV serotypes has been shown to occur as soon as 2.5 hrs post co-infection in replication complexes within HeLa cells (Egger & Bienz, 2002). Once cells were co-infected with PV1 (Mahoney strain) and PV2 (Sabin strain), type-specific fluorescent riboprobes were used to visualise the genome regions within replication complexes by fluorescent *in situ* hybridisation (FISH). Recombination was subsequently demonstrated by RT-PCR and sequencing.

Multiple levels of recombination have been identified in natural PVs outside the laboratory: intraserotypic recombinants (within each PV serotype) and interserotypic recombinants (combining PV1, PV2, or PV3) have been isolated from OPV vaccinees (Cammack *et al.*, 1988; Cuervo *et al.*, 2001; Furione *et al.*, 1993; Yang *et al.*, 2005); wild-type PV serotypes 1, 2, and 3 have been shown to recombine with each other (Dahourou *et al.*, 2002); wild-type and OPV strains have recombined (Georgescu *et al.*, 1995; Guillot *et al.*, 2000), and lastly, vaccine-derived PVs have recombined with unknown HEV-C strains (Kew *et al.*, 2002; Rakoto-Andrianarivelo *et al.*, 2007; Shimizu *et al.*, 2004). As yet, there have been no reported cases of interspecies recombinants involving PV and other HEV species. Interspecies recombination is rare (Simmonds & Welch, 2006), and so far only one event has been suggested between HEV-B and HEV-C viruses (Bolanaki et al., 2007).

The trivalent nature of the live attenuated OPV has, as a consequence, provided the best opportunity for interserotypic recombination to occur. High titres of all three PV serotypes are introduced into the one host at the same time, and target the same tissues. Recombination events between the PV serotypes of OPV recipients have been documented with higher frequencies between serotypes 2 and 3, which are considered to be less genetically stable than serotype 1 (Cammack *et al.*, 1988; Furione *et al.*, 1993; Paximadi *et al.*, 2006). Better molecular methods have also

resulted in more cases of intraspecies recombinants being identified from PV outbreaks. For example, phylogenetic analysis revealed that the source of the 2000-2001 outbreak of poliomyelitis in Hispaniola was due to recombinants formed between OPV derived PVs and the non-structural regions derived from an unknown enterovirus strain (species C) (Kew *et al.*, 2002). Two other reports of cVDPV in Madagascar and the Philippines, implicated vaccine/non-polio recombinants (Rousset *et al.*, 2003; Shimizu *et al.*, 2004). These and other reports have confirmed that live attenuated strains of PV originating from vaccinees are able to recombine with circulating non-polio species C enteroviruses creating neurovirulent recombinant viruses capable of causing serious disease. Outbreaks attributed to PV or recombinant forms of PV are more likely to occur in regions where vaccine coverage is sub-optimal and the circulation of other enteroviruses is high (Kew *et al.*, 2005).

Certain enteroviruses are thought to have emerged as a result of early recombination events. This includes evidence of speciation of PV from a coxsackievirus A (species C) ancestor that occurred through mutation of the capsid region. This speciation event involved a cellular receptor switch in order to possibly escape selective pressure imposed by the host (Jiang *et al.*, 2007). Extensive analysis of HEV B species has also indicated recombination events involving other unknown species B virus sequences and this has occurred within the HEV B species group many times to create a complex mosaic of RNA genomes in a short, eight year period (Simmonds & Welch, 2006). Surprisingly, recombination events were observed not only between VP1 and $3D^{pol}$, but between VP1 and VP4 virus sequences. Frequencies of recombination (between VP1 and 3D regions) were also related to time periods (i.e. virus isolation dates). Approximately 40% of species B isolates from the same year were recombinant, and this rose to 70% for isolates collected two to three years apart. Almost all isolates were recombinant in nature when the collection period exceeded three years. It was also noted that enterovirus diversity occurred asymmetrically between structural and non-structural coding regions, particularly in echovirus (E30) strains (McWilliam Leitch *et al.*, 2009). Phylogenetic analysis shows that while VP1 sequences diverged by natural genetic drift, $3D^{pol}$ sequences clustered in groups and interspersed with other species B serotypes, further demonstrating the 'modular' evolution of structural and non-structural genome

regions. This illustrates the dynamic nature of species B evolution, where fluctuating outbreaks of E30 activity have been associated with new genomic lineages that have replaced previously circulating genotypes (Oberste et al., 1999).

The location of recombination sites in enteroviruses appear to occur throughout the genome with most crossover events occurring in the 5' NCR and in the 2A, 2B, 2C, and 3D protein encoding regions (Savolainen-Kopra et al., 2009). Recombination sites in the capsid region predominantly occur between PV serotypes 2 and 3, and usually at the C-terminal end of the protein region (Blomqvist *et al.*, 2010; Tao *et al.*, 2010). Interserotypic recombination in the capsid region is rare, possibly due to the sensitivity of structural constraints that are important in order to maintain the correct capsid shell.

Although recombination has been responsible for diverging viruses from ancestral lineages, the role of recombination in allowing virus populations to respond to selective pressures is unclear. At a genetic level this would include recombination in the genome that is 'selected for' by specific events, such as immunological responses from the host, differing temperature ranges of host cells and the influence of environmental chemicals.

Methods of RNA virus recombination

As mentioned briefly, genetic recombination is also responsible for and drives RNA virus evolution and this has been well documented in enteroviruses. This phenomena generates novel genomes that may encourage a selective advantage over the original parental genomes (Nagy & Simon, 1997). Genetic recombination involves the exchange of sequences between two non-segmented genomes and this can occur via two distinct methods. Non-homologous recombination occurs in genetic material that does not share sequence homology and this is relatively infrequent in RNA viruses (Kirkegaard & Baltimore, 1986, Lai, 1992). This process is a replication-independent mechanism of breakage and rejoining of segments of genetic material and is more commonly seen in DNA viruses (Lai, 1992). Until recently, only a few studies had investigated replication-independent recombination, however reports have described this process in PVs (Gmyl *et al.*, 1999; Gmyl *et al.*, 2003), BVDV (Gallei *et al.*, 2004), and bacteriophage Qβ (Chetverin *et al.*, 1997). It was suggested that

recombination in PV may occur as the joining of broken RNA molecules by transesterification (Gmyl *et al*., 1999), although this would occur less often. Homologous recombination by contrast, is the result of crossovers between two similar or closely related RNA molecules containing extensive sequence similarity. This form of recombination is known as 'copy-choice' and is the consequence of template switching by the RdRp during negative-strand synthesis as demonstrated by Kirkegaard and Baltimore (1986). 'Copy choice' was first proposed as the mechanism for recombination in PV (Copper et al., 1974), and to date, the majority of studies of recombination in RNA viruses have supported this model (reviewed in Agol, 1997).

RdRp 'copy choice' by template switching

Replication-dependent recombination in plant and animal viruses requires RNA templates and viral replicase for successful template switching to occur. The simplest model suggests that two RNA templates are necessary, including the primary RNA template (donor) which binds RdRp to initiate replication, and the acceptor RNA template that receives the RdRp mid-elongation (figure 1.5A). The crossover sites or 'breakpoints' occur at positions similarly matched in sequence so that the resultant recombinant RNA templates retain the same sequence length and structural integrity as the parental RNA molecules (Lai, 1992). This is commonly referred to as precise recombination. Imprecise recombination, on the other hand, results in changes including nucleotide mismatches, deletions, or insertions at the junction site, or close to it (figure 1.5B) (Lai, 1992; Nagy & Simon, 1997). Depending on the modifications that have occurred, a viable virus may still result from this form of genetic exchange. 'Copy choice' template switching may not always occur between homologous RNA strands, but crossover sites in coding regions are somewhat constrained by reading frame, recognition sequences for proteolytic processing, and protein function, which could conceal such events (Wimmer et al., 1993).

Template switching during replication is also thought to be responsible for the generation of naturally occurring deletion mutants, or DI particles (Lazzarini et al., 1981). They can be generated naturally by multiple passage at high multiplicity of infection (MOI), (i.e. 16 to 18 passages for PV1 Mahoney strain), and deletions in PV genomes are located in the capsid encoding region (Cole *et al.*, 1971; Kajigaya *et*

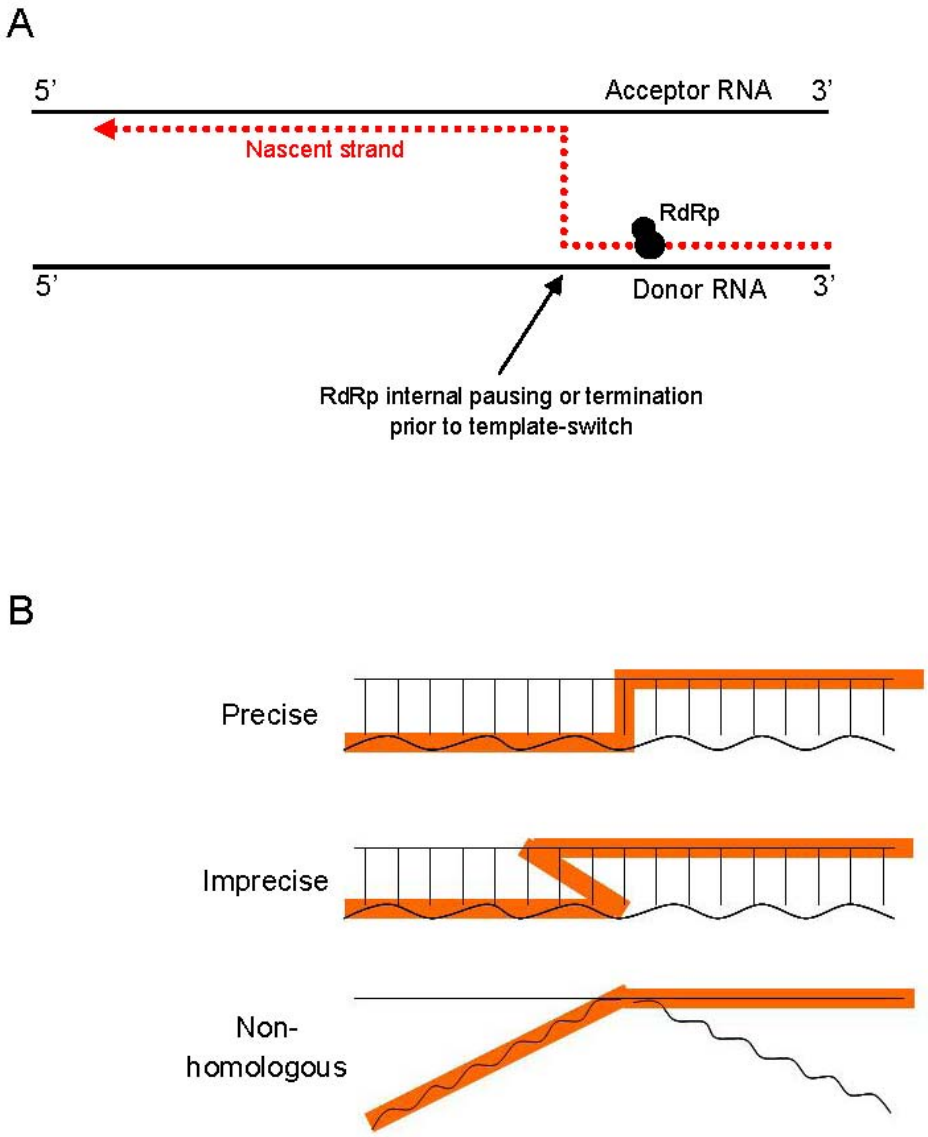


Figure 1.5 Current replication-dependent template switching models for RNA recombination. (A) Template switching during replication. The RdRp exchanges one template (donor) for another (acceptor) at corresponding positions in the genome. There are several possibilities as to why the RdRp is removed from the donor template (see text). (B) Diagrams depicting precise and imprecise recombination events during replication. Non-homologous recombination is also demonstrated, although this may not be dependent on replication. (Source: Lai, 1992).

al., 1985;(Cole *et al.*, 1971; Kajigaya *et al.*, 1985; Kuge *et al.*, 1986; Wimmer *et al.*, 1993). It was observed that all deletions in isolated DI particles of PV occurred in frame (Kuge *et al.*, 1986), suggesting that some viral proteins were required to be encoded *in cis* for genome replication. Poliovirus DI particles can therefore initiate their own replication but cannot produce progeny virions as they do not synthesise capsid proteins – these must be provided by a homologous helper virus.

Since the same mechanisms for RNA recombination are thought to be responsible for the generation of DI particles, Kuge and colleagues performed comparative sequence analysis of PV1 DI particles to determine the underlying genetic mechanisms behind their generation (Kuge *et al.*, 1986). Selection pressures may have influenced the location of deletions, as these were limited to between nucleotide positions 1226 and 2705 of the capsid coding region. There was no evidence of sequence homology between immediate donor and acceptor template sites, and it was proposed that deleted regions were 'looped out' by secondary RNA structures, meaning that the RdRp skipped these regions during replication. As PV DI particles are rarely formed or identified, further study of the mechanisms behind their formation are limited.

What mechanisms influence recombination?

At the genomic level, intrinsic signals present on the donor or acceptor RNA templates may influence switching of the RdRp between templates. At present, it is suggested that RNA signals either pause or terminate the RdRp, both of which cause the release of RdRp from the RNA template (King, 1988; Nagy & Simon, 1997). Possible signals include: A/U-rich and U-rich sequences that promote RdRp slippage, local strong hairpin structures present on either template, the formation of stable heteroduplexes between templates, and mis-incorporations (Carpenter *et al.*, 1995; Makino *et al.*, 1986; Pilipenko *et al.*, 1995; Tolskaya *et al.*, 1987). If local RNA signals are essential for template switching of the RdRp, then it is unclear why there is no consistency of crossover sites for the same PV strains reported in the literature.

Researchers have succeeded in achieving genetic recombination of enteroviruses *in vitro* (in particular in PV) in an attempt to define the preferred region/s (but not specifically the RNA signals) for recombination. Duggal and Wimmer (1999)

demonstrated that selective pressure (e.g. temperature) influenced the regions at which crossover occurred between two PV RNA strands (Duggal & Wimmer, 1999). Recombinants produced at 34°C showed crossover sites predominantly in the capsid region while an increase in assay temperature (37°C - 40°C) shifted the crossover sites towards the non-structural genes. Sequence similarity also appears to be important for successful recombination and is detectable as a viable recombinant observed using *in vitro* methods. Kirkegaard & Baltimore, (1986) demonstrated that interserotypic recombination was 100 times less likely to occur than intraserotypic recombination in PV. This was measured using serotypes containing genetic markers (one was guanidine dependent and the other temperature sensitive) which were separated by 190 nts in the genome. Recombination frequencies (intra- and interserotypic crosses) were given as the yield of recombinant virus divided by the sum of yields of parental viruses. Although very few events have involved exchanges of sequences encoding regions of the capsid, the majority of studies have found that recombination producing viable virus occurs more frequently in the non-structural region of the genome, particularly in the 2A-2C region of enteroviruses although the precise regions of recombination or necessary requirements for high titre resultant progeny are not presently known (Chen *et al.*, 2007; Kew *et al.*, 2002; Oberste *et al.*, 2004b). Many constraints can be imposed on a virus population at several stages of viral infection and replication in the host, although the knowledge of their roles in causing recombination is undefined.

Identifying virus recombination

Sequence analysis coupled with phylogenetic techniques has proven to be effective in identifying and characterising several recombination events among RNA viruses from field isolates (Gao *et al.*, 1998; Revers *et al.*, 1996). Although these methods offer precise locations of putative crossover regions, these studies require full-length genome recombinant and non-recombinant reference sequence information, which is expensive and time consuming to acquire (Oberste *et al*., 2004b; Simmonds & Welch, 2006; (Zoll et al., 2009). Recombinant viruses are also identified by sequencing regions at either end of isolated genomes and identifying incongruent regions by phylogenetic analysis (Oberste *et al.*, 2004c; Simmonds & Welch, 2006). Although now considered an outdated technique, groups used ribonuclease T_1 oligonucleotide fingerprinting to detect structurally unique patterns once viral

genomes were cleaved and separated in two dimensions (Agol *et al.*, 1985; Kew & Nottay, 1984). Likewise, the use of genetic markers and Random Fragment Length Polymorphism (RFLP) have proven successful in identifying recombinant genomes in viral populations (Paximadi et al., 2006).

1.4 Aims

From this brief review of the literature, it is clear that the occurrence of recombination is well established in enteroviruses and other RNA viruses. Recombination acts to progress the evolution of viruses, and must provide advantages to viruses since these recombinants are associated with current disease outbreaks. Although groups have attempted to determine the molecular triggers for template switching in PVs, it is striking that there is no consensus between "hot spots" in these genomes and little agreement as to the intrinsic RNA signals responsible. The overall aim of this study was to investigate the mechanism of recombination and to characterise recombinant viruses obtained during the process.

Specific Aims

- To develop an experimental method for generating a population of natural interserotypic PV recombinants with little or no remaining parental virus in cell culture.
- To facilitate the identification of early recombinant genomes without extensive selection/passage.
- To determine the underlying mechanisms of template switching at the nucleotide and protein level, by characterising recombinants generated by the above method.
- To investigate whether viruses generated by this method acquired growth advantages as a result of recombination.
- To determine if recombination between enterovirus species is possible in an artificial *in vitro* setting.

CHAPTER TWO: Materials and Methods

2.1 Cell Culture and Virological Methods

Cell maintenance

Human cervical (HeLa), human embryonic rhabdomyosarcoma (RD), murine fibroblast (L929), human embryonic lung (MRC-5), and African green monkey kidney (Vero) cells were grown as monolayers in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 100 μg/ml of streptomycin, 100 U/ml of penicillin, 2 mM L-glutamine, and 10% heat inactivated (HI)-foetal bovine serum (FBS). Baby hamster kidney cells stably transfected with a T7 RNA polymerase expression plasmid (BsrT7) were maintained in Glasgow Minimum Essential Medium (GMEM) supplemented with 100 μg/ml of streptomycin, 100 U/ml of penicillin, 2 mM L-glutamine, and 10% HI-FBS. Antibiotic G418 was also added to BsrT7 cell medium to select cells containing the T7 RNA polymerase gene under control of the cytomegalovirus (CMV) promoter and the neomycin resistance gene (Buchholz *et al.*, 1999). Intracellular adhesion molecule-1 expressing RD (RD-ICAM-1) cells were maintained in DMEM supplemented with 100 μg/ml of streptomycin, 100 U/ml of penicillin, 2 mM L-glutamine, and 10% HI-FBS (Shafren *et al.*, 1997). All cells were passaged in the presence of trypsin-ethylenediaminetetracetic acid (EDTA).

Plasmid DNA and RNA transfection of mammalian cell lines

Purified DNA or RNA (amounts as specified elsewhere) in H_2O was made up to 250 μl in serum and antibiotic free medium and mixed with an equal volume solution of Lipofectamine 2000 (3:1 Lipofectamine 2000 to μg DNA), (Invitrogen) in similar medium. The mixture was incubated according to the manufacturer's instructions and added to cell monolayers in 6-well tissue culture plates.

Short interfering RNA transfection of mammalian cell lines

Transfections were performed in triplicate in 24-well format according to the manufacturer's instructions (Invitrogen). Briefly, 0.5 μl Lipofectamine 2000 was diluted in 50 μl serum- and antibiotic-free medium and incubated at room temperature (RT) for 5 minutes (mins), during which 25 picomoles (pmol) of a short

interfering (si)RNA oligomer (Invitrogen) in H_2O was made up to 50 µl in similar medium. Both mixtures were combined and after 20 mins incubation at RT, the solution was added to each well preincubated with 0.4 ml of fresh medium containing 10% HI-FBS.

siRNA transfection of mammalian cells by electroporation

HeLa cells from two T175 flasks were pelleted at 2000 revolutions per minute (rpm) for 5 mins and resuspended in serum-free DMEM twice before quantitating by haemocytometer. Cells were pelleted again and resuspended in sterile phosphate buffered saline (PBS). A Gene Pulser Xcell Electroporation System (Bio-Rad) was utilised for electroporations. 1.5 x 10^6 cells were transferred to a chilled 4 millimetre (mm) cuvette (Molecular BioProducts) together with 200 pmol siRNA and incubated on ice for 5 mins. Cells were electroporated using the following settings: voltage 160 volts (V), resistance $\infty\Omega$, capacitance 500 microfarads (µF), and transferred to a 6-well plate containing 2 ml/well fresh medium with 10% HI-FBS.

Virus infections

Cell monolayers were infected with virus at the stated MOI 24 to 48 hr post-transfection, after removal of transfection media. Virus was absorbed onto monolayers for 30 mins at 37ºC/5% carbon dioxide CO_2/air. Virus supernatant was removed and replaced with medium supplemented with or without 10% HI-FBS. Virus supernatant was removed 24 to 48 hours post-infection, upon completion of full cytopathic effect (CPE; appearance of cell rounding or detached cells).

Luciferase assay

Supernatant was removed from transfected cell monolayers, cells were rinsed with PBS and lysed using 200µl or 500µl 1 x Glo Lysis Buffer (Promega®) per well in a 12- or 6- well plate respectively. The oxidation reaction was catalysed by the addition of 100 µl cell lysate (luciferase enzyme) to 100 µl RT *Bright-Glo*™ Luciferase Assay System (Promega®) substrate and shielded from the light for 5 mins. Luciferase activity was measured using a luminometer (Turner Biosystems) with values expressed in relative light units (RLU).

Tissue culture infectious dose$_{50}$ ($TCID_{50}$)

Virus from RNAi infections were titred by $TCID_{50}$ which measures the quantity of virus required to infect 50% of inoculated microplate wells per sample (Minor, 1985). Each well of 96-well plates was seeded with 3.75 x 10^4 HeLa cells, allowing for four replicates of titration per sample. In a separate 96-well round-bottomed plate, a serial 10-fold dilution of sample with DMEM/10% FBS/penicillin-streptomycin (PS) from 10^{-1} to 10^{-12} was performed. Media was removed from confluent cells, 100 µl of the appropriate virus dilution was added to each well, (including 100 µl DMEM/10% FBS/PS to control wells) and incubated for 1 hour at 37ºC/5% CO_2. An extra 100 µl DMEM/10% FBS/PS was added per well and plates were incubated for four days. Plates were stained with crystal violet for one hour and rinsed with tap water before drying upside down on paper towel. Virus titre was expressed as $\log_{10}$ $TCID_{50}$/ml (Reed & Muench, 1938).

Plaque assay

Cells were seeded in 6-well plates (5 x 10^5 cells per well) and grown to 90% confluency. Ten-fold dilutions of virus stock were made in medium/10% FBS/PS. Once medium was removed from seeded wells, cells were inoculated with 500 µl of prepared virus and incubated for 30 mins at 37°C in the presence of 5% CO_2/air to allow absorption to occur. Plaque overlay media (refer to section 2.3 for contents) was added to each well and plates were incubated for two to three days at 37°C in the presence of 5% CO_2/air. Cells were stained with crystal violet solution and re-stained post-removal of the plaque overlay media. Plaques were counted and the virus titre expressed as plaque forming units per millilitre (PFU/ml).

Immunofluorescence assay

Sterile 15 mm glass coverslips were placed in tissue culture plate wells before cell seeding. Coverslips were gently washed twice with PBS in wells before cells were fixed to coverslips with 4% paraformaldehyde for 10 mins. Coverslips were removed from wells and mounted upside down on glass slides (Menzel-Glaser) with mounting medium for fluorescence containing 4',6-diamidino-2-phenylindole (DAPI) (Vectashield). 4',6-diamidino-2-phenylindole is a fluorescent stain that binds to DNA

in fixed cells (Kapuscinski, 1995). Coverslips were attached to slides by nail polish and the presence of green fluorescent protein (GFP) was detected using a Leica SP2 upright confocal microscope with an ultraviolet (UV) filter.

Replication kinetics assay

Replication kinetics of wild-type and recombinant viruses was compared in single-step growth experiments performed in 12-well plates to determine the kinetics of virus growth (PFU over time) as a measure of viral fitness. HeLa cells were infected at an MOI of 10 for 30 mins before cells were washed twice with sterile PBS and incubated with fresh DMEM/10% FBS/PS. Plates were incubated at 37ºC/5% CO_2/air and supernatant harvested at various time points over 24 hrs. Virus concentrations were determined by plaque assay.

Plaque purification by limiting dilution

Recombinant virus from RNA transfections was biologically cloned by limiting dilution in 96-well microtitre plates. Virus dilutions were performed using a multichannel pipette in a round bottomed 96-well microtitre plate (Sterilin) prior to transfer onto fresh cell monolayers. Fifty-five microlitres of serum-free media was added to each well prior to addition of diluted virus. Approximately 55 µl diluted virus supernatant was added to each well in the top row. After mixing by pipetting, 55 µl of solution was transferred to the next row (already containing 55 µl serum-free media). Pipette tips were discarded after the mixture was mixed in each row and these two-fold dilutions continued down the plate. Fifty microlitres was transferred from each well onto a fresh cell monolayer in flat bottomed 96-well plates seeded with 2.6 x 10^4 HeLa cells/well. The plates were left to allow for virus absorption for 30 mins at 37°C in the presence of 5% CO_2/air. An extra 200 µl serum-free media was added to each well and plates were incubated at 37ºC/5% CO_2/air. After 4 days, the supernatant was transferred to a new microtitre plate and cell monolayers were stained by crystal violet for the presence of complete CPE. RNA was extracted from the highest dilution of virus supernatant per column and RT-PCR was carried out to amplify the crossover region of the recombinant virus.

Statistical analysis (chapter three)

All statistical analysis used SPSS for Windows, Rel. 17.0.0, 2008, Chicago: SPSS Inc with a two sided, non-directional test of the hypothesis with significance set at 5%. Homogeneity of the variance and a normal distribution was not assumed due to the relatively small number of data points in some of the experiments. Individual experiments were not prospectively powered to show statistical significance and therefore treatment differences are considered exploratory. The non-parametric, Kruskal Wallis test was used to test each micro (mi)RNA or siRNA experiment as per the following hypothesis: H_0 = median of the PV serotype 1 or 3 wells = median of the siRNA/siRNA test wells. H_a = median of the PV serotype 1 or 3 wells ≠ median of the miRNA/siRNA test wells. If the Kruskall Wallis test indicated a significant difference between treatments at 5% significance, individual miRNA or siRNA were compared to the PV serotype 1 or 3 virus controls by the non-parametric Mann-Whitney U test. Descriptive summaries are presented with graphs including the geometric mean ± standard error from assays performed in triplicate unless otherwise specified. A significance value of $P < 0.05$ is marked with an asterisks symbol.

Statistical analysis (chapter five)

A parametric, one way analysis of variance (ANOVA) (Dytham, 2003) was used to test whether there was significant difference between the mean point accepted mutation (PAM) values at the crossover sites and the mean PAM values for the entire region. H_0 = mean PAM values at the cross over regions = mean PAM values for the entire sequence. H_a = mean PAM values at the cross over regions ≠ mean PAM values for the entire sequence. The analysis included a two sided, non directional test of the hypothesis with significance set at 5%. A normal distribution was assumed and the homogeneity of variance tested by Levene's statistic ($P = 0.87$).

2.2 Molecular Genetic Techniques

Plasmid DNA extractions from Escherichia coli using commercial kits

Overnight cultures (5 ml and 100 ml) of transformed *E. coli* in Luria-Bertaini (LB) broth with the appropriate antibiotic for selection were pelleted in preparation for minipreps and midipreps respectively. Small scale isolations were carried out using GeneJET™ Plasmid Miniprep Kit (Fermentas) according to the protocol provided.

Plasmid DNA was eluted in 40 µl distilled (d)H_2O. Medium scale isolations were performed according the instructions supplied with the QIAfilter™ Plasmid Midi Kit (Qiagen) and DNA was resuspended in 500 µl dH_2O. A list of plasmids/replicons used during this project is outlined in table 2.1 at the end of this chapter.

Plasmid DNA extractions from Escherichia coli using alkaline lysis

Constituents for Solutions I, II, and III are detailed in Section 2.3 Stock Solutions and Buffers. Single colonies were picked from agar plates and used to inoculate 500 µl LB broth with appropriate antibiotic selection in 1.5 ml microcentrifuge tubes. Tubes were placed in a tube rack which was then wrapped with cling film, and allowed to shake at 250 rpm overnight at 37°C. Fifty microlitres of culture was removed the following day for storage and the remaining culture was centrifuged in the same tube for 2 min at maximum speed. The pellet was resuspended in 100 µl ice-cold cleared lysate buffer (Solution I). Following resuspension, 200 µl lysis mix (Solution II) was added and the tubes gently inverted 4 to 6 times to disrupt the bacterial membranes and release cell contents, at which time 150 µl of potassium acetate (Solution III) was added to neutralise NaOH from the previous step and precipitate genomic DNA and sodium dodecyl sulphate (SDS). The tube contents were mixed gently by inversion and centrifuged at 13 000 rpm in a bench top microcentrifuge for 5 min. Once the supernatant was removed carefully to a new microcentrifuge tube, the contents was combined with 500 µl 100% ethanol and inverted 3 to 4 times. Following incubation at RT for 2 mins, the tubes were centrifuged for 10 min at 13 000 rpm in order to pellet the precipitated DNA. The supernatant was carefully removed and the pellet washed with 200 µl 100% ethanol and centrifuged again for 1 min. The ethanol supernatant was removed and the pellet left to air dry at RT. The pellet was resuspended in 25 µl H_2O and allowed to stand at RT for 10 min.

Virus RNA extractions from cell culture supernatant

RNAeasy Mini Kit (Qiagen) was used for RNA extraction from supernatant recovered from virus infections or transfections according to the manufacturer's instructions. RNA was stored at -80°C and when required, was thawed in an ice bath.

PCR product clean up through column

High Pure PCR Cleanup Micro Kit (Roche Applied Science) was used for DNA purification from PCR product according to the manufacturer's instructions. Sterile dH_2O was used in place of supplied elution buffer. DNA was stored at -20°C until required.

Extraction of DNA from agarose gel

Deoxyribonucleic acid fragments were extracted from agarose gel using the DNA Extraction Kit (Fermentas) according to the manufacturer's instructions. The DNA was eluted in 10 to 30 µl dH_2O and quantitated by spectrophotometry.

Transformation of E. coli with plasmid DNA

Between 1-3 µl ligation mixture was added to 50 µl ice-thawed α-Select Chemically Competent *Escherichia coli* (Bioline) and stored in an ice bath for 20 mins. The mix was placed in a 42°C waterbath for 30 secs before being returned to ice for 30 secs. Five hundred microlitres of SOC medium (Sambrook *et al.*, 2000) was added before incubating the tube in a 37°C shaker (225 rpm) for 1 hour. One hundred fifty microlitres of cells were plated onto LB agar plates supplemented with appropriate antibiotic selection and incubated overnight at 37°C.

A-tail reaction before ligation into blunt/T vectors

Single deoxyadenosine (A) overhangs were attached to the 3' end of amplified fragments by a shorter PCR cycle including deoxyribonucleotide adenosine trisphosphate (dATP) only and *Taq* DNA Polymerase (Fermentas) according to "Amplification of DNA fragments (up to 3kb) - Polymerase Chain Reaction (PCR)" later in the chapter. Briefly, 7 µl of column purified PCR product was added to a standard PCR mixture (with only dATP) in a 10µl reaction. The mixture was incubated at 70°C for 30 mins and purified again through a column as previously stated.

Ligation of DNA fragments

Amplicon or digested DNA fragments were ligated into vectors at a molar ratio of 3:1 (insert:vector) in a 10 µl reaction containing 1 x Ligation Buffer (400 mM Tris-HCl, 100 mM $MgCl_2$, 100 mM dithiothreitol (DTT), 5 mM adenosine trisphosphate

(ATP) (pH 7.8 at 25°C)) and 5 U T4 DNA Ligase (Fermentas). Reactions were incubated at 4°C overnight and used directly without further purification for bacterial transformations.

DNA sequencing

Four hundred nanograms (ng) plasmid DNA or 70 ng amplicon was added to 5.5 pmol sequencing oligonucleotide in a total reaction volume of 10 µl. DNA was sequenced by Molecular Biology Services at University of Warwick using an ABI PRISM 3130xl Genetic Analyser. Sequences were analysed using the Lasergene® v6.0 sequence analysis package by DNA*®.

Restriction enzyme digestion

Digestion of DNA was carried out using 1 U or 10 U of restriction enzyme in a solution containing 1 x the specific reaction buffer supplemented, and if required, with 100 µg/ml bovine serum albumin (BSA). Incubation temperature and subsequent thermal inactivation were carried out according to the manufacturer's instructions. The digestion reaction was incubated from 1 to 4 hrs before being run on a 1% weight per volume total (w/v) agarose gel for size/banding pattern analysis. The restriction sites used to linearise plasmids during this project are outlined in table 2.2 at the end of this chapter.

In vitro reverse transcription (cDNA synthesis)

Reverse transcription reactions were carried out using Superscript II Reverse Transcriptase (Invitrogen). Ten microlitres of purified RNA (at unknown concentration) was incubated in a mixture containing 100 pmol oligo dT (Invitrogen), 10 mM dNTP mix and 1.5 µl H_2O for 5 min at 65ºC. Following a 2 min cool on an ice bath, 0.2 µmoles DTT, 1 x Superscript Buffer (250 mM Tris-HCl, pH 8.3 at RT, 375 mM KCl, 15 mM $MgCl_2$), and 20 U RiboLock RNase Inhibitor (Fermentas) was added and incubated for 2 mins at 46ºC. Two hundred units of Superscript II was added to the reaction before a final 50 min incubation at 46ºC and reaction termination at 70ºC for 15 mins. The cDNA mixture was stored at -20ºC while not in use.

Amplification of DNA fragments (up to 3kb) - PCR

A master mix was prepared containing 1 x *Taq* Buffer with $(NH_4)_2SO_4$ (750 mM Tris-HCl (pH 8.8 at 25.5°C), 200 mM $(NH_4)_2SO_4$, 0.1% Tween-20), 2.5 mM each deoxyribonucelotide triphosphate (dNTP), 30 pmoles of the relevant forward and reverse oligonucleotide, and 2.5 U *Taq* DNA Polymerase (Fermentas) in a 50 µl reaction volume.

Thermal cycling comprised 25-30 cycles as follows: one cycle of denaturation for 2 mins at 95°C, 24-29 cycles of denaturation for 1 min at 95°C, annealing for 30 sec at 55°C (or different according to oligonucleotide conditions) and elongation for 1 min/kilo base pairs (kps) at 72°C, and a last cycle of 5 min at 72°C. A list of oligonucleotides is recorded in table 2.3 at the end of this chapter.

Amplification of DNA fragments (full genome) - PCR

This method was applicable for amplification of full-length genomes. A master mix was prepared containing 1 x PCR Buffer for KOD XL DNA Polymerase (contents proprietary), 0.2 mM of each dNTP, 50 pmoles of the relevant forward and reverse oligonucleotide, and 2.5 U KOD XL DNA Polymerase (Novagen) in a 50 µL reaction volume. Three µl of template was added following first strand synthesis of cDNA from virus RNA extracted from tissue culture supernatant.

Thermal cycling comprised the following: 30 cycles of denaturation for 30 seconds (secs) at 94°C, annealing for 10 secs at 61°C and elongation for 6 mins at 72°C, and a final cycle of 10 mins at 74°C.

Amplification of DNA fragments –PCR gradient

Gradient PCR was used in instances where the annealing temperature of oligonucleotides was undetermined. The PCR mixture was set up according to protocols outlined in this chapter. Thermal cycling conditions replaced one annealing temperature with a range of temperatures (usually in increments ranging from 40 - 60ºC across the heating block) to assess the hybridisation ability of oligonucleotides. Products from different cycling conditions were analysed on an agarose gel and optimal annealing temperature determined.

Site-directed mutagenesis

Point mutations were introduced into virus encoding plasmids using QuikChange® Site-directed Mutagenesis Kit, or QuikChange® II XL Site-directed Mutagenesis Kit (Stratagene). Briefly, 20 to 40-mer oligonucleotides were designed for opposite strands of the plasmid and incorporated nucleotide mutations necessary for the project. The plasmids were amplified by PCR (using the mutagenic oligonucleotides) according to the provided protocol. Competent *E. coli* cells provided with the kit were transformed with amplified plasmid and individual clones sequenced to confirm the presense of specifically introduced mutations.

Overlap extension PCR

This method was applied to incorporate multiple point mutations when required. Polymerase chain reaction oligonucleotides complementary to both DNA strands were designed so that incorporated mutations were at least 8 base pairs (bps) from the 5' and 3' ends of the oligonucleotide. Two DNA fragments, overlapping in the region defined by the mutagenic oligonucleotides were amplified using appropriate external oligonucleotides. The gel purified fragments underwent linear amplification (in the absence of oligonucleotides) in order to obtain a complete double stranded (ds) amplicon using the following reaction mix and cycling conditions. Approximately 120 ng of each amplicon was mixed with 1 x PCR Buffer for KOD XL DNA Polymerase, 0.2 mM of each dNTP, and 2.5 U KOD XL DNA Polymerase (Novagen) in a 25 µl reaction volume. Thermal cycling comprised the following:

95°C	2 min	
95°C	1 min	5 cycles (* + 5°C per cycle)
45°C*	30 sec	
72°C	1 min/kbp	
95°C	1 min	
72°C	2 min	
4°C	soak	

Both smaller PCR fragments with approximately 40 bases overlap were subject to final assembly amplification using the original external oligonucleotides used to amplify the smaller fragments. Polymerase chain reaction conditions are stated in

"Amplification of DNA fragments (up to 3kb) - PCR" using *Taq* DNA Polymerase (Fermentas).

In vitro transcription

Linearised plasmid for RNA transcription was first prepared by ethanol precipitation. Briefly, 2 x volume of 100% ice cold ethanol was mixed with the digestion reaction and incubated at -20°C for 30 mins. The cDNA was pelleted for 15 mins at 13000 rpm before being air dried at RT and resuspended in 10 µl diethylpyrocarbonate (DEPC) treated H_2O.

Between 1 - 2 µg of linearised plasmid DNA suspended in DEPC H_2O was incubated in a reaction mixture containing 1 x T7 transcription buffer (Fermentas), 20 nmoles of each ribonucleotide triphosphate (rNTP), 20 U RiboLock RNase Inhibitor (Fermentas), 0.2 µmoles DTT, and 20 U T7 RNA polymerase (Fermentas), to a total volume of 26.5 µl. The reaction was incubated in a 37°C waterbath for 4 hrs. Residual DNA template was digested with 2 U DNase Turbo (Ambion) at 37°C for 15 mins. Between 1.5 and 2 µL of RNA was confirmed on a 1% agarose gel before column purification using RNeasy Mini Kit (Qiagen) and quantification on a spectrophotometer.

2.3 Stock Solutions and Buffers

Crystal violet

0.5 g crystal violet powder in 20 ml 100% ethanol, 880 ml dH_2O containing 0.9 g NaCl and 100 ml 40% formaldehyde.

Mini-prep solutions

Solution I: 50 mM glucose, 10 mM EDTA, 25 mM Tris-HCl pH 8.0 (store at 4°C)

Solution II: 0.2 M NaOH, 1% SDS w/v

Solution III: 5 M postassium acetate pH 4.8 (store at 4°C)

DEPC-treated water

0.1% DEPC to dH_2O, store overnight at RT with a loose lid in a fume hood, autoclave and aliquot into single use volumes.

6 x DNA agarose gel loading buffer

25 mg bromophenol blue to 3 ml 100% glycerol. Make up to 10 ml with dH_2O. Store at RT.

Guanidine hydrochloride

Guanidine hydrochloride powder dissolved in H_2O at 200mM.

Plaque overlay medium

10% v/v Minimum Essential Medium (EMEM) with Earle's salts (10x), 1% v/v L-glutamine, 3% v/v 7.5% sodium bicarbonate, 2% v/v FCS, 1% v/v PS, and 30% v/v 2% agar.

Agar for plaque overlay medium

2% w/v bacto-agar (Dibco) in distilled water. Microwaved to dissolve powder and stored at RT.

Penicillin/streptomycin

Filter sterilised 10000 U/ml of penicillin and 10000 µg/ml in distilled water.

DMEM, GMEM/G418, 10 x EMEM, trypsin/EDTA, and sterile PBS

Supplied by the Media Preparation facility in School of Life Sciences, University of Warwick.

2.4 List of DNA plasmids

Table 2.1 List of subgenomic replicon and full-length clones used in this study

Plasmid Name	Description	Reference
pE7Luc	E7 replicon with virus capsid replaced by luciferase, derived from pT7 E7	Chris Bull and Minghui Ao (in house)
pEV70Luc	EV70 luciferase-encoding subgenomic replicon with ribozyme and T7 promotor	Sheila Waugh's thesis (in house)
pRibo-CAV21-NaeIDel-backbone (pRiboCAV21ΔP1)	CVA21 backbone vector lacking complete capsid region with *Age*I deletion	Claire Blanchard's thesis (in house)
pRLucWT[a]	PV1 replicon with virus capsid replaced by luciferase	Andino *et al*., 1993
pT7Rep3-L	PV3 replicon with virus capsid replaced by luciferase derived from pT7FLC	Barclay *et al*., 1998
pT7FLC	Full-length PV3 Leon P3/Leon/37 infectious clone	Goodfellow *et al.*, 2000
pT7FLC/SL3	pT7FLC with mutated CRE, unable to synthesise positive-sense RNA	Goodfellow *et al*, 2000
pT7 E7	Full-length E7 infectious clone	Lindberg *et al*, 1997
PV3BKCVA21P1[b]	Recombinant full-length PV3 backbone with PV3 P1 replaced by CVA21 Coe P1	Claire Blanchard's thesis (in house)
CVA21BKPV3P1[b]	Recombinant full-length CVA21 backbone with CVA21 P1 replaced by PV3 Leon P1	Claire Blanchard's thesis (in house)
CVA21BKSabin1P1[b]	Recombinant full-length CVA21 backbone with CVA21 P1 replaced by PV1 Sabin strain P1	Claire Blanchard's thesis (in house)
pT7CB4A/rep1	Replicon based on CBV4 infectious clone, capsid replaced by CAT gene	Barclay *et al*., 1998
pRibo-CVA21	Full-length CVA21 Coe strain in vector containing ribozyme, derived from pCVA21	(Hughes *et al.*, 1989)
pT7E30GB37	Full-length E30 strain GB37 in pGEMT-Easy vector, T7 promoter, amp resistant	This project
pT7E7ΔCRE	pT7E7 with mutated CRE, unable to synthesis positive-sense RNA	This project

[a] Kind gift from Ian Goodfellow

[b] Plasmids constructed by Claire Blanchard, University of Reading (unpublished).

2.5 Restriction Enzymes to Linearise Plasmids

Table 2.2 Restriction sites used to linearise plasmids for RNA transcriptions

Plasmid	Restriction Enzyme	Linearised Size (bp)
pE7Luc	*Xho*I	9726
pEV70Luc	*Sal*I	10396
pRibo-CAV21-NaeIDel-backbone	*Xho*I	10429
pRLucWT	*Apa*I	~10500
pT7Rep3-L	*Sal*I	10231
pT7FLC	*Sal*I	11128
pT7FLC/SL3	*Sal*I	11128
pT7 E7	*Not*I	10328
pT7 EV70+Ribo	*Sal*I	11155
PV3BKCVA21P1	*Sal*I	11131
CVA21BKPV3P1	*Sal*I	10426
CVA21BKSabin1P1	*Sal*I	10435
pT7CB4A/rep1	*Mlu*I	8473
pT7FLC3GFP	*Sal*I	9304
pT7E30 GB37	*Sal*I	10493
pT7E7ΔCRE	*Not*I	10328

2.6 List of Oligonucleotides

Table 2.3 Oligonucleotides used throughout this study

Name	Database	Sequence	Target	Identical location
MJREV	O-AH-7C	ATAAGAATGTCGAC(TX29)CCGCACCGAATGCGGAGAATTTACCC	E30	Last 26 nt 3' end
E30-1FClaI	O-AH-6C	ACGCCTCGAGATCGATTTAATTAAAACAGCCTGTGGGTTGWWCCCAC	E30	1-27
E7FcreMUT	O-AH-5B	GTAACTATATCCAATTTAAATCCAAGTGCCGTATCGAACCC	E7	4390-4451
E7RcreMUT	O-AH-6B	GGGTTCGATACGGCACTTGGATTTAAATTGGATATAGTTAC	E7	4451-4390
E7-3782F	O-AH-3E	GACTATGTCGAACAACTCGG	E7	3782-3801
M13R	O-AH-5D	CAGGAAACAGCTATGAC	pUC19	-
PV1/3-3280F	O-AI-4G	AACACRTMMGWGTCTGGTGC	PV1, PV3	3280-3299
GEN-4615R	O-AG-6B	RTADCCRTCRAAGTGWKMYGG	PV1, PV3, CVA21	~4615R
PV3-2995F	O-AG-5F	GCAAACATCTTCCAACCCGTCC	PV3	2995-3016
PV3-5191R	O-AG-1G	GGAGTCAACTGCTTGAAGCAGA	PV3	5191-5170
PV1-3004F	O-AG-6F	GCAAACCTCATCAAATCCATCA	PV1	3004-3025
PV1-5200R	O-AG-9F	GGAGTCAACTGCTTGGAGCAAG	PV1	5200-5179
Oligo dT	-	TTTTTTTTTTTTTTT	Poly(A)-tail	-
E7-6F	O-AB-8C	CAGATCATGTACATACC	E7	2873-2889
E7-5247R	O-AH-2F	GACACGAATGTGGTTAGTGC	E7	5274-5255
RiboSmaIE30F	O-AI-3A	CTCGTACCCGGGTTCTTAAAACAGCCTGTGGGTTG	E30	-
E305UTRSacIR	O-AI-4A	ATGTTTTTGGCCAGAGCTCCCATTTTGTTGTGTTGAG	E30	-
SacILucF	O-AI-2A	TGATGGGAGCTCTGGCCAAAAACATAAAGAAAG	Luciferase	-
BstZ17ISmaILucR	O-AI-6A	CTCGTACCCGGGGTTCGTATACGGCACGGCGATCTTTCCG	Luciferase	-
E30-2551F	O-AH-2H	AGTGGAGACAGGGCACAC	E30	2551-2568

CHAPTER THREE: RNA-mediated Interference to Enrich Recombinants

3.1 Introduction

Although virus recombination is well documented, isolating recombinants from picornavirus populations is relatively difficult due to the high titre of progeny (Romanova *et al.*, 1980). Recombination between PV serotypes was reported to occur at a frequency of 7.6×10^{-6}, which is 100-fold lower than recombination within serotypes (Kirkegaard & Baltimore, 1986). The majority of literature identifies and characterises recombinant RNA viruses that have been the attributing cause of disease outbreaks and individual cases of pathogenesis. Unfortunately, less analysis has concentrated on studying the mechanisms behind RNA recombination, particularly in positive-strand RNA viruses. As described in detail in the introduction, researchers have succeeded in recapitulating genetic recombination of enteroviruses *in vitro* (in particular PV) in order to try and define the preferred region/s of recombination. Simply co-infecting cells with two strains/serotypes creates recombinants naturally, although this method is inefficient in that a progeny may only contain 1 to 10 recombinant viruses in a population of 10^6 to 10^7 virus particles. The parental viruses will always dominate during a co-infection and make it difficult to characterise replication competent recombinants that are present at low levels. In another method, recombinant crosses are performed involving partners with genetically defined selection markers (e.g. temperature and guanidine sensitivity/resistance) that limit growth ordinarily. Although the latter method is successful in generating small ranges of recombinant viruses for further study, these methods of selection are only applicable to viruses where such genetic markers are known. In most cases, the reversion rate of such genetic markers would exceed the rate of recombination. It is time consuming to select natural virus mutants that escape applied inhibitors followed by full genome sequencing to identify the location of associated mutations. These types of studies are therefore limited to well-characterised RNA viruses such as PV, and to those viruses with extensive sequence information and phenotypic analysis.

In order to study recombination in any number of enteroviruses available, regardless of extensive analysis on virus phenotypes etc, an alternative generic approach was

investigated. It was decided to attempt to select recombinants by the application of an RNA-mediated interference (RNAi) selection strategy. This method did not require extensive information about input virus partners beforehand, only that full genome sequences were available. Numerous enteroviruses have been sequenced in their entirety, with most sequences available in GenBank (National Center for Biotechnology Information, [NCBI]).

The native cellular mechanism of RNAi has emerged to become a successful gene-silencing technique for controlling sequence-specific degradation of homologous RNA in cultured cells (Fire *et al.*, 1998). Ribonucleic acid-mediated interference is widespread and is a phenomenon found in plants, animals, and fungi. It also plays an important role in defending cells against viruses and other invading genes (van Rij, 2008). Briefly, small non-coding RNA molecules mediate post-transcriptional gene silencing (PTGS) and can exist in several forms. MicroRNA are small double stranded (ds)RNA derived from endogenous RNA-coding genes in the genome that, once processed and exported to the cytoplasm, bind specifically to mRNA in a sequence dependent manner to regulate the expression of cellular proteins. MicroRNA either induce degradation of the target mRNA or inhibit translation from the mRNA template. Similarly, siRNA are small interfering RNAs derived from long exogenous dsRNA (such as invading viruses) and target the cleavage of foreign RNA by an endonuclease in a sequence dependent manner (Elbashir *et al.*, 2001a; Hutvágner & Zamore, 2002). Although the primary transcripts and processing of miRNA and siRNA differ, both culminate in the cytoplasm to converge at the RNA-inducing silencing complex (RISC) to target RNA degradation via the same process.

Genome-encoded miRNA originate as large stemloop structures (primary miRNA) in the nucleus and are processed into shorter fragments by an RNase III enzyme Drosha and double stranded RNA-binding domain (dsRBD) (reviewed in Bushati & Cohen, 2007). The ~ 70 nt precursor miRNAs (pre-miRNAs) are transported to the cytoplasm by Exportin-5 (Yi *et al.*, 2003) where they are cleaved by the same process as precursors to siRNA. Briefly, pre-miRNAs or synthesised dsRNA molecules comprising > 26 nt RNA duplexes are processed by the ribonuclease (RNase) III-like enzyme Dicer into miRNA or siRNA respectively, both approximately 22 nt in length (Hutvágner *et al.*, 2001). One strand in the duplex

(usually the strand with the least stable 5' end, labelled the guide strand) is incorporated into the RISC (Schwarz *et al.*, 2003) which contains Argonaute proteins capable of endonucleolytic cleavage (Liu *et al.*, 2004). The miRNA/siRNA guides RISC to its target by bp interactions, which ultimately recognises and degrades sequence-specific mRNA or foreign nucleic acids.

Several methods exist for performing RNAi *in vitro* and these are summarised in table 3.1. Chemically synthesised siRNA can be generated by the same methods as DNA oligonucleotides by commercial companies and introduced directly into cells. Expression plasmids and viral vectors (including adenoviral, retroviral, and lentiviral) can express siRNA or short-hairpin RNA (shRNA) off RNA polymerase II or III promoters that are later processed in cells. RNAi is a valuable research tool and is has been used for large-scale screens of gene function with miRNA libraries readily available for many host systems that target all known genes.

Thousands of human miRNAs have been discovered, with the best known probably being miR-122. This miRNA is a liver-specific miRNA that regulates cholesterol and fatty-acid metabolism in mice (Esau *et al.*, 2006). Hepatitis C virus is known to recruit miR-122, which enhances virus replication as a result (Jopling *et al.*, 2005). It is not surprising therefore, that miR-122 has become an ideal therapeutic target for HCV infection. This is an unusual example of a non-viral miRNA benefitting virus replication. In contrast, several laboratories have demonstrated the effectiveness of partially inhibiting picornavirus propagation by specific siRNA (Liu *et al.*, 2005; Saleh *et al.*, 2004; Schubert *et al.*, 2005; Sim *et al.*, 2005) with some groups observing a reduction of up to two $\log_{10}$ of virus titre as a result of RNA sequence targeted degradation. In an elegant experiment, Gitlin and colleagues inserted another prominently studied miRNA, let-7, into PV cDNA prior to transfections with transcribed RNA (Gitlin *et al.*, 2005). Transfected HeLa cells express high levels of let-7 naturally which controls developmental pathways (regulating cell proliferation and differentiation) (Hutvágner *et al.*, 2001). Not only was virus replication reduced as a direct result of expression of let-7 miRNA in cells (down approximately 4 $\log_{10}$), but it was also concluded that only positive-stranded PV RNA was targeted by anti-PV siRNA.

Table 3.1 Summary of RNAi and delivery methods

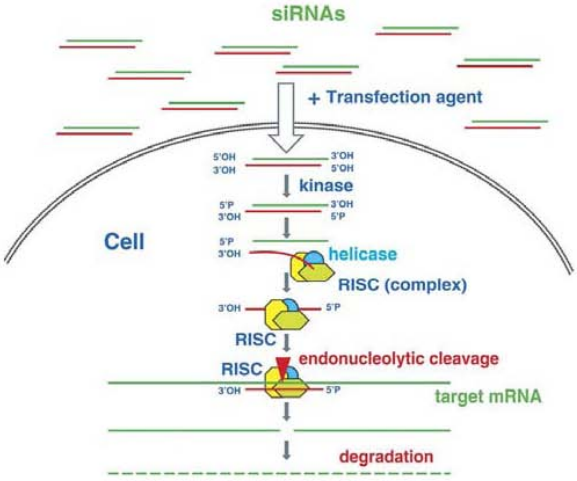
RNAi Source	Delivery Methods	Expression
chemically synthesised siRNA (non-replicating)	lipid-mediated transfection electroporation	Transient
plasmid-expressing miRNA, shRNA (self-replicating)	lipid-mediated transfection electroporation	Transient
viral vectors containing siRNA, shRNA:		
Adenovirus	infection (dividing and non-dividing cells)	transient (does not integrate)
Lentivirus	infection (dividing and non-dividing cells)	transient and stable (can integrate*)
Mouse stem cell virus (retrovirus)	infection (dividing cells)	transient and stable (can integrate)
Baculovirus	infection (dividing and non-dividing cells)	transient and table (can integrate)
PCR-generated siRNA or shRNA Expression Cassettes (*in vitro* transcribed siRNA or shRNA transcripts are later processed in the cell	lipid-mediated transfection electroporation	Transient

* Integrate into chromosomal DNA for long-term, consistent expression
Sources: Ambion website (www.ambion.com/techlib/presentations/siRNA.pdf) and SuperArray Bioscience Corporation (http://www.sabiosciences.com/newsletter/rnai.html)

Antiviral siRNAs can interfere with virus replication by targeting sequence dependent RNA degradation to reduce virus yields in cells, and this process was investigated in this project as a way to study virus recombination further. Figure 3.1 provides schematic representations of siRNA action within the cell cytoplasm, as well as potential sites of siRNA targeting during the virus multiplication life cycle within cells. There are several methods by which short RNA-molecules can be introduced into cells, and the remainder of the study will focus on miRNA and siRNA.

The overall aim was to investigate the sequence determinants and genome locations of recombination between two serotypes (and potentially two species) of enteroviruses by employing RNAi to select and enrich recombinant viruses from a mixed virus infection. The application of RNAi would be practical because of its ability to reduce virus titres in cells protected by anti-viral siRNA. It has been demonstrated in the studies mentioned above, that virus replication is inhibited to certain degrees by short RNA molecules that are designed specifically to target infecting and newly synthesised virus RNA. In the first stage of this project, several anti-PV miRNA expression plasmids and chemically synthesised anti-PV siRNA duplexes were tested for their ability to reduce virus titres during a single PV infection. It was known that a two $\log_{10}$ reduction in virus yield was the minimum level acceptable as this is the greatest reduction reported for antiviral siRNA so far (Gitlin *et al.*, 2002; Lee *et al.*, 2007). It was also considered the minimum level that might impact on the overall yield of progeny to facilitate identification of recombinants. If the indicated minimum level of virus reduction was achieved for a single PV serotype infection, then the ultimate purpose of the project was tested. The aim was to reduce parental virus titres during a dual infection by using short RNA molecules that targeted opposing ends of each infecting parent virus genome. At the same time, newly synthesised recombinant viruses would escape these specific anti-viral measures due to the recombinant nature of their genomes – effectively enriching recombinants in the population. A diagrammatic representation of this is provided in figure 3.2. It was expected that reducing the high numbers of parent viruses in the progeny would effectively moderate competition between both parents and the resulting recombinants. This would aid in the retention of a diverse range of natural recombinants that could exhibit various levels of fitness.

A



B

(Source A & B: Colbere-Garapin *et al.*, 2005)

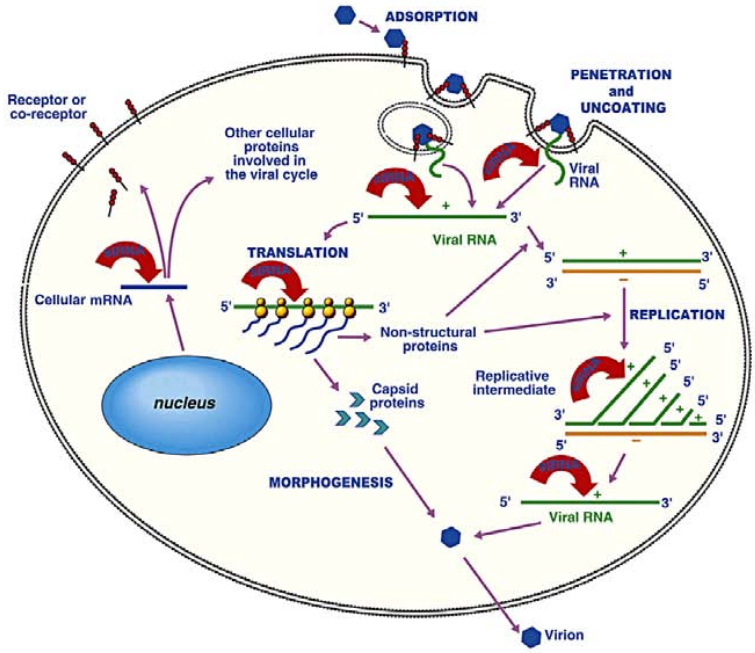


Figure 3.1 RNAi silencing of mRNA in cells. Schematic representation of mRNA silencing by siRNA and potential sites of virus RNA targeting by siRNA during the virus cycle within the cell cytoplasm during infection. (A) Synthetic siRNAs are phosphorylated once transfected into cells, are unwound by a helicase, and the antisense strand of siRNA guides RISC to target mRNA. The target mRNA is cleaved by an endo-ribonuclease in the RISC, resulting in mRNA degradation. (B) Multiplication of virus in the cell upon infection. There are several instances during the life cycle with potential siRNA targets to degrade virus RNA. (Source: Colbere-Garapin *et al.*, 2005).

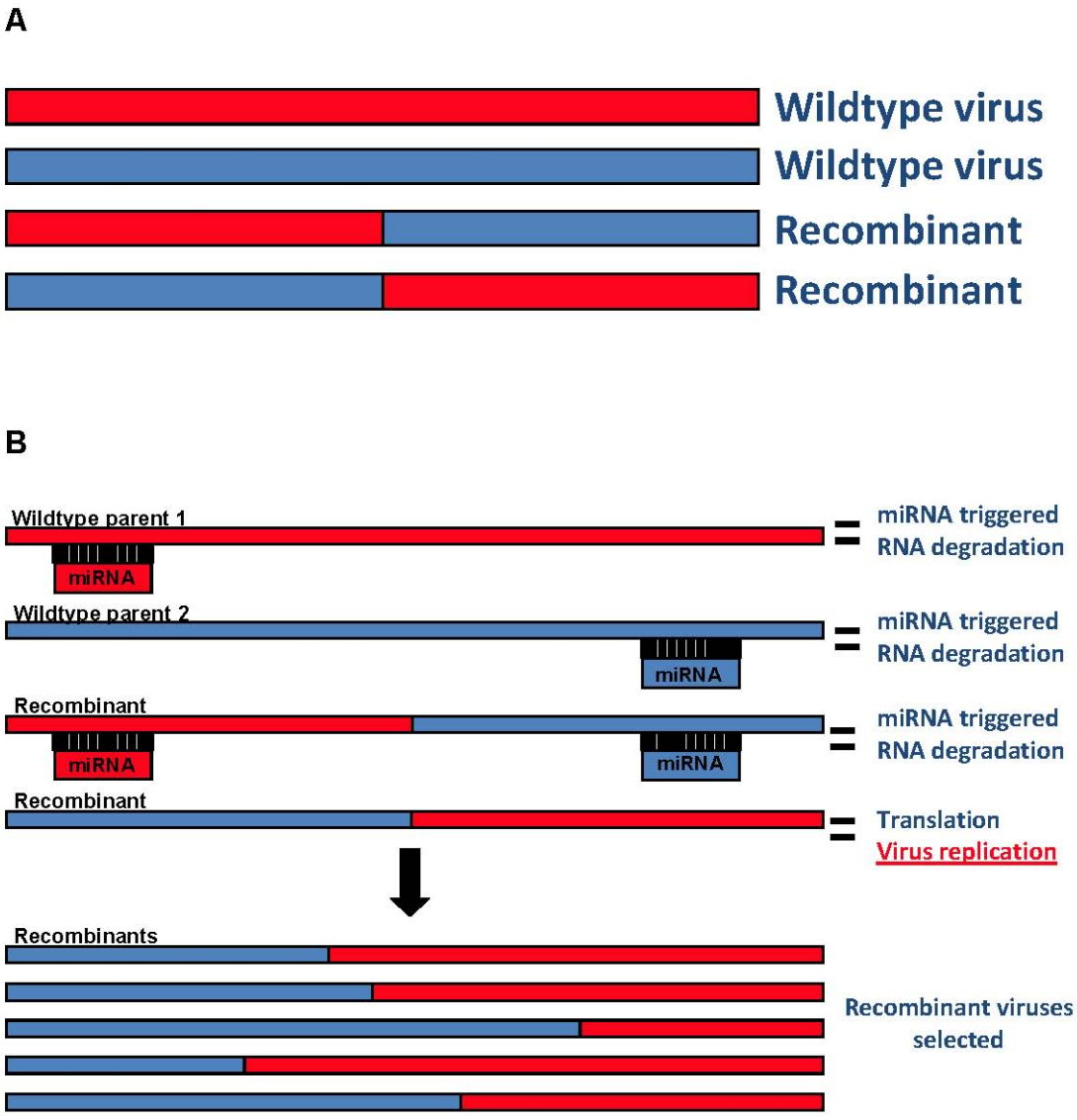


Figure 3.2 RNAi regimes to enrich recombinants in a virus population. (A) A simplified representation of an expected mixed virus population after a dual infection with two similar viruses coloured red and blue. (B) Diagrammatic representation of RNAi strategy to target mRNA degradation of infecting parental viruses and specifically orientated recombinant genomes.

It is only very recently that the first report into the use of RNAi to study RNA virus recombination was published (Gao *et al.*, 2011). This study involved the use of HIV-1 subtype-specific siRNAs to enrich inter-subtype HIV-1 recombinants and inhibit parental HIV-1 replication during a dual infection. This group was successful in generating replication-competent recombinants and identifying a range of recombination breakpoints, some of which were similar to those found in natural isolates from HIV-1 patients.

Aims

The specific aim of this part of the project was to investigate the antiviral activity of specifically designed miRNA/siRNA against PV1 and PV3 replication *in vitro*, using different delivery methods. Antiviral activity was measured by determining the virus yield with and without the presence of miRNA/siRNA in infected cells. If RNAi successfully reduced virus production in HeLa cells during single virus infections, then this method would be expanded to incorporate dual infections to generate recombinant virus.

3.2 Construction of anti-PV miRNA Expressing Plasmid

Initially, a method involving the use of plasmid-derived/encoded miRNA sequences was tested. This method mimics the presence of endogenous miRNA produced naturally in cells and processes the short stem loops in the same way for incorporation into RISC. MicroRNA molecules were designed with high complementarity to virus sequences of interest and miRNA duplex-expressing plasmids were transfected into cells prior to virus challenge. Inhibition of virus replication was calculated by comparing virus titres in miRNA treated cells versus cells left untreated (mock transfected). A “scrambled miRNA” was also included in the studies. This scrambled miRNA was designed as an oligonucleotide control with no specificity to virus or cellular mRNA. This was to determine whether the transfection of cells with miRNA expressing plasmids affected normal cell function.

James Taylor designed miRNA against VP1, VP3, 3C, and 3D in PV type 1 (Mahoney) and 3 (Leon) miR 2 to miR 12 (table 3.2). These sequences were targeted

Table 3.2 Synthesised miRNA oligonucleotides (DNA) and corresponding nucleotide positions in PV1 and PV3 genome

Name	Target	Nucleotide (nt) sequence	Nt location[a]	Region
Cntrl-scr(F)*	-	5'TGCTGGGTTATCATAACCTCACAGCCGTTTTGGCCACTGACTGACGGCTGTGATTATGATAACC	-	-
Cntrl-scr(R)*	-	5'CCTGGGTTATCATAATCACAGCCGTCAGTCAGTGGCCAAAACGGCTGTGAGGTTATGATAACCC	-	-
miR 2(F)	PV1	5'TGCTGTGTACAACATGTCTGGTTTGCGTTTTGGCCACTGACTGACGCAAACCACATGTTGTACA	2632	VP1
miR 2(R)		5'CCTGTGTACAACATGTGGTTTGCGTCAGTCAGTGGCCAAAACGCAAACCAGACATGTTGTACAC		
miR 3(F)	PV1	5'TGCTGTAGACTCTGACCTTGACCTATGTTTTGGCCACTGACTGACATAGGTCAGTCAGAGTCTA	2677	VP1
miR 3(R)		5'CCTGTAGACTCTGACTGACCTATGTCAGTCAGTGGCCAAAACATAGGTCAAGGTCAGAGTCTAC		
miR 4(F)	PV3	5'TGCTGATCAGAACTCCATCATTCGTCGTTTTGGCCACTGACTGACGACGAATGGGAGTTCTGAT	5716	3C
miR 4(R)		5'CCTGATCAGAACTCCCATTCGTCGTCAGTCAGTGGCCAAAACGACGAATGATGGAGTTCTGATC		
miR 6(F)	PV3	5'TGCTGAATAGGTCTGGATCGCATCCAGTTTTGGCCACTGACTGACTGGTGCGCCAGACCTATT	6607	3D
miR 6(R)		5'CCTGAATAGGTCTGGCGCATCCAGTCAGTCAGTGGCCAAAACTGGATGCGATCCAGACCTATTC		
miR 7(F)	PV3	5'TGCTGTTCACAATCAGAACTCCATCAGTTTTGGCCACTGACTGACTGATGGAGCTGATTGTGAA	5722	3C
miR 7(R)		5'CCTGTTCACAATCAGCTCCATCAGTCAGTCAGTGGCCAAAACTGATGGAGTTCTGATTGTGAAC		
miR 8(F)	PV1	5'TGCTGATCAAATTCAGGCAGCGCACAGTTTTGGCCACTGACTGACTGTGCGCTCTGAATTTGAT	1832	VP3
miR 8(R)		5'CCTGATCAAATTCAGAGCGCACAGTCAGTCAGTGGCCAAAACTGTGCGCTGCCTGAATTTGATC		
miR 9(F)	PV1	5'TGCTGAACCGAACCCTATACATATTTCCGTTTTGGCCACTGACTGACGGAAATGTAGGGTTCGGTT	1960	VP3
miR 9(R)		5'CCTGAACCGAACCCTACATTTCCGTCAGTCAGTGGCCAAAACGGAAATATGTATAGGGTTCGGTTC		

Name	Target	Nucleotide (nt) sequence	Nt location[a]	Region
miR 10(F)	PV1	5'TGCTGATCCATGGCACTACCATAGTAGTTTTGGCCACTGACTGACTACTATGGGTGCCATGGAT	2257	VP3
miR 10(R)		5'CCTGATCCATGGCACCCATAGTACGTCAGTCAGTGGCCAAAACTACTATGGTAGTGCCATGGATC		
miR 11(F)	PV1	5'TGCTGTATGGTTTGCCGATACGTGGTGTTTTGGCCACTGACTGACACCACGTAGGCAAACCATA	2285	VP3
miR 11(R)		5'CCTGTATGGTTTGCCTACGTGGTCGTCAGTCAGTGGCCAAAACACCACGTATCGGCAAACCATAC		
miR 12(F)	PV1	5'TGCTGAAAGAGGGACGACTATTCTAGGTTTTGGCCACTGACTGACCTAGAATACGTCCCTCTTT	2354	VP3
miR 12(R)		5'CCTGAAAGAGGGACGTATTCTAGGTCAGTCAGTGGCCAAAACCTAGAATAGTCGTCCCTCTTTC		

* Designed and constructed by Kym Lowry

[a]Numbering according to PV1 Mahoney or PV3 Leon, 5' end of the target sequence

Underlined nucleotides - inserted linker and loop sequences (by Invitrogen) to generate secondary stem loop structure.

as they flanked the recombination region found naturally in virus isolates, i.e. 2A to 2C region near the middle of the genome (as shown in figure 3.2B). As part of this study, all miRNAs were sequenced and concentrations verified. The majority of miRNA inserts were designed using the web-based tool provided by Invitrogen[1]. The poliovirus sequence of interest was submitted in an online form; in this case, regions VP1 and 3C of both PV1 and PV3, and the website automatically calculated 10 potential miRNA target sequences with high probability of target binding to the submitted sequence. Although the exact algorithms for calculating preferred binding sites for miRNA to RNA template was not available to the user, several guidelines have been established mainly by Tom Tuschl's group (Elbashir *et al.*, 2001a; Elbashir *et al.*, 2002; Elbashir *et al.*, 2001b). Suggested guidelines for choosing siRNA target sites are as follows: the targeted region is usually 21 nt in length, avoids intronic regions, targets beginning with AA are ideal as 3' overhanging uracil dinucleotides are effective, siRNA/miRNA oligonucleotides to target mRNA should be selected outside regions of secondary structure, avoid sequences outside the 30-50% GC content range, and avoid stretches of > 4 nucleotide repeats. The oligonucleotides designed using the Invitrogen website conformed to these parameters. Basic local alignment search tool (BLAST) was also performed later to avoid off-target binding to other sequences (virus and host genome). It was also confirmed by BLAST[2] that each set of PV target sites differed between both serotypes, in order to eliminate cross-reactivity of miRNA to the heterologous virus.

In total, 11 pairs of miRNA oligonucleotides targeting PV1 and PV3 were separately cloned into pcDNA6.2™-GW+EmGFP-miR expression vector provided in the kit according to the manufacturer's instructions. Each pair of oligonucleotides consisted of a forward and reverse complementing miRNA sequence that was annealed prior to ligation into the vector. The top and bottom strand oligonucleotides annealed to generate a ds fragment with 4 nt overhangs necessary for directional cloning into the commercial vector. A plasmid containing a scrambled control was also constructed, having used the scrambling miRNA/siRNA design tool provided by the Promega website[3]. In short, a miRNA was submitted online and the program scrambled the

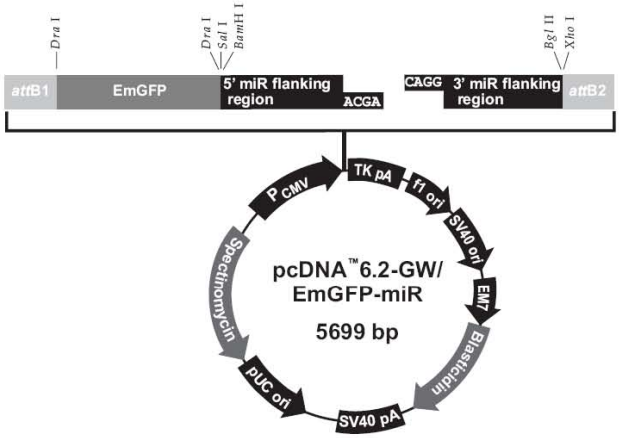
[1] (https://rnaidesigner.invitrogen.com/rnaiexpress/)
[2] (http://blast.ncbi.nlm.nih.gov/Blast.cgi)
[3] (http://www.promega.com/siRNADesigner/)

sequence in such a way that no potential 'off targets' to mammalian genomes existed. Basic local alignment search tool results also identified no potential 'off targeting' sites in PV. This acted as a control to indicate whether the presence of random miRNA effected virus production and cellular function.

A commercial expression vector system was employed to deliver virus specific miRNA to cells by lipid-mediated transfection. In the first instance, the BLOCK-iT™ Lentiviral Pol II miR RNAi Expression System Kit was used to achieve this. Specific miRNAs are designed by the researcher and cloned into the vector backbone provided in the kit. This kit allows for the eventual construction of a replication-incompetent lentivirus that distributes miRNA sequence of interest to both dividing and non-dividing mammalian cells. This initial expression clone contains the ds oligo encoding the pre-miRNA sequence (designed specifically by the user), which is expressed in cells using an RNA polymerase II promoter, specifically the human CMV immediate early promoter (figure 3.3A). The vector allows for co-cistronic expression of Emerald GFP to measure the efficiency of DNA transfection in cells by fluorescent cell imaging. Figure 3.3B illustrates the role of miRNA processing into RISC and its potential to target RNA degradation of virus RNA in cells.

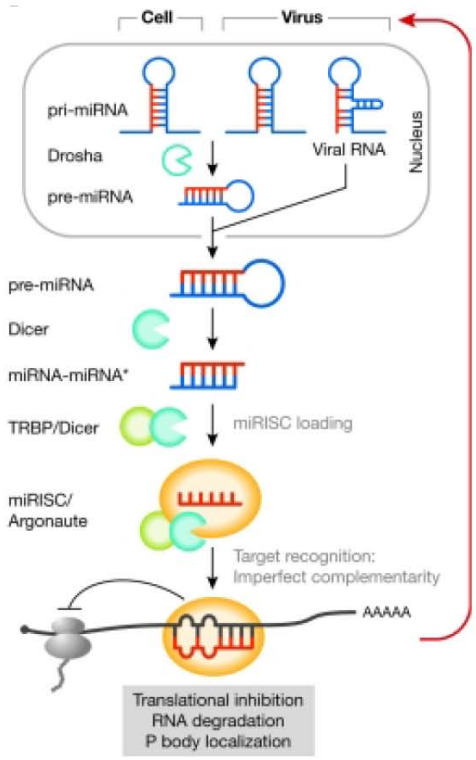
The first step was to determine if transient expression of antiviral miRNA in cells was capable of reducing virus production after virus infection in the same cells. If a reproducible reduction in virus yield by specific miRNA oligos was observed, the miRNA would be incorporated into a lentiviral vector in order to establish a stable transfectant in which antiviral miRNA are constitutively expressed in cells. Stably transfected cells would have integrated the nucleic acid into their genome and once expanded these cells could permanently express the miRNA of interest, making it ideal to perform RNAi experiments over the longer term. During virus infections, low levels of virus RNA will escape RNAi, and recombinants in a certain orientation (figure 3.2B) will be enriched. Mammalian cells that constitutively express antiviral miRNAs can be utilised for further recombinant experiments – for example, the presence/absence of external stimuli (temperature, virus concentrations, and mutagens) may alter recombination events.

A



(Source: http://tools.invitrogen.com/content/sfs/vectors/pcdna6_2gw_emgfp_miR_map.pdf)

B



(Source: van Rij, 2008)

Figure 3.3 The expression vector kit used to deliver virus specific miRNA to cells. (A) The vector incorporates the coding sequence of Emerald GFP to allow for tracking of plasmid transfection efficiency. Double stranded miRNA oligonucleotides are cloned into the vector according to the manufacturer's instructions. (B) Cellular or virus pre-miRNA processing in the nucleus, followed by RNA degradation once miRNA are incorporated into the RISC.

Statistical analysis using both Kruskal Wallis and Mann-Whitney U tests were performed to determine whether reductions in virus titre due to treatment with miRNA were stastically significant (see Materials and Methods for full details).

3.3 Transient Expression of miRNA to Reduce Virus Replication

miRNA plasmids express EmGFP in transfected HeLa cells

HeLa cells were transfected separately with three miRNA expression clones to confirm the efficiency of transfection by Lipofectamine 2000 (Invitrogen) in this highly transfectable cell line. Approximately 3.5 x 10^5 cells cultured on sterile glass coverslips in 6-well plates were transfected with 2 μg plasmid DNA. Instructions supplied with Lipofectamine 2000 were followed regarding concentrations of plasmid DNA required for all transfections. After 48 hrs, the coverslips were removed, washed, fixed, and mounted using medium for fluorescence with DAPI (see Materials and Methods). Transfection efficiency was shown by direct fluorescence of GFP in the cytoplasm of cells, and this was interpreted to indicate the co-expression of anti-viral miRNA in the same cells. Untreated, control scrambled (cntrl-scr), and miRNA 7 plasmid transfected cells are shown in figure 3.4. Between 40-50% of cells were transfected and levels of fluorescence varied between cells. Similar fluorescence results were seen for the remaining miRNA (data not shown).

Anti-PV1 and PV3 miRNA targeted RNA degradation of PV1 and PV3

Initially, the anti-PV miRNA encoded by these plasmids was analysed for the ability to inhibit virus replication by specifically targeting virus RNA degradation. By following the protocol outlined for transfections in 24-well plates, 500 ng of plasmid DNA was transiently transfected into HeLa cells in triplicate using Lipofectamine 2000 according to the manufacturer's instructions (Invitrogen). The following individual plasmids containing ds miR pairs were transfected into cells: cntrl-scr, miR 2, miR 3, miR 4, miR 6, miR 7, miR 8, miR 9, miR 10, miR 11, and miR 12. A mock transfection was performed each time containing Lipofectamine 2000 alone in the absence of plasmid. Twenty-four hrs post-transfection, the medium was removed and cells were infected with an MOI of 0.1 with either PV1 (Mahoney) or PV3 (Leon) for 30 mins. Once virus was removed, medium was replaced and

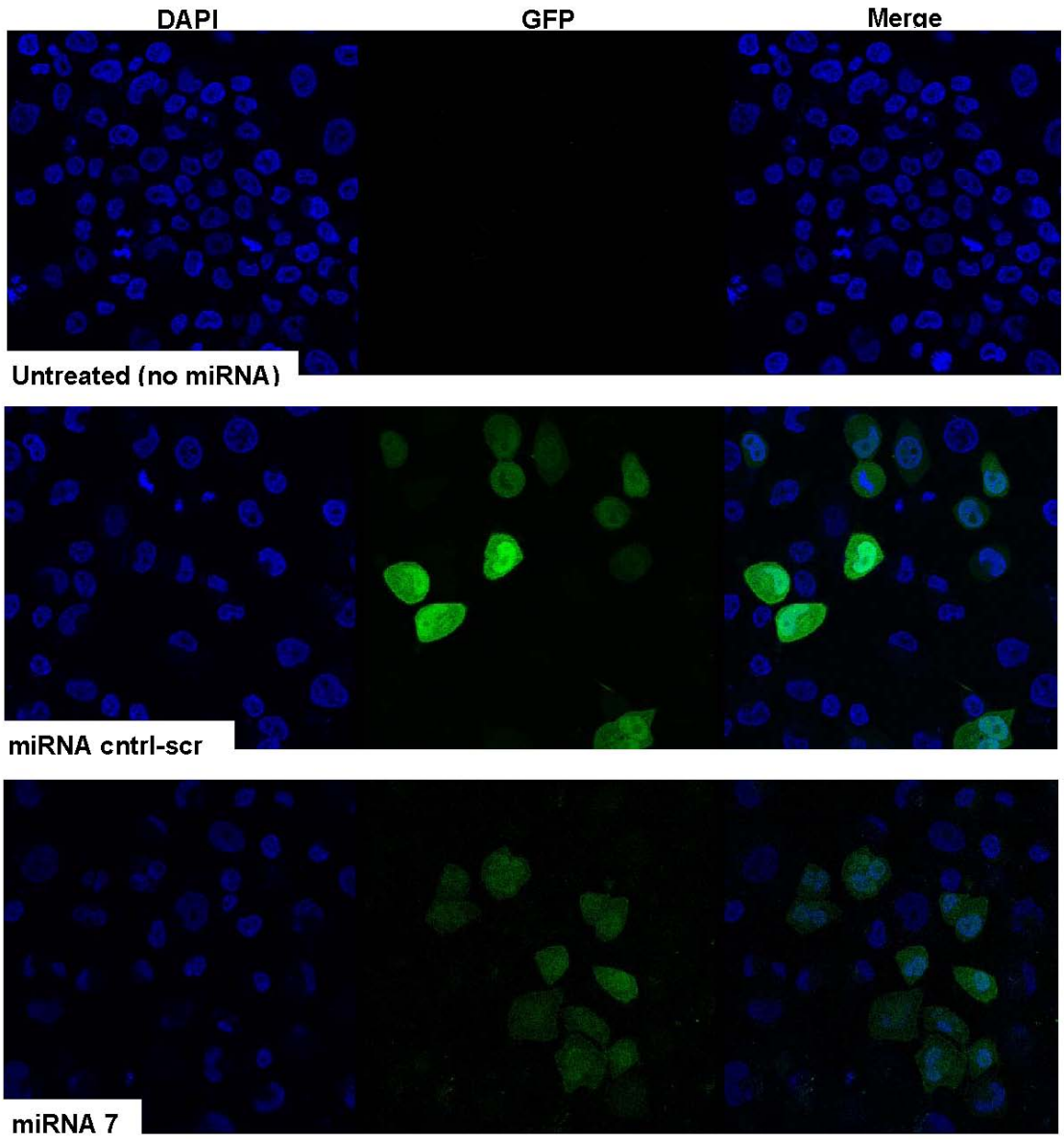


Figure 3.4 Immunofluorescence of GFP/ miRNA expressing plasmids in cells. Immunofluorescence of untreated/cntrl-scr/miRNA 7 HeLa cells at 24 hrs post-transfection. The nuclei were stained with DAPI and the presence of GFP in the cell cytoplasm was examined under UV light.

plates were incubated at 37°C/5% CO_2/air. Supernatant containing virus was harvested from cells 24 hrs post-infection and titred by $TCID_{50}$.

The plasmid expressing cntrl-scr miRNA displayed minimal disruption to PV1 and PV3 production (figure 3.5). This was expected since this was designed to have no sequence specificity to virus. This also suggested that miRNA expression plasmids used in the study did not affect normal cell function, as transfected cells appeared healthy (as seen by fluorescence microscopy) and produced similar titres of virus output. Transient expression of specific miRNAs, on the other hand, showed highly variable inhibition of virus production and results indicated no sequence specific effect of antiviral miRNAs on its targeted virus.

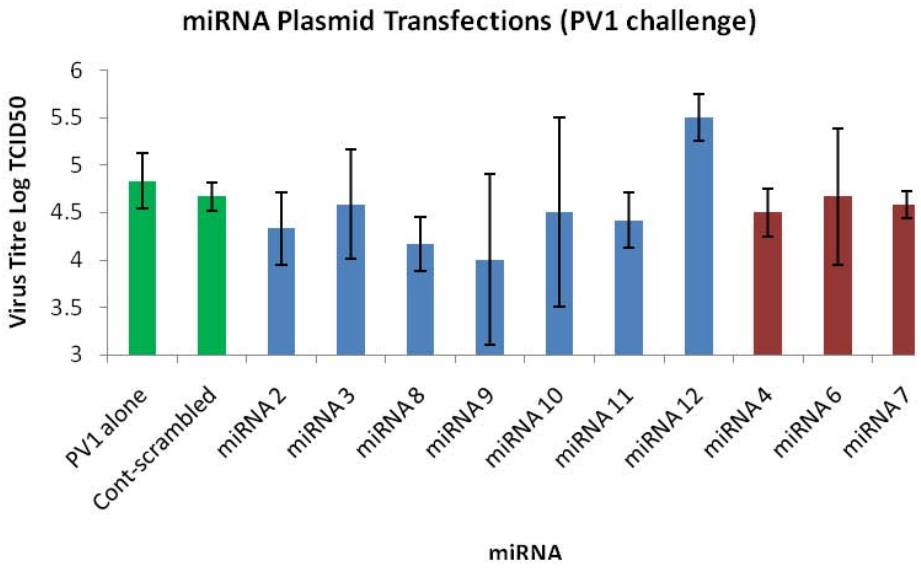
All anti-PV1 and anti-PV3 miRNAs affected PV1 titres to varying levels, with reductions in virus titres ranging from 0.75 to 0.25 $\log_{10}$ $TCID_{50}$ (miRNA 9 and miRNA 3 respectively) (figure 3.5A). Unexpectedly, the PV1 titre increased by > 0.5 $\log_{10}$ in the presence of miRNA 12, a result which remains unexplained. Anti-PV3 miRNA 4 exhibited subtle cross-reactive antiviral activity with just less than 0.5 $\log_{10}$ $TCID_{50}$ reduction in titre compared to the PV1 control titre. A wide range of readings within replicates was noted, in particular for cells treated with miRNA 9 and 10. The excessive standard deviation (see error bars) meant that the difference in average virus titre may not have been significant.

By referring to figure 3.5B, both anti-PV1 and anti-PV3 miRNA affected the titre of PV3 in HeLa cells to varying degrees, from approximately 0.75 $\log_{10}$ (anti-PV miRNA 3) to 0.1 $\log_{10}$ $TCID_{50}$ (anti-PV1 miRNA 8). Although statistical analysis showed that virus titres were reduced significantly compared to the control, these reductions were still less than one $\log_{10}$. There was no evidence of serotype specific reduction due to miRNA, as was demonstrated by cross-reactivity between PV1 and PV3 miRNA.

Intra-precision assays with miRNA expression plasmids

To demonstrate intra-assay variability in miRNA transfection/virus titre, each well of the triplicate in the aforementioned experiment (miRNA 9 and 10 with PV1

A



B

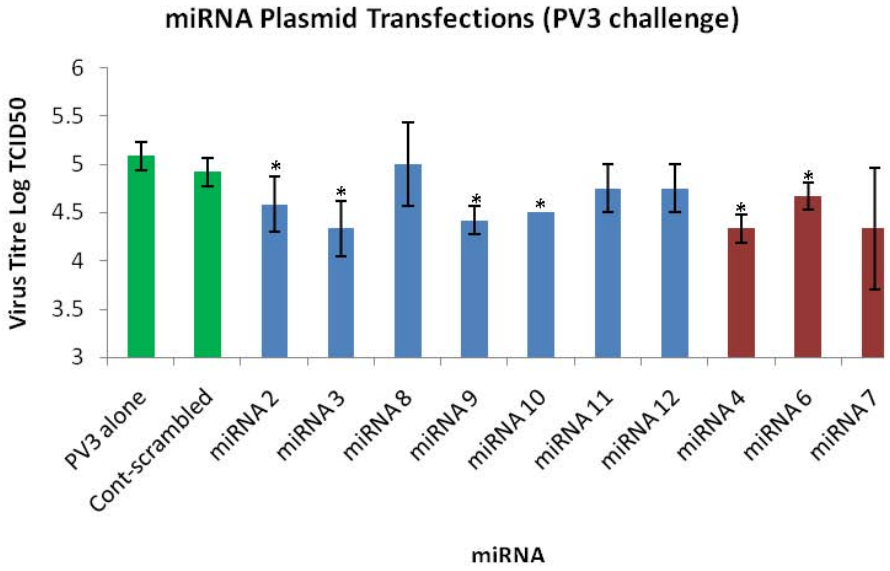


Figure 3.5 Effects of miRNA on PV replication. Representative results showing the effects of transient expression of miRNA expressing plasmids on (A) PV1 and (B) PV3 virus replication. Twenty-four hrs post-transfection, HeLa cells were infected with MOI 0.1 of either PV1 or PV3 for a further 24 hrs before virus was harvested and titrated. The graphed data are the geometric mean from triplicates, error bars indicate standard deviation. The blue bars indicate anti-PV1 miRNA, the red bars represent anti-PV3 miRNA. miRNA 10 has a SD of 0, depicted by the lack of error bar at three significant figures. Asterisk; $P < 0.05$ by Mann-Whitney U test with the test miRNA compared to the PV1 or PV3 alone, respectively.

challenge) was graphed separately (figure 3.6). This showed very wide transfection and infection variability between each well on the same plate performed simultaneously in one experiment. The virus titre from each well was determined once by $TCID_{50}$. Intra-assay variation had a maximum value of two TCID log_{10} difference in virus titre within cells transfected with the miRNA 10 expression plasmid. Given that two log_{10} was set as the minimum target for significant knockdown, an intra-assay variation of up to two TCID log_{10} was considered too large.

To conclude, in those that were reduced, the log_{10} reduction was not considered extensive enough for use in recombination experiments. If a one log_{10} $TCID_{50}$ reduction from 10^5 was achieved using miRNA, the progeny would still contain 10^4 $TCID_{50}$ parent virus able to escape RNAi and re-infect surrounding cells. Also, cross-reactivity was observed between both sets of anti-viral miRNAs to PV1 and PV3 production, indicating a lack of specificity in the assay. Even if stable transfectants were constructed using the miRNAs, it was believed that intracellular virus RNA would still escape targeted degradation with cross-reactivity remaining an issue. Lastly, intra-assay variation was high as indicated by a large standard deviation for some results performed in triplicate. As an alternative approach, it was decided to investigate the direct transfection of siRNA to determine if this suppressed viral yield sufficiently for the planned recombination experiments.

3.4 Transient Expression of siRNA to Reduce Virus Replication

Antiviral activity of chemically synthesised siRNA duplexes

An alternative RNAi delivery system was therefore employed whereby chemically synthesised ds siRNA molecules were directly transfected into cells. These were designed to provide anti-PV siRNA in cells. The synthetic siRNA oligonucleotides were designed to target regions within the VP1 and $3D^{pol}$ of PV1 and PV3 and were also designed using the web-based tool provided by Invitrogen (table 3.3). Of note, only one siRNA target overlapped with a miRNA target in the same virus region between both experiments. The BLOCK-iT™ RNAi Designer for siRNA used the same proprietary algorithms with 100% homology to their target, so that upon processing, their activity would result in cleavage of that mRNA target. The

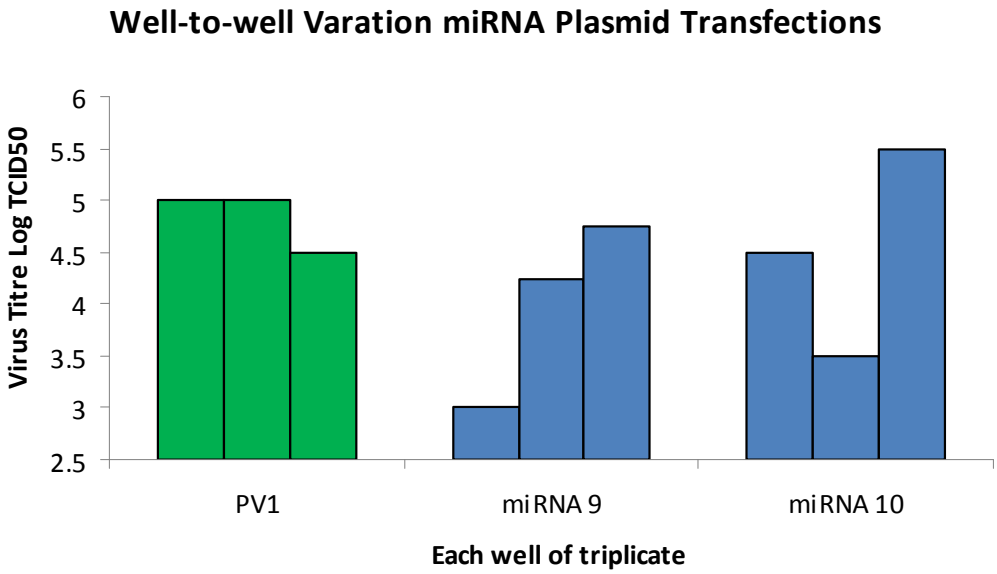


Figure 3.6 Intra-assay variation using miRNA. Representative results indicating intra-assay variation within a triplicate dataset. HeLa cells were transfected in triplicate with anti-PV1 miRNA expressing plasmid and infected with MOI 0.1 PV1 24 hrs later. Virus supernatant was harvested and serially diluted 24 hrs post-infection. Column (PV1) in graph was mock transfected, and infected with PV1 in the same manner as those with anti-PV1 miRNA.

Table 3.3 Synthesised siRNA oligonucleotides (RNA) and corresponding nucleotide positions in PV1 and PV3 genome

Name	Target	Nucleotide sequence	Nucleotide location[a]	Region
Cont-scrambled	-	5' UUAUCAUAACCUCACAGCCgg	-	-
MET-2C[b]	PV1/3	5' AGAGCAAACACCGUAUUGAgc	~4460-4480	CRE/2C
si-2599PV1	PV1	5' CCCUUCUGAUACAGUGCAAac	2599-2619	VP1
si-3146PV1	PV1	5' GCACUAGGUGACUCCCUUUau	3146-3166	VP1
si-6336PV1	PV1	5' GCUACCCUUAUGUAGCAAUgg	6336-6356	3Dpol
si-6439PV1	PV1	5' CCUCCCACUGGUGACUUAUgu	6439-6459	3Dpol
si-PV13D*	PV1	5' GUGCGAUCCAGAUUUGUUUug	6619-6639	3Dpol
si-7055PV1	PV1	5' GCUGACAAAUCAGCUACAUuu	7055-7075	3Dpol
si-2769PV3	PV3	5' CCGGGCACAGAAACUAUUUgc	2769-2789	VP1
si-2962PV3	PV3	5' GCACCCACACCAAAGUCAUgg	2962-2982	VP1
si-6211PV3	PV3	5' GGACCAUUAUGCUGGACAAcu	6211-6231	3Dpol
si-6243PV3	PV3	5' GGAUAUCAGCACAGAGCAAau	6243-6263	3Dpol
si-PV33D*	PV3	5' AUGCGAUCCAGACCUAUUCug	6610-6630	3Dpol
si-6681PV3	PV3	5' CAGGAUACGACGCAUCACUua	6681-6701	3Dpol

[a]Numbering according to PV1 strain Mahoney or PV3 strain Leon
[b] Adapted from Lee *et al*., 2007
* Adapted from Gitlin *et al.,* 2005

difference between miRNA and siRNA was that the output products for miRNA incorporated a small sequence to form a stem loop between the target sequence and appropriate nucleotide overhangs so that correct miRNA structure was achieved. Output siRNA duplex targets consisted of 21-mer target sequences without additional inserted sequences. Similar parameters were required for siRNA design (GC content, outside regions of secondary structure, etc.) as for miRNA which were discussed in detail earlier. Poliovirus type 1 siRNAs were further checked by BLAST to ensure that they did not contain high sequence identity to untargeted PV3 sequence and *vice versa*. An siRNA targeting the highly conserved enterovirus CRE was also included in these experiments and was named MET-2C (Lee *et al.*, 2007). All 21-mer oligonucleotide duplexes contained two 3' nt overhangs matching the target sequence. To assess the antiviral effects of these siRNA individually, HeLa cells were transfected with individual siRNAs using Lipofectamine 2000 reagent and infected with either PV1 or PV3 at an MOI of 0.1 to determine if single siRNAs could inhibit virus production in cells and to determine whether there was a virus serotype specific response. Transfection and infections of HeLa cells were performed in triplicate, and each triplicate titred once by $TCID_{50}$ assay.

The effects of anti-PV1 and anti-PV3 siRNA on virus replication were quantified by comparing virus titre in the presence or absence of 25 pmol/well siRNA at 24 hrs post-infection. This was the recommended dose for siRNA according to the number of target cells using Lipofectamine 2000. The effect of single siRNAs transfected into cells prior to virus infection was calculated by comparing virus titres to those of untreated cells to give a $\log_{10}$ $TCID_{50}$ virus yield. The geometric mean for each triplicate was graphed, together with the standard deviation (as error bars) and statistical significance marked with an asterisks symbol where appropriate.

As expected, prior transfection with the scrambled control siRNA (cont-scrambled) had little or no effect on both PV1 and PV3 replication (figure 3.7). This proved that the scrambled siRNA sequence had little homology to an untargeted virus sequence, and also did not affect normal cell function since resulting virus titres were similar to mock transfected cells. With regards to antiviral PV1 siRNA (blue bars in figure 3.7), all but one was able to reduce PV1 yield in transfected cells, with the highest and statistically significant reduction caused by si-3146PV1 (> one $\log_{10}$),

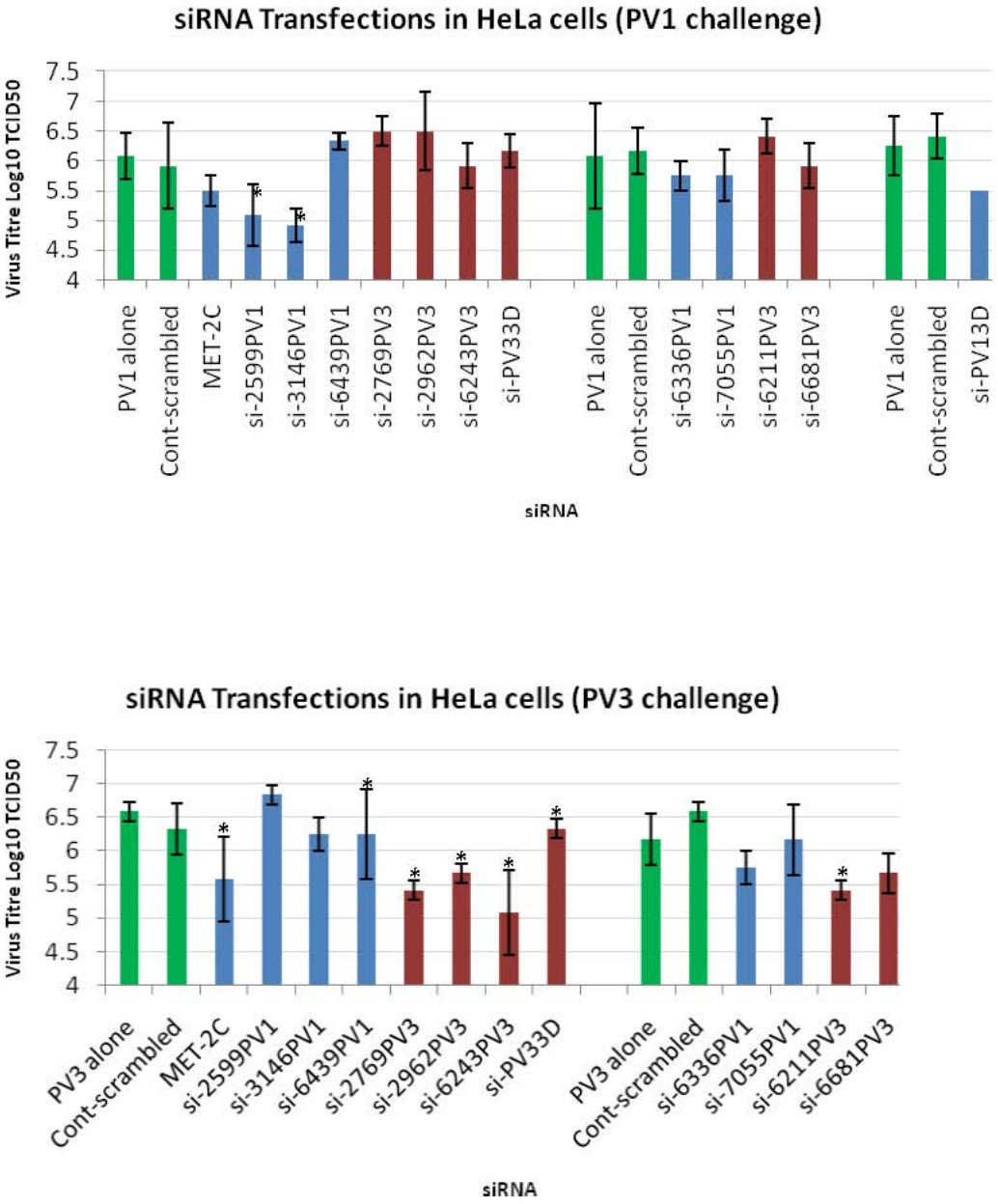


Figure 3.7 Effects of siRNA on PV replication. Representative results showing the effects of transient siRNA expression in HeLa cells on PV1 and PV3 replication. Twenty-four hrs post-transfection, HeLa cells were infected with MOI 0.1 of either PV1 or PV3 for a further 24 hrs before virus was harvested and titrated. The graphed data are the geometric mean from triplicates, error bars indicate standard deviation. Blue bars indicate anti-PV1 siRNA, red bars represent anti-PV3 siRNA. These experiments were performed on separate days, as indicated by the gap in the x-axis. The siRNA si-PV13D has a SD of 0, depicted by the lack of error bar at three significant figures. Asterisk; $P < 0.05$ by Mann-Whitney U test with the test miRNA compared to either the PV1 or PV3 control, respectively. The statistical analysis did not include siRNA si-PV13D.

followed closely by si-2599PV1 (one $\log_{10}$). In contrast, an increase in virus yield was seen in cells transfected with si-6439PV1 (0.25 $\log_{10}$). Two anti-PV3 siRNA (si-6243PV3 and si-6681PV3) had a subtle effect on PV1 replication, reducing average virus yield by no more than 0.25 $\log_{10}$.

All anti-PV3 siRNA reduced PV3 titres (red bars in figure 3.7), with the highest reduction caused by si-6243PV3 (1.5 $\log_{10}$). MET-2C (conserved CRE sequence for PV1 and PV3) also reduced PV3 by 1 $\log_{10}$. Although si-6439PV1 was calculated to reduce the PV3 titre significantly, the range of error bars suggested that this result required further testing of replicates. In fact, variation within several triplicates was high again as indicated by the standard deviation, but this did not affect the statistical significance of some average virus yield reductions (calculated on the median by non-parametric methods).

The MET-2C siRNA was able to reduce both PV1 and PV3 titres when transfected into HeLa cells prior to virus challenge, indicating that this small siRNA oligonucleotide was able to interact with the stem loop structure of the CRE. This was expected as the siRNA targeted the conserved 2C-CRE region of enteroviruses essential for positive-strand synthesis and was able to inhibit virus replication, even if only in a limited capacity.

Unfortunately, antiviral siRNA were not able to protect HeLa cells from virus challenge to the levels first expected. Yield reductions of ≤ 1.5 $\log_{10}$ did not satisfy previous targets of 2 $\log_{10}$ reductions to begin recombination experiments involving dual virus infections. Modifications to experimental conditions were therefore employed in an attempt to improve the level of virus inhibition achieved.

Dose-responses of siRNA duplexes

After demonstrating the inhibition of PV1 and PV3 replication by certain specifically targeted siRNAs, it was necessary to determine the effectiveness of different concentrations of siRNA on virus replication. The expectation was that it should be possible to optimise the siRNA concentration to maximise the reduction in virus yield. A dose-response experiment was also performed to ascertain whether increased concentrations of siRNA would induce a non-specific cytotoxic effect in

HeLa cells when combinations of siRNA were used to inhibit virus replication during co-infection experiments. In the first instance, MET-2C was used as it was originally designed to cross-react with the conserved PV1 and PV3 CRE target and published data supported this (Lee *et al.,* 2007). HeLa cells were transfected with 3 to 100 pmol per well of MET-2C 24 hrs prior to PV1 and PV3 challenge at an MOI of 0.1. Mock and control (with cont-scrambled) transfections were also performed in the same assay. All transfections/infections were performed in triplicate and virus titres assayed once by $TCID_{50}$ assay. Figure 3.8 shows virus titres of MET-2C transfected cells after PV1 and PV3 challenge. Although the virus titre was only reduced by approximately 1 $\log_{10}$ at best (12.5 pmol MET-2C against PV3), both PV1 and PV3 results showed a dose-response curve for virus titres across the range of MET-2C concentrations. The highest and lowest concentrations of siRNA had a smaller influence on virus production, whereas the highest reduction to virus yield was caused by the 12.5 pmol MET-2C siRNA, roughly a midway concentration. High siRNA concentrations may have had a reduced affect due to lower transfection efficiencies if more siRNAs aggregated with cationic lipid prior to introduction into cells. High siRNA concentrations may also have disrupted cell processes that limited defences.

To determine whether this result was reproducible with another siRNA, cells were transfected with increasing concentrations of si-6243PV3 prior to PV3 challenge at an MOI of 0.1. Previously this siRNA was shown to exhibit significant antiviral activity against PV3 (figure 3.7). Results with varying concentrations of si-6243PV3 showed that although the mean inhibition showed a subtle dose-response effect, the overlap of the standard deviation values (by error bar) indicated that there was no substantial difference between the effect of the concentrations (figure 3.9). This was despite the fact that 100 pmol of si-6243PV3 reduced PV3 replication significantly (by non-parametric methods using median values). Overall, the reduction was still less than one $\log_{10}$. Results indicated that an increase in siRNA concentration would not greatly improve anti-viral activity with a plateau effect evident. As no statistically significant difference was seen with the range tested, further studies continued with the use of 25 pmol siRNA, the recommended concentration according to the Lipofectamine 2000 protocol.

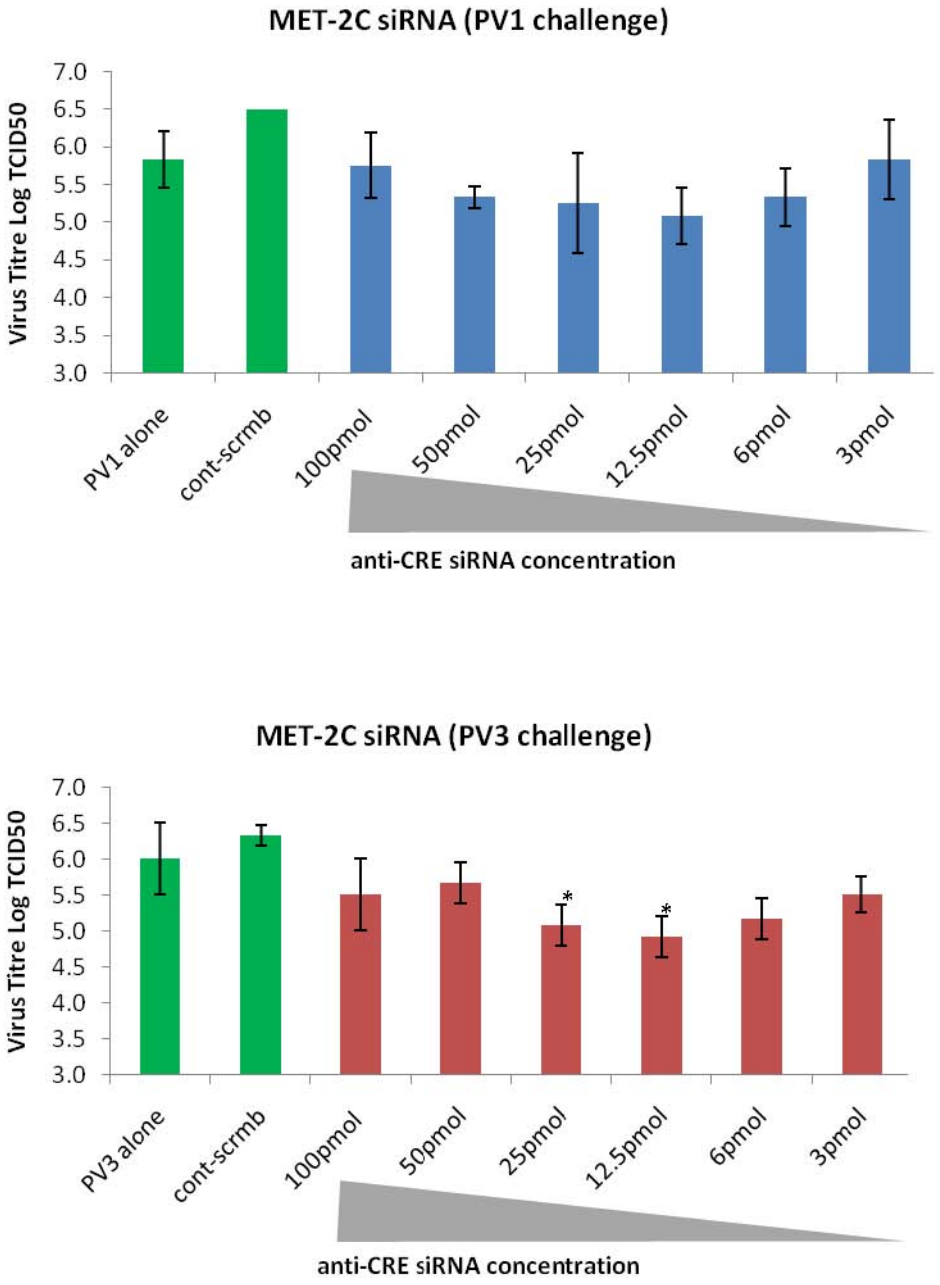


Figure 3.8 Effects of MET-2C siRNA dose-response on PV replication. Antiviral activity of different concentrations of anti-CRE MET-2C (conserved for both PV1 and PV3) on PV replication. Twenty-four hrs post-transfection with 3 to 100 pmol MET-2C per well in triplicate, HeLa cells were infected with MOI 0.1 PV1 or PV3 for a further 24 hours before virus was harvested and titrated. The graphed data are the geometric mean from triplicates, error bars indicate standard deviation. The cont-scrmb in the top graph had an SD value = 0, explaining the lack of error bar. Asterisk; $P < 0.05$ by Mann-Whitney U test with the test siRNA compared to the PV1 or PV3 alone, respectively.

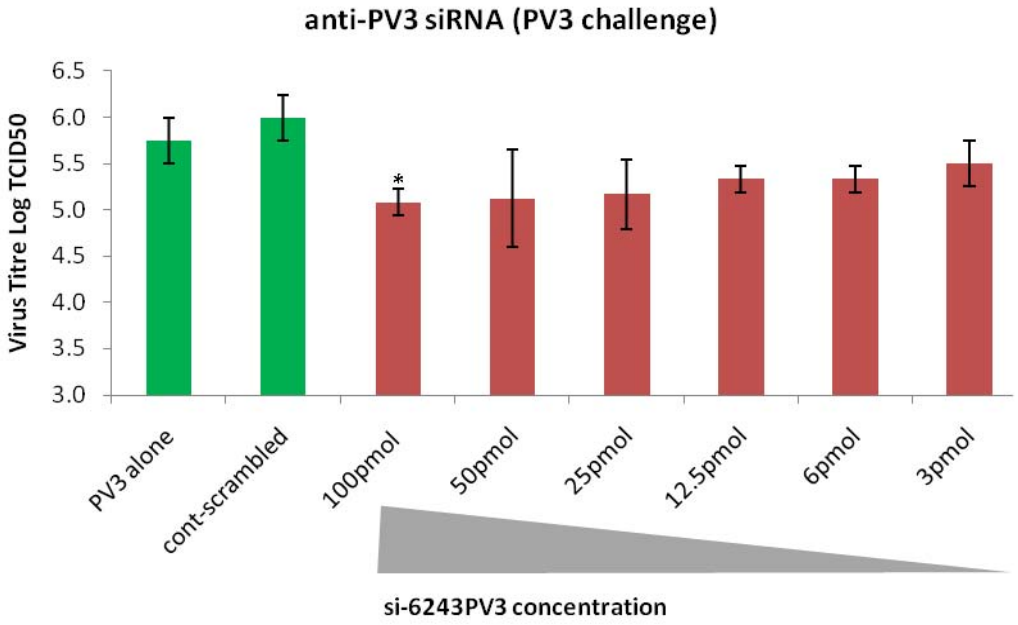


Figure 3.9 Effects of si-6243PV3 siRNA dose-response on PV3 replication. Antiviral activity of different concentrations of si-6243PV3 on PV3 replication. Twenty-four hrs post-transfection, HeLa cells were infected with MOI 0.1 PV3 for a further 24 hrs before virus was harvested and titrated. The graphed data are the geometric mean from triplicates, error bars indicate standard deviation. Asterisk; $P < 0.05$ by Mann-Whitney U test with the test siRNA compared to the PV1 or PV3 alone, respectively.

Accessing the anti-viral activity of combinations of siRNA duplexes

Although single anti-virus siRNAs were able to reduce the yield of virus production by as much as one $\log_{10}$, greater reduction (> 2 $\log_{10}$) was both expected and desirable for the selection of recombinant virus populations. The delivery of siRNA to all cells is important for efficient virus silencing, and it was likely that virus was escaping targeted degradation in cells that missed siRNA delivery. If only 40-50% of HeLa cells were transfected with siRNA (an assumption based on the transfection efficiencies seen with miRNA expressing plasmids), then this left half the number of cells unprotected against virus challenge. And although only an average of 9.5% of cells were infected initially at an MOI 0.1, this number increased with subsequent rounds of re-infection over a 24 hr period. Based upon the expected siRNA transfection level, approximately 50% of infected cells would be untransfected and therefore not able to display any suppression of parent viruses. This was evident by the scale of the reduction achieved with a single siRNA suggesting untreated cells were replicating parental virus to 5 to 6 TCID $\log_{10}$ (figure 3.7). Combinations of siRNA could improve virus inhibition both by potentially leaving more cells transfected (hence protected) and by the cumulative effect of multiple inhibitory siRNAs. This could also limit the ability of virus to escape RNA degradation by the introduction of spontaneous nucleotide mutations in siRNA target regions. It was expected that the use of two or more siRNAs in co-transfections could increase the effectiveness of virus inhibition by producing an additive or even synergistic effect. For example, two siRNA molecules co-transfected together would be expected to reduce virus progeny more than by a single siRNA molecule alone. An even greater reduction in virus titre (i.e. more than reduction of two single siRNA molecules added together) would indicate a synergistic result.

Combinations of siRNA were transfected into HeLa cells simultaneously to determine if there was an additive, synergistic, or no antiviral effect. Each transfection contained 25 pmol of single siRNAs, so that a co-transfection with four siRNA contained a total of 100 pmol per reaction. Assays were performed in triplicate in 24-well plates and each triplicate was assayed once by $TCID_{50}$ assay. As shown in figure 3.10, the virus titre was reduced up to one $\log_{10}$ (compared to virus alone) when cells were co-transfected with more than one virus-specific siRNA molecule which mirrored results from individual siRNA transfections (figure 3.7).

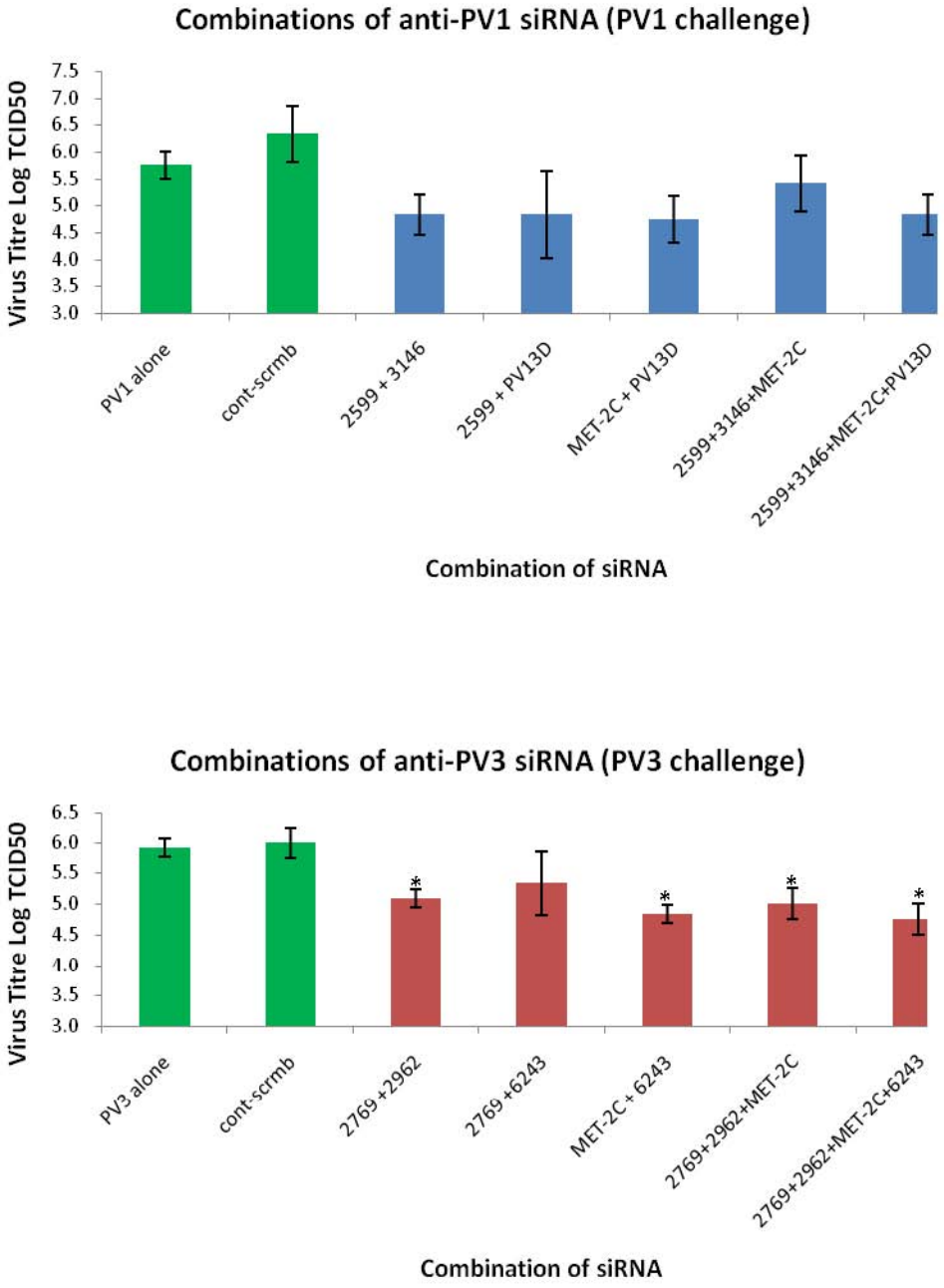


Figure 3.10 Effects of combinations of siRNA on PV replication. The inhibitory effect of mixed-siRNAs on PV1 and PV3 replication (top and bottom respectively). Two, three, or four siRNAs were transfected simultaneously into HeLa cells. Twenty-four hrs post-transfection cells were infected with PV1 or PV3 (MOI 0.1) for a further 24 hrs before virus was harvested and titrated. The plotted data are the geometric mean from triplicates, error bars indicate standard deviation. The blue bars indicate anti-PV1 siRNA, the red bars represent anti-PV3 siRNA. Asterisk; $P < 0.05$ by Mann-Whitney U test with the test siRNA compared to the PV1 or PV3 alone, respectively.

However, there was no evidence of an additive or synergistic effect when cells were pre-treated with more than one siRNA. This will be discussed further at the end of the chapter.

Optimising transfection methods for improved induction of RNAi

It is not clear how long siRNA duplexes are able to induce RNAi and protect cells from virus infection. Since siRNA are unable to replicate within a cell, the concentration of siRNA is effectively diluted as the cells divide and therefore become less effective at inducing RNAi as a consequence. It was therefore necessary to perform a series of preliminary experiments in which the incubation period of either the siRNA transfection or virus infection was varied in order to optimise the reduction in virus titre by anti-viral siRNAs. The si-3146PV1 siRNA was selected, as this siRNA was shown to inhibit PV1 replication in previous experiments (figure 3.7). Briefly, HeLa cells were transfected with 25 pmol si-3146PV1 siRNA per well, together with 0.5 µL Lipofectamine 2000. The assay was performed in duplicate, in 24-well plates and included the control scrambled siRNA and negative 'mock' transfection. Cells were infected with PV1 at an MOI of 0.1 at 6, 12, 24, or 48 hrs post-transfection. Virus was harvested from all wells after a further 24 hrs and titred by $TCID_{50}$ assay. Alternatively, HeLa cells were transfected as described above for 24 hrs only and infected with PV1 in the same manner, although virus was harvested after 8, 12, 24, or 48 hrs and subsequently titred by $TCID_{50}$ assay. Figure 3.11 (both A and B) showed that anti-PV1 siRNA si-3146PV1 provided little or no protection to HeLa cells from PV1 infection. This experiment demonstrated variability between assays performed at different times, and although this has not been addressed in the results thus far, it indicated the lack of consistency with siRNA experiments (i.e. low reproducibility). No virus titre reduction was observed after 48 hrs infection, indicating that virus was able to replicate uninhibited in cells with little or no siRNA protection. This suggested transfection efficiency and perhaps stability of siRNA in cells was not optimal or the siRNA was diluted as cells replicated over this time period. The greatest reduction in virus titre (i.e. best protection) was detected after a 12 hr virus harvest, which equated to one complete virus cycle. The non-specific scrambled control however, provided better protection than the virus specific siRNA.

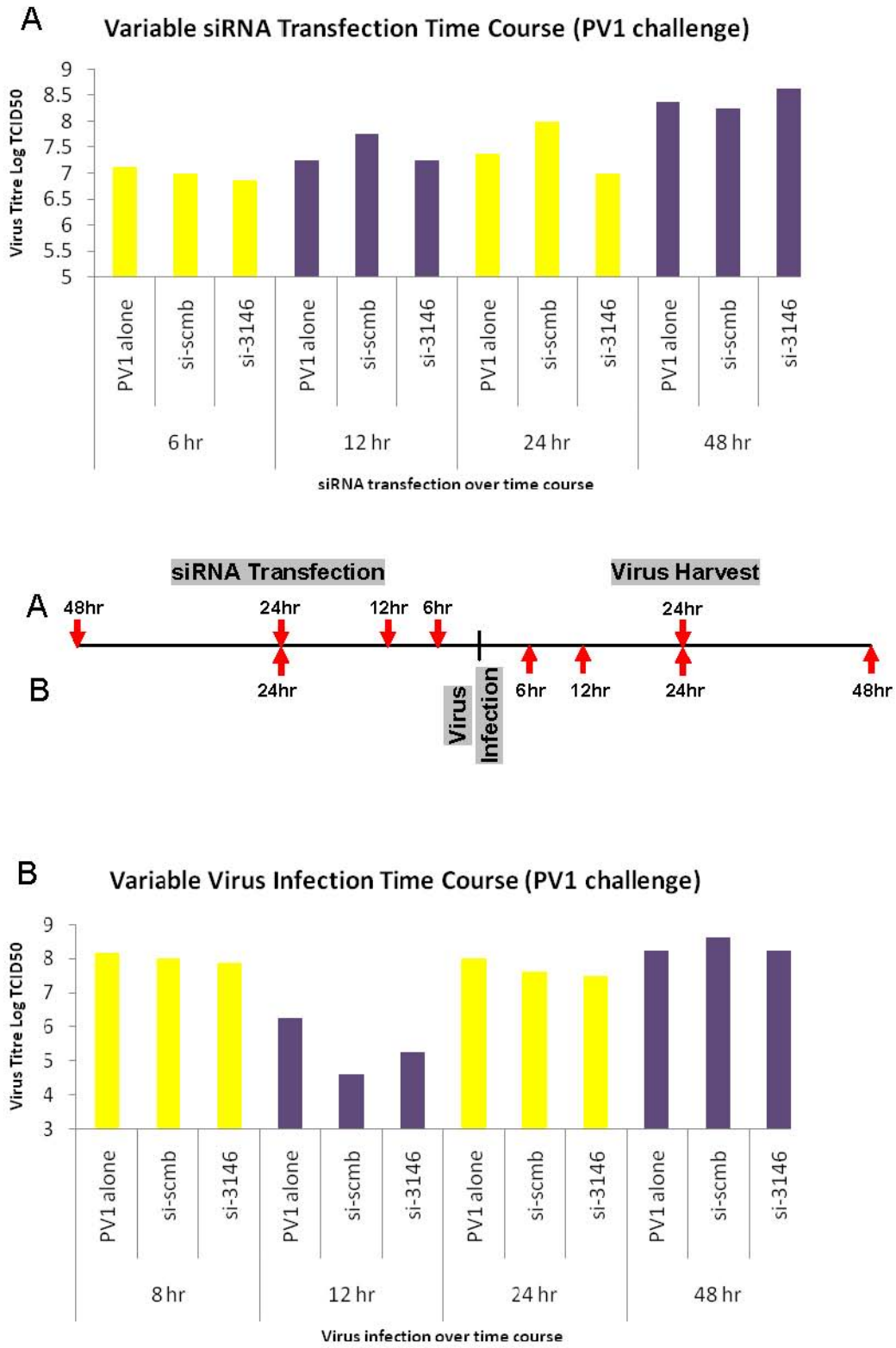


Figure 3.11 Optimisation of siRNA transfection and infection time points. (A) This graphs shows the changes in virus titre when HeLa cells are transfected with si-3146PV1 from 6 to 48 hrs prior to PV1 challenge for 24 hrs. (B) The graphed data depicts virus titre after variable infection times after HeLa cells were transfected with si-3146PV1 for 24 hrs. Geometric means of duplicates shown.

Ultimately, this experiment did not show that any particular transfection/infection regime significantly influenced the inhibitory effect of the siRNA.

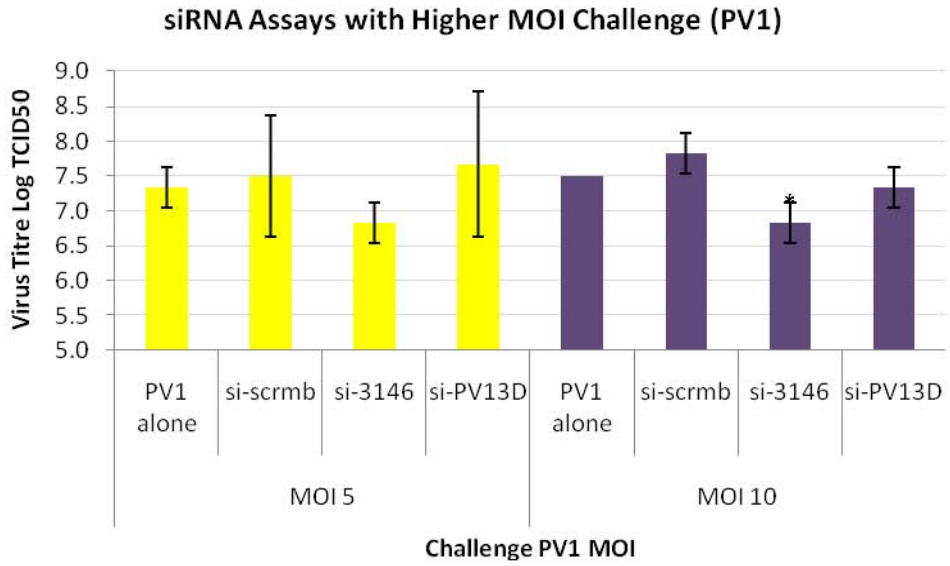
Increasing MOI of virus infection post-siRNA transfection

For further evaluation of the effectiveness of siRNA targeted degradation of propagating viruses, a selection of siRNA were tested against virus challenge at a higher MOI. This was essential as cells require successful co-infection to encourage recombination between the viruses, and this would only be achieved using higher MOI of parental viruses for future recombination experiments. To this end, two siRNAs that were shown to work individually against PV1 (si-3146PV1 and si-PV13D) and PV3 (si-2962PV3 and si-6243PV3) were selected and transfected into HeLa cells with controls used as previously described. Transfections were performed in triplicate in 24-well plates with 25 pmol of each siRNA. After 24 hrs, cells were infected with either PV1 or PV3 at MOI 5 or 10 and incubated for 24 hrs. Virus was harvested and titred once by $TCID_{50}$. The overall virus titres were generally comparable regardless of initial MOI suggesting that the infectious cycle had peaked at 24 hrs (figure 3.12). The scrambled control siRNA had little or no effect on virus titre which was expected and transfection with si-3146PV1 protected HeLa cells from PV1 challenge at both MOIs, but only by approximately 0.5 $\log_{10}$ which was calculated to be statistically significant using non-parametric methods. Likewise, si-2962PV3 reduced PV3 titres in both experiments significantly, although progeny was only reduced by approximately one $\log_{10}$. The experiments suggested that anti-virus siRNA were unable to reduce the titre of propagating viruses to any useful degree (> two $\log_{10}$), especially when higher quantities of virus was used to challenge transfected cells. This was not a surprising result, considering that an average of 99% of cells would initially become infected by at least one virus particle when cells were infected at an MOI 5 or 10.

3.5 Discussion

Since its discovery in *C. elegans*, RNAi has become an efficient and commonly used tool to regulate gene expression and study cell processes in detail (Fire *et al.*, 1998). Virologists have adopted this system as an antiviral tool with mixed results. The

A



B

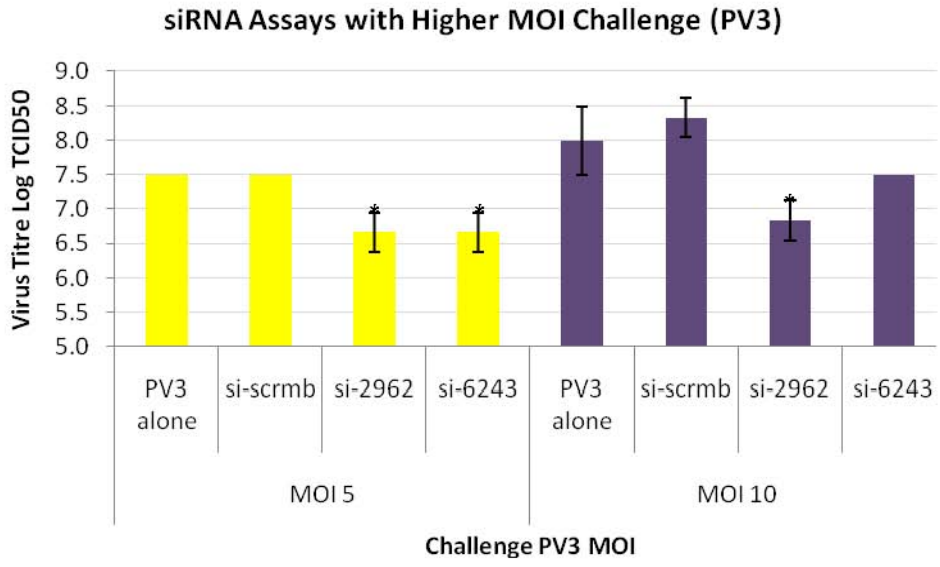


Figure 3.12 Challenging siRNA transfected HeLa cells at higher MOIs of PV. (A) Representative results showing the effects of transient siRNA in HeLa cells on PV1 replication. Cells were transfected with 25 pmol virus specific siRNA and subsequently infected with PV1 at an MOI of either 5 or 10. Virus was harvested after 24 hrs and titrated by $TCID_{50}$. Geometric mean and standard deviation of triplicates are shown. (B) Repeated with anti-PV3 siRNA and challenge with PV3. In graph B, the bars without error bars have a SD = 0. Asterisk; $P < 0.05$ by Mann-Whitney U test with the test siRNA compared to PV1 or PV3 alone, respectively.

majority of groups directly challenged infectious virus with anti-viral siRNA and when presented as plaque forming units, demonstrated reductions in virus titre by 1 to 2 $\log_{10}$ (Gitlin *et al.*, 2002; Lee *et al.*, 2007). Although reductions in virus load may be statistically significant, there still remains in excess of 10^5 to 10^7 PFU/ml virus in the progeny. This excess of parental viruses was not considered acceptable when attempting to identify recombinant progeny viruses.

The present work suggested that transient expression of miRNAs provided unpredictable inhibition of virus replication and results demonstrated little sequence-specificity of the plasmid generated miRNA to its targeted virus. This included the antiviral activity of scrambled controls and empty vectors, which has been observed in other studies (Wang *et al.*, 2004) but was not presented in this chapter. Intra- and inter-assay variation was found to be excessive (data not shown) indicating that this method was not reproducible and appropriate given the desired target knock down of $> 2 \log_{10}$. Collectively, results suggested that RNAi using plasmid expression systems may be unreliable in that (1) transfection efficiency of plasmids was lower than necessary for complete cell protection against virus infection, (2) specific miRNAs cross-reacted with heterologous virus, therefore no virus specificity could be gained in the assay; (3) an interferon response may have been indirectly triggered by the presence of 60-70 base pair dsRNA duplexes made by the miRNA plasmids, even though there was no overall reduction in viral yield; (4) although reductions in virus yield were sometimes statistically significant compared to the controls, the amount of reduction did not warrant subsequent recombination experiments.

It has been suggested that nonspecific interferon responses could be activated by plasmid-mediated miRNA expression, as longer ds RNA molecules are generated in cells by plasmids including those used in this study, and this was briefly investigated but without success (Elbashir *et al.*, 2001a; Gitlin *et al.*, 2002; Watanabe *et al.*, 2007). This was probably not an issue with these experiments, as some anti-viral miRNA had no effect on virus titre regardless. The project moved to using shorter siRNA duplexes to overcome potential complications involving interferon pathways.

The second approach relied on introducing chemically synthesised siRNA duplexes designed against regions encoding VP1 (capsid), CRE and $3D^{pol}$ (RdRp) of PV1 and

PV3, into mammalian cells. Individual siRNAs reduced virus yield by up to one $\log_{10}$ and showed specificity to the targeted virus. Similar outcomes were noted when pre-treated cells were challenged with higher concentrations (MOI 5 and 10) of virus.

There was no evidence of additive or synergistic antiviral activity when combinations of siRNAs were used as would be expected from this type of experiment. The lack of an effect was similar to recent results with other picornaviruses (Ahn *et al.*, 2005; Wu *et al.*, 2008). One possibility was that transfection reagent/siRNA complexes that formed were unevenly distributed to cells during the transfection process. Subsequently, a proportion of cells in the population lacked the protective action of siRNA, allowing for full replication of infecting virus. Alternatively, only a small proportion of virus RNA may have been accessible intracellularly in the first place.

Binding of an individual siRNA molecule may also change the secondary structure of RNA, making the genome more or less accessible for local binding of another siRNA. Experimentally this could be verified by various RNA secondary structure studies including selective 2' hydroxyl acylation analysed by primer extension SHAPE (Merino *et al.*, 2005) but these additional investigations were not feasible given the scope of this project.

Depending on the transfection efficiency of siRNA into mammalian cells (not tested), it has been postulated that a proportion of cells in the population lacked the protective action of siRNA, therefore allowing full replication of infecting virus effectively "by-passing" the siRNA in the experimental system (Ong *et al.*, 2006; Pijlman *et al.*, 2008). The transfection efficiency of siRNA in HeLa cells was not determined in these experiments, but could be calculated using a commercial fluorescently labelled siRNA.

Electroporation provided an alternative method of delivering siRNA into cells, which could enhance the distribution of siRNA to a greater proportion of cells. An experiment was performed in which HeLa cells were electroporated with combinations of siRNA, using conditions outlined in materials and methods on the Gene Pulser Xcell Electroporation System (Bio-Rad), before PV challenge. There

was no observable difference in virus titre between the mock transformed controls and in the presence of anti-PV siRNA (data not shown). However, to be confident with these results, the assay would need to be repeated and possibly optimised further.

RNAi may not have been an effective antiviral method from the start if it is accepted that mRNA degradation triggered by RNAi occurs away from areas in the cell involved in virus replication. It has been postulated that key components of RISC are somehow linked with processing bodies (P bodies) within cells, and disruption of P bodies can affect the efficiency of RNAi as a result (Jakymiw *et al.*, 2005; Sen *et al.*, 2005). Dougherty *et al* have recently shown that P bodies are disrupted during PV infection and this correlates with the build up of viral proteinases in cells (Dougherty *et al.*, 2011). According to earlier papers, it was not clear whether infecting viruses were associated with P bodies, but evidence suggested that infection of cells with PV could disrupt the formation of stress granules which are often docked to P bodies (Beckham & Parker, 2008; White *et al.*, 2007). The association between RNAi and virus replication in cells remains to be determined, as virus replication occurs on rosette-like membranous vesicles (induced by virus and cellular proteins) which are shown to be 'closed' entities whereby no exchange of virus protein, RNA, or membranes is evident (Bienz *et al.*, 1980; Egger *et al.*, 2000; Teterina *et al.*, 1997). It would be unlikely that the RISC/siRNA complex could interact with this replication complex in order to associate with target virus RNA. Taken together, this suggests that P bodies and virus replication complexes are separate entities within cells and may not associate with each other, meaning that induction of RNAi *in vitro* may not have reduced virus replication in cells. This is yet to be proven and was not included in the scope of this project.

In conclusion, exogenously introduced RNAi was not an effective method for reducing virus replication *in vitro* (in this case, parental viruses upon co-infection). As demonstrated, virus titre reduction was not altered sufficiently (i.e. a minimum of > two $\log_{10}$) by the use of different siRNA delivery methods into cells. Thus, the relatively inefficient reduction in virus replication can be explained by a lower than expected transfection efficiency. Detectable recombination events are rare (frequency of 10^{-5} to 10^{-6} *in vitro*) and recombinants would be extremely difficult to isolate from

assays that only had a one $\log_{10}$ reduction in parental viruses in the final virus progeny.

This study could be altered to limit the concentration of virus progeny. Infectious virus RNA (and not virus particles) could be co-transfected simultaneously with sequence-specific siRNA into a permissive cell line, e.g. L929 or BsrT7 cells. As the cell lines lack the necessary receptors for enteroviruses, they would not be able to reamplify viruses that have escaped RNAi in neighbouring cells. Low levels of virus RNA will always escape RNA degradation and form wild-type virus, but the process should allow for a higher proportion of recombinants to be selected using virus specific siRNA when following the original RNAi design (figure 3.2B).

Ultimately, this approach to studying virus recombination was abandoned and an alternative strategy involving the use of parental genomes bearing defined lesions was investigated.

CHAPTER FOUR: *In vitro* Recombination

4.1 Introduction

The result of many separate recombination events have been observed and characterised predominantly from isolates obtained after virus outbreaks and natural infections (Furione *et al.*, 1993; Kew *et al.*, 2002; Oberste *et al.*, 2004a; Simmonds & Welch, 2006). However, little is actually known about the mechanisms (viral and host) that establish the location of recombination breakpoints across the virus genome. Early *in vitro* and *in vivo* studies have investigated minor molecular mechanisms that influence or correspond to recombination events but there is scope for more exploratory analysis (Pilipenko *et al.*, 1995; Romanova *et al.*, 1986; Tolskaya *et al.*, 1987). The primary challenge in studying recombinant viruses is their isolation from progeny, particularly when parental strains dominate the population and can out-compete the few functional recombinant viruses present. As an example, recombinant viruses have been shown to represent a fraction of the total virus population, with interserotypic recombinants estimated to have frequencies of 1×10^{-6} (Jiang et al., 2007, Kirkegaard & Baltimore, 1986). The development of a method that increases recombinant viruses in a population, and allows for easier isolation of such viruses, will aid in understanding the underlying molecular mechanisms of recombination.

An attempt to enrich recombinants from a virus population using RNAi proved unsuccessful (see chapter three) due to high titres of parental virus that could not be reduced by RNAi methods. Extremely high yields of parental virus would make identification of recombinant viruses difficult in this system, especially following a dual infection requiring high MOIs. A model that only produced recombinant virus in the absence of any background parental virus would be ideal.

Recombination has been demonstrated from crosses in the presence of selection, using parent viruses that contain easily tractable point mutations that allow them to escape (or perish under) a single selection pressure (reviewed in chapter one). As a result, recombinants detected to date have ‘inherited’ the genetic mutations allowing them to survive under conditions that restrict growth of the parents. Although such

studies have been successful in generating recombinants, these recombinants would have had to arise more frequently than revertants in the original parent strains. If viruses originally evolve point mutations to escape selective pressure (e.g. guanidine resistance, temperature sensitivity, etc.), then it is probable that reversions of the point mutation can restore normal function to viruses. If point mutations occur slightly less than one per genome replication cycle, e.g. 2.1×10^{-4} to 5.9×10^{-4} for picornavirus RdRp, then this is more likely to occur than generating one interserotypic recombinant (10^{-6}) during a dual infection (Crotty *et al*., 2001; Kirkegaard & Baltimore, 1986). If indeed rates of reversions are higher than recombination events, then this type of recombination experiment may have low levels of non-recombinants. Therefore, it would be ideal to incorporate two defective partners that are absolutely incapable of reverting to original infectivity without the help of recombination. For example, using virus RNA molecules with large deletions so that only portions of the genome are added, or destroying replication elements within the genome. This would create a tightly regulated process with no background from breakthrough parents.

Therefore, a new *in vitro* method for generating chimeric viruses was formulated to simplify recovery of recombinant viruses in the absence of selection altogether. This system used two defective virus RNA molecules (herein described as "parents") that were mutated so that neither parent RNA, when introduced into a cell alone, could generate a progeny virus. A subgenomic replicon was the obvious choice of a replicating RNA containing a large irreversible deletion (i.e. the virus parent encoding the capsid had been replaced by a reporter gene). When transfected into cells, this RNA would be unable to encode capsid proteins required to form virus particles, but would provide the non-structural region of the genome for potential crosses. A second partner would therefore be required to produce the capsid encoding region necessary, but contain irreversible defects in the non-structural region so that virus could not be generated from this partner alone. A good candidate for the second partner was a full-length virus RNA with mutations introduced in the CRE that disrupts positive-strand RNA synthesis (Goodfellow *et al*., 2000). This was deemed a semi-replication competent RNA partner, as only negative RNA strands are synthesised upon transfection into cells.

The system relied on introducing both parent virus RNAs to a single cell where subsequent recombination (donation of a full capsid and functional CRE by either parent) created functional, recombinant progeny which could be cloned and studied. As RNA was used for both parents, and not infectious virus, the method did not require cells susceptible to infection. It only required that cells were permissive to virus replication. Both RNA molecules retained the ability to replicate (although only partially for one partner), and for this reason, the aim was to investigate replication-dependent recombination in enteroviruses.

The proposed system allowed recombination with any enterovirus partner to be investigated, as long as virus sequence was available in plasmid form (cDNA). With the aid of reverse genetics, any enterovirus cDNA could be modified to incorporate mutations or large deletions that inhibited the production of infectious virus.
As this approach relied on cell culture, recombinants were generated and characterised before they were subjected to significant selection pressures (i.e. immune system and selective pressures that would result from infecting multiple host cell types).

This system was therefore used to determine whether recombination was possible within the cell between two RNA partners belonging to different species. This was possible by by-passing such restrictions as virus attachment to cells and infection process. The intention was to emphasise whether restrictions on recombination were imposed by virus factors, host factors, or both. This method was used to characterise recombinants and compare them to those naturally selected recombinants from patients and environmental samples, and possibly analyse mechanisms that influence the RdRp template switching process.

Aim

The aim was to develop a method to generate live, recombinant PV *in vitro* in the absence of infectious output parental viruses. If such an approach yielded infectious PV in the first instance, then the intention was to expand on input parental RNA sequences to include other non-poliovirus enteroviruses.

4.2 Proposed Methodology

A replication-dependent method was developed to generate natural populations of recombinant viruses *in vitro* without the presence of parent virus in the progeny. To make this system free of non-recombinant background virus, two PV RNA molecules containing lesions rendering them unable to generate viable virus on their own were co-transfected into either mouse fibroblast (L929) or baby hamster kidney (BsrT7) cells. Both cell lines were permissive for virus RNA replication but lacked the necessary PV receptor necessary for infection, therefore any recombinant progeny virus generated was unable to re-infect neighbouring cells. In theory, the joining of a 5' capsid encoding region with 3' non-structural parts of the PV genome should result in the generation of a complete infectious PV genome that could be identified upon passage to a susceptible cell line.

A diagrammatic representation of this method is provided in figure 4.1. The first virus genome was a full-length RNA molecule with introduced mutations ("lesions") in the CRE. Mutations in the CRE were lethal to the generation of viable virus but still allowed for synthesis of negative RNA strands (Goodfellow et al., 2003), which complemented this process, since recombination has been shown to occur during negative RNA strand synthesis (Kirkegaard & Baltimore, 1986). The second RNA molecule was a subgenomic replicon that had the capsid region replaced with a reporter gene (e.g. luciferase). This was entirely replication competent but could not generate a viable virus without capsid proteins. As defects existed in each partner RNA molecule, only recombined RNA could generate virus, and recombination could only occur in the region between the virus capsid sequence and the CRE sequence (i.e. viral sequences 2A, 2B and part of 2C), which in the case of PVs is where the majority of recombination is reported to occur in natural isolates (refer to chapter one).

Figure 4.2 provides a diagram showing the method used to generate and isolate viruses. As mentioned briefly, two parent RNAs were simultaneously transfected into either L929 or BsrT7 cells. As cells blocked reamplification of newly generated recombinant viruses, this removed competition between recovered viruses. As a consequence, a range of possible recombinant viruses could be recovered regardless of growth advantages as this assay essentially yielded viruses that had only

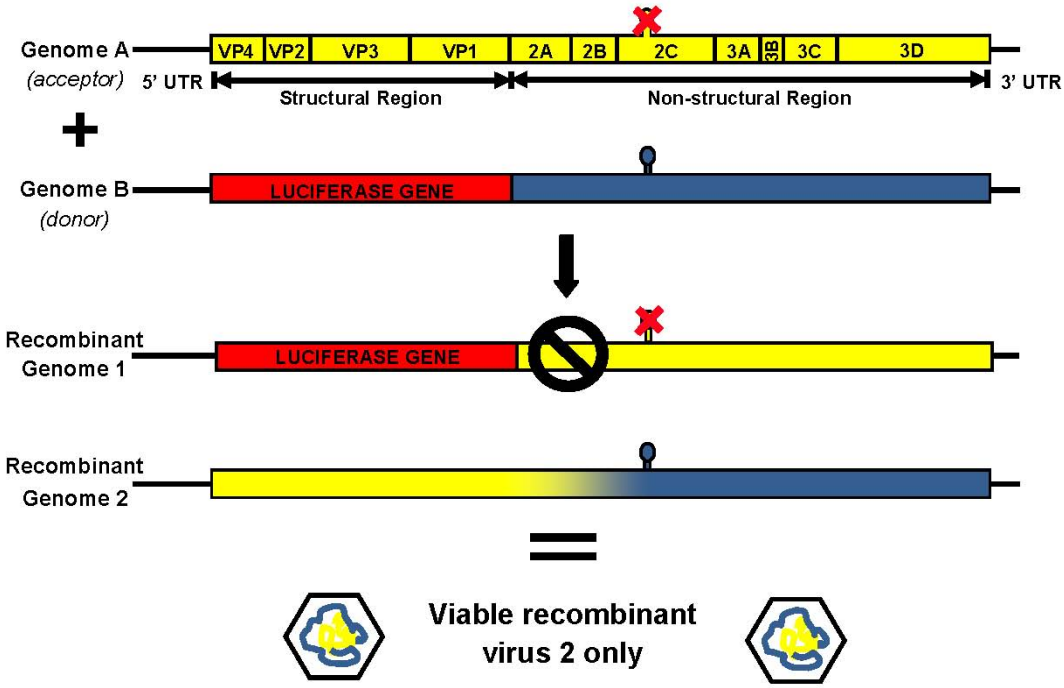


Figure 4.1 Generating recombinant virus in the absence of infectious parental virus. Two RNA molecules containing "lesions" rendering them unable to generate viable virus on their own were co-transfected into L929 mouse cells. Viable infectious virus was recovered only when both RNA molecules were present and able to recombine. Genome A encoded a full-length virus cDNA with mutated CRE that inhibited positive-strand synthesis during virus replication. Genome B contained a reporter gene (such as luciferase) in place of the capsid region in a subgenomic replicon. Crossover regions were targeted to 2A, 2B, and 2C before the CRE.

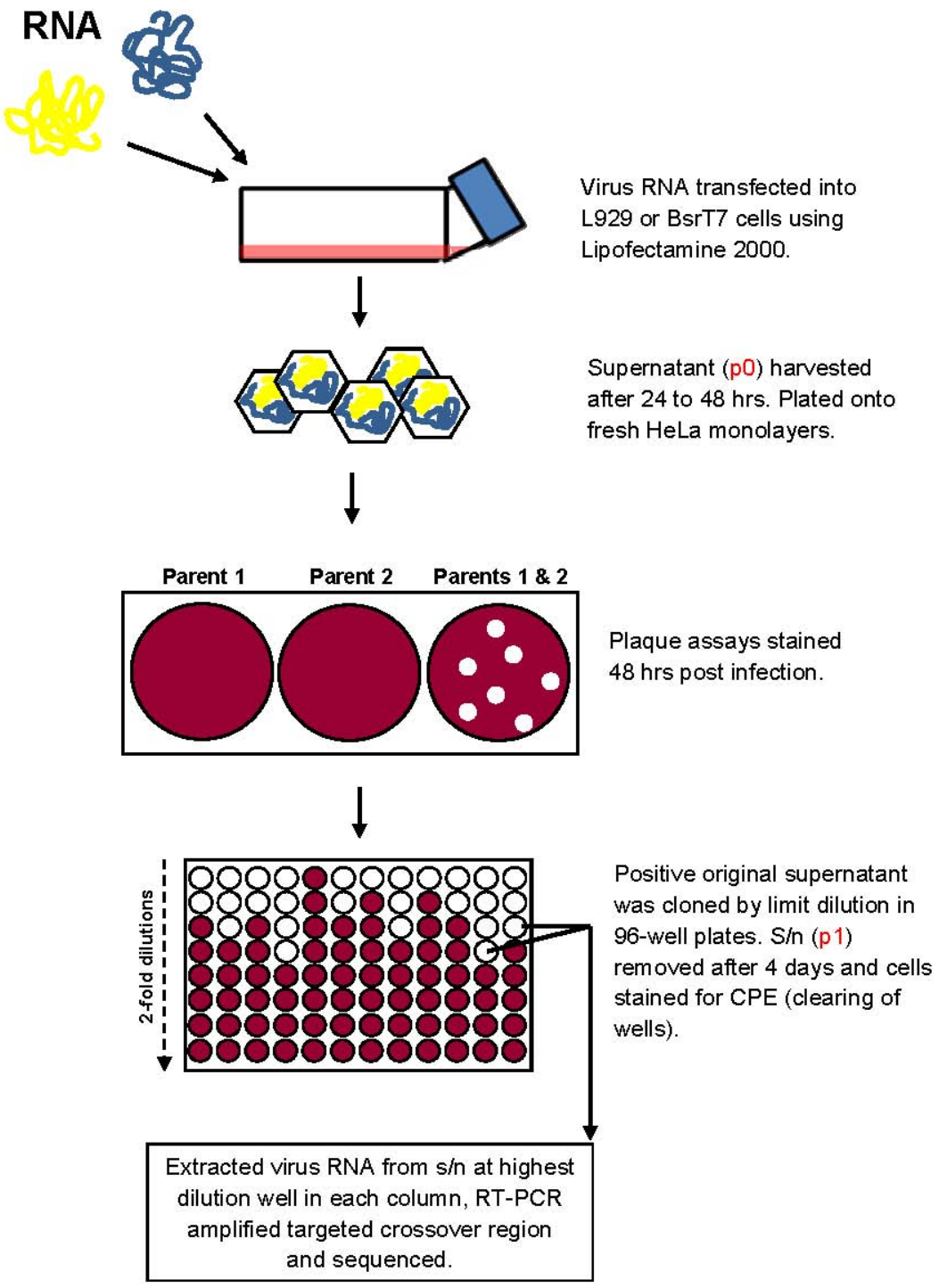


Figure 4.2 Method used to generate recombinant viruses by co-transfection. Two virus RNA partners were co-transfected into L929 or BsrT7 cells. Twenty-four to 48 hrs post-transfection, supernatant (s/n) was harvested and plated undiluted onto fresh HeLa cells and overlaid with overlay medium. Cells were stained with crystal violet after 48 hrs and supernatant positive for viable virus was diluted onto fresh HeLa cells in 96-well plates. Supernatant was harvested, cells stained again, and virus RNA was extracted and amplified by RT-PCR.

experienced a single round of replication. Viable viruses capable of causing CPE could be detected in co-transfection supernatant by plaque assay using mammalian cells (HeLa cells) susceptible to PV infection.

The next step involved isolating individual viruses in the progeny to characterise the nature of infectious genomes recovered. It was necessary to limit the use of RT-PCR for identifying viruses in a mixed population as some DNA polymerases are known to switch from one template to another during the amplification process (Pääbo *et al.*, 1990). This could introduce false crossover sites into the sequences of recovered, recombinant viruses. Consequently, PVs were biologically cloned by limiting dilution in 96-well plates using a method adapted from $TCID_{50}$ assays (refer to Materials and Methods). Virus genomes could be analysed by standard molecular methods thereafter.

4.3 Recovery of Intraserotypic Recombinants

Recovery of PV3/PV3 intraserotypic recombinants from co-transfections

To test this experimental system, two homologous virus RNA partners were used to establish whether the model would generate viable recombinant viruses. If combining two PV3 RNA molecules sharing 100% sequence identity was unsuccessful, then this approach would require further modification. Both PV3 subgenomic replicon pT7Rep3-L and full-length PV3 CRE mutant pT7FLC/SL3 plasmids were linearised with *Sal*I and RNA transcribed using T7 polymerase. Two micrograms RNA per well was singly or co-transfected into L929 cells in 6-well plates with Lipofectamine 2000. In the first instance, different ratios of partner RNA were transfected into cells, 1:1, 1:4, and 4:1 (replicon: full-length CRE mutant). As a control to demonstrate the dependence of replication on this recombination process, an additional co-transfection was performed in the presence of 4 mM guanidine hydrochloride, a potent inhibitor of PV replication (Caliguiri & Tamm, 1968). In addition, a luciferase expression control using the luciferase replicon with and without the addition of 4 mM guanidine hydrochloride was performed simultaneously to assess the effectiveness of transfections and replication in cells, and this was implemented for all subsequent replicon transfections. Since observations at 24 hrs showed no signs of CPE due to the lack of virus re-infection,

supernatant was left a further day and harvested 48 hrs post-transfection. As concentrations of potential virus were unknown, 500 µl of undiluted supernatant was used to inoculate 6-well HeLa plates for plaque assays. Plates were incubated for 48 hrs, stained with crystal violet, and photographed.

When either pT7Rep3-L or pT7FLC/SL3 RNA was transfected into L929 cells alone, supernatant harvested from L929 cells did not yield plaques in HeLa cell monolayers, although in some experiments drying artefacts were observed in some wells (both test and control) (figure 4.3A). When both RNA partners were transfected into L929 cells simultaneously, viable infectious virus was detected in harvested supernatant regardless of input RNA partner ratio (figure 4.3A). As undiluted co-transfection supernatant was used to inoculate HeLa cells for plaque assays, it was evident that the resulting virus titre was very limited. Although plaque numbers were difficult to count in 1:1 and 1:4 wells, plaque numbers at a 4:1 of RNA indicated the recovery of 84 PFU/ml of infectious virus.

No virus was produced in the presence of guanidine hydrochloride which confirmed that RNA replication was required during the process of virus recovery. The guanidine hydrochloride control did not rule out the occurrence of RNA breakage and ligation as a form of recombination in cells but did suggest that RNA required continuous replication to produce enough RNA copies for continued virus propagation in the L929 cells.

To determine the efficiency of this model for producing virus from two genome 'halves', full-length infectious PV3 RNA (pT7FLC) was transfected into L929 cells using the same experimental conditions described above. Approximately 1.05×10^5 PFU/µg wild-type RNA was generated. This was 2500-fold more than the concentration of virus recovered from pT7Rep3-L and pT7FLC/SL3 co-transfections, which were two homologous RNA partners. Evidently, if this virus was in fact the result of genetic recombination in cells, then this illustrates the infrequent occurrence of RNA recombination.

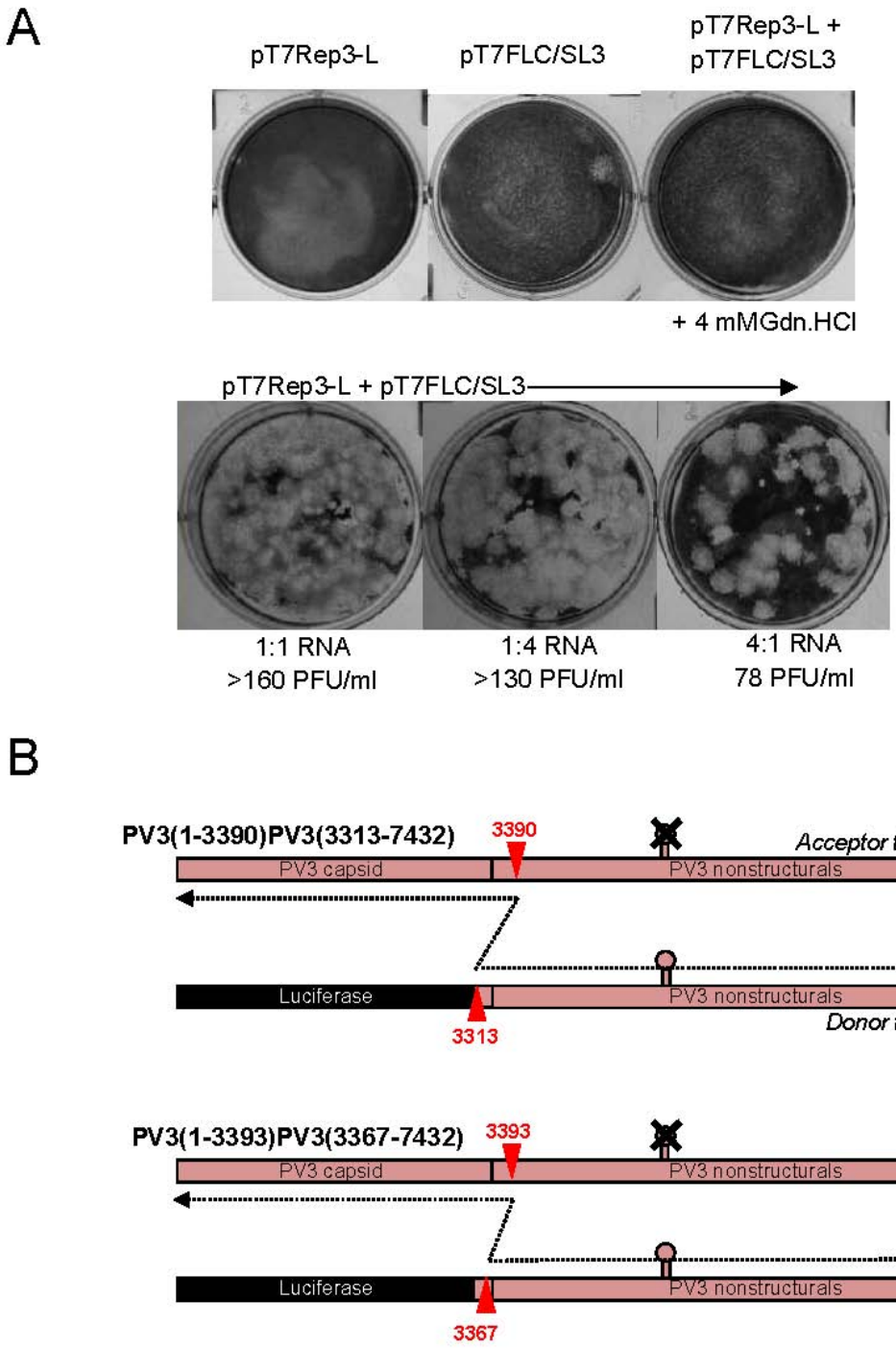


Figure 4.3 Virus recovered from L929 cells co-transfected with PV3 RNA molecules. (A) Crystal violet stained HeLa cells infected with undiluted transfection supernatant and overlaid with plaque assay overlay medium. Control wells containing one RNA partner each and when both pT7Rep3-L and pT7FLC/SL3 RNA partners were cultured with and without guanidine hydrochloride. Plaques were visible after co-transfection with both partners at various RNA ratios. (B) Diagrams outlining imprecise intraserotypic recombinants PV3(1-3390)PV3(3313-7432) and PV3(1-3393)PV3(3367-7432) with insertions at the crossover junctions.

Characterisation of intraserotypic PV3 recombinants

The virus population generated from co-transfections above were further analysed to confirm that infectious virus was indeed PV3. Several PV3/PV3 viruses were biologically cloned by limiting dilution and sequenced in order to define genome sequence between capsid and mutated CRE in the 2C region. To enable the generation of viable virus, a recombination event would have occurred in this region of the virus sequence. The crossover region was amplified using oligonucleotides PV3-2995F and PV3-5191R (table 2.3). The majority (n=13) of the sequenced PV3/PV3 clones contained no changes to PV3 sequence between the capsid and CRE. However, two viruses contained imprecise junction sites with additional PV3 sequence identified (figure 4.3B). The first clone, PV3(1-3390)PV3(3313-7432) contained an extra 78 nt sequence in viral sequence 2A, while the second clone, PV3(1-3393)PV3(3367-7432), had 27 nts inserted also in the 2A region. Crossover positions were determined from full-length PV3 sequence strain Leon (GenBank accession number X00925).

The identification of imprecise crossover sites was unexpected considering that two identical molecules of virus RNA were used, apart from differing by the 'defects' used in each to force recombination.

4.4 Recovery of Interserotypic Recombinants

Recovery of PV3/PV1 interserotypic recombinants from co-transfections

As this model was able to produce viable virus from two RNA molecules encoding the same virus (PV3), the next step was to demonstrate *in vitro* recombination between two different serotypes that are known to recombine, e.g. PV serotypes 1 and 3 (Cuervo *et al.*, 2001; Paximadi *et al.*, 2006). Poliovirus serotype 1 subgenomic replicon pRLucWT and full-length PV3 CRE mutant pT7FLC/SL3 plasmids were linearised with *Apa*I and *Sal*I, respectively, and transcribed into RNA using T7 polymerase. The RNA partners were transfected into BsrT7 cells and again, different ratios of RNA were combined during co-transfections to determine whether an excess of one viral partner affected the outcome of recombination. Ribonucleic acid concentrations 1:1 (1 µg:1 µg), 1:4

(0.4 µg:1.6 µg) and 4:1 (1.6 µg:0.4 µg) of replicon to full-length CRE mutant were used. The BsrT7 and L929 cells were used interchangeably throughout experiments and did not exhibit cell-specific recombination outputs (data not shown).

As expected, BsrT7 cells transfected with either pRLucWT or pT7FLC/SL3 RNA alone did not generate infectious virus. No plaques were present in these control wells using HeLa cell monolayers (figure 4.4A).

However, plaques were obtained in co-transfection supernatant when pRLucWT and pT7FLC/SL3 were introduced into cells simultaneously at 1:1 and 1:4 ratios. Figure 4.4A shows the plaques present in HeLa cells when undiluted co-transfection supernatant was inoculated onto cells. The absence of virus recovered from the 4:1 RNA ratio (replicon: CRE mutant) was not surprising, since it has been shown that the frequency of recombination depends on an excess of acceptor template concentration (Jarvis & Kirkegaard, 1992). Previously, a lower titre of virus was recovered from 4:1 (pT7Rep3-L: pT7FLC/SL3) co-transfections (figure 4.3A).

The method described here has been proven to generate interserotypic PV recombinants from two virus cDNA molecules, and this is discussed in chapter five. There was no evidence of background contaminants or virus using this method and populations of recombinants were recoverable as a result of a tight selection process (no revertant parent viruses). There was little or no host selection applied to recombinants once generated, as cells were not susceptible to re-infection. However, viruses were selected at the genome level, particularly for continued viability. Recombinant genomes had to remain in-frame for correct protein translation, and conserve cleavage regions for appropriate post-translational processing. The complete characterisation of 100 biologically cloned PV3/PV1 recombinants was performed and is discussed in detail in chapter five.

4.5 Recovery of Intraspecies Recombinants

Recovery of intraspecies C recombinants from co-transfections

Both intra- and interserotypic PV recombinants have being reported, and a few intraspecies crosses between PV and species C viruses have been identified in

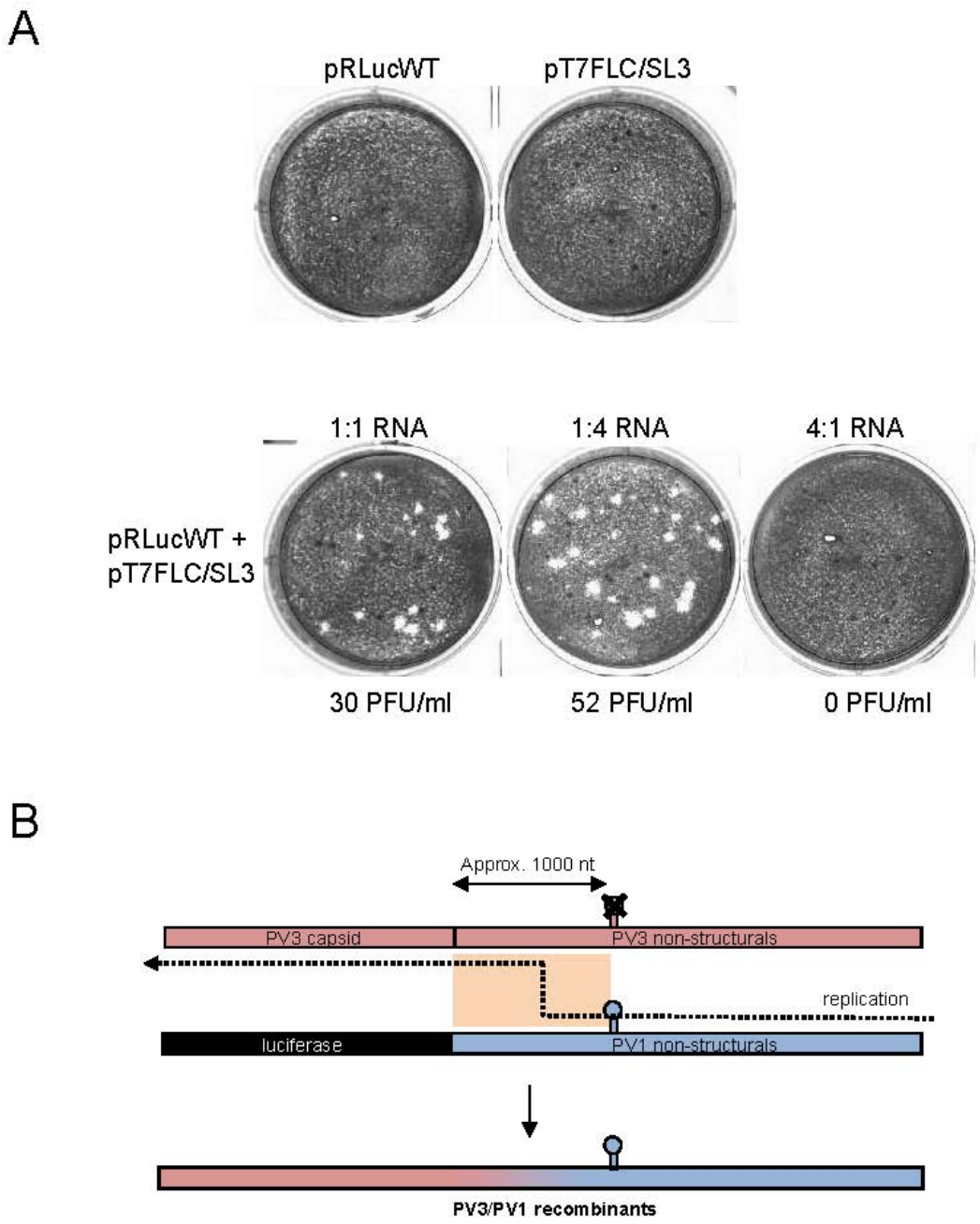


Figure 4.4 Virus recovered from BsrT7 cells co-transfected with PV1 and PV3 RNA partners. (A) Crystal violet stained HeLa cells infected with undiluted transfection supernatant and overlaid with plaque assay overlay medium. Control wells containing one RNA partner were clear; plaques were present at ratios 1:1 and 1:4 of combined RNA partners. (B) Schematic of PV3/PV1 recombinants generated by co-transfection, with RdRp template switching occurring between luciferase and mutated CRE (within the shaded region).

clinical isolates (Kew *et al.*, 2002; Rakoto-Andrianarivelo *et al.*, 2007). However, widespread intraspecies recombination has been described for other members of enteroviruses, justifying the need to attempt similar experiments using the method developed in this study (Oberste *et al.*, 2004a; Oprisan *et al.*, 2002; Santti *et al.*, 1999; Simmonds & Welch, 2006).

To examine whether this model could yield intraspecies recombinants, the available PV subgenomic replicons were co-transfected with artificially constructed recombinants containing coxsackievirus A (CVA)21 non-structural regions. A series of full-length, intraspecies recombinant clones containing CVA21 with PV1 or PV3 segments were constructed by Dr. Claire Blanchard using basic molecular cloning techniques during her PhD (figure 4.5). The plasmids CVA21BKSabin1P1 and CVA21BKPV3P1 contained a PV1 and PV3 capsid respectively, in a pRiboCVA21 backbone. Plasmid PV3BKCVA21P1 was constructed by introducing a CVA21 encoded capsid into a PV3 backbone vector derived from pT7FLC/REP3 (Barclay *et al.*, 1998). Exchanges of P1 sequence were performed whilst retaining phenylalanine/glycine (in the case of CVA21) and tyrosine/glycine (in the case of PV3 and PV Sabin-1) amino acid sequences at the VP1/2A junctions to maintain correct processing (Claire Blanchard's thesis, University of Reading, 2004). Interestingly, the transfected chimeric T7 RNA transcripts did not cause CPE in HeLa or RD-ICAM cells meaning that they did not produce viable virus. VP1 protein expression was detected in cells transfected with RNA derived from recombinant cDNAs indicating correct translation and processing by heterologous proteases of RNA. Although all constructs translated and processed-proteins correctly (as shown by *in vitro* and *in vivo* transcription/translation assays), they were not capable of producing infectious virus, explained by possible defects in virus maturation, assembly, or possible lack of compatible packaging signals. The constructed recombinants appeared to be replication-competent, even though levels of newly synthesised RNA were not as high as those observed in positive control samples. As a CVA21 replicon and full-length CVA21 CRE mutant were not available during the current project, the aforementioned clones provided replication-competent CVA21 RNA for co-transfection experiments with PV RNA using this model. Restriction sites used to linearise plasmids are outlined in table 2.2. The three constructs were

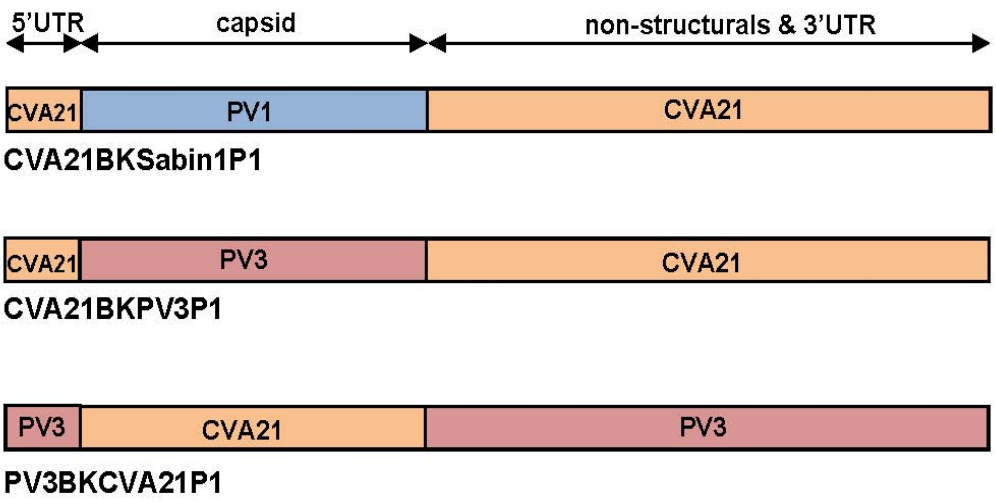


Figure 4.5 Intraspecies C recombinants constructed by Dr. Claire Blanchard. The genomic origin of recombinants constructed previously – the capsid sequence has been removed from the full-length infectious cDNA and replaced with another species C virus capsid sequence.

co-transfected with either pRLucWT (PV1) or pT7Rep3-L (PV3) subgenomic replicons.

Using similar assay conditions to those employed previously, a total of 2 μg of 1:1 RNA partners was co-transfected into L929 or BsrT7 cells. Likewise, individual partner RNA was also transfected into cells to confirm their inability to create viable virus alone. Supernatant was harvested 48 hrs post-transfection and inoculated undiluted onto HeLa cells or RD-ICAM cells in the case of reactions containing CVA21 coding capsid regions. RD-ICAM cells are modified RD cells expressing ICAM-1 receptor required by CVA21 for cell attachment and infection.
Figure 4.6 illustrates results of plaque assays in HeLa cells only. Each RNA partner transfected into cells alone did not produce virus as indicated by clear HeLa monolayers – neither the replicons nor constructs were able to generate virus on their own.

However, virus was recovered from the following co-transfections: pRLucWT + CVA21BKPV3P1, pRLucWT + CVA21BKSabin1P1, pT7Rep3-L + CVA21BKPV3P1, and pT7Rep3-L + CVA21BKSabin1P1 (figure 4.6). No CPE was evident with RD-ICAM cells when supernatant was harvested from cells co-transfected with either replicon together with PV3BKCVA21P1 (data not shown). It was not surprising that virus was not recovered from these co-transfections, as it was shown previously that chimeras generated with CVA (species C) capsids and the replication proteins of PVs were either dead or weakened (Jiang *et al.*, 2007). This phenomenon is discussed in more detail at the end of this chapter.

Virus containing supernatants harvested from pRLucWT + CVA21BKSabin1P1 and pT7Rep3-L + CVA21BKPV3P1 co-transfections were retained for future investigation. Recovered viruses were confirmed as recombinants by sequencing and crossover junctions between RNA partners were characterised and are discussed in detail in chapter five.

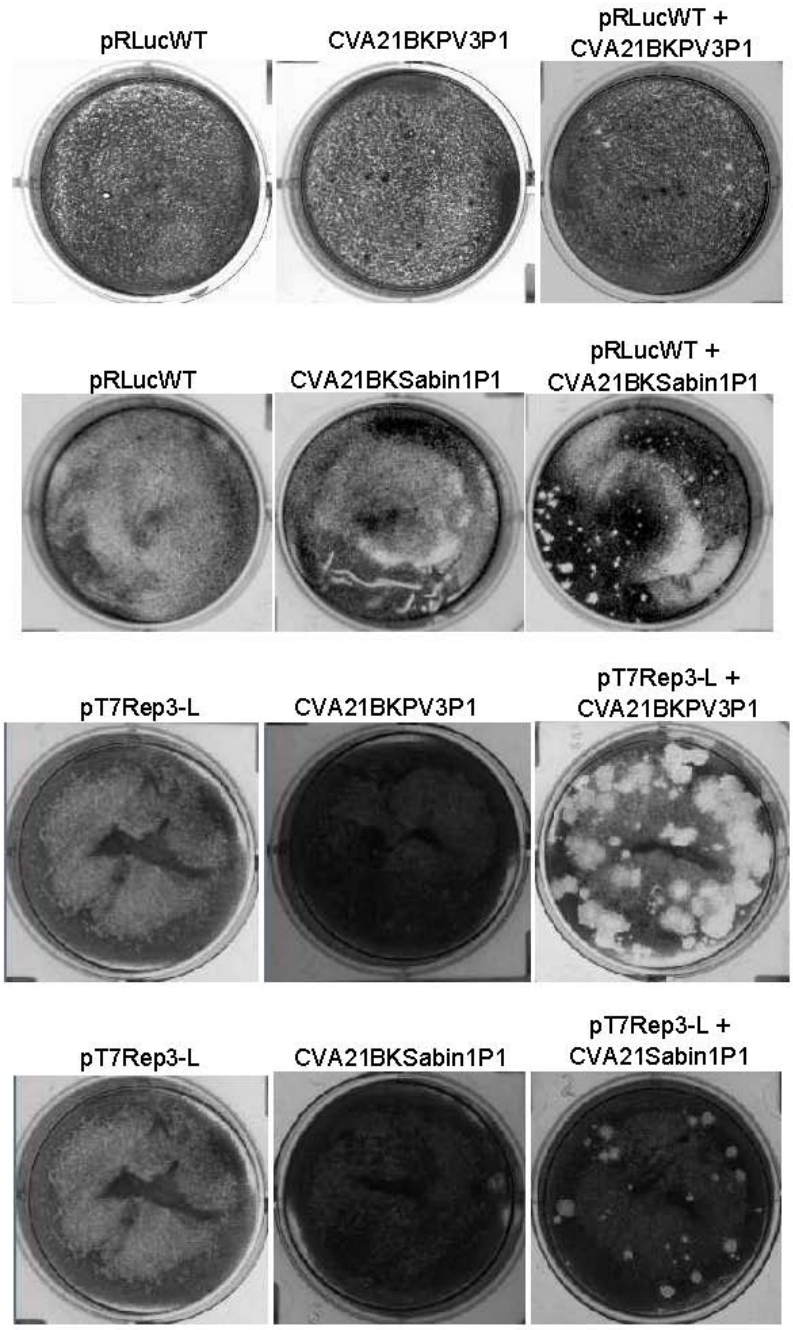


Figure 4.6 Virus recovered from intraspecies RNA co-transfections. These combined PV subgenomic replicons and Dr. Blanchard's constructed intraspecies recombinants (CVA21BKSabin1P1, CVA21BKPV3P1, and PV3BKCVA21P1). Transfection supernatant was plated undiluted onto a HeLa monolayer and overlaid with plaque overlay medium. Cells were stained 48 hrs post-infection. All control wells with one RNA partner only were negative. Plaques are evident in co-transfected wells (subgenomic replicon and full-length intraspecies recombinants).

4.6 Co-transfections with RNA Partners Belonging to Species B

A method has been described that generated infectious virus from various defective PV and RNA partners containing the CVA21 non-structural region by co-transfections. The final intention was to determine whether similar assays could be performed between RNA partners from different enterovirus species to yield interspecies recombinants.

Before this was attempted, it was reasonable to repeat earlier co-transfection reactions with two homologous partners from an enterovirus species other than species C. Therefore, the aim was to validate the system by attempting the recovery of a species B recombinant.

To test this, two echovirus (EV)7 RNA partners were used. Subgenomic replicon pT7E7luc (table 2.1) and full-length pT7E7ΔCRE (construction details in chapter six) were both used for the assay. A luciferase reporter gene replaced the capsid region in pT7E7luc and the CRE was mutated in a full-length E7 infectious clone. Although the replication ability of pT7E7ΔCRE was not tested, it was assumed that this behaved in the same manner as fully characterised pT7FLC/SL3. The CRE structure and location is conserved in both viruses and mutations introduced in pT7E7ΔCRE mimicked those in the PV3 virus equivalent. Transfections with pT7E7ΔCRE RNA alone did not yield infectious virus.

Plasmids pT7E7luc and pT7E7ΔCRE were linearised with *Xho*I and *Bsp*EI respectively and RNA transcribed using T7 polymerase. A total of 2 µg of RNA (0.4 µg pT7E7luc + 1.6 µg pT7E7ΔCRE) was transfected into BsrT7 cells with Lipofectamine 2000 in 12-well plates. Each partner was transfected individually for controls; a practice used for all controls. Transfection supernatant was harvested after 48 hrs and inoculated undiluted onto fresh RD cells in order to perform plaque assays.

Echovirus 7 encoding RNA did not cause CPE in RD monolayers when transfected into cells alone, an expected result (figure 4.7). Virus was detected, however, when both RNA partners were co-transfected into cells simultaneously. Virus plaques were present in RD cell monolayers. To confirm that E7 was generated in this assay,

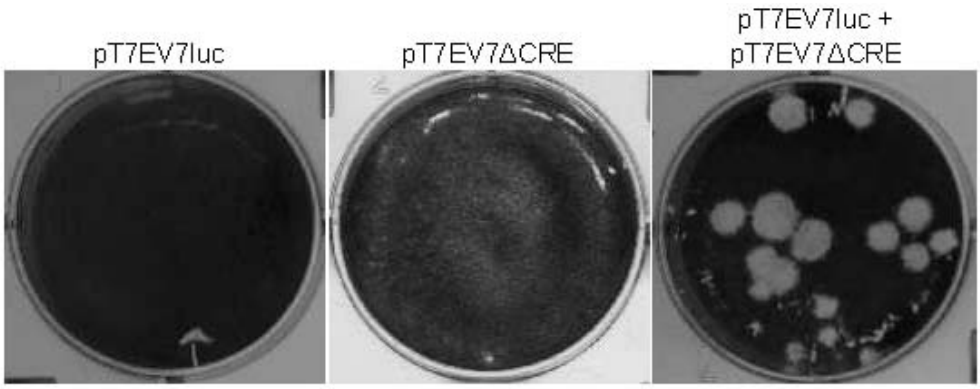


Figure 4.7 Virus recovered from pT7E7luc + pT7E7ΔCRE co-transfections. Crystal violet stained RD cells infected with undiluted supernatant and overlaid with plaque assay overlay medium. Control wells contained individual RNA partners; dual transfection reaction was positive for plaques.

individual viruses were cloned by limiting dilution in 96-well plates layered with RD cells and several virus samples were sequenced. Briefly, virus RNA was extracted from clonal populations of virus dilutions and used to generate cDNA with Superscript II Reverse Transcriptase (Invitrogen). Oligonucleotides E7-6F andE7-5247R (table 2.3) were used to amplify the region between both "lesions" in the partners used by PCR using *Taq* DNA Polymerase (Fermentas). Eight biologically cloned viruses were confirmed as being E7 by sequencing a portion of the PCR product. Due to time constraints, the entire targeted region of recombination was not sequenced; therefore conclusions could not be drawn as to whether crossover sites were precise or imprecise. In conclusion, this method was deemed suitable for generating viable virus from both species B and species C RNA partners.

Co-transfections with interspecies RNA partners

To perform co-transfection assays between RNA partners from different enterovirus species, several plasmids available during the project were used. All RNA partners available, as well as the combinations used in co-transfections, are outlined in table 4.1.

Details of plasmid features and references are summarised in table 2.1. The restriction enzymes used to linearise plasmids prior to RNA transcription are outlined in table 2.2. All co-transfections were performed in either L929 or BsrT7 cells and 1 µg of each partner was introduced into cells by Lipofectamine 2000 in 12-well plates.

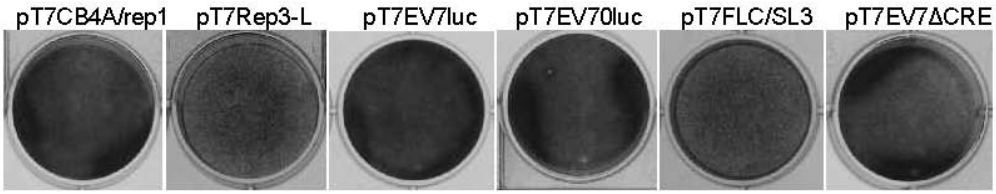
After attempts in which the majority of possible RNA partners were co-transfected into cells, it was only possible to detect virus in supernatant harvested from one co-transfection assay (figure 4.8). A limited number of plaques were observed when cells were co-transfected with pT7CB4A/rep1 and pT7FLC/SL3 RNA partners. This virus population was biologically cloned and several viruses were sequenced across the targeted region to determine whether they contained a combination of coxsackie virus species B (CVB)4 and PV3 sequence. However, only PV3 sequence was detected over the region, indicating contamination during this experiment. No plaques were observed in any other plaque assays.

Table 4.1 Combinations of interspecies co-transfections performed

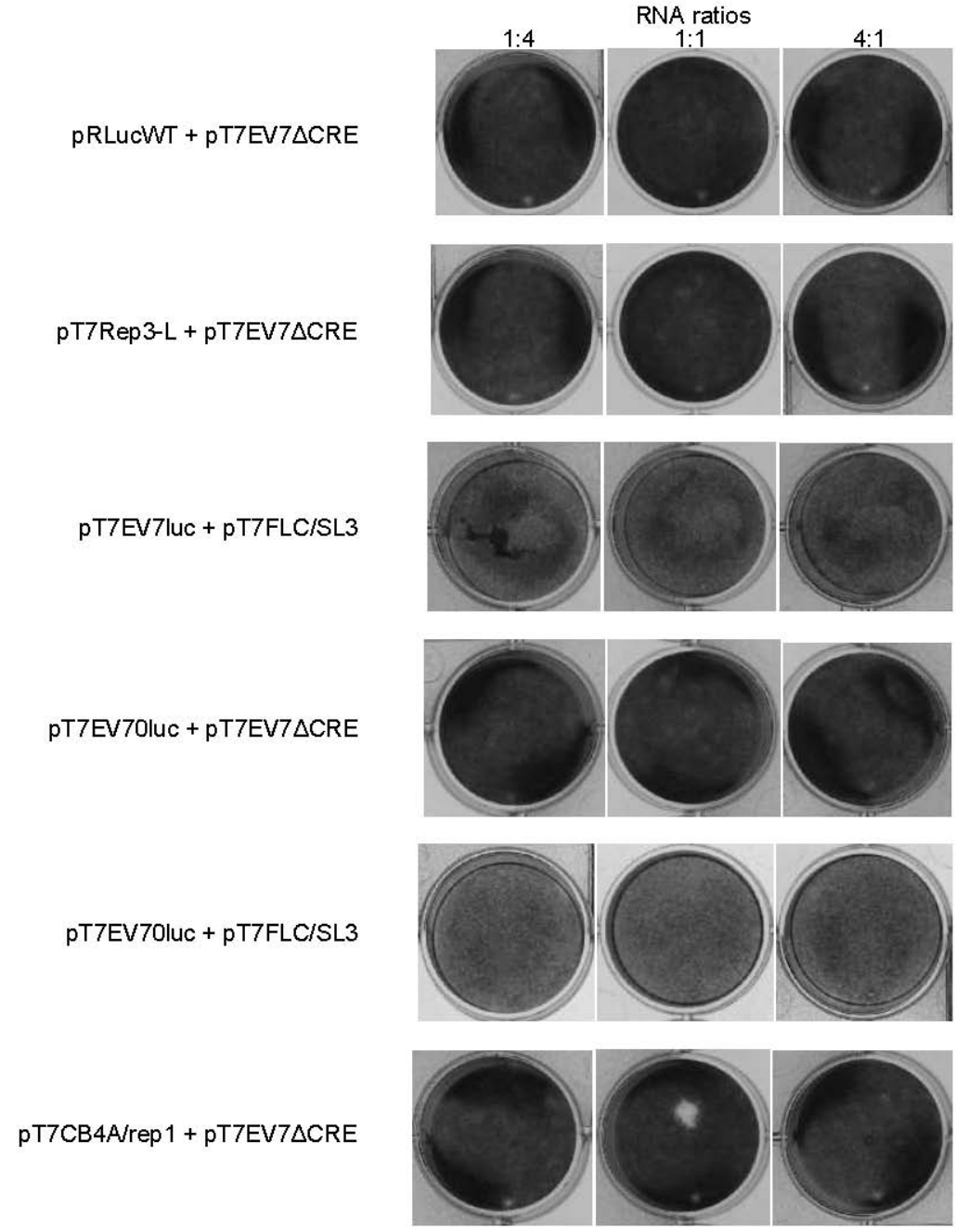
	Full-length clones	pRiboCAV21	pT7LeonFLC/SL3	pT7E7◊CRE	PV3BKCAV21P1	CAV21BKPV3P1	CAV21BKSabin1P1
Replicons							
pRLucWT		0	++	0	0	++	++
pT7Rep3-L		0	++	0	0	++	++
pE7Luc		0	0	++	0	0	0
pEV70Luc		0	0	0	0	0	NT
pRiboCAV21ΔP1		0	0	0	0	0	0
pT7CB4A/rep1		NT	0	0	NT	NT	NT

(0) no virus detected, (++) virus detected, (NT) not tested.

Controls



Cotransfections



Continued over the page.

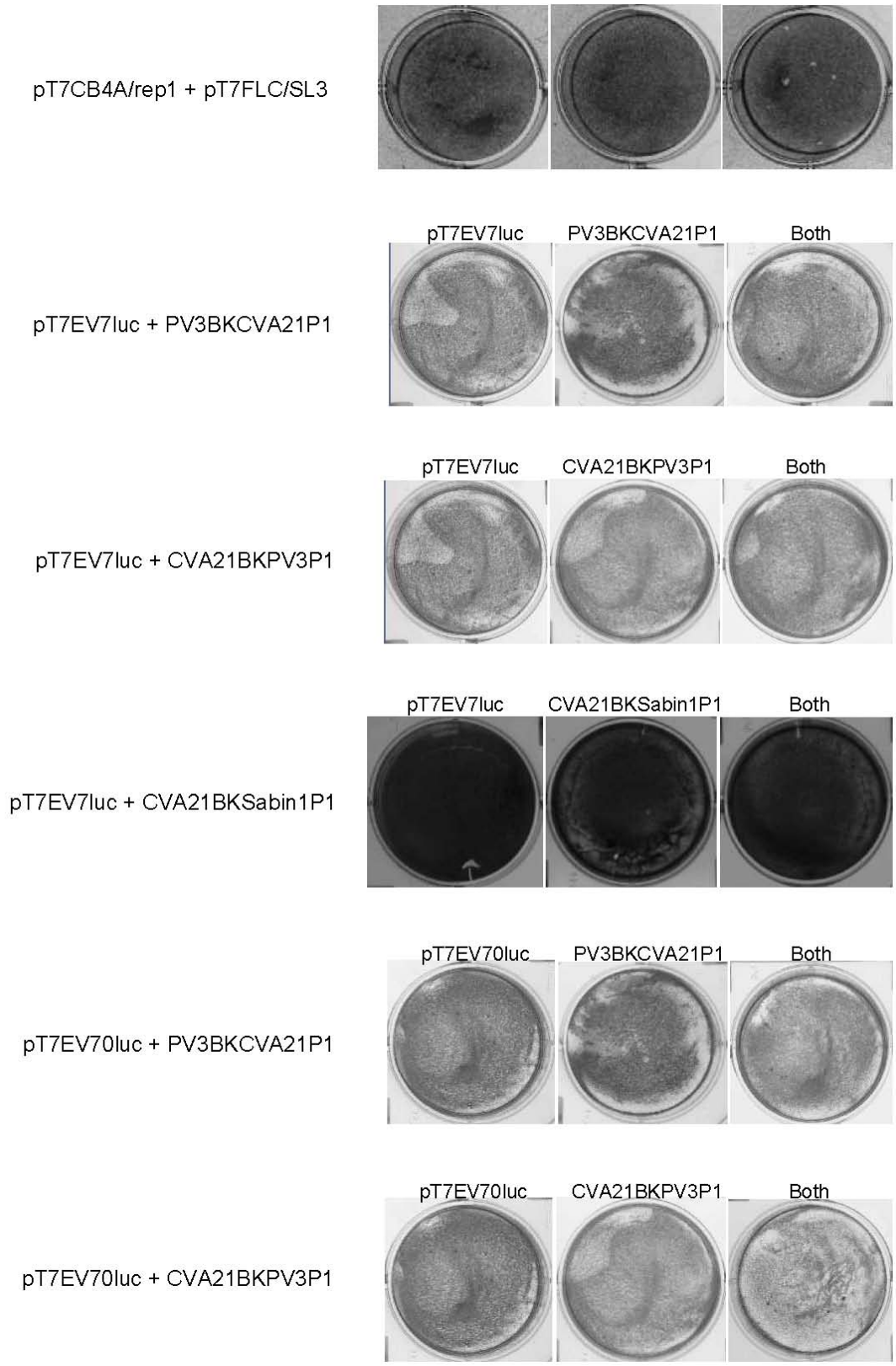


Figure 4.8 Plaque assays of interspecies co-transfection supernatant. Crystal violet stained cells infected with undiluted transfection supernatant and overlaid with plaque assay overlay medium. Control wells contain one RNA partner each, co-transfections are labelled with 'both' or those labelled with the RNA ratio.

Recombination between RNA from different species was not observed using this method. Either interspecies recombination *did* occur but the presence of recombinant virus was below the level of assay detection, recombination occurred but the resulting recombinants were not viable, or it was indeed not biologically possible in the first place. If interserotypic recombination occurred at a frequency 100-fold lower than intraserotypic recombination, then it is not unexpected that recombination between species has an even lower frequency (King, 1988; Tolskaya *et al.*, 1983). If this frequency was reduced 100-fold again, the frequency of interspecies recombination could be as low as 1×10^{-7}, or one virus particle in the typical yield from a standard flask of cells.

4.7 Discussion

An *in vitro* approach for generating recombinants of PV and other enterovirus species in the absence of parental virus background was demonstrated. This robust approach was replication-dependent and captured what is believed to be the initial products of genome recombination. This effectively limited post-recombination amplification of viruses and removed potential additional mutations (i.e. insertions, deletions, further recombination) in chimeric viruses that would lead to misrepresentation of results.

Initially, co-transfections using different ratios of parental RNA molecules was investigated and the majority of RNA ratios resulted in the generation of infectious virus. As mentioned already, the concentration of acceptor template (CRE mutants in this case) had a subtle effect on recombination efficiency: more plaques were evident when acceptor template was in excess. Jarvis and Kirkegaard observed this phenomenon previously and suggested that recombination was more likely to occur when RNA template was concentrated in replication complexes (Jarvis & Kirkegaard, 1992). A 1:1 ratio of replicon to full-length CRE mutant was applied thereafter (1 µg of each partner; molar ratio of ~ 1.16:1) because this ratio of RNA gave consistent results and spared RNA for further experiments. Approximately 2.9×10^{11} molecules of replicon and 2.5×10^{11} molecules of full-length genome were co-transfected into approximately 9.5×10^{5} BsrT7 or L929 cells to yield a range of

viable recombinant virus titres (from 6 to 160 PFU/ml), depending on the sequence identity of the RNA partners.

A striking observation was the inability to recover intraspecies recombinant viruses containing a CVA21 capsid with PV replication proteins. This phenomenon of incompatibility was investigated using CVA20 and PV as models (Jiang *et al.*, 2007; Liu *et al.*, 2010). Both these studies implicated a role of protein $2C^{ATPase}$ in virus morphogenesis and found that a single mutation in either VP3 (capsid protein) or $2C^{ATPase}$ could rescue virus production weakly, and mutations in both regions could restore virus to the wild-type phenotype. Therefore, incompatibilities between CVA21 VP3 and PV1/3 $2C^{ATPase}$ could explain the lack of these chimeric viruses identified during these experiments.

This model allowed expansion and incorporation of interspecies partners. A series of co-transfection experiments were performed incorporating interspecies partners and no viable virus was detected as a consequence. It seems that interspecies recombinants, with a junction site in the centre of the genome, have either remained undiscovered, or are not viable due to incompatibilities in RNA-RNA, RNA-protein, or protein-protein interactions during replication. This will be discussed further in chapter seven.

The primary restriction to the use of this method is the availability of infectious cDNAs for the parental genomes. Virus partners can be mutated to inhibit production of live virus while maintaining partial or complete replicating ability (qualities required for RNA partners in this method). Moreover, changes to recombination types as well as frequencies can be examined further by introducing mutations into the RdRp of the RNA partners to alter replication fidelity, a concept that has been explored previously using PV and brome mosaic virus (Arnold *et al.*, 2005; Arnold & Cameron, 1999; Figlerowicz *et al.*, 1997).

The next step would be to sequence a large population of recombinants to characterise crossover sites. Several questions remain including the suggested causes for RdRp pausing during elongation, (which may encourage template switching) and whether precise or imprecise template switching is sequence dependent. These

questions could be addressed by producing RNA partners with specific mutations introduced to the RdRp sequence and observing the frequency of recombinant changes when incremental changes to the viral sequence between the luciferase and CRE junctions was introduced. The possibility that RNA partner heterogenicity resulted in reduced recombination frequencies, as observed using PV1/PV1 and PV1/PV2 recombinants in the approach described by Kirkegaard, may suggest a hypothesis for why fewer intraspecies recombinants are observed in nature (Kirkegaard & Baltimore, 1986). However, this experiment should be initially expanded using an appropriate statistical method to compare the observed and expected frequencies (i.e. Chi square if normality is assumed) using a larger number of recombinants. In this instance, the null hypothesis would be that the observed frequency of recombinants in the PV3/PV3 experiments is equal to the observed frequency of recombinants in the PV3/PV1 experiments.

CHAPTER FIVE: Characterisation of Recombinants

5.1 Introduction

Chapter four described and demonstrated a new system for generating recombinant species C enteroviruses in the absence of parental viruses in the population. The results revealed that enteroviruses could be generated by RNA recombination between two genomic virus RNA molecules that were at least partially replication-competent but unable to form virus on their own. The replication-dependent method yielded intraserotypic, interserotypic, and intraspecies C recombinant viruses, the identity of which were confirmed by sequencing.

Many papers describe the occurrence of recombination and how such viruses shape enterovirus evolution, but few of these examine the underlying mechanisms of recombination in detail with the exploratory analysis of many crossover points (Pilipenko *et al.*, 1995; Tolskaya *et al.*, 1987). It is difficult to capture early recombination events when viruses are subjected to selection pressures involving significant passage or competition. In contrast, this experimental model allowed for the identification of what are believed to be early recombinants.

The next step was to analyse junctions from a population of recombinants generated using this method. An analysis of the junctions would allow speculation on the mechanism by which these early recombinants were generated; for example, following RdRp pausing on the donor template strand. This could be investigated by looking at the following possible correlations, including: U-rich or AU-rich stretches, local secondary structures in RNA sequence such as hairpins, heteroduplexes, and mis-incorporations. Recombination could occur with or without the presence of such RNA features and the aim was to determine the molecular mechanisms that influenced these events.

Aims

The aim was to perform a large scale experiment by repeating the interserotypic recombination experiment outlined in chapter four with the intention of performing an exploratory analysis of crossover sites for an extensive population of chimeric

PV1 and PV3 recombinants (100 biological clones in total). This number was considered sufficient to allow a representative number of different recombinants to be identified with the intention of determining if there were specific patterns to recombination junctions observed. Intraspecies C recombinants first generated and discussed in chapter four were also characterised to determine the crossover positions between input RNA partners. By carrying out detailed RNA sequence analysis, the purpose was to determine whether the donor or acceptor virus sequence influenced RdRp template switching and to characterise the fitness of chimeric viruses before and after further selection. The final aim was to establish whether recombinants generated by this method reflected those found naturally and whether this method provided an accurate 'snap shot' of RNA recombination in mammalian cells.

5.2 PV3/PV1 Recombinants

Generation of PV3/PV1 recombinant viruses

Briefly, two virus RNA molecules containing "lesions" rendering them unable to generate viable virus on their own were co-transfected into permissive cells. Poliovirus serotype-1 subgenomic replicon pRLucWT and full-length PV3 CRE mutant pT7FLC/SL3 plasmids were linearised with *Apa*I and *Sal*I respectively, and transcribed into RNA using T7 polymerase (Fermentas). A 25 cm^2 flask of L929 mouse cells was prepared the day prior and co-transfected with 2600 ng of each RNA partner using Lipofectamine 2000. Almost equimolar ratios of both template were used (subgenomic replicon: full-length CRE mutant) and calculated to be approximately 1.16:1. The transfection supernatant was harvested after 48 hrs and 500 µl of this supernatant was used undiluted to inoculate a single well of a 6-well plate containing HeLa cells in order to detect virus by plaque assay. Plates were stained after 48 hrs with crystal violet and recombinant virus identified by the presence of plaques in the HeLa monolayers. Although controls were not performed in this instance, the transfection of single RNA partners in L929 cells had not resulted in the recovery of infectious virus on several previous occasions.

Stained HeLa cell monolayers demonstrated the presence of plaques, interpreted as viable virus (figure 5.1A). Supernatant containing virus at an initial titre of 86 PFU/ml was subsequently biologically cloned by limiting dilution in 96-well

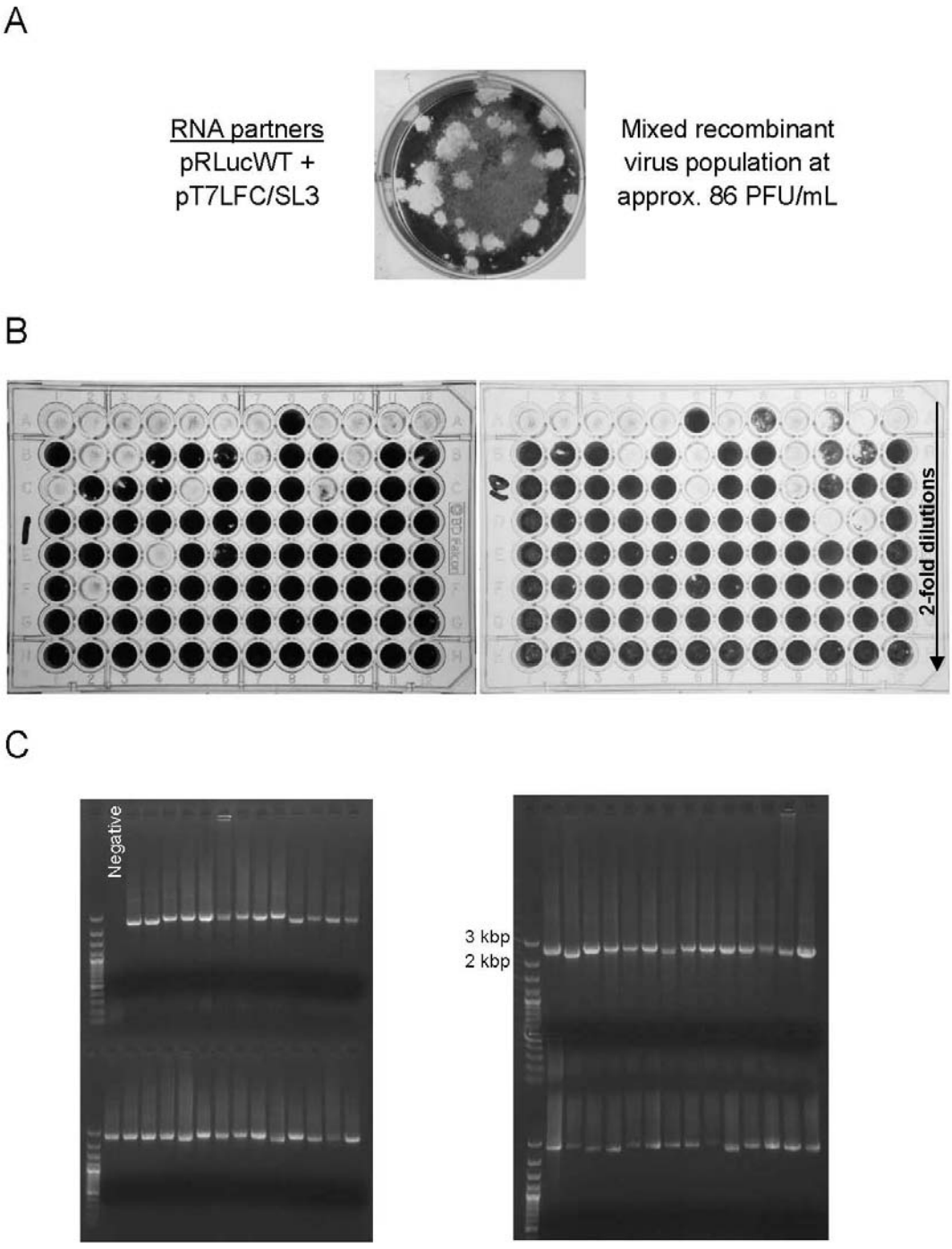


Figure 5.1 Identification and biological cloning of 100 PV3/PV1 recombinant viruses. These were recovered from co-transfections in L929 mouse cells. (A) Co-transfection supernatant was plated undiluted onto a fresh HeLa cell monolayer and overlaid with plaque overlay medium. Cells were stained 48 hrs post-infection with crystal violet. Viable virus is evident by the presence of plaques. (B) Biologically cloning recombinant viruses in transfection supernatant by limiting dilution. (C) Gel electrophoresis of RT-PCR product of recombination region. Ethidium bromide stained agarose gels showing a heterogeneous mixture of amplified virus sequences from cloned recombinants.

plates (see Materials and Methods). Briefly, 40 µl of transfection supernatant passage 0 (p0) was mixed with 70 µl serum-free DMEM/PS in each well of row A across 96-well plates. This seeded each well with approximately 3.44 PFU of recombinant virus before doubling dilutions were performed down the column with further medium so that 55 µl virus mix remained in each well. Fifty microlitres of each virus dilution was transferred to fresh 96-well plates containing monolayers of confluent HeLa cells. Post-infection, virus supernatant (now p1) was removed and stored and cells were stained for CPE using crystal violet (figure 5.1B). Cytopathic effect was evident by clear wells after staining, indicating that virus had infected cells and destroyed the HeLa monolayer. The highest dilution of supernatant to cause CPE was retained for further analysis.

Genetic characteristics of PV3/PV1 recombinants

The target was to analyse the crossover junctions of 100 PV3/PV1 biological clones (generated as described in chapter four, figure 4.2) in order to provide an exploratory investigation of recovered virus. The aim was to observe a range of variation in progeny, and considering the workload involved, this was a manageable number of viruses to study. Thus, a statistical power calculation or subsequent statistical analysis was not applied. The targeted region of recombination was sequenced for each biologically cloned virus. Virus RNA was extracted from p1 supernatant and cDNA was synthesised using Superscript II Reverse Transcriptase (Invitrogen) with an oligo dT oligonucleotide. The PCR was used to amplify the targeted recombination region, from the 3' end of the VP1 region to the start of the CRE in the 2C region. Oligonucleotides PV3-2995F and PV1-5200R (table 2.3), were used to amplify PV3/PV1 recombinant cDNA using KOD XL DNA Polymerase (Novagen) as they were specific for each parental 'donor' part of the genome. A small portion of RT-PCR product was loaded into a 1% agarose gel for electrophoresis (figure 5.1C). Each band in the agarose gel represented the crossover region from a biologically cloned virus. The first observation was that the single amplified product was not of a uniform size, and that the vast majority were larger than the expected target of 2205 bps. The expected target was calculated as the distance between both oligonucleotide pairs when PV1 and PV3 sequences were aligned together. This suggested that recombination was imprecise in nature with extra sequence inserted at the junction site. The RT-PCR product was column

purified, sequenced and analysed to identify junction sites between the PV3 and PV1 parental sequences. Identifying sequences belonging to either PV1 or PV3 manually without the aid of recombination software was straightforward. The sequence length between the 3' end of the capsid coding region and start of the CRE (crossover target) was the same in both viruses (1058 nt) and shared 77% nt sequence identity.

In total, 110 virus clones were sequenced. Of these, 10 were removed from further analysis as they did not contain a clonal population of virus as was evident from the sequence chromatogram (figure 5.2). This occurred despite such samples apparently generating a single sized PCR fragment. Low levels of contaminating virus may have been below levels of detection on an agarose gel.

The recombinant viruses consisted of a PV3 structural and PV1 non-structural region, henceforth referred to as PV3/PV1. Therefore, the PV1 virus RdRp was able to switch templates during replication and combine both RNA partners to form viable virus with a PV3 capsid region replacing the luciferase gene (refer to chapter four, figure 4.4B, for a depiction of this process). Sequence analysis of 100 recombinants identified 16 distinct crossover junctions within the ~1 kb separating the "lesions" used for selection. Table 5.1 lists the distinct groupings of crossover junctions in the recombinants recovered including the approximate genome positions at which point template switching occurred and the number of extra nucleotides incorporated at the crossover site where applicable. Only one group from 16 contained a precise junction site, and this was situated in the 5' terminus of the 2C region [PV3(1-4277)PV1(4278-A tail)]. All remaining recombination events were imprecise and incorporated extra sequence at the junction site. Figure 5.3 shows a sequence chromatogram for one virus clone (105B); PV3 sequence ended at position 306 in the chromatogram sequence, the following 137 bps belonged to the 3' terminus of the luciferase gene, and the remainder of the sequence was PV1 from position 444. This equated to a recombinant virus PV3(1-3491)Luc(137 nt)PV1(3380-A tail) with an extra 249 nt sequence incorporated at the junction site. Correct numbering and crossover positions were determined from a Clustal alignment (Saitou & Nei, 1987) containing full-length PV1 strain Mahoney and PV3 strain Leon (GenBank accession numbers V01149 and X00925 respectively). In every instance, the extra nucleotides inserted at the junction site were divisible

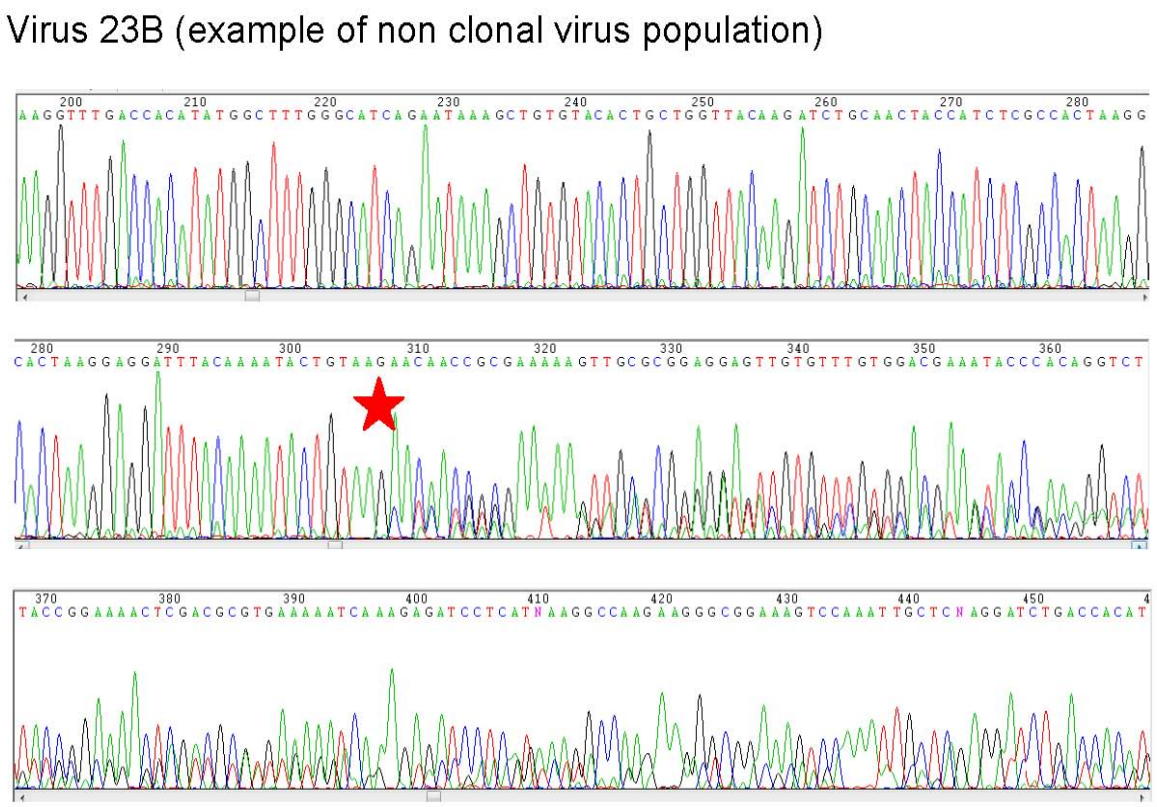


Figure 5.2 An example of a non-clonal recombinant virus mixture. Virus RNA was extracted from p1 supernatant after biological cloning by limiting dilution in 96-well plates. The region of targeted recombination from 2A to start of the CRE in 2C was amplified by RT-PCR and sequenced. The chromatogram shows that dual peaks are present in sequence from position 308 (red star) indicating the presence of more than one template type in the reaction.

Table 5.1 Characterisation of PV3/PV1 recombinants by crossover region from single co-transfection

Name*	Clone Name	Position of crossover junctions PV3	PV1	Luciferase sequence[#]	Total extra nucleotides	No. clones[~]
PV3(1-3431)luc(134)PV1(3380-A tail)	**11A**	3431	3380	134	186	12
PV3(1-3449)luc(107)PV1(3380-A tail)	**1C**	3449	3380	107	177	45
PV3(1-3454)luc(174)PV1(3380-A tail)	**36B**	3454	3380	174	249	2
PV3(1-3465)luc(205)PV1(3380-A tail)	**34D**	3465	3380	205	291	1
PV3(1-3491)luc(137)PV1(3380-A tail)	**105B**	3491	3380	137	249	2
PV3(1-3505)luc(168)PV1(3380-A tail)	**53A**	3505	3380	168	294	2
PV3(1-3612)PV1/3(447-462)PV1(3386-A tail)	**25A**	3612	3386	PV1/3 447-462	243	8
PV3(1-3621)luc(79)PV1(3380-A tail)	**67B**	3621	3380	79	321	1
PV3(1-3669)luc(1)PV1(3380-A tail)	**16B**	3669	3380	1	12	1
PV3(1-3850)PV1(3848-A tail)	**9C**	3850	3848	-	3	1
PV3(1-3851)PV1(3813-A tail)	**33B**	3851	3813	-	39	1
PV3(1-3868)PV1(3797-A tail)	**83B**	3868	3797	-	72	6
PV3(1-3894)PV1(3742-A tail)	**51G**	3894	3742	-	153	1
PV3(1-3901)PV1(3782-A tail)	**58A**	3901	3782	-	120	2
PV3(1-3917)PV1(3693-A tail)	**14C**	3917	3693	-	225	1
PV3(1-4277)PV1(4278-A tail)	**4E**	4277	4278	-	0	14

* Numbering according to Clustal alignment of poliovirus type 1 Mahoney and poliovirus type 3 Leon strains (GenBank accession numbers V01149 and X00925 respectively)

[#] No. of nts from the 3' end luciferase fused into PV1 sequence at position 3380 of genome

[~] No. of clones per crossover in single co-transfection

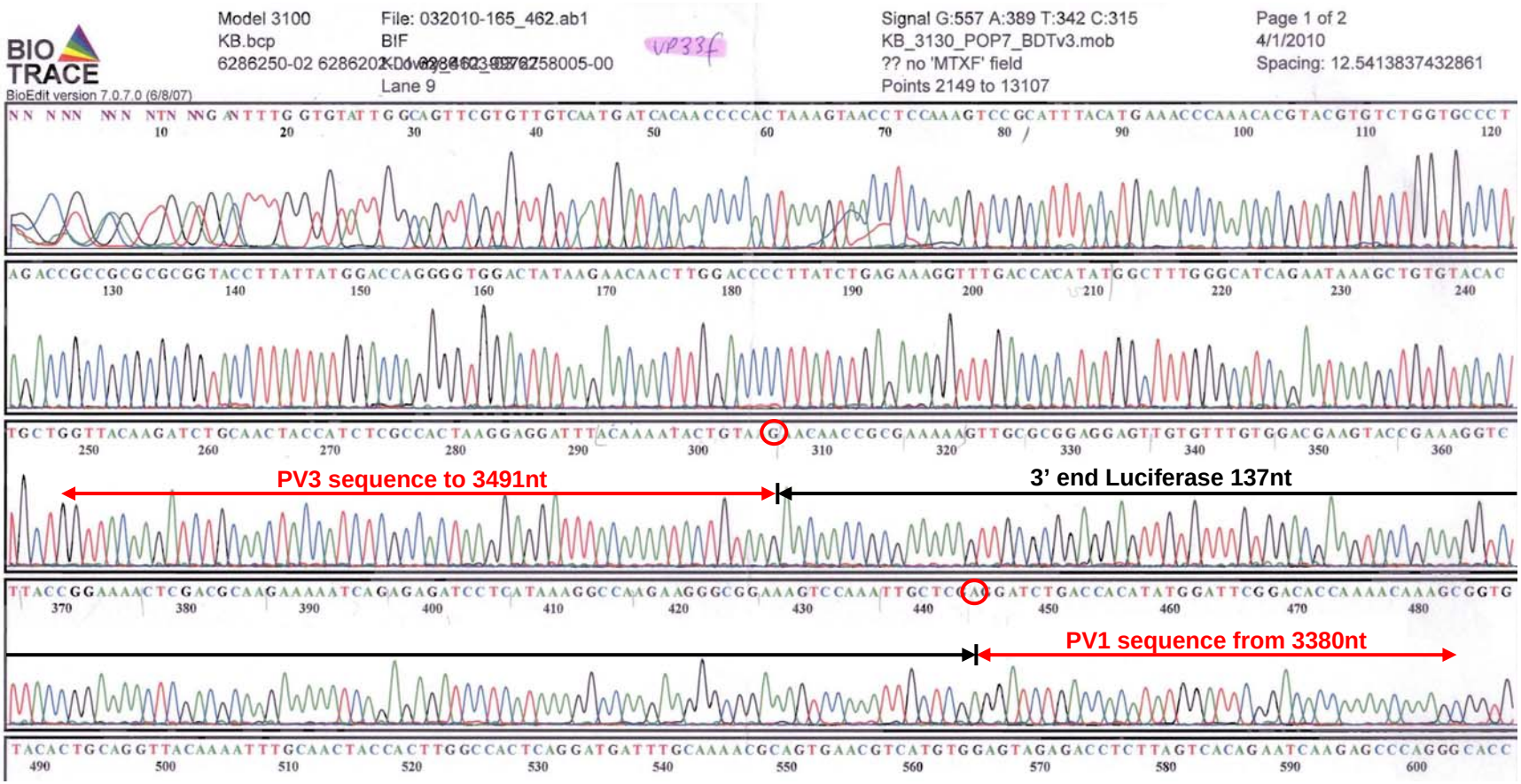


Figure 5.3 Chromatogram of sequenced recombinant virus 105B. Red arrows indicate PV sequence, and black arrow highlights luciferase sequence incorporated at the crossover junction. Luciferase sequence is inserted between nts 306 and 444 according to chromatogram base pair numbering. The chromatogram demonstrates that a clonal population of virus has been sequenced.

by three and therefore maintained an in-frame translation product for all viruses, as would be expected.

It was not appropriate to analyse the frequency by which the crossovers occurred in the entire set of 100 clones, as distinct recombinants may have arisen either from a single event in a single cell that replicated abundantly, or the same event may have occurred in several cells to produce the same frequencies.

Figure 5.4 provides a diagrammatic view of all 16 crossover junctions identified in the PV3/PV1 recombinants. The recombinants grouped into those containing extra PV sequence, those containing extra PV and luciferase sequence, and the single virus with a precise junction in 2C. In every recombinant, the coding region for 2A and 2B was derived in its entirety from one or the other parental genomes, e.g. none of the recombinants relied exclusively on a chimeric 2A or 2B coding region. Recombinant 4E, of which 14 clones were recovered from 100 sequences, was the only virus with a chimeric protein (2C) derived from the crossover event. The presence of a precise junction was an exception rather than the rule, indicating that this result could be a rare event during RNA recombination early in cell infection. It also remains to be determined whether this virus was advantaged as a result, as several virus clones were recovered in the set.

Five recombinants (11A, 34D, 53A, 67B, and 4E) in the set contained a single nucleotide at the recombination junction which was identical for PV1 and PV3 (figure 5.5). In other words, it was not possible to define whether the exact crossover occurred before or after the nucleotide *or* which parental genome the nucleotide was derived from. It was not clear whether one shared nucleotide provided evidence for a dependency of RNA sequences in determining regions of template switching, as this was found in 31% of junctions. If recombination occurred randomly, 25% of junctions would be predicted to show single shared nucleotides (Gallei *et al.*, 2004), although this would rely on sequence that consisted of equal proportions of each nucleotide. It was later determined that the PV regions were slightly A-rich in composition. The potential role of donor and template sequences was investigated and is discussed in detail later in this chapter.

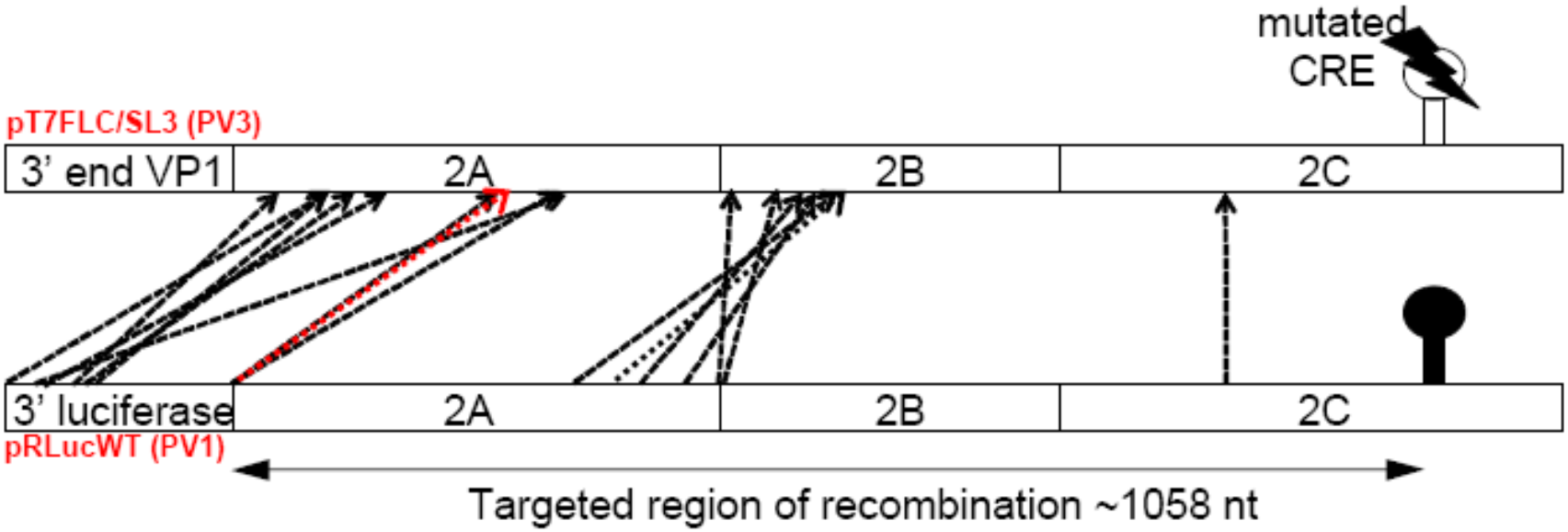
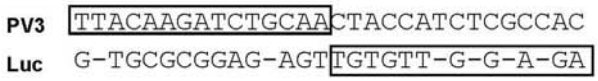


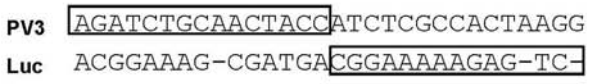
Figure 5.4 Sixteen distinct crossover junctions identified from 100 cloned PV3/PV1 recombinants. The targeted region of recombination is shown in both pRLucWT and pT7FLC/SL3 RNA partners. PV1 RdRp initiated RNA synthesis from the 3' end of the genome, and switched to the PV3 acceptor genome as indicated by the arrows. The RdRp continued to copy to the extreme 5' end of the PV3 genome. The red arrow indicates clone PV3(1-3612)PV1/3(447-462)PV1(3386-A tail), which contained a 15 nt insertion derived from the 5' NCR.

PV3(1-3431)luc(134)PV1(3380-A tail) 11A

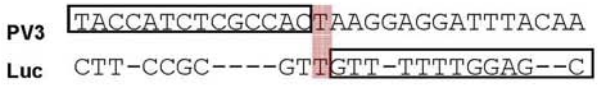
PV3(1-3449)luc(107)PV1(3380-A tail) 1C



PV3(1-3454)luc(174)PV1(3380-A tail) 36B



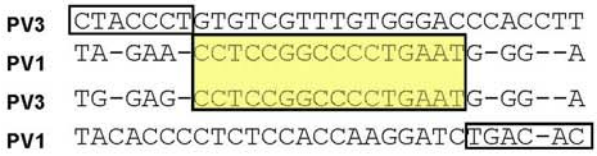
PV3(1-3465)luc(205)PV1(3380-A tail) 34D



PV3(1-3491)luc(137)PV1(3380-A tail) 105B

PV3(1-3505)luc(168)PV1(3380-A tail) 53A

PV3(1-3612)PV1/3(447-462)PV1(3386-A tail) 25A



PV3(1-3621)luc(79)PV1(3380-A tail) 67B

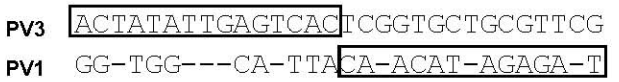
Continued over page.

PV3(1-3669)luc(1)PV1(3380-A tail) 16B

PV3(1-3850)PV1(3848-A tail) 9C

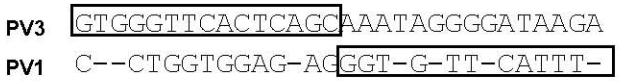
PV3(1-3851)PV1(3813-A tail) 33B

PV3(1-3868)PV1(3797-A tail) 83B



PV3(1-3894)PV1(3742-A tail) 51G

PV3(1-3901)PV1(3782-A tail) 58A



PV3(1-3917)PV1(3693-A tail) 14C

PV3(1-4277)PV1(4278-A tail) 4E

Figure 5.5 Sites of crossover junctions in recovered PV3/PV1 recombinants. Boxed regions indicate recombinant sequence; red boxes show nt identity at crossover junctions. Yellow box indicates inserted sequence from IRES.

One recombinant recovered from the set contained a 15 nt sequence insert at the crossover junction (25A). Interestingly, the inserted sequence originated from the 5' NCR, specifically the IRES, and was conserved between PV1 and PV3. The reason for this insertion remained unclear, and it was difficult to speculate on spatial locations of both regions of the genome during replication, particularly when the genome would have been unwound by the virus replicase during the recombination event. The inserted sequence originated from between stem loops IV and V of the IRES, with half the sequence base paired in stem loop V and the other half single stranded (figure 5.6) (Kafasla *et al.*, 2010). As this figure provided a secondary structure for IRES folding only, it was not possible to predict the location of the sequence origin at a tertiary structure level.

In another independent co-transfection experiment with pRLucWT and pT7FLC/SL3 RNA partners in L929 cells, four distinct groups with unique crossover sites were recovered from 31 biologically cloned viruses (table 5.2). The infectious recombinant viruses demonstrated different crossover junctions to those previously identified. Once again, the crossover sites were either precise, or imprecise with inserted sequence at the junction (figure 5.7). The precise recombination junction sites occurred at the 3' terminus of 2A and the very start of the CRE in 2C and were located in conserved amino acid regions. Recombination at conserved sites probably accounted for the fact that chimeric proteins were obtained in this experiment, and it seems that these recombinants adhere to the same general rule observed with the previous recombinant set, in which a complete 2A and 2B region was obtained from one parent during template switching. The recombinant 2A protein was not chimeric by amino acid sequence. It also appeared that chimeric 2C proteins are tolerated by recombinants. From two independent experiments, it was becoming apparent that there were favoured 'regions' for template switching, but not necessarily specific sites from which the RdRp exchanged templates.

To conclude, the unexpected nature of recombinants recovered from this assay, including 'extraneous' sequence insertions, warranted further investigation to determine whether these were rarely observed 'early' events of virus RNA recombination or simply the product of this artificial system. Examples of such

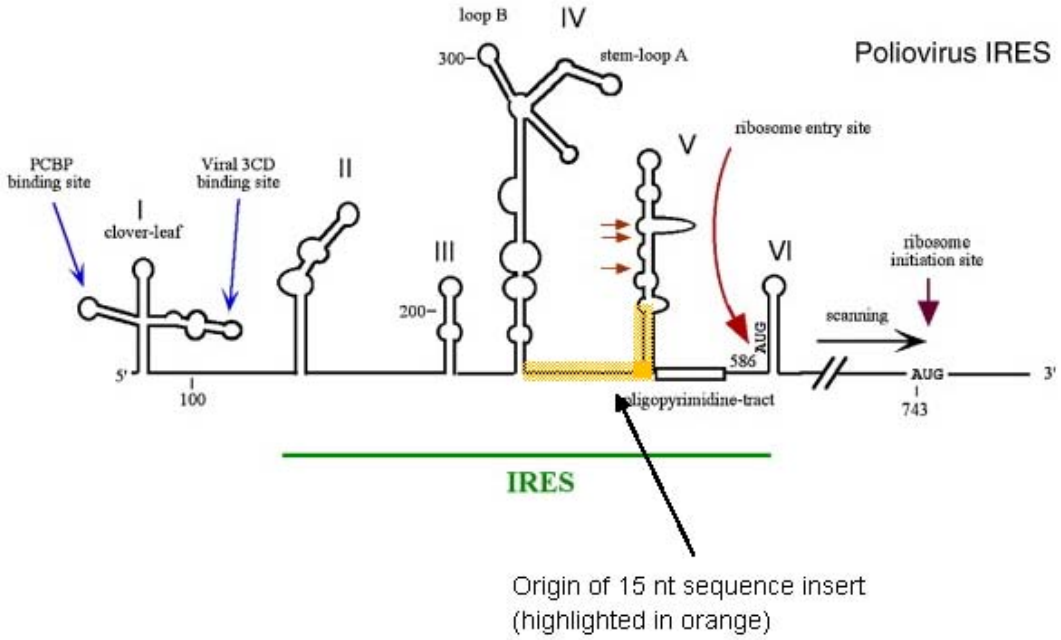


Figure 5.6 Origin of the inserted 15 nt sequence in recombinant 25A at the crossover site. This fragment of sequence is conserved between PV1 and PV3. (Source: Balvay *et al.*, 2009).

Table 5.2 Characterisation of PV3/PV1 recombinants by crossover region from an independent co-transfection in L929 cells.

Name*	Clone Name	Position of crossover junctions PV3	PV1	Luciferase sequence[#]	Total extra nucleotides	No. clones[~]
PV3(1-3553)luc(120)PV1(3380-A tail)	**28C**	3553	3380	120	294	3
PV3(1-3623)PV1(3383-A tail)	**32A**	3623	3383	-	240	1
PV3(1-3824)PV1(3825-A tail)	**35C**	3824	3825	-	0	14
PV3(1-4473)PV1(4474-A tail)	**44B**	4473	4474	-	0	13

* Numbering according to Clustal alignment of PV serotype 1 Mahoney and PV serotype 3 Leon strains (GenBank accession numbers V01149 and X00925 respectively)

[#] No. of nts from the 3' end luciferase fused into PV1 sequence at position 3380 of genome

[~] No. of clones per crossover in single co-transfection

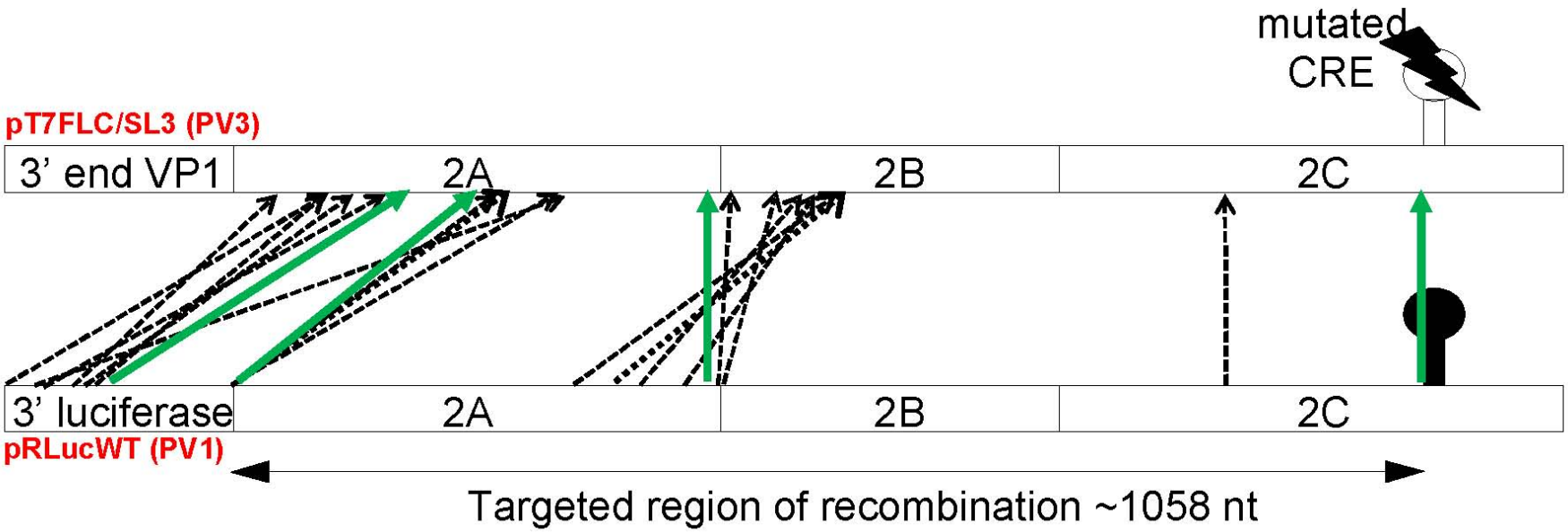


Figure 5.7 Four distinct crossovers identified from 31 cloned PV3/PV1 recombinants in a repeated, independent experiment. The targeted region of recombination is shown in both pRLucWT and pT7FLC/SL3 RNA partners. Poliovirus serotype-1 RdRp initiated RNA synthesis from the 3' end of the genome, and switched to the PV3 acceptor genome as indicated by the arrows. The RdRp continued to copy to the extreme 5' end of the PV3 genome. Green arrows represent the four distinct crossover sites from this second experiment.

imprecise recombination events are not reported for natural isolates. The nature and size of inserts will be discussed later in this chapter.

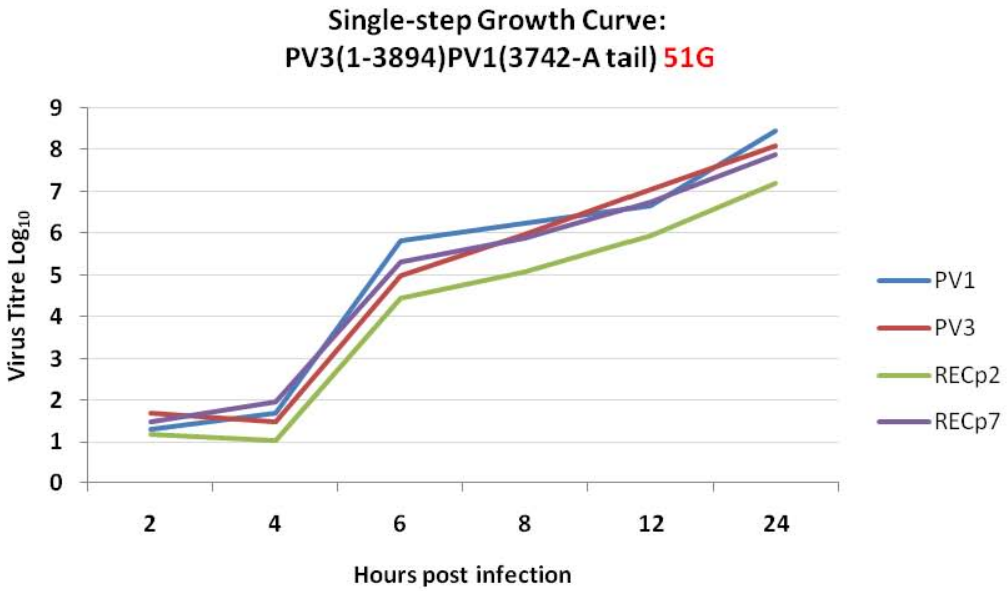
Growth kinetics in HeLa cells (single-step growth curves)

The system allowed the recovery of viruses with minimum subsequent selection, so it may have been sensitive enough to capture initial events during the virus replication cycle. It was necessary to determine whether the viruses demonstrated similar growth characteristics compared to the parent viruses from which they were derived. Two PV3/PV1 viruses were selected for further analysis. Both were selected because of the substantial stretch of inserted sequence at the junction site and also because template switching occurred in either virus region $2A^{pro}$ or 2B. The first, PV3(1-3894)PV1(3742-A tail), henceforth referred to as 51G, acquired a 153 nt insert consisting of PV sequence at the crossover junction. The second, PV3(1-3454)Luc(174)PV1(3380-A tail), referred to as 36B hereafter, gained an extra 249 nts at the junction site, with 174 nt derived from the luciferase reporter gene in pRLucWT. Both viruses were subjected to five further passages (p3 to 7) through susceptible HeLa cells to determine whether changes would occur to these viruses over the short term that would affect the nature of the additional sequences retained within the genome. Briefly, HeLa cells were infected with the selected viruses (p2) at an MOI of 10 for 30 mins at 37°C. Virus was removed and cells were washed twice with sterile PBS to remove excess virus. New serum-free DMEM was added to cells and plates were incubated at 37°C. Supernatant was removed after 24 hrs and a quarter of this supernatant was used undiluted for each subsequent inoculum onto fresh HeLa cells.

Single-step growth curves were performed in HeLa cells (figure 5.8). Cells were infected at an MOI of 10 with PV1, PV3 or the selected recombinants after two and seven passages (p2 to 7) through HeLa cells. Virus supernatant was harvested at various time points up to 24 hrs post-infection and titrated by plaque assay.

Unpassaged 51G (p2) demonstrated reduced fitness (one $\log_{10}$ from 6 to 24 hrs) compared with wild-type viruses. The recombinant shared the same replication kinetics with both parent viruses after five further passages through HeLa cells,

A



B

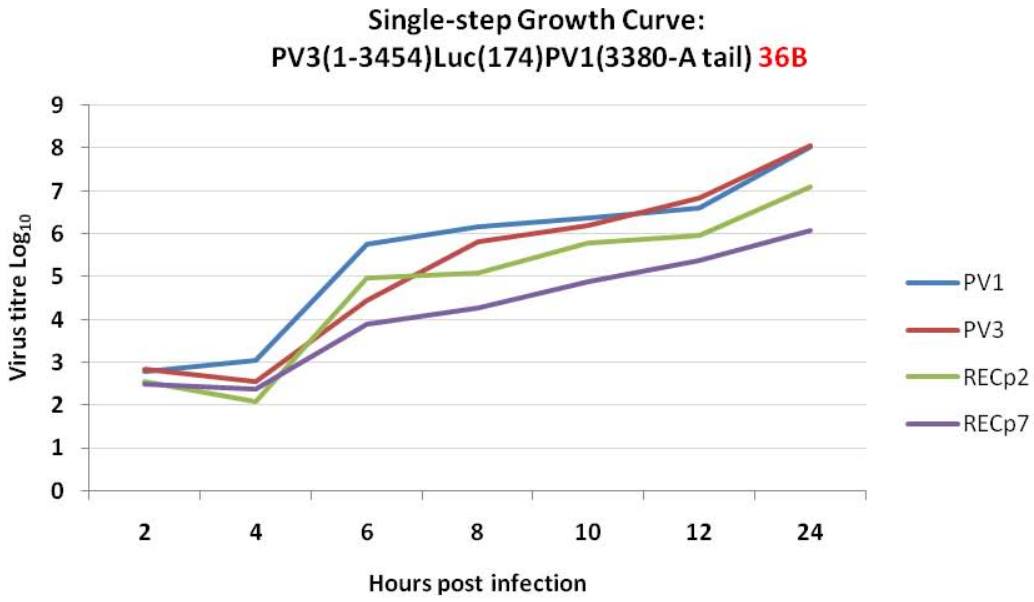


Figure 5.8 Single-step growth curves. These were performed by comparing PV parents with two PV3/PV1 recombinants generated from co-transfections with pT7FLC/SL3 and pRLucWT RNA partners. (A) Growth curve comparing PV3(1-3894)PV1(3742-A tail) to PV1 and PV3 parents. (B) Growth curve comparing PV3(1-3454)Luc(174)PV1(3380) to both parents from which it was derived.

indicating that this growth characteristic of the passaged virus was indistinguishable from both wild-type parents (from which the original virus RNA was derived). This was not the case for recombinant 36B. The serially passaged virus seemed to decline in fitness over the short term, indicating that either the extra inserted sequence was a hindrance, or a mixed population of viruses was forming which confounded virus replication. These points were addressed by performing competition assays with parent viruses as well as characterising recombinant viruses after continuous passage.

Plaque phenotypes were compared between parent viruses (PV1 and PV3) and PV3/PV1 recombinants at passages 2 and 7 (figure 5.9). The recombinants displayed small and irregular plaques at p2, whereas the plaque phenotypes started to reflect the same size and regularity of parents PV1 and PV3 following subsequent rounds of replication in HeLa cells (p7). Plaque size and morphology were altered following passage in cell culture, indicating that these recombinant viruses were changing/adapting through subsequent rounds of replication. These findings suggest that insertions in the centre of the virus genome modulate virus plaque size phenotype, and not the chimeric nature of the recombinant.

In summary, there was a small yet noticeable difference between recombinant viruses before and after serial passage through cells. This suggested that passaged virus was changing as a result of further rounds of replication, and it was essential to investigate this further.

Competition assays between recombinants and wild-type parents

As briefly mentioned above, competition assays were also performed as a way of determining if 'atypical' recombinants would survive in a mixed infection if both parents were also present. Were the recombinant viruses fitter than parents as a result of the exchange of genetic material? This perhaps more accurately reflects the natural conditions under which recombinants arise.

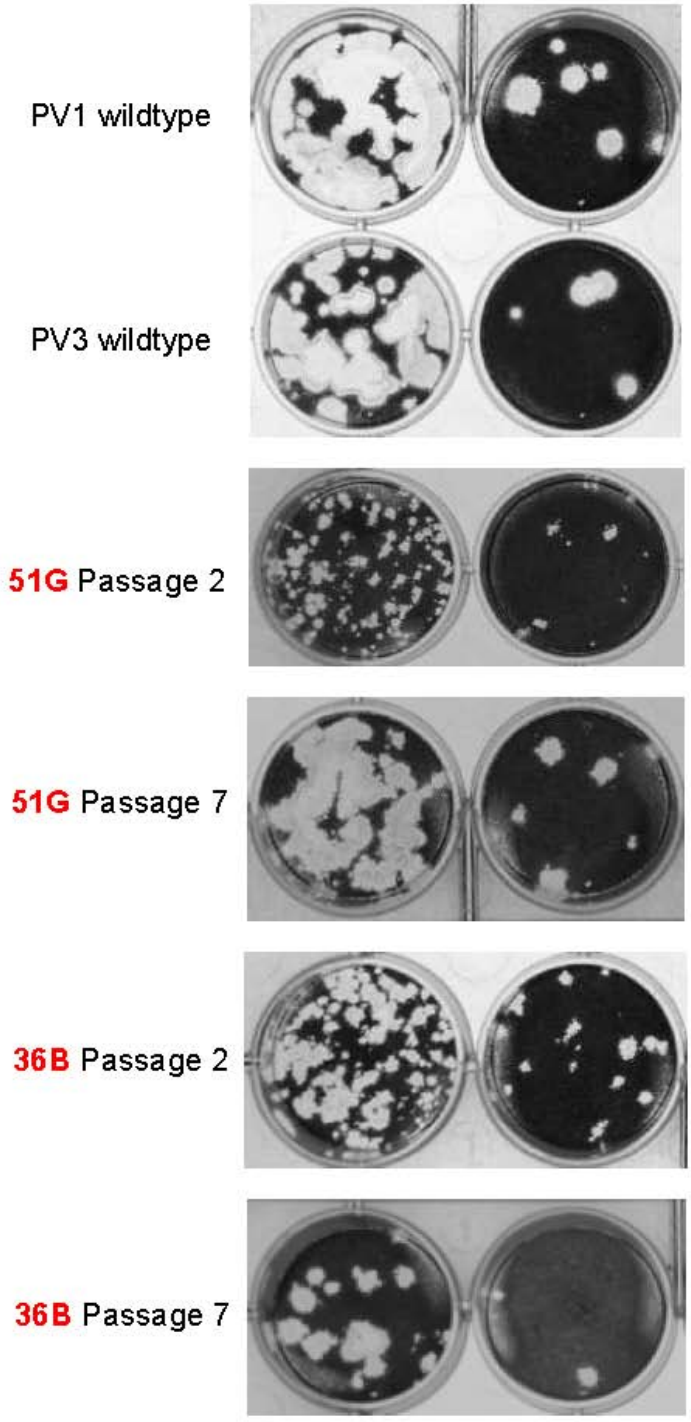


Figure 5.9 Plaque phenotypes of parent and two selected recombinant viruses. HeLa cells were infected with a 10-fold dilution series of virus and overlaid with plaque assay overlay medium. Cell monolayers were stained after 48 hrs with crystal violet. Wild-type parents PV1 and PV3, and recombinants 51G and 36B, after 2 and 7 further passages through HeLa cells. Recombinant virus plaques change morphology after five subsequent rounds of replication, to more closely reflect those of PV1 and PV3.

HeLa cells were infected with a virus mix containing an MOI of 10 of each recombinant and both parents PV1 and PV3. The resulting supernatant was passaged through HeLa cells five times and the recombination target region was amplified again by RT-PCR. Generic PCR oligonucleotides PV1/3-3280F and GEN-4615R (table 2.3) were used to amplify both parents and recombinants and specific products were discernable by band size. As the expected band size for the parent target region was approximately 1335 bp, larger bands would originate from recombinant viruses with sequence inserted in the amplified region.

Two DNA bands were evident at p2 belonging to either recombinant or wild-type viruses on 1% agarose gels (figure 5.10), with original virus mix used to inoculate cells labelled p2. Mixed virus populations were present in both p2 and p3 RT-PCR product, as dual bands existed in both lanes. After five passages (to p7), gel bands derived from recombinants were indistinguishable from the wild-type bands – suggesting that recombinant viruses had effectively disappeared. There might be two simple explanations for this: (1) recombinant virus was literally outcompeted by parent viruses and ceased to survive further rounds of infection, or (2) recombinants remained in the virus mix but were altered by the removal of additional sequence. If the latter occurred, the same pattern would emerge if the recombinants were passaged individually. As an aside, it is not clear why the 'recombinant' DNA band was less intense than the smaller 'wild-type' DNA band in p2 PCR product. Both virus populations were infected at an MOI of 10, and infectious virus titres were originally determined by plaque assay. Either PCR conditions were not optimal for amplifying recombinant cDNA, or there was a significantly lower particle-to-PFU ratio for the PV3/PV1 recombinants.

Continuous passage of recombinants alters crossover junctions

Above mentioned recombinant viruses only were serially passaged 10 times through HeLa cells and the amplification of the crossover region was repeated to determine if recombinants were altered. Cells were infected at an MOI of 10 with p2 recombinant virus. Virus RNA was extracted from inoculum (p2) and supernatant after 1, 5, and 10 passages (p3, 7 and 12 respectively) and amplified by RT-PCR as described above.

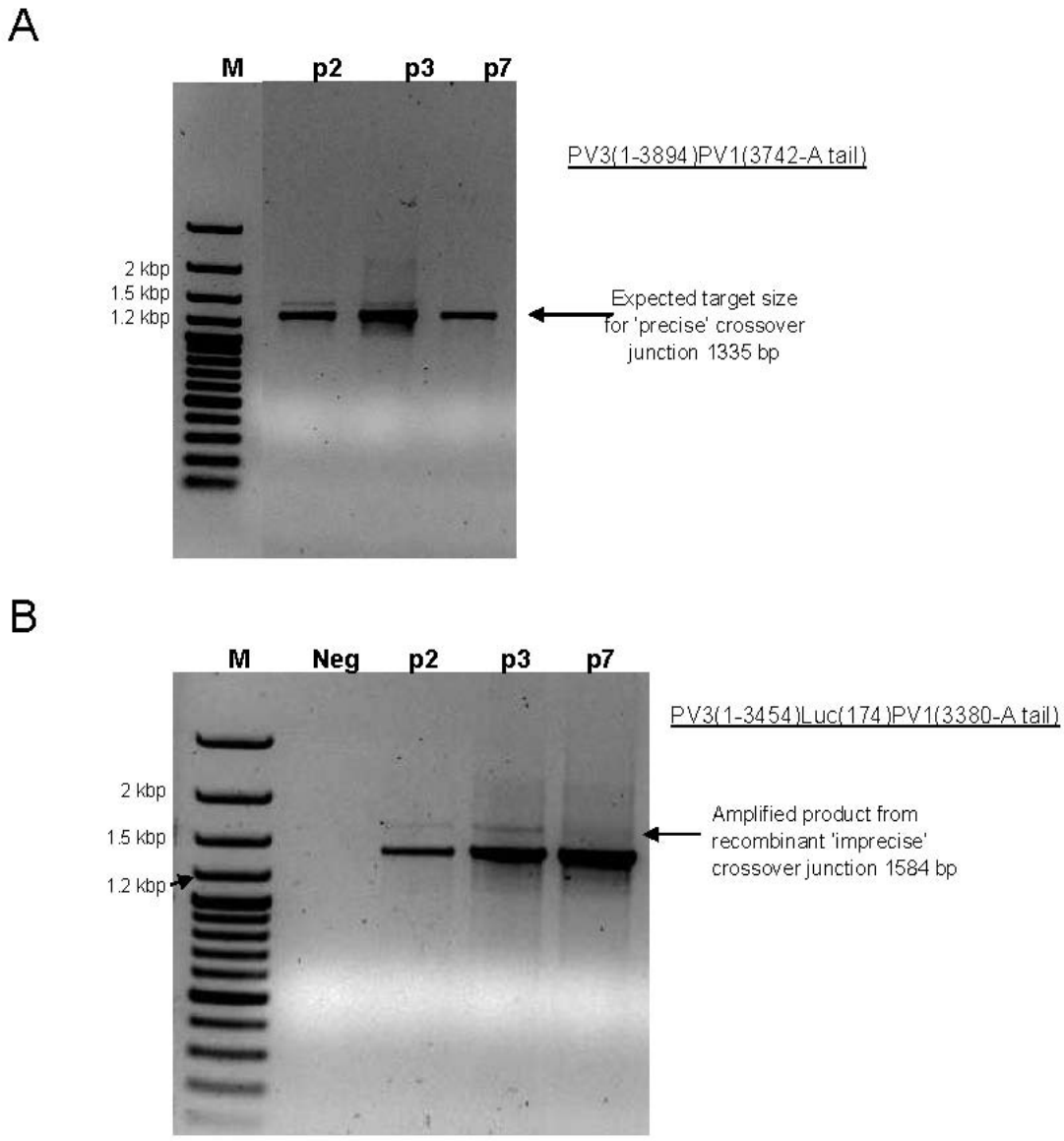


Figure 5.10 Competition assays between recombinant and wild-type viruses in the same infection. HeLa cells were infected with a virus mix containing an MOI of 10 of each recombinant and both parents PV1 and PV3. The resulting supernatant was passaged through HeLa cells five times and the recombination target region was amplified again by RT-PCR. Passage 2 was the initial virus mixture inoculated onto cells. 'Neg' refers to the template control (H_20) in PCR. Both recombinant and parent viruses are present in p2 and p3 lanes, but the recombinant virus (indicated by larger band) disappears altogether in gel (A), while the majority of virus remaining in gel (B) is indistinguishable from the product generated from entire PV1 or PV3 (1335 bp).

PCR product was resolved on a 1% agarose gel by electrophoresis and stained with ethidium bromide. The predominant amplified product changed in 51G virus preparations between 4 and 7 passages through cells, as indicated by the reduced band size from 1488 bp to 1335 bp (figure 5.11). By p7, the majority of amplified product from virus supernatant was the same size as the expected wild-type product, and this was the band present after a further 5 passages. The same was evident for recombinant 36B, although removal of extra sequence at the crossover junction progressed more slowly (from passage 7 onwards). After 10 passages, the predominant, amplified product from supernatant reflected the wild-type band size, although larger PCR product was still present. It is quite possible that a complete change in virus band size would occur with subsequent rounds of replication in HeLa cells but this was not tested.

There were no intermediate DNA bands visible between 'recombinant' and 'wild-type' PCR product on agarose gels, suggesting that these were either below the level of detection, perhaps not resolved well on an agarose gel, or did not exist. To conclude, it was clear that the nature of the recombination crossover site changed after further rounds of replication, and reflected the size expected for precise junctions. Removal of additional sequences presumably occurred by an additional recombination event, either intramolecularly or with a co-infecting genome – therefore in *cis* or *trans*. If the passaged recombinants had indeed changed as a result, it was of interest to ascertain the nature of these new crossover sites.

Characterisation of junction sites of serially passaged PV3/PV1 recombinants

The aim was to serially passage recombinants at high MOIs rapidly through HeLa cells to determine the effect on growth kinetics and plaque phenotype. After continuous passage of PV3/PV1 recombinants 51G and 36B through HeLa cells, it was clear that the distance between both PCR oligonucleotides became smaller, indicating that the crossover junction (originally containing inserted sequence) was visibly reduced. Virus recovered after 12 passages was biologically cloned again before repeating RT-PCR and sequencing of the crossover regions. This was to determine if changes had occurred at the recombination junctions and to establish the fate of these sizeable insertions.

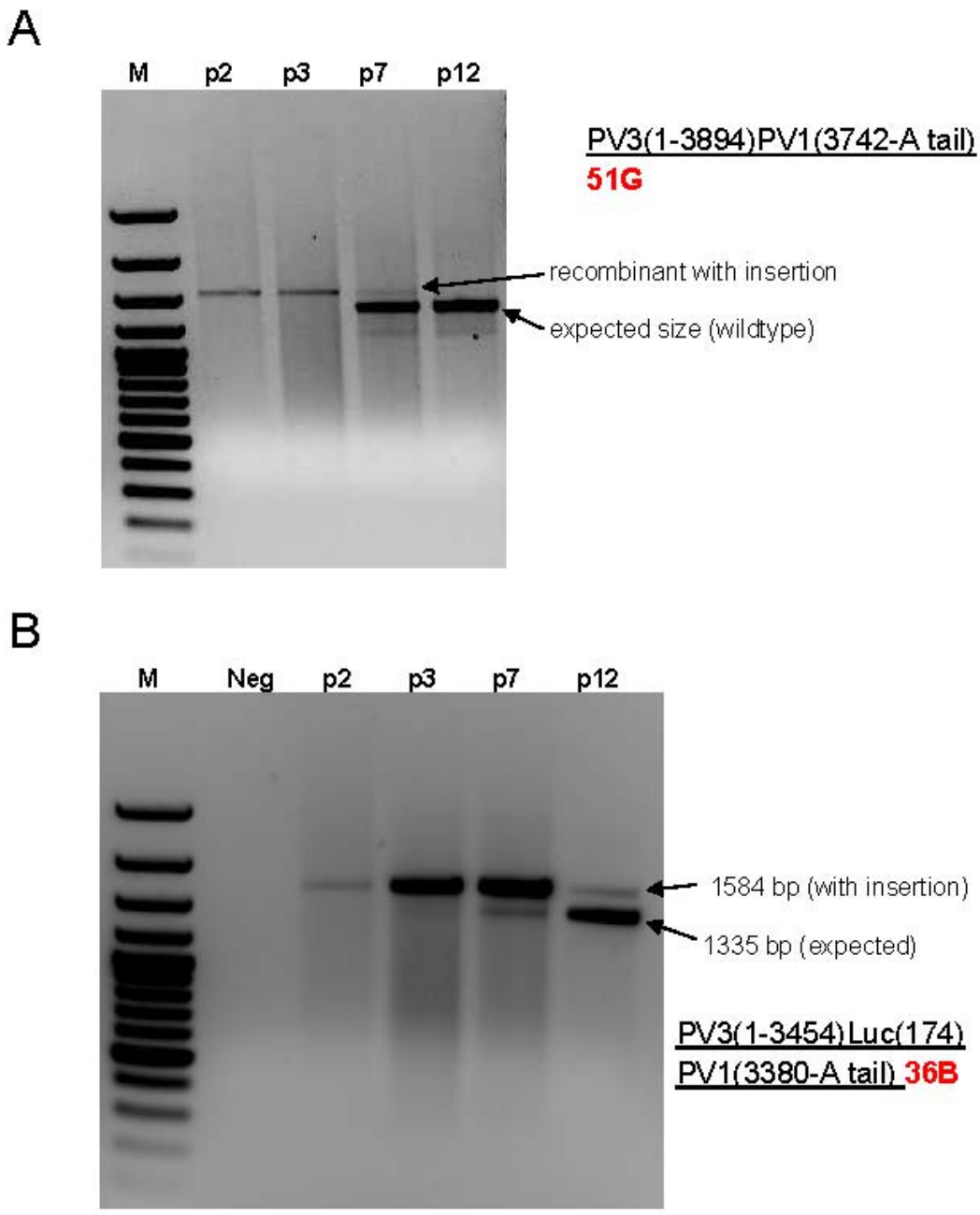


Figure 5.11 Changes in crossover regions from passaged PV3/PV1 recombinants. Following five subsequent passages through HeLa cells, RT-PCR product of amplified region in recombinant viruses was run on a 1% agarose gel and stained with ethidium bromide. (A) Recombinant 51G crossover region gradually reduced in size from 1488 bp (with insert) to 1335 bp. This was the length of the precise crossover region. (B) Likewise, 36B crossover regions contained two populations after 5 passages, with the majority of the original recombinant with insert still present. Following 10 subsequent passages, most virus in the population contained a precise junction site with 1335 bp.

Virus from p12 supernatant material was biologically cloned in 96-well plates and prepared for RT-PCR and sequencing as stated earlier. Oligonucleotides PV3-2995F and GEN-4615R were used to amplify the recombination region (table 2.3). RT-PCR product was run by electrophoresis on a 1% agarose gel (figure 5.12) and several clones were analysed for each original recombinant.

Visibly, the once heterogeneous collection of virus clones for each recombinant (figure 5.1C) in the initial population appeared uniform in size and were the length of the calculated target region without inserted sequence (figure 5.12). The RT-PCR product recovered from all clones was sequenced to determine the nature of the altered crossover regions. Eight virus clones were analysed for the 51G recombinant and 15 clones for the 36B recombinant.

By p12, recombinant virus clones contained precise junctions between PV3 and PV1 sequence with no insertions or deletions (relative to the parental viruses) detected at these sites. New virus labelling and junction sites are outlined in table 5.3, and the nt sequences of PV3, PV1, and each clone were aligned and compared in figure 5.13A and B.

Progeny virus derived from the 51G recombinant no longer contained extra PV sequence at crossover junctions, and interestingly, the eight clones analysed fell into four distinct groups with unique crossover sites between the PV3 and PV1 template. Therefore, mixed populations of recombinant viruses were recovered from originally cloned virus. It was not possible to pinpoint the exact crossover site in each virus, as this occurred within a small region of sequence conserved between donor and acceptor template, hence the approximate sites of the recombination junctions are shown in brackets, i.e. PV3(3820-4)PV1 in table 5.3.

Similarly, the extra inserted sequence originating from luciferase and PV at the crossover junction was removed upon further passage of recombinant 36B. Five separate recombination junction sites were identified from virus clones and figure 5.14 depicts the original recombinants and subsequent passaged virus clone junction sites across the target region. Once more, exact crossover points were not

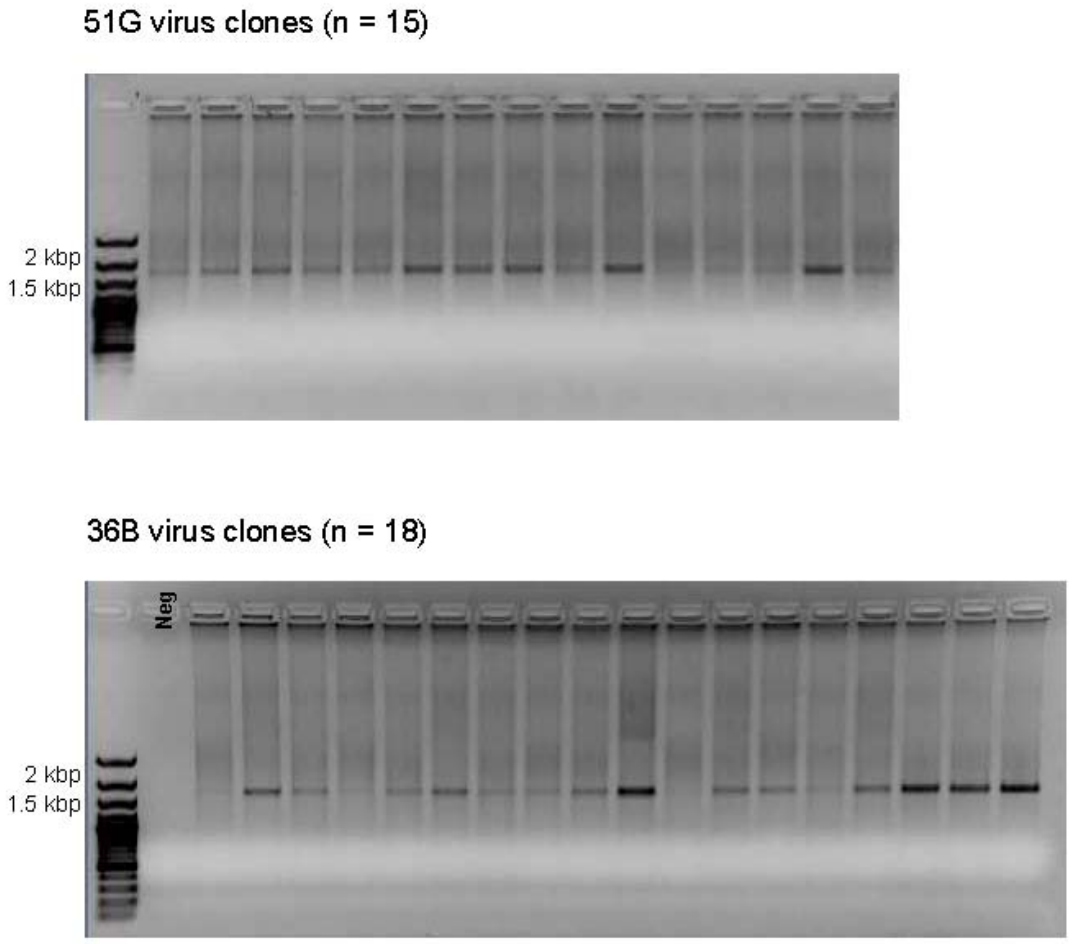


Figure 5.12 Crossover regions of PV3/PV1 recombinants after serial passage through HeLa cells. Gel electrophoresis of RT-PCR product of recombination region. Ethidium bromide stained 1% agarose gels showing the uniform size of amplified virus fragments from both biologically cloned recombinants. The expected band size of approximately 1620 bp was observed in all samples and these were sequenced to determine crossover junctions between PV3 and PV1 sequence. Both PCRs were performed simultaneously, with the single negative control (H_2O) included on the 36B clone agarose gel. The PV3/PV1 recombinant plasmid controls were not available for inclusion in PCR assays. The forward oligonucleotide (PV3-2995F) was specific for PV3 sequence only.

Table 5.3 Characterisation of PV3/PV1 recombinants by crossover region after continuous passage through HeLa cells

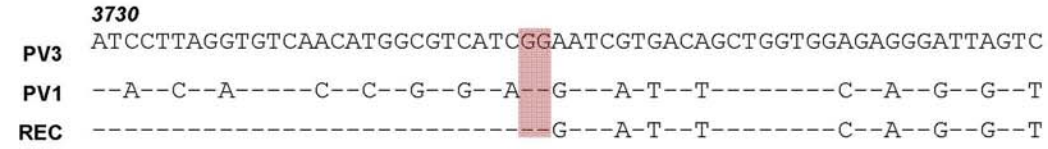
Name*	Position of crossover junctions PV3	PV1	Luciferase sequence	Total extra Nucleotides	No. Clones
PV3(1-3894)PV1(3742-A tail)[#] **51G**	3894	3742	0	153	-
PV3(3757-8)PV1	3757-8	3757-8	0	0	1
PV3(3784-5)PV1	3784-5	3784-5	0	0	2
PV3(3796-7)PV1	3796-7	3796-7	0	0	2
PV3(3820-4)PV1	3820-4	3820-4	0	0	3
PV3(1-3454)luc(174)PV1(3380-A tail)[#] **36B**	3454	3380	174	249	-
PV3(3386-98)PV1	3386-98	3386-98	0	0	1
PV3(3523-9)PV1	3523-9	3523-9	0	0	1
PV3(3538-9)PV1	3538-9	3538-9	0	0	9
PV3(3634-54)PV1	3634-54	3634-54	0	0	2
PV3(3640-54)PV1	3640-54	3640-54	0	0	2

* Numbering according to Clustal alignment of PV serotype 1 Mahoney and PV serotype 3 Leon strains (GenBank accession numbers V01149 and X00925 respectively)

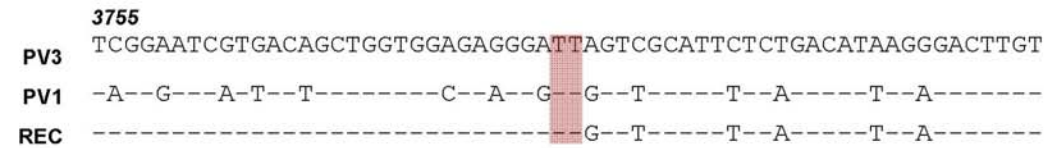
[#] Original PV3/PV1 recombinant subsequently passaged through cells. Four recombinants beneath represent new recombination crossover sites from recovered virus

A

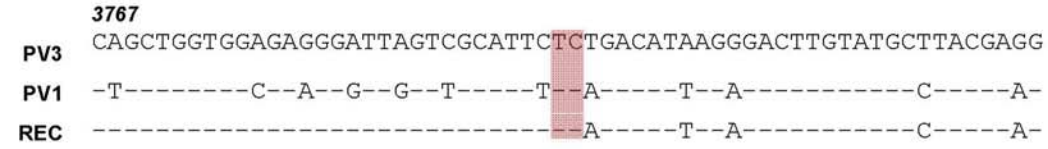
PV3(3757-8)PV1



PV3(3784-5)PV1



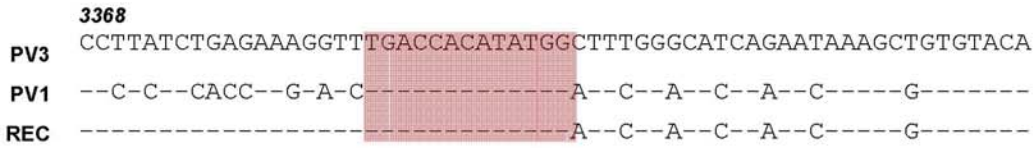
PV3(3796-7)PV1



PV3(3820-4)PV1

B

PV3(3386-98)PV1



Continued over page.

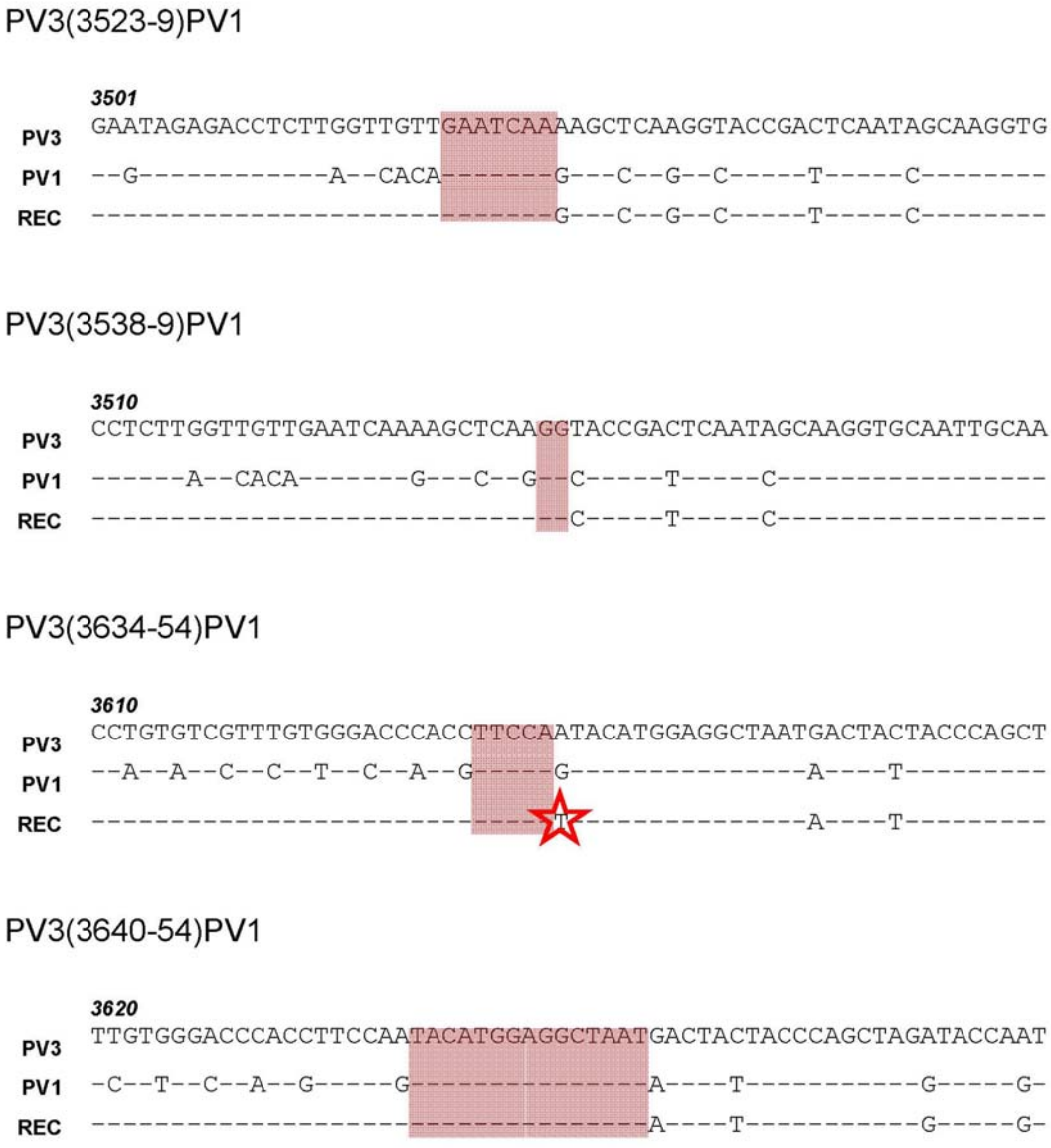


Figure 5.13 Sites of the recombination junctions in the passaged recombinants 51G and 36B. Numbering according to Clustal alignment of PV serotype 1 Mahoney and PV serotype 3 Leon strains (GenBank accession numbers V01149 and X00925 respectively). Dashes depict nt matches to PV3 sequence and red shading indicates the approximate region of template switching. (A) Four unique crossover sites identified from clones after recombinant 51G was passaged through HeLa cells. (B) Crossover sites of five unique recombinants after serial passage of 36B through HeLa cells. The red star highlights the nt mis-incorporation.

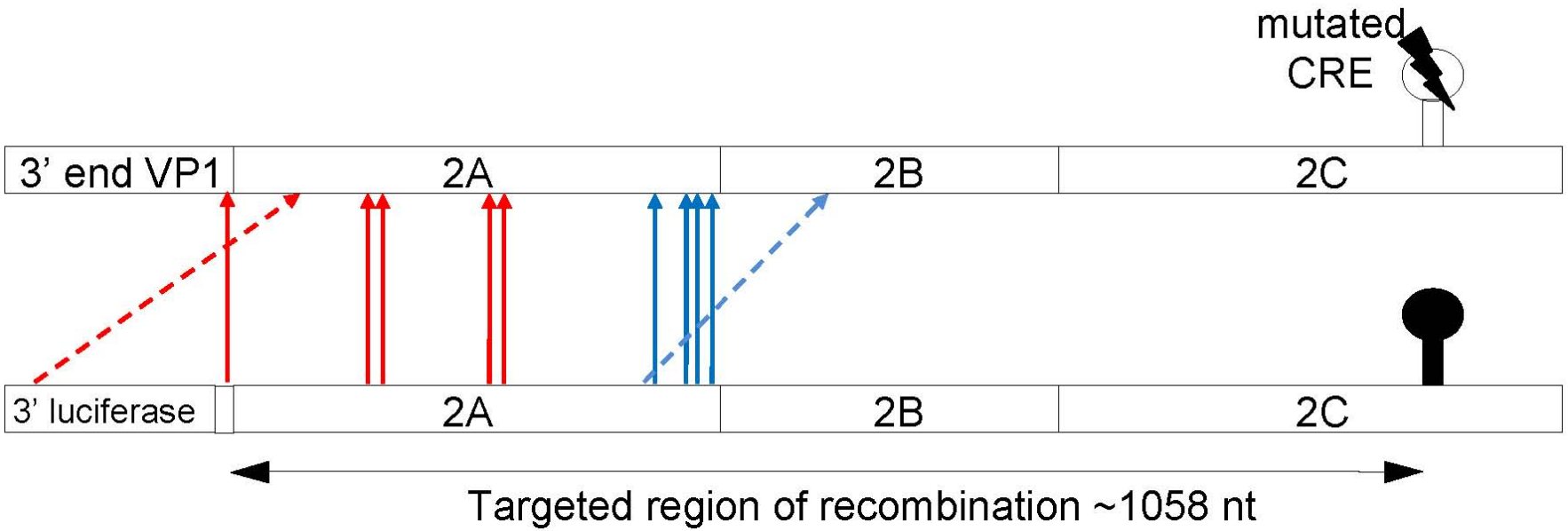


Figure 5.14 New crossover sites for serially passaged recombinants. Viruses are colour-coded as follows, **51G** [PV3(1-3894)PV1(3742-A tail)] and 36B [PV3(1-3454)Luc(174)PV1(3380-A tail)]. Blue arrows represent precise junction sites for biologically cloned viruses derived from 51G. Red arrows indicate the same for 36B clones. The latter viruses could not have evolved from the original PV3/PV1 recombinant because the new crossovers occurred outside available sequence, and this is discussed in the chapter further.

located for these clones as they occurred within a region of shared sequence identity. A mis-incorporation occurred at the site creating more ambiguity, as is shown in figure 5.13B in clone PV3(3634-3654)PV1.

Unfortunately, in the case of the 36B parental virus, all but one newly acquired precise crossover regions did not fall within the available PV3 and PV1 sequences that were present in the primary recombinant. On this basis, it was impossible that the serially passaged recombinants evolved from the original selected recombinant. This was clear in figure 5.14, in which the solid red arrows, indicating new precise junctions, fell outside the junction sites for donor and template sequence. The passaged viruses must have arisen from another primary recombinant, not from the one originally selected. To confirm that the correct primary recombinant was utilised, material from the initial inoculum (p2) was reamplified and sequenced to validate the exact location of template switching. Indeed, the equivalent result was obtained. The RdRp left PV1 donor template 174 nt in from the 3' luciferase gene and joined the PV3 acceptor template from position 3454 nt. A possible explanation was the presence of undetectable levels of contaminating recombinant/s with different crossover sites at early stages that survived through subsequent virus passages. This contaminating recombinant may have had either a growth or stability disadvantage (leading to more rapid evolution) compared to the predominant recombinant as well. This may have contributed to continuous recombination events providing unexpected crossover sites.

For the nine new recombinants, it was not possible to pinpoint the exact nucleotide positions for recombination junctions. Between two to 15 nts at the sites were identical for both partners (refer to figure 5.13). The fact that 100% of the recombination junctions contained ambiguous nucleotides (compared to 31% of original recombinants recovered) demonstrated the role of the sequences in determining crossover sites. These extremely short regions of sequence identity may have guided RdRp during template switching, as previously reported (Tolskaya *et al.*, 1987).

All original PV3/PV1 recombinants contained a complete 2A and 2B sequence derived in its entirety from one or the other parent. The imprecise nature of junction

sites meant that parts of 2A and 2B were also incorporated into the infectious genome. After continuous passage of two recombinants with these attributes, it was shown that these viruses removed/deleted extraneous sequences at the junction site, and in the process, generated chimeric 2A protein and more significantly, restored the length of the genome. According to this data, it was suggested that precise recombination events that produced chimeric proteins within genomes was a rare event during early stages of virus replication. Only one precise crossover event occurred initially, represented by clone 4E, which contained a chimeric 2C encoding region.

5.3 Intraspecies Recombinants

Characterisation of intraspecies C recombinants recovered from co-transfections

As discussed in chapter four, co-transfections were performed using different serotypes of RNA partners encoding the CVA21 non-structural region. As a CVA21 subgenomic replicon was not available for co-transfections, non-infectious PV recombinants containing the CVA21 non-structural region were used instead. Virus was recovered from the following crosses: pRLucWT + CVA21BKPV3P1, pRLucWT + CVA21BKSabin1P1, pT7Rep3-L + CVA21BKPV3P1, and pT7Rep3-L + CVA21BKSabin1P1 (figure 4.6). Supernatant containing virus was harvested from transfected cells and biologically cloned as previously described.

Biologically cloned virus derived from partners pRLucWT + CVA21BKSabin1P1 and pT7Rep3-L + CVA21BKPV3P1 co-transfections were examined further to confirm the existence of recombinant virus and to determine the location of crossover junctions.

All amplified PCR products (figure 5.15) were sequenced for each group of viruses using oligonucleotides outlined in table 2.3. Recombinant virus resulting from partners pRLucWT + CVA21BKSabin1P1 contained the following fragments of virus cDNA across the whole genome when the composition of input RNA partners was taken into account; CVA21/PV1/CVA21/±Luc/PV1. Co-transfections with RNA partners pT7Rep3-L + CVA21BKPV3P1 resulted in recombinant virus with the following genome organisation, CVA21/PV3/CVA21/±Luc/PV3.

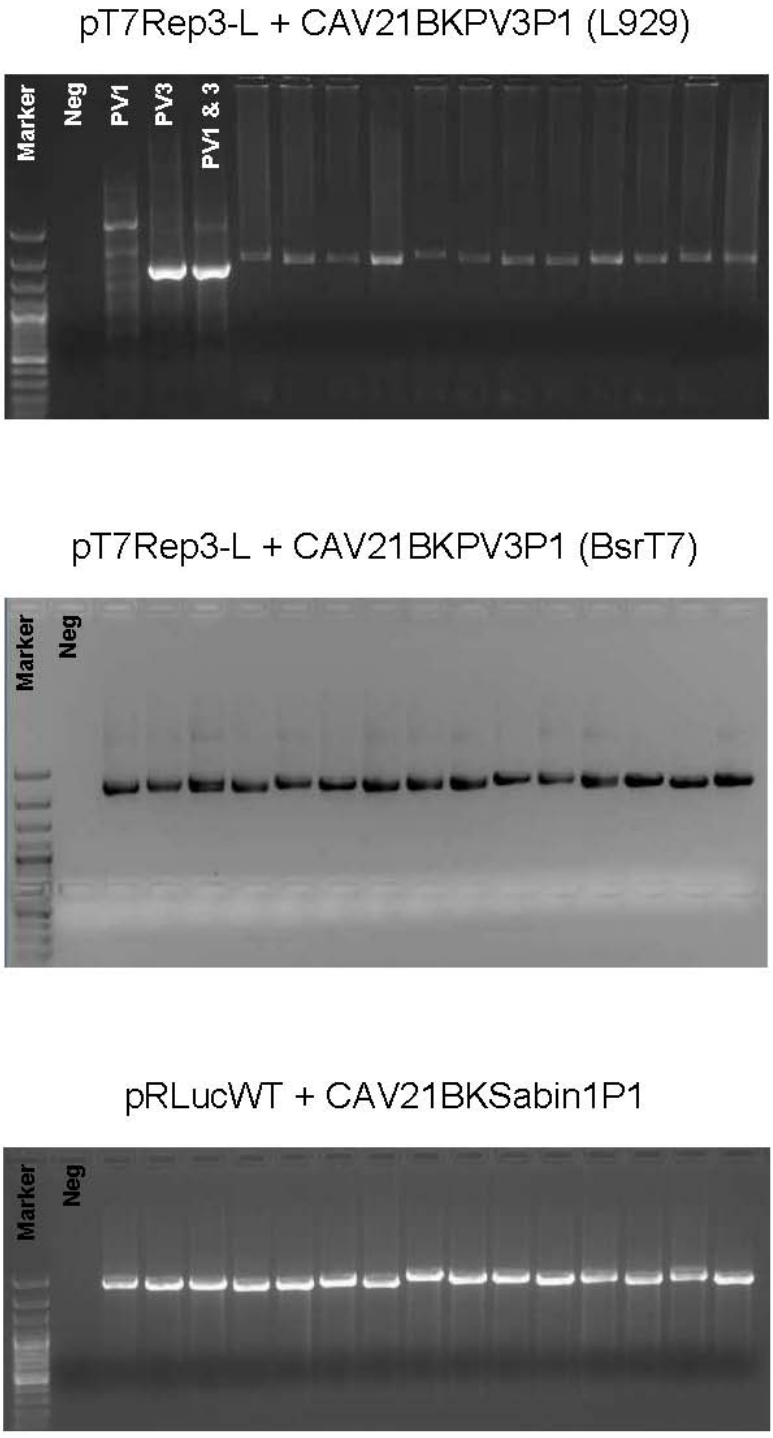


Figure 5.15 RT-PCR product of biologically cloned intraspecies C recombinants. These were recovered from co-transfections in L929 and BsrT7 cells. Gel electrophoresis of RT-PCR product of recombination region. Ethidium bromide stained agarose gels showing a heterogeneous mixture of amplified virus sequences from cloned recombinants. Controls in top gel (PV1 and PV3 plasmids) were performed to check specificity of virus specific oligonucleotides. Positive controls were not included in the bottom two PCR assays.

Crossover positions were determined from a cDNA Clustal alignment (Higgins & Sharp, 1988) with PV1 strain Mahoney, PV3 strain Leon, and CVA21 strain Coe (GenBank accession numbers V01149, X00925, and D00538 respectively). All virus clones sequenced, regardless of the RNA partner origin, contained imprecise junction sites and several distinct recombinants were recovered from each co-transfection performed. All recombinants contained extra sequence at the site of the crossover.

Table 5.4 outlines the RNA partners involved in crosses and resulting recombinant viruses identified through sequencing. The unexpected result was the number and nature of extra nts that were incorporated into these infectious recombinants. As an example, CVA21/PV1/CVA21/±Luc/PV1 recombinants contained between 123 and 513 nts of extra sequence at the junction sites and more importantly, sequence derived from the luciferase gene was incorporated. Nine unique recombinants were identified from 16 clones sequenced, with template switching sites occurring in the 2A region of CVA21. Figure 5.16 illustrates the virus sequence fragments that formed recombinant CVA21(1-711)PV1(712-3403)CVA21(3404-3413)Luc(318)PV1(3385 to A-tail).

Six unique recombinants were identified from 15 CVA21/PV3/CVA21/±Luc/PV3 clones recovered from co-transfections in L929 cells. Similarly, four unique recombinants emerged from five clones derived from BsrT7 cells. Again, crossover sites clustered in the 2A region of CVA21. All recombinants contained extra sequence at the crossover site, ranging from 18 to almost 300 nt, and this occurred irrespective of transfection cell type (L929 or BsrT7). Interestingly, the four unique recombinants recovered from BsrT7 cells did not contain luciferase, although conclusions could not be drawn from this due to the low number of clones examined. The majority of crossover sites were different between the cell lines indicating random RdRp movement between virus templates that differed with each experiment.

Regardless of crossover site locations, recombinants recovered from both RNA partner crosses contained a complete set of 2A and 2B virus proteins from the subgenomic replicon (CVA21BKSabin1P1 + pRLucWT) and (CVA21BKPV3P1 + pT7Rep3-L) (figure 5.17). With reference to figure 5.17B, the

Table 5.4 Recombinant viruses recovered from *three* intraspecies co-transfections

RNA Partners	Recombinant Name*	Luciferase sequence	Total extra nucleotides
pRLucWT +	CVA21(1-711)PV1(712-3403)CVA21(3404-3413)Luc(318)PV1(3385-7460)	318	345
CVA21BKSabin1P1	CVA21(1-711)PV1(712-3403)CVA21(3404-3419)Luc(135)PV1(3385-7460)	135	168
(BsrT7 cells)	CVA21(1-711)PV1(712-3403)CVA21(3404-3423)Luc(227)PV1(3385-7460)	227	264
	CVA21(1-711)PV1(712-3403)CVA21(3404-3425)Luc(84)PV1(3385-7460)	84	123
	CVA21(1-711)PV1(712-3403)CVA21(3404-3456)Luc(443)PV1(3385-7460) **10G**	443	513
	CVA21(1-711)PV1(712-3403)CVA21(3404-3498)Luc(227)PV1(3385-7460)	227	339
	CVA21(1-711)PV1(712-3403)CVA21(3404-3508)Luc(43)PV1(3385-7460)	43	165
	CVA21(1-711)PV1(712-3403)CVA21(3404-3552)Luc(227)PV1(3385-7460)	227	393
	CVA21(1-711)PV1(712-3403)CVA21(3404-3691)Luc(67)PV1(3385-7460) **12E**	67	372
pRep3-L +	CVA21(1-711)PV3(712-3403)CVA21(3404-3420)Luc(84)PV3(3313-7460)	84	192
CVA21BKPV3P1	CVA21(1-711)PV3(712-3403)CVA21(3404-3421)Luc(107)PV3(3313-7460) **2B**	107	216
(L929 cells)	CVA21(1-711)PV3(712-3403)CVA21(3404-3428)Luc(183)PV3(3313-7460)	183	300
	CVA21(1-711)PV3(712-3403)CVA21(3404-3516)PV3(3346-7460)	-	171
	CVA21(1-711)PV3(712-3403)CVA21(3404-3519)PV3(3307-7460)	-	213
	CVA21(1-71)PV3(712-3403)CVA21(3404-3554)PV3(3393-7460)	-	162
pT7Rep3-L +	CVA21(1-711)PV3(712-3403)CVA21(3404-3420)PV3(3403-7460)	-	18
CVA21BKPV3P1	CVA21(1-711)PV3(712-3403)CVA21(3404-3469)PV3(3378-7460)	-	93
(BsrT7 cells)	CVA21(1-711)PV3(712-3403)CVA21(3404-3505)PV3(3353-7460)	-	153
	CVA21(1-711)PV3(712-3403)CVA21(3404-3520)PV3(3338-7460) **19E**	-	183

* Numbering according to Clustal alignment of PV serotype 1 Mahoney, PV serotype 3 Leon, CVA21 Coe strains (GenBank accession numbers V001149, X00925, and D00538 respectively).

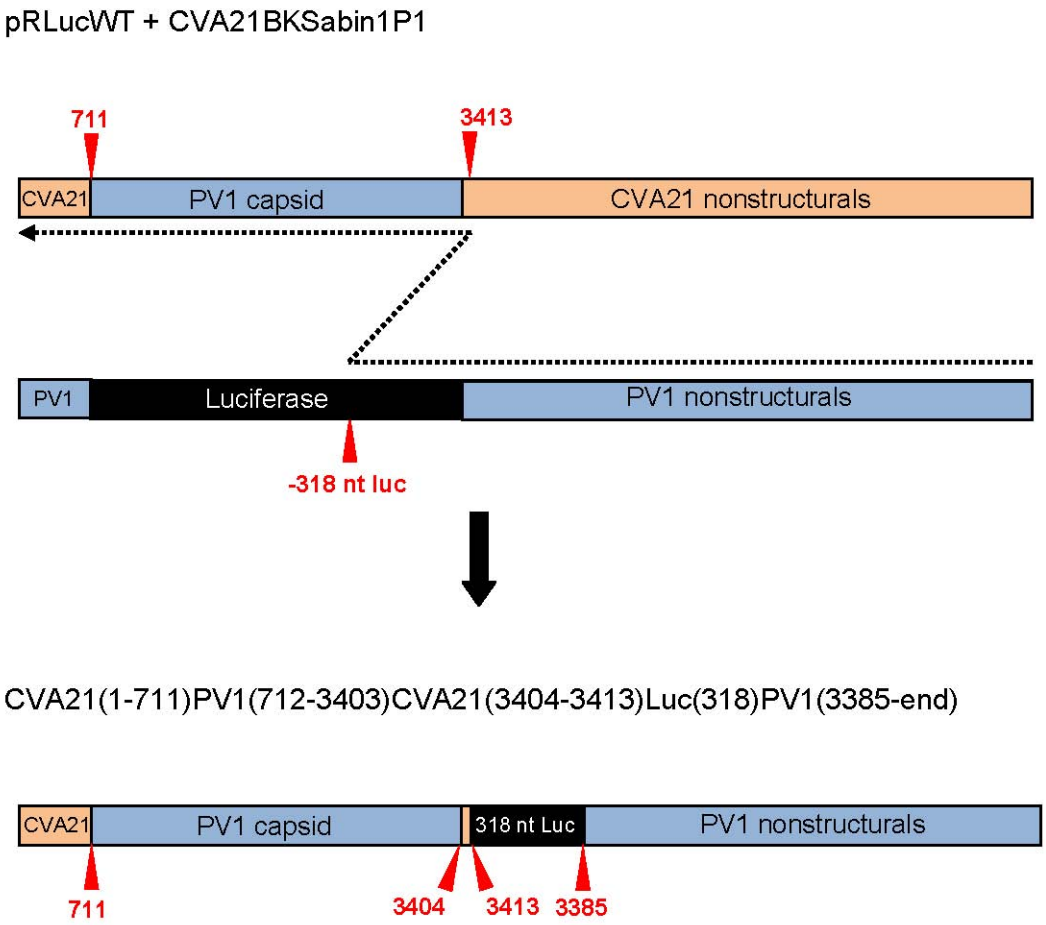


Figure 5.16 Single intraspecies recombinant formed from RNA partners. RNA partners pRLucWT and CVA21BKSabin1P1 were co-transfected into BsrT7 cells. Supernatant was removed 48 hrs post-transfection and inoculated onto fresh HeLa cells to plaque. Virus in transfection supernatant was biologically cloned and individual viruses amplified by RT-PCR across the targeted region of recombination. Both constructed and natural recombination junction sites are represented with red numbering.

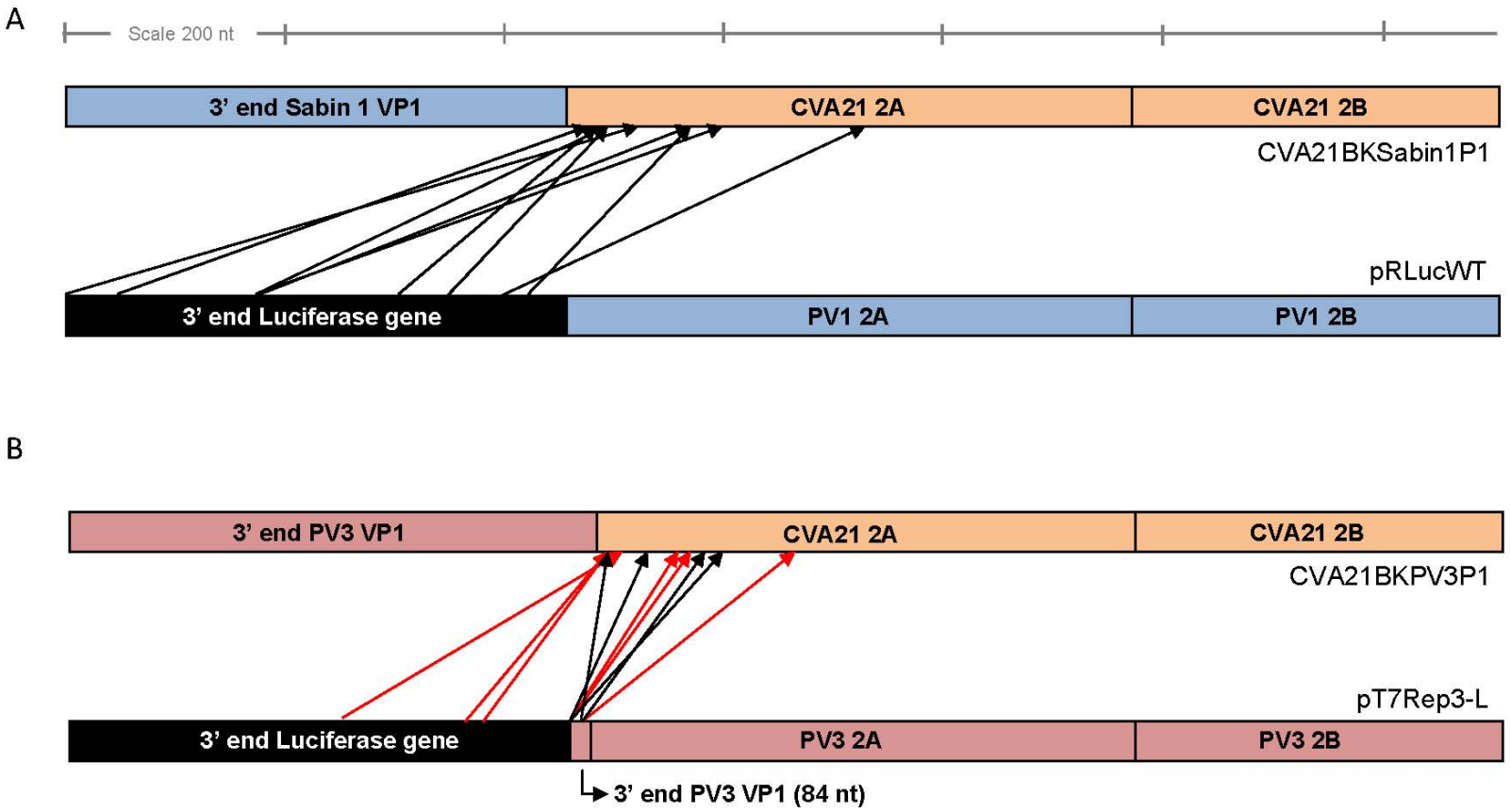


Figure 5.17 Crossover junctions of intraspecies C recombinants formed from replicons and Dr. Blanchard's constructed recombinants. Arrows indicate crossover sites that occurred between pRLucWT and CVA21BKSabin1P1 RNA partners. The RdRp left the donor template (pRLucWT) within the luciferase gene, and rejoined the acceptor template early in the CVA21 2A region. (B) Crossover sites between pT7Rep3-L and CVA21BKPV3P1. Red arrows indicate recombinants recovered in L929 cells, black arrows in BsrT7 cells.

RdRp always switched from pT7Rep3-L in either the small remaining region of PV3 VP1 (of which only 84 nts were present) or the 3' terminus of luciferase, but never from the 2A region. These subgenomic replicons were constructed in such a way that a small portion of VP1 remained after the reporter gene and before the 2A region. This was to ensure that the appropriate virus protease cleavage site was retained. Even though the number of recombinant viruses studied was small, no clones were recovered which showed the RdRp leaving the replicon in the 5' portion of the 2A coding region. Although template switching may have been random at the nucleotide level, it was possible that viruses with complete virus proteins and not those with chimeric CVA21/PV proteins were detected because of viable protein function. Once again, the recombination events adhered to some general rules already observed; the crossover sites were unlikely to be precise, the inserted sequence was no more than 513 nts in length, and chimeric 2A or 2B coding regions were not observed.

Genetic stability of the recombinants

Chimeric viruses were recovered after a single round of replication, thus were not exposed to significant purifying selection that could further modify virus sequence. For this reason, four intraspecies recombinants with inserted sequence at the crossover junctions located in 2A were subjected to 10 further passages through HeLa cells to determine if the recombinants removed the additional inserted sequence over the passages. For brevity, the recombinants (2B, 10G, 12E, and 19E) were coded and identified in table 5.4. Briefly, HeLa cells were infected with the following viruses (p2) at an MOI of 10 for 30 mins at 37°C. Cells were washed twice with sterile PBS once virus was removed and replaced with serum-free DMEM. Supernatant was removed after 24 hrs and a quarter of this supernatant was used undiluted for each subsequent inoculum onto fresh HeLa cells. After selected passages, RNA was extracted directly from the harvested supernatant using RNAeasy Mini Kit (Qiagen) and RT-PCR performed using oligonucleotides bordering each junction site. The crossover regions were amplified using oligonucleotides PV3-2995F, PV1-3004F, GEN-4615R, or PV1-5200R (table 2.3) according to virus cDNA partners used and sequenced in order to locate the crossover junctions.

The different sized cDNA fragments in the targeted crossover region were discernable by gel electrophoresis using a 1% agarose gel and provided a semi-quantitative measure as to whether the additional sequences was removed upon serial passaging.

Visually, the length of crossover region altered over serial passages through HeLa cells (figure 5.18). The recombination region became smaller as indicated by the faster migration of bands and this became increasingly reflective of the expected product size (for precise junction sites). Interestingly, three bands representing the crossover site were present in the virus mix during passages on the gel for clone 10G. This mix of band sizes appearing after the first passage suggested that the extraneous sequence was removed in a stepwise fashion from the crossover region in at least two distinct fragments. In comparison, clone 12E removed the entire fragment abruptly within five passages with only a single phenotype emerging (figure 5.18 top right). Agarose gel results imply that the majority of recombinant viruses recovered were unstable due to extraneous sequence inserted at the junction crossover sites and these insertions self-corrected over time after subsequent passage through cells. However, recombinant 2B proved to be an exception. The crossover region remained the same length after 10 passages through HeLa cells, indicating neither an advantage nor disadvantage associated with the additional 216 nt sequence. It would be of interest to determine if this virus maintained its stability while in competition with its parental viruses. If the inserted sequence was retained and the recombinant grew as well as the wild-type parents, then the junction sites may be of interest for future research purposes.

To conclude, recombinant species C viruses with CVA21 were serially passaged through HeLa cells at high MOIs in an attempt to apply small selective pressures on the viruses and determine whether the unexpected nature of crossover sites were eventually corrected. This indeed proved to be the case from three of the four clones tested. Although viruses were not sequenced again after 10 subsequent rounds of passage, semi-qualitative changes were observed in crossover lengths using gel electrophoresis. Due to time restraints, it was not possible to assess the chimeric nature of progeny virus. Ideally, the effect of complex selective pressures on these ‘aberrant’ recombinant viruses would be assessed *in vivo*.

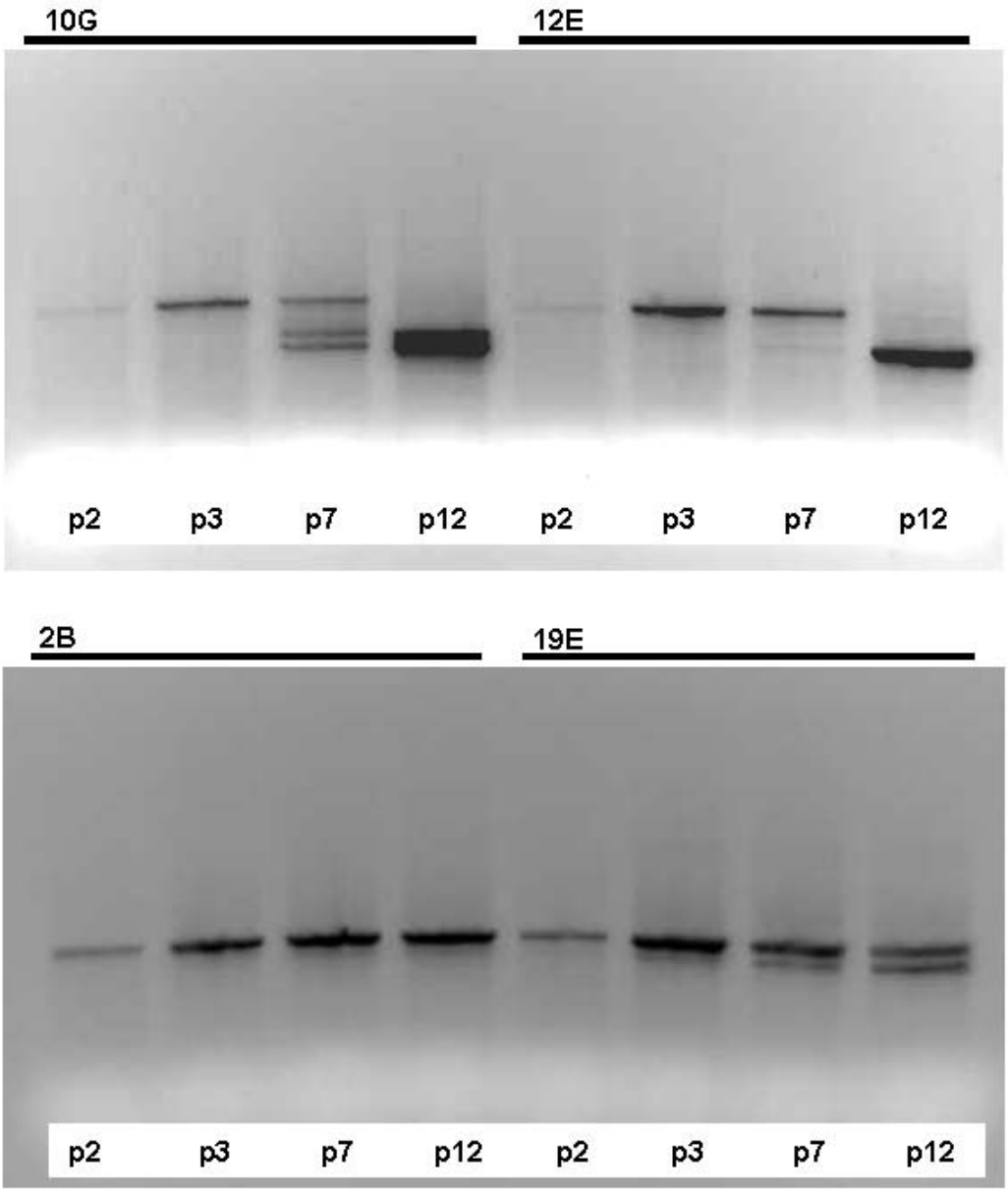


Figure 5.18 Changes in crossover region of serially passaged intraspecies recombinants. HeLa cells were infected with p2 virus supernatant at an MOI of 10 and serially passaged over 24 hr time points, 10 more times. Virus RNA was extracted from supernatant at these passages and amplified by RT-PCR. The target region of recombination was amplified across the passages and PCR product was run on ethidium bromide stained 1% agarose gels. Band shifts were evident in later passages of virus, indicating that the extra sequence at crossover junctions was being removed from viruses. The smaller band correlates in size to the expected target sequence (1056 bp). Full recombinant names/details are provided in table 5.4.

5.4 Comparative Bioinformatic Analysis of RNA Partner Sequences

This method of generating recombinant PV was originally developed to study the fundamental features of early virus RNA recombination within the cell. It was speculated that the virus RdRp pauses during RNA synthesis (Wu *et al.*, 1995), and removes the nascent RNA with it to another template before continuing with RNA synthesis. The primary objectives were to determine whether this occurred at the nucleotide level by studying RNA sequences at crossover junctions. From then on, it was necessary to gain insight into recombination at the protein level – to determine whether functional requirements directly influenced these events. Survival of most randomly constructed recombinants at RNA level may be restricted by stringent structural requirements of virus proteins so that chimeric proteins have a distinct disadvantage. This could explain why recombination events rarely occur in the enterovirus capsid region, as it requires strict protein-protein assembly for correct particle formation and virion encapsidation.

Initially, comparative analysis was performed on the genomic RNA sequences of the two PV partners PV1 (pRLucWT) and PV3 (pT7Leon) as this large scale experiment provided exploratory data for study. All PV3/PV1 recombinants contained a crossover event within a region of 1058 nts encompassing the 2A and 2B encoding regions, as well as the first 319 nts of 2C before the CRE. Initial recovered recombinants' crossover sites were found to be unevenly distributed across this region (on the donor template), with preferences towards luciferase (56.25%), the 2A region (31.25%), 2B region (6.25%), and 2C region (6.25%) (figure 5.4). However, after serial passages of recombinants in cell culture, the now precise crossover sites among the progeny viruses occurred exclusively in the 2A region. Passaged virus from the 51G recombinant contained new junction sites at the 3' terminus of 2A (within a range of approximately 65 nt), while new viruses from passaged 36B had scattered junctions sites in the first half of 2A (over approximately 295 nts). The crossover sites occurred within small identical nucleotide regions of both parental genomes, sometimes only 2 nts in length.

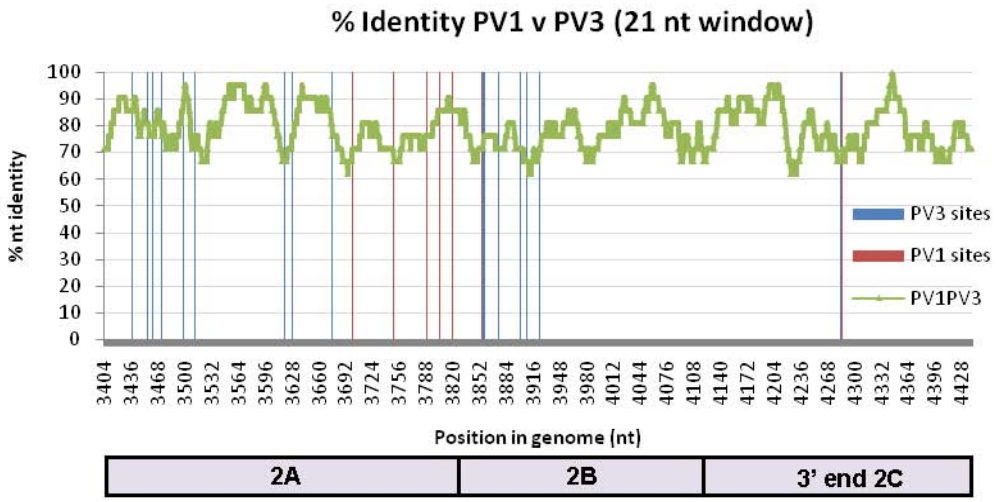
To analyse whether sites of RNA recombination correlated with RNA partners sharing high sequence identity, a pairwise comparison was performed using similarity scanning analysis run by bioinformatics software Simmonic 2005, version

1.75 between the targeted regions of pRLucWT and pT7Leon (Simmonds, 2006). Both imprecise and precise recombination sites (from initial recombinants and passaged progeny respectively) were superimposed along the regions of nucleotide identity (figure 5.19). Nucleotide similarity between PV1 and PV3 ranged from 62-100% in these regions (the average being 78.9%), and no apparent link was evident between the occurrence of precise recombination junctions and regions of high sequence identity. In relation to local sequence identity – if recombination required sequence specificity, then it must be very short, particularly if between 1 and 15 nts are shared between both templates as determined by data from this study.

Pausing or termination of RdRp has been implicated in the process of template switching, and this is thought to be mediated by sequence and/or local secondary structure (reviewed in Nagy & Simon, 1997). Short direct repeats (generally U-rich or AU- rich sequences) were strongly preferred for switches between donor and acceptor templates by RdRp (King, 1988, Pilipenko et al., 1995). It was postulated that the RdRp may 'slip' on AU-rich stretches that created low thermodynamic stability. Analysis of PV1 (donor) nucleotide composition indicated that the region of interest was slightly A-rich (33%) compared to a 24% composition of U nts. Similarly, A and U compositions were 31% and 27%, respectively across the PV3 acceptor template. Junction sites were plotted together with mononucleotide frequencies across PV1 and PV3 target regions. Frequencies were calculated in 21 nt sliding windows and advanced by 1 nt, thereby generating thousands of values across the genome that could be graphed (figure 5.20). Less than 50% of crossovers were flanked by A-rich regions, and only one followed a U-rich region.

To increase the resolution of sequence motifs adjacent to crossover sites, the presence of small sequences of identical nts, such as AAAA, GGG, or TTTTTT, called homopolymers was investigated. To determine whether small homopolymeric stretches were situated close to junction sites, all possible homopolymers of 3-10 nts were identified computationally (DJE, personal communication). Locations of all $\geq$ 3 nt homopolymers in the PV1 and PV3 region of interest were plotted along the x-axis. The different homopolymers were assigned arbitrary values to allow the separation and visualisation of the sequences across the region (figure 5.21).

A



B

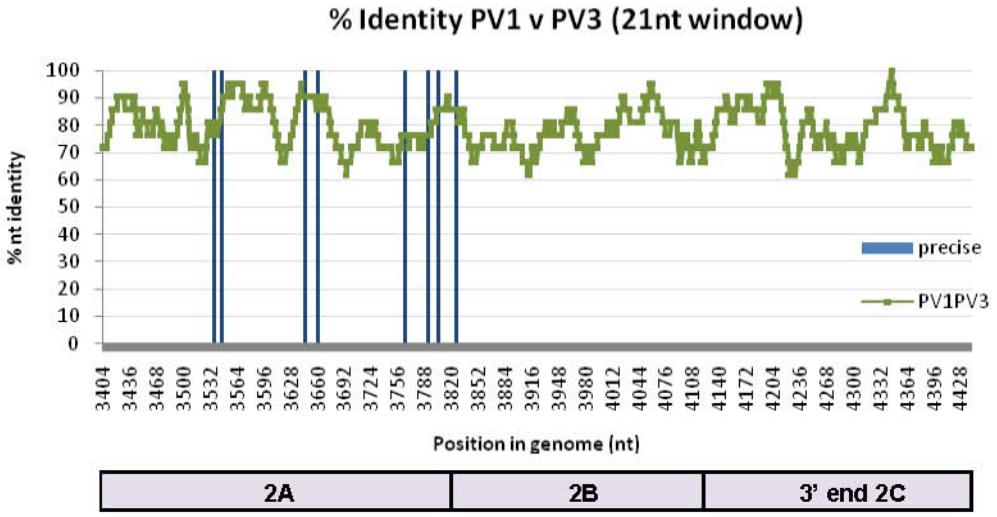
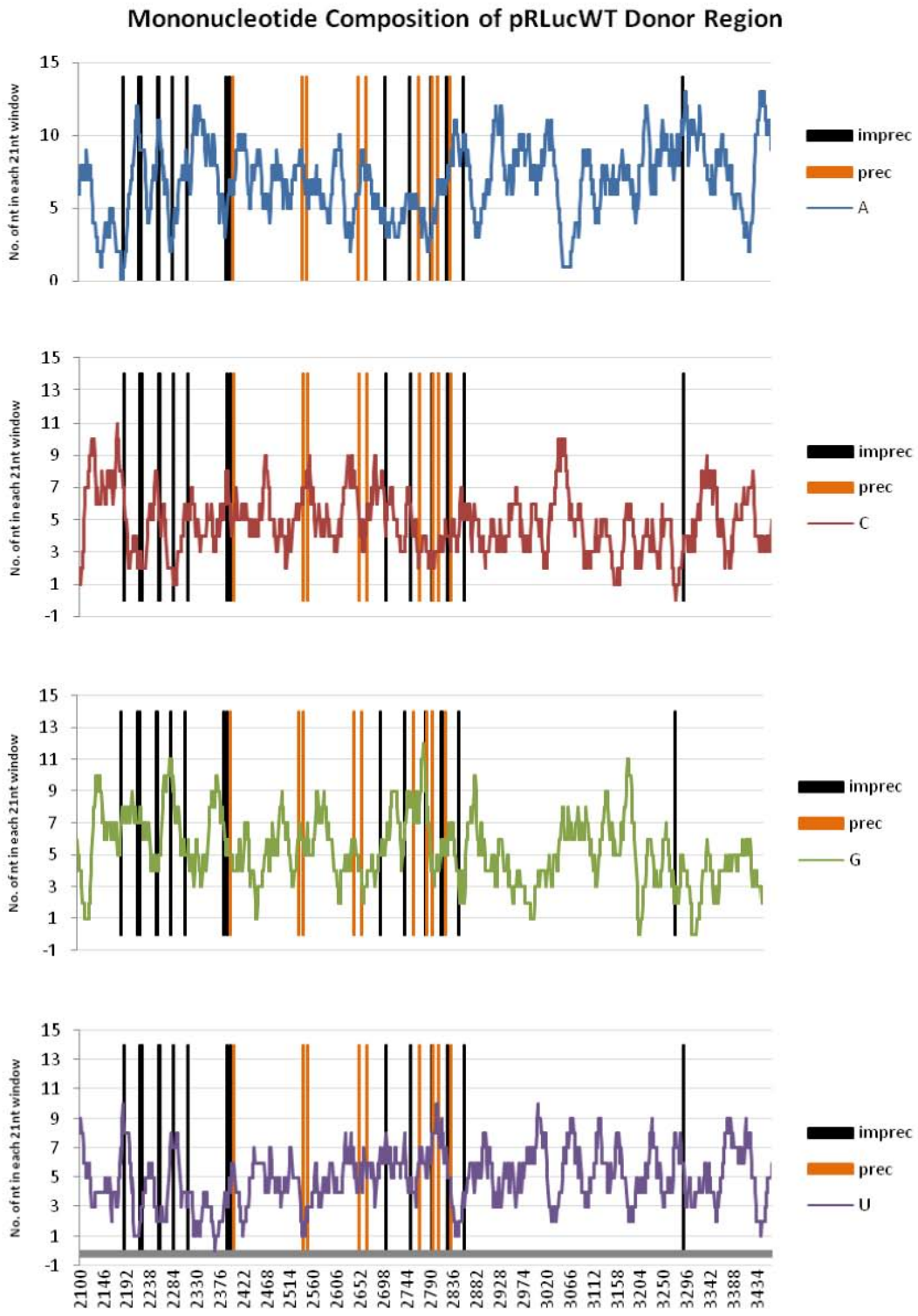


Figure 5.19 Imprecise and precise recombination junctions and correlation to sequence identity between RNA partners. Target regions of pRLucWT and pT7Leon were scanned for stretches of nucleotide sequence similarity. In this case the number of matches in a sliding window size of 21 nts was presented as a percentage. Junction sites were given an arbitrary value of 100% and are represented by columns at the genome position. (A) Both PV1 and PV3 junction sites identified from original pRLucWT and pT7Leon co-transfections were plotted along with nucleotide identity. (B) Precise junction sites identified from passaged recombinant progeny correspond to the exact genome position of PV1 and PV3.

A



Continued over page.

B

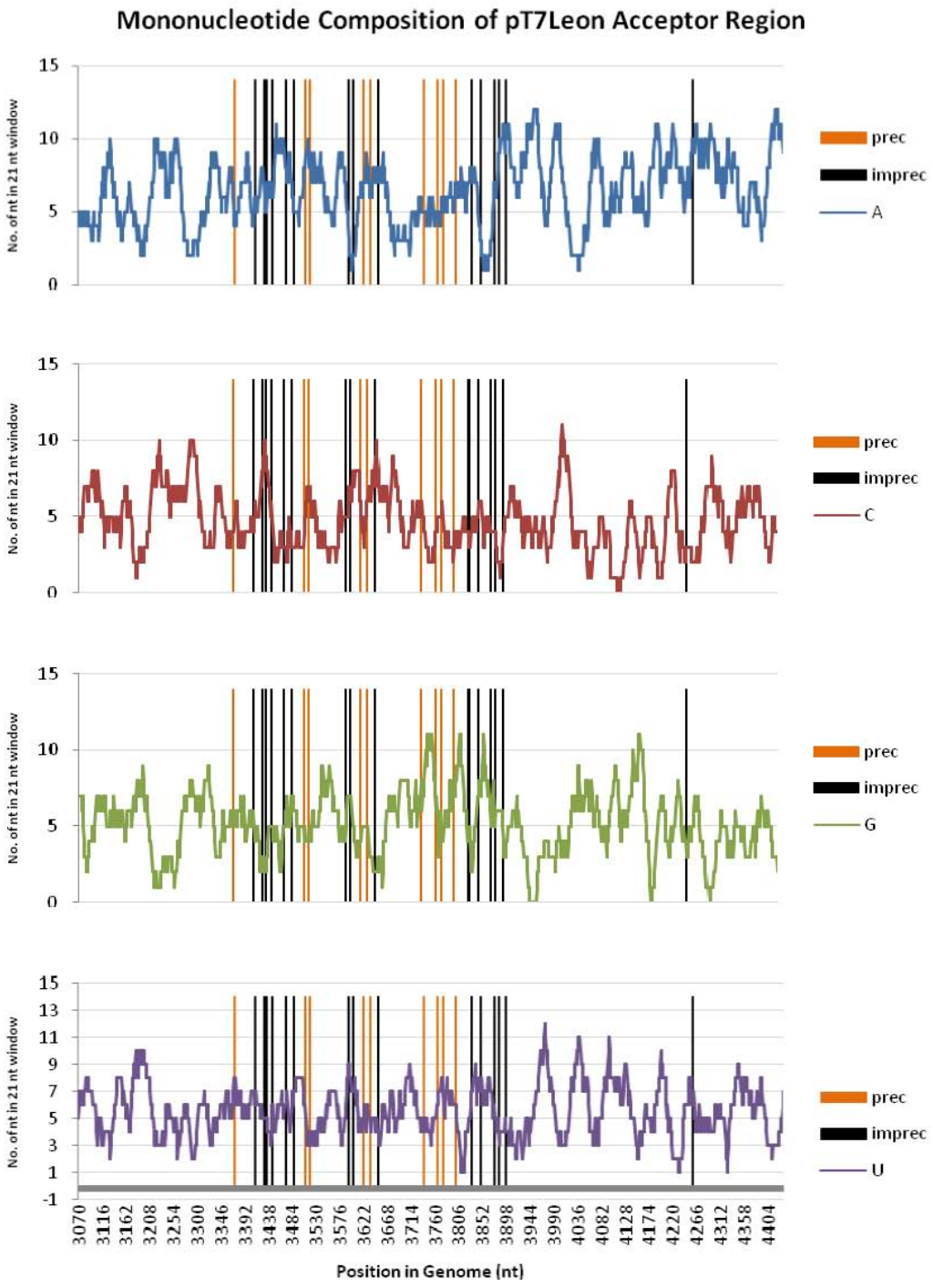
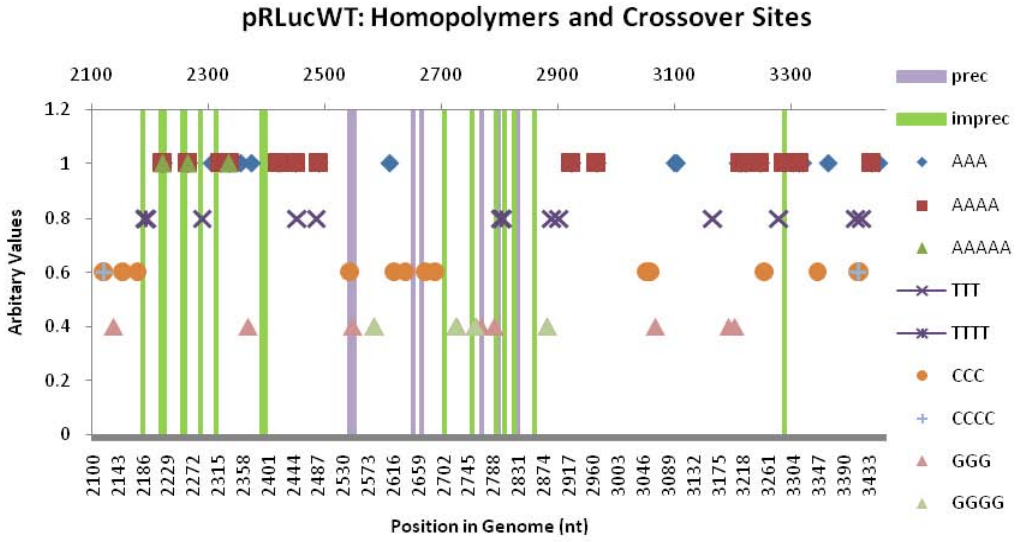


Figure 5.20 Mononucleotide composition of RNA partners. (A) donor (pRLucWT) and (B) acceptor (pT7Leon) template regions of targeted recombination. Composition scans performed in Simmonic 2005 V1.75 across a 21 nt window advancing by 1 nt each time. Donor and acceptor template nt composition (average should be 5.25 per nt) with precise and imprecise junction sites superimposed onto graphs.

A



B

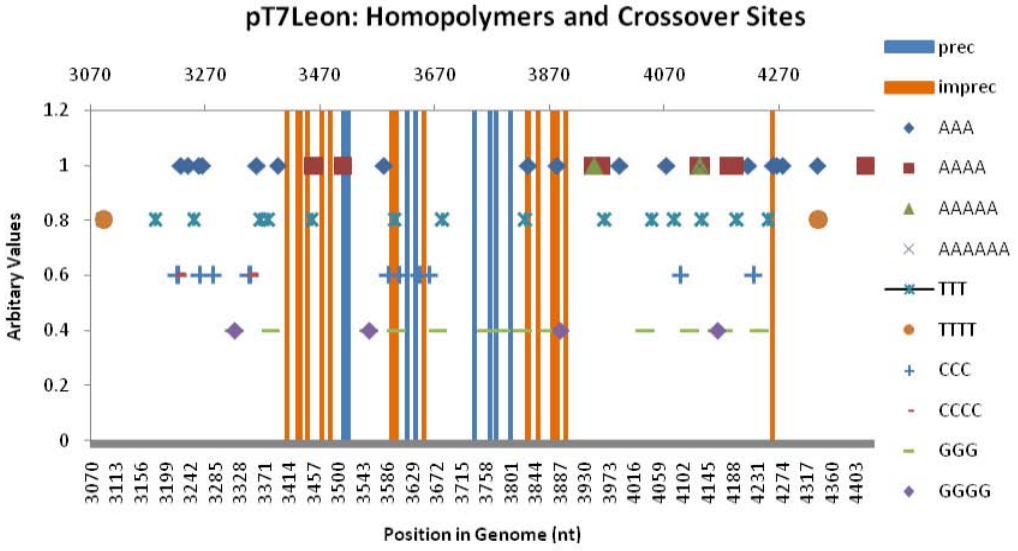


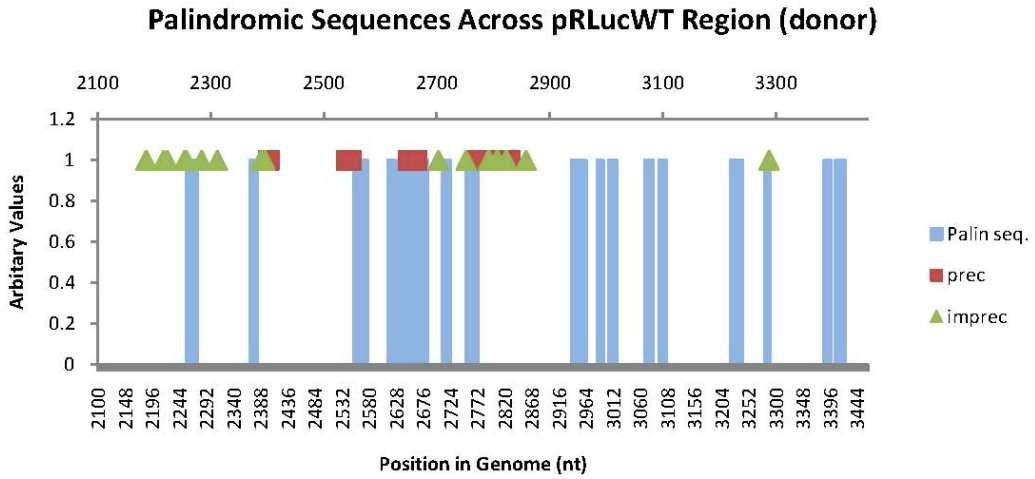
Figure 5.21 Comparison of junction sites and homopolymers located in recombination regions of PV1 and PV3. All features in the legend have been assigned an arbitrary number in order to graph locations on X- and Y- axes. Data generated by Prof. David Evans. (A) Location of all homopolymers (≥3 nts) and crossover sites in donor pRLucWT PV1 sequence. (B) Homopolymers and crossover sites were plotted along genome axis of acceptor template pT7Leon. Y-axis values were applied to each catergory to allow them to be readily visualised.

There was no apparent clustering of sequence motifs or repeated nucleotides close to crossover junctions, particularly along the donor template where RdRp would initially detach. Therefore, factors other than nucleotide identity or the existence of homopolymers were responsible for the occurrence of template switching.

The targeted region of recombination is relatively unstructured with the CRE in the 2C region the only identifiable conserved secondary structure (Goodfellow *et al.*, 2000; Simmonds *et al.*, 2004). To determine if recombination occurred at locations within or flanked by sequences with the potential to form local secondary structures, programs were developed to identify small palindromic sequences with the potential to form perfectly matched hairpin structures in RNA during synthesis. Palindromic sequences with a minimum length of 11 nts were identified (a 4 bp duplex and 3 nt loop). The longest sequences potentially forming stem loops were located in pT7Leon (positions 4165 nt and 4229 nt) with 12-7 (6 bp duplex and 7 nt loop) and 10-9 (5 bp duplex and 9 nt loop) configurations, and at genome position 2559 nt in pRLucWT (10-9). Graphs in figure 5.22 suggested that recombination between both templates was not affected by the presence of potential minor stem loop structures.

Furthermore, Minimal Free Energy Differences (MFED) were calculated across the region. This method predicted native folding free energy of the native RNA secondary structure, and compared this to the free energies calculated from 999 random scrambles of the same region (maintaining codon composition, codon order, and dinucleotide frequencies). This was performed within 101 nt sliding windows which advanced by 3 nt, and was expressed as the percentage difference between native and the mean value of the randomised scores (Zuker, 1989; Zuker, 2003) (Simmonics 2005 Version 1.75). Since the accepted model suggests that recombination occurs during negative-strand synthesis, particular attention was paid to structures formed in positive-sense RNA strands. This was to determine whether RNA secondary structure in the template strand contributed to the polymerase pausing mid-elongation of negative-strand synthesis. Although minimum and maximum values were high, the average MFED values ranged from 1.48% to -2.75% for pRLucWT (donor) and pT7Leon (acceptor) sense strands respectively to denote unstructured RNA sequence (figure 5.23).

A



B

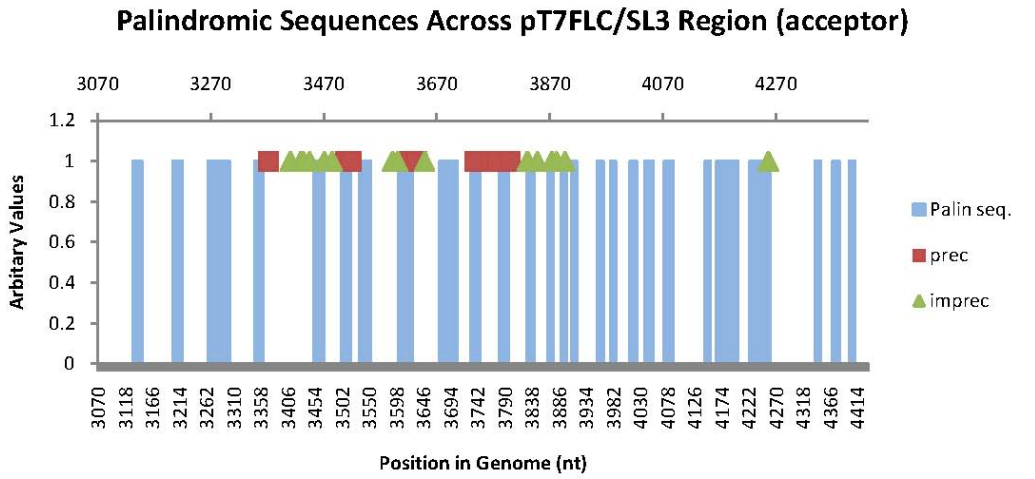
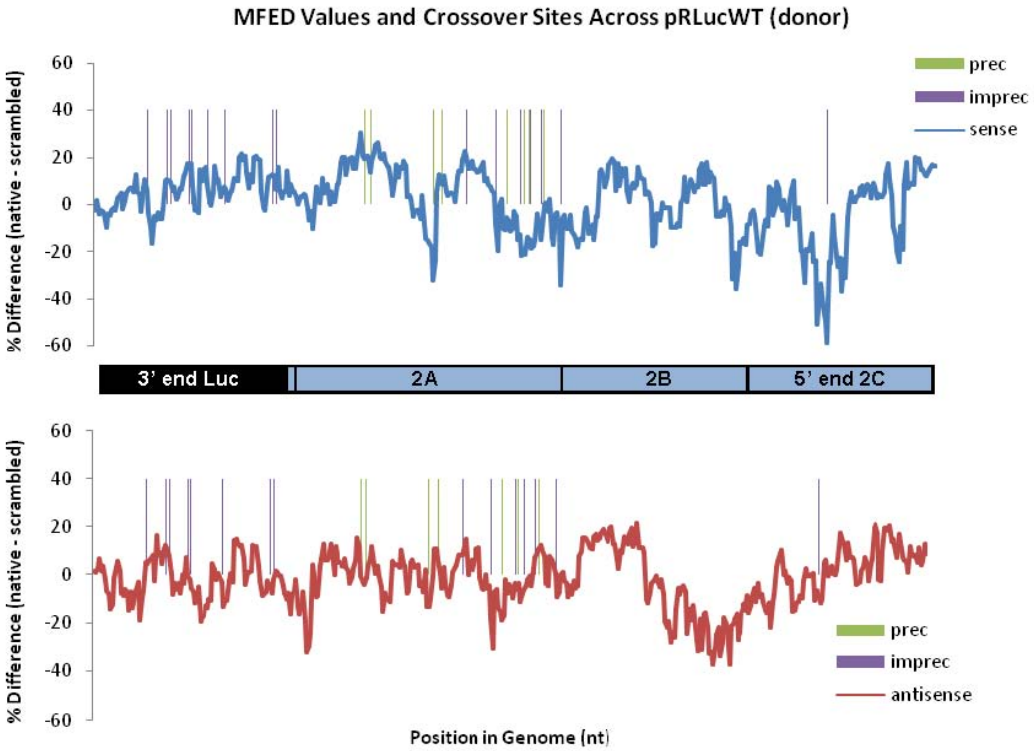


Figure 5.22 Palindromic sequences identified within targeted recombination region of PV1 and PV3. These are represented by the blue bars at the nucleotide location. Palindrome combinations ranged from a minimum length of 11 nts (4 bp duplex with 3 nt loop) to 19 nt (6 bp with 7 nt loop) and were plotted along the region of recombination, together with precise and imprecise crossover locations from characterised recombinants (squares and triangles). (A) Crossover sites in pRLucWT PV1 (donor) superimposed onto axis showing locations of palindromic sequence. (B) Similar for pT7FLC/SL3 PV3 (acceptor) sequence. Y-axis values were applied to each catergory to allow them to be readily visualised.

A



Continued over page.

B

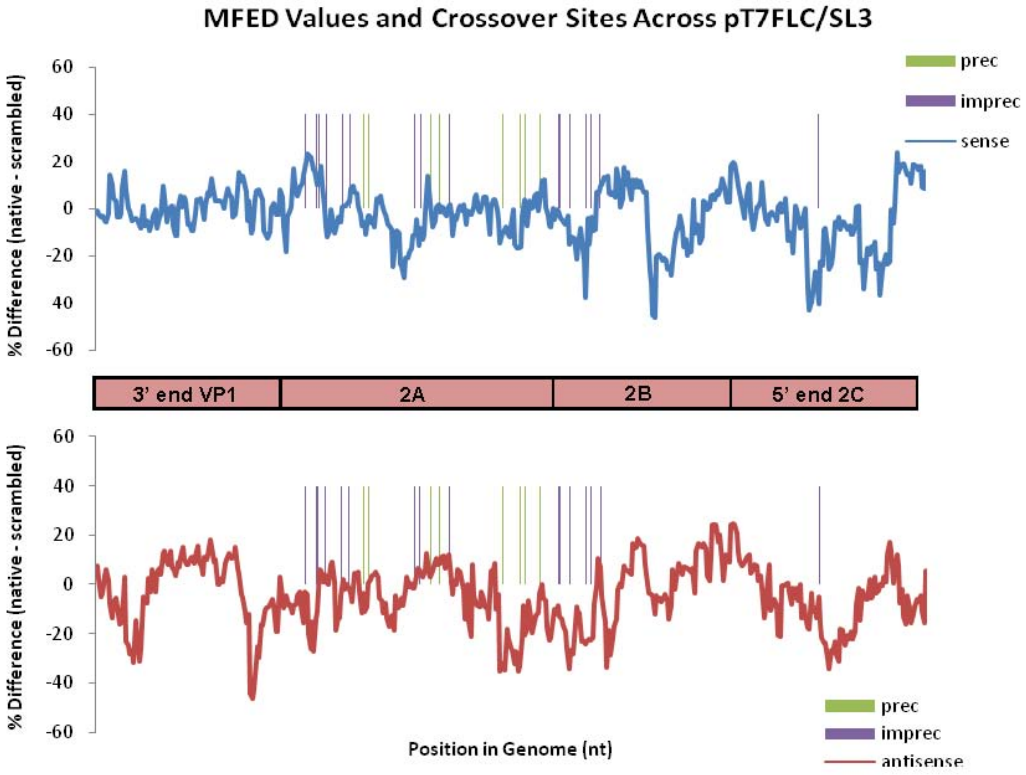


Figure 5.23 Local secondary structures using MFED analysis for PV1 and PV3 regions of recombination. (A) Minimal free energy difference readings throughout recombination region for donor pRLucWT RNA, both sense and antisense values (B) for acceptor pT7FLC/SL3 RNA. The precise and imprecise sites of template switching are added to the graphs as green and purple lines respectively.

Plotted crossover sites did not correlate visually with 'all peaks' or 'all troughs' in the graphs. The absence of stable local RNA secondary structure indicated that crossover sites did not correspond to structural elements during negative-strand synthesis.

RNA sequence and structural analysis above did not explain the location of crossover sites identified in PV3/PV1 recombinants. Junction sites were not evenly distributed throughout the available sequence space as expected, and the original recombinants incorporated a complete set of non-structural regions (namely 2A and 2B) together with fragments of sequence caused by the imprecise nature of junction sites. As already mentioned, a general pattern started to emerge that governed recombination events; the imprecise nature of most crossover junctions, the length of inserted sequence (≤ 513 nts), the lack of sequence identity associated with the junction site, and the incorporation of a complete 2A or 2B coding regions from either parent. It was possible that these were the consequence of early selection, and any other recombination event deviating from these 'rules' would not generate a viable virus.

The requirement for structural compatibility in chimeric proteins posed a reason for selective localisation of junctions after virus passage. Junction sites would be selected for those changes that did not interfere with normal virus protein function by disrupting structural integrity. Amino acid similarity scans were performed using a program in Simmonic 2005 Version 1.75 that scanned pairs of sequences using a Point Accepted Mutation (PAM) matrix developed by Margaret Dayhoff (Dayhoff *et al*., 1978). This compared closely related amino acid sequences and scored regular and rare substitutions according to chemically related amino acids. For example, amino acids that replaced each other regularly would score positively while amino acids that rarely replaced each other (from different groups) would score negatively. High scores were still given if conservative amino acids were interchanged (scores of 10 indicated similarities of sequence approaching 100%). Amino acid sequences were highly conserved within the targeted regions of PV1 and PV3 (averaging 9.64, SD = 0.57), suggesting that chimeric proteins would not alter protein structure significantly (figure 5.24). The average PAM value for the nine crossover sites at the approximate positions 3410, 3524, 3539, 3635, 3641, 3758, 3785, 3785, 3797 and 3821 was 9.61, SD = 0.47. There was no significant difference ($P = 0.87$) between

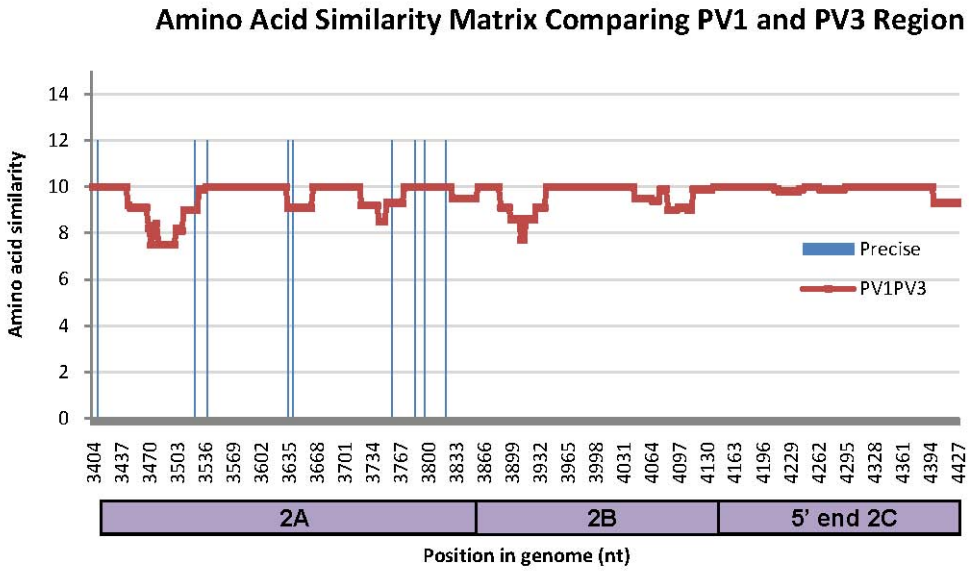


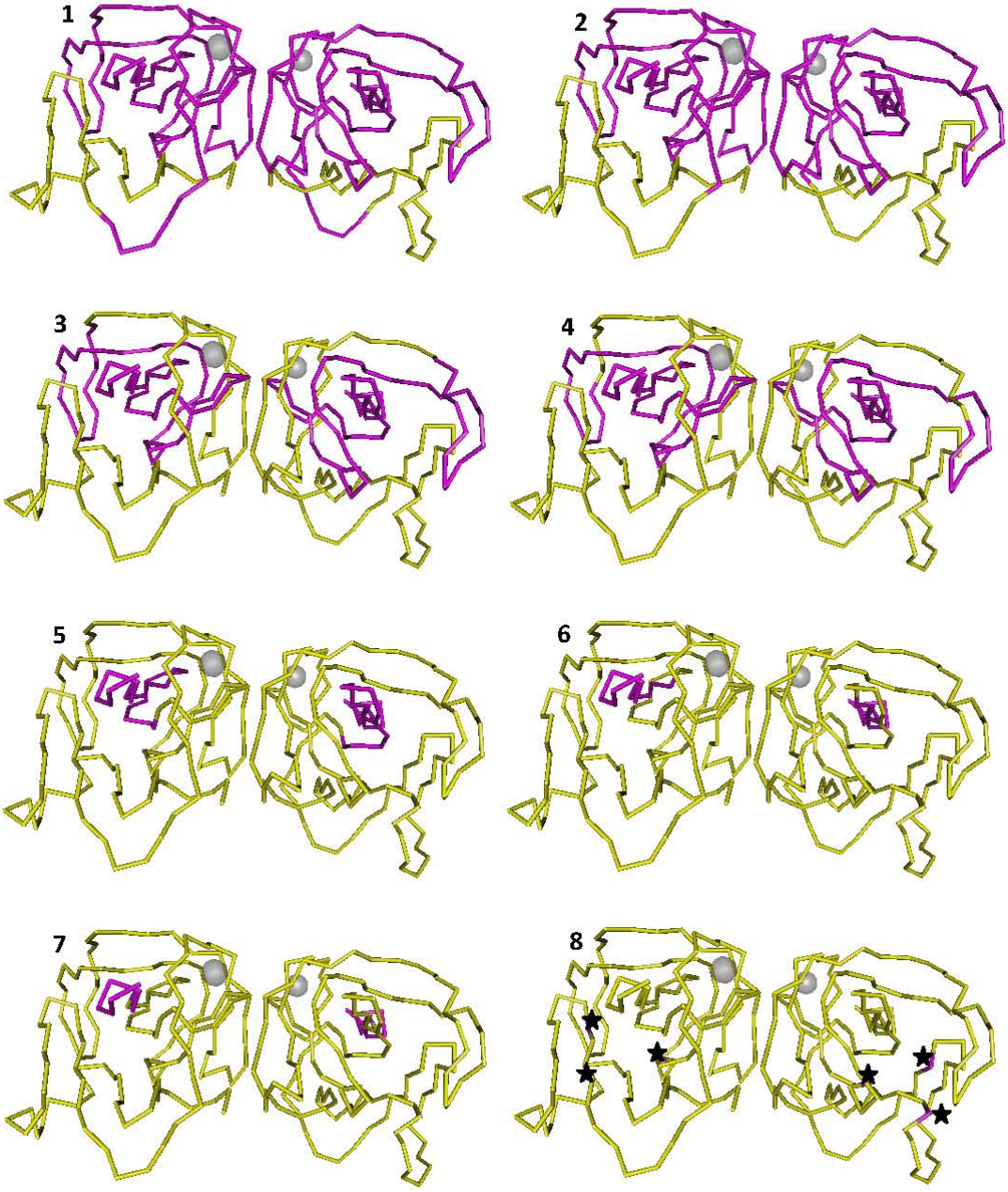
Figure 5.24 Amino acid similarity scan between PV1 and PV3. Analysis was performed using Simmonics 2005 Version 1.75. Comparisons of pairs of amino acid sequences were calculated within a sliding window size of 10 amino acids with a 3 nt overlap to generate data points across the recombination region. The precise crossover sites forming chimeric PV proteins were plotted according to the location within the region and were given arbitrary values to aid in visualisation. Amino acid similarity plots were calculated by PAM matrix, and this compared closely related amino acid sequences and scored regular and rare substitutions according to chemically related amino acids. Amino acids that replaced each other regularly scored positively while amino acids that rarely replaced each other (from different groups) scored negatively. High scores were still given if conservative amino acids were interchanged (scores of 10 indicated similarities of sequence in the window approaching 100%).

the mean PAM values of the crossover sites (mean = 9.61, 95% confidence interval [CI] 9.25, 9.97) and the mean PAM value for the entire sequence (mean = 9.64, 95% CI 9.58, 9.70). Considering the high degree of conservation, it was surprising that such precise chimeric recombinants were not identified in the early events that were captured.

To test whether precise crossover sites occurred away from sequences that encoded active regions of virus proteins, the positions of conserved domains of $2A^{pro}$ (as determined by crystallography of CB4 $2A^{pro}$) were superimposed onto recombinant junction sites of PV1 and PV3 (Baxter *et al.*, 2006). Poliovirus $2A^{pro}$ encodes a proteinase that cleaves virus proteins *in cis* and cellular proteins *in trans*, and similarly to CVB4 $2A^{pro}$, consists of a catalytic triad (H20, D38, C109), a four-stranded antiparallel β-sheet as the N-domain, a C-terminal domain containing a six-stranded antiparallel β-barrel, helical loops, an interdomain loop, and a C-terminal domain essential for virus RNA replication (Baxter *et al.*, 2006; Li *et al.*, 2001; Petersen *et al.*, 1999). The location of precise crossover junctions forming chimeric PV3/PV1 virus proteins was superimposed onto the crystal structure of HRV2-$2A^{pro}$ as well as onto PV3 sequence encoding purported motifs/domains (figure 5.25). The crystal structures in figure 5.25A demonstrate that recombination events did not occur at the 'interacting' faces of the 2A dimer, meaning the chimeric protein was either all PV3 or all PV1 in that region. Although 50% of crossovers occurred within β-strands, this would not significantly disrupt structures as PV1 and PV3 share high sequence identity in these regions. This must not have affected functionality of the virus protease as the resulting chimeric viruses remained infectious in cell culture. The potential cleavage sites of selected imprecise recombinants (14C, 16B, 25A, 34D, and 67B) were mapped according to the new recombinant polyprotein structure (appendix 1). A complete $2A^{pro}$ present in each construct should have been able to cleave excess protein from the polyprotein to retain normal viral characteristics.

The recombinant nature of intraspecies recombinants was investigated briefly. Results reflected those observed with PV3/PV1 recombinants. Recombinants with a complete set of virus proteins, and not those with chimeric forms due to junction sites, were the initial products of recombination. According to all documented crossover events in figure 5.17, a complete PV 2A encoding protein was included in

A Eight chimeric PV3/PV1 locations overlaid on dimer crystal structure of HRV2 2A protein. PV3 portion of protein (gold), PV1 portion of protein (purple).



Continued over page.

B

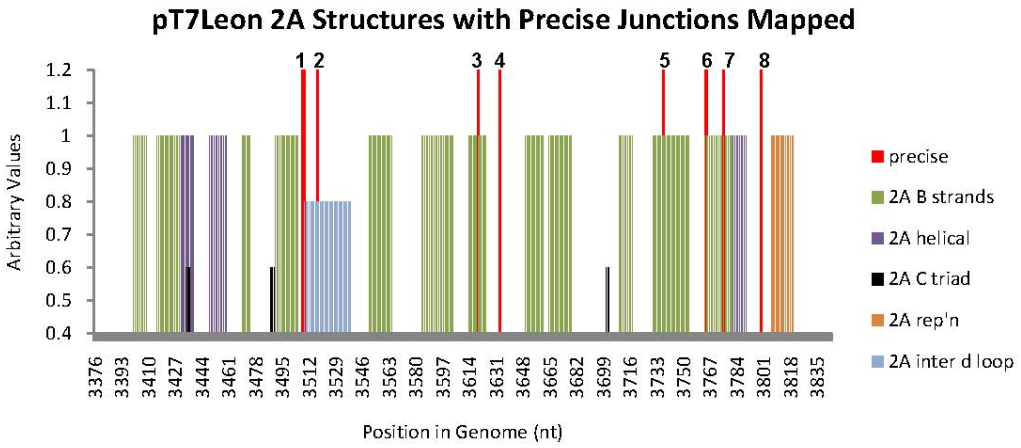


Figure 5.25 Comparing 2A^{pro} domain structures to precise junction sites. (A) Crystal structure of the dimer 2A proteinase from HRV2. The chimeric nature of the proteins (as derived from recombinants in this region), are colour-coded. Poliovirus serotype 3 portions are represented by gold tubes, while purple tubes highlight PV serotype 1 sequence. The catalytic triad is depicted by small black stars in structure 8 (His-18, Asp-35, and Cys-106). (Source: NCBI/Cn3D 4.1). (B) Location of structural features of PV3 2A^{pro} in relation to precise recombination junction sites of passaged recombinant progeny. The PV structural domains were extrapolated from the solution structure of CVB4 2A^{pro} (Baxter *et al.,* 2006) and mapped according to nt positions in the recombination region. The eight precise junction sites of recombinants are numbered 1 to 8 (also reflected in the individual dimers in part A). Y-axis values were applied to each category to allow separation of structural features.

the recombinant, to match the PV VP1 sequence, therefore the fragments of incorporating CVA21 2A were extraneous and served no purpose. Coxackievirus A21 genome shares approximately 80% amino acid sequence identity to PV, and the similarity scans in figure 5.26 emphasised divergence that would suggest issues of incompatibility within the protein (Xiao *et al.*, 2001).

Comparing natural isolate crossovers with those obtained during this project

Recombination events with insertions at the junction site are rarely reported in natural enterovirus isolates, whether from patients or from samples collected during sewerage surveillance etc. This is not surprising, since human isolates have undergone extensive rounds of replication and selection while in the host. This magnitude of selection pressure could repair mutations (such as insertions or deletions) at the crossover site, as was demonstrated during this project. The location of precise junction sites in PV recombinants reported in the literature, and from those acquired from this study, were plotted together for comparison.

Figure 5.27 depicts the approximate positions of several recombination junction sites derived from human isolates, and those recovered from two recombinants that were serially passaged through HeLa cells during this project. There is limited overlap of sites between both sets of recombinants, and sites appear to distribute evenly over the region. Recombination 'hot spots' have not been accentuated from this small exercise, and not all reported recombinants could be included if sequence data was not provided.

Correlation between donor and acceptor sites, and the size of insertions

A Spearman rank-order correlations coefficient analysis (Dytham, 2009) was performed on recovered populations of recombinants to determine whether the location of the junction site on the acceptor template was influenced by the location of the donor site for template switching. Locations of donor and acceptor sites were plotted along the targeted region of recombination only (being the same sized fragment for both templates) applied to the 16 distinct crossover junctions identified from one co-transfection with PV1 and PV3 partners (pRLucWT and pT7FLC/SL3) (table 5.1), and the three co-transfection experiments involving CVA21 containing RNA partners (table 5.3). The 2-tailed analysis of correlation ($P < 0.05$) assumed that

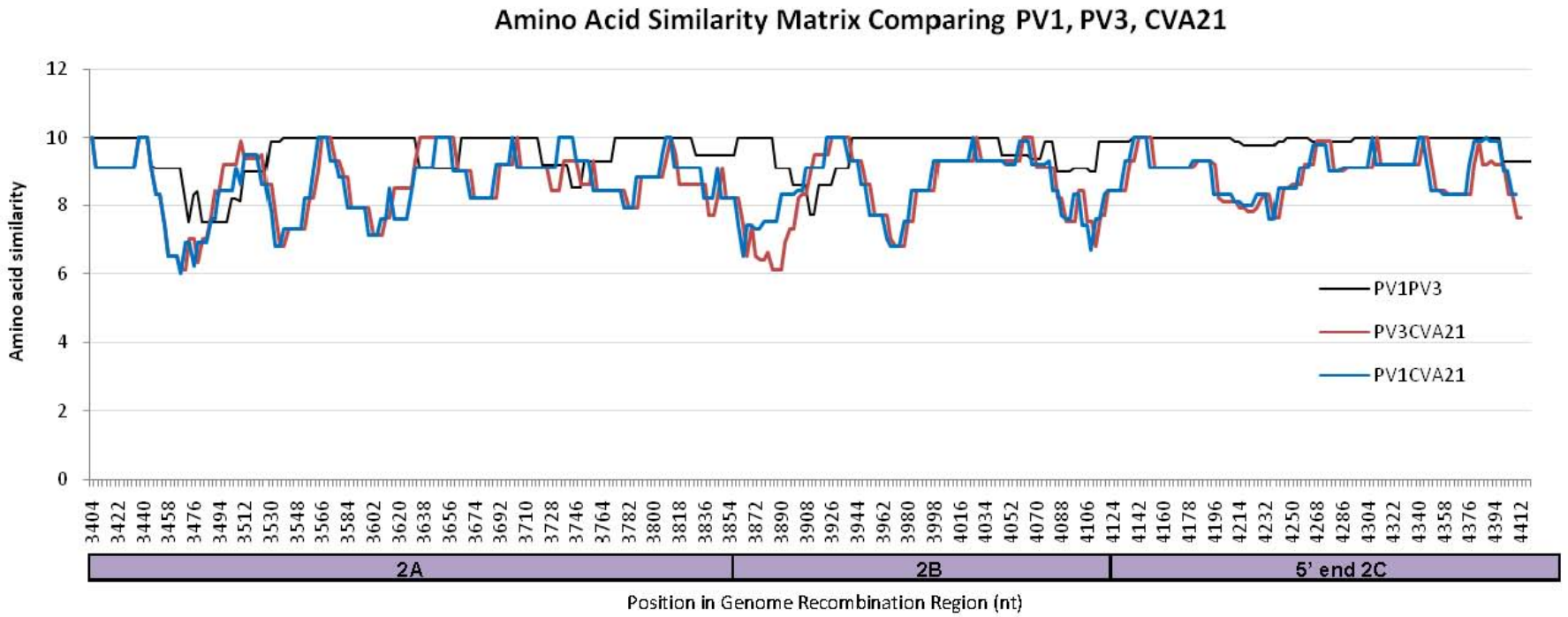


Figure 5.26 Amino acid similarity scans comparing PV1, PV3, and CVA21 protein sequences across the recombination region. Scans were performed using bioinformatics software Simmonic 2005 Version 1.75. Amino acid similarity plots were calculated by PAM matrix, and this compared closely related amino acid sequences and scored regular and rare substitutions according to chemically related amino acids. Amino acids that replaced each other regularly scored positively while amino acids that rarely replaced each other (from different groups) scored negatively. High scores were still given if conservative amino acids were interchanged (scores of 10 indicated similarities of sequence in the window approaching 100%).

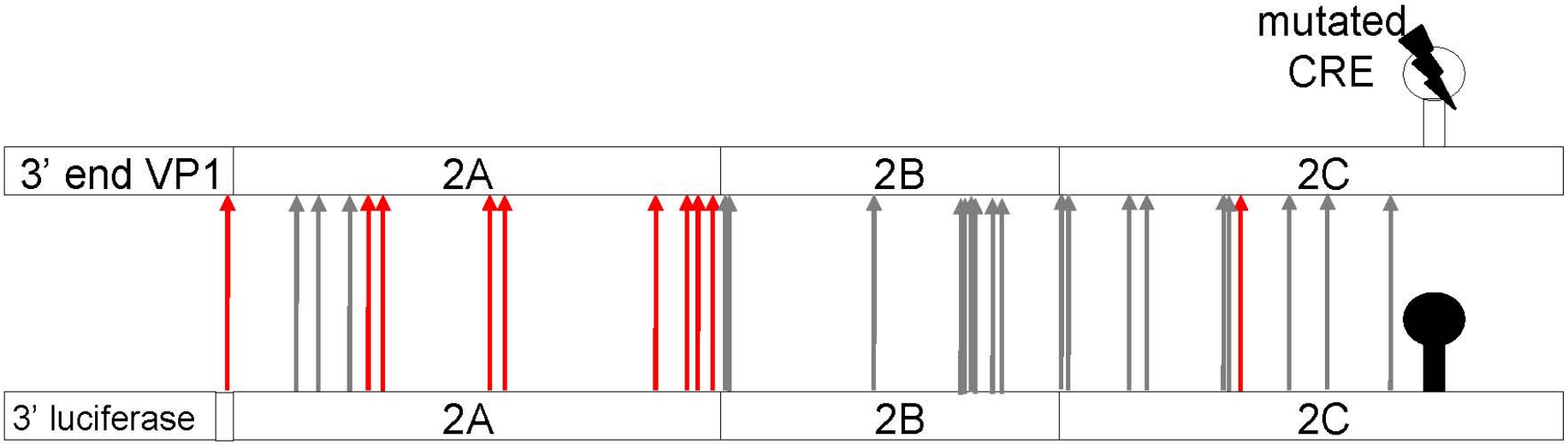


Figure 5.27 Crossover sites in PV3 and PV1 recombinants from this study and from the literature. New crossover sites for serially passaged recombinants 51G and 36B are shown in red arrows. Sites from natural isolates reported in the literature are shown with grey arrows, those in red are sites from this study.

the recombination sites were on a continuous scale (range ~1400 nt between the luciferase and CRE sequences) and did not assume a normal distribution. The hypotheses tested a null-hypothesis (H_0) = there was no significant relationship between donor and acceptor sites and an alternate hypothesis (H_a) = there was a significant relationship between donor and acceptor recombination sites.

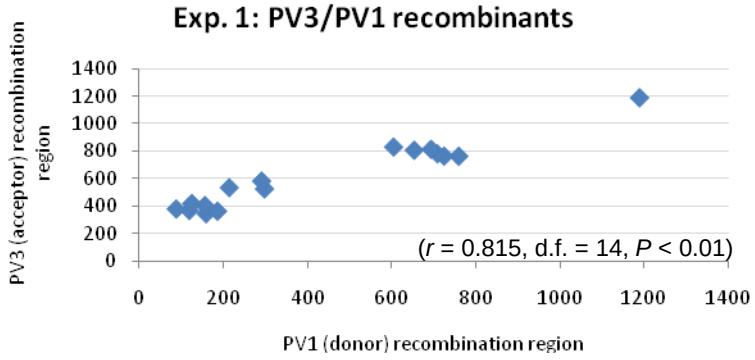
Figure 5.28 shows scatter graphs depicting the relationship between donor and acceptor sites for the recombinant virus populations analysed, as well as their *r* values.

The results suggested a significant positive association between PV3 and PV1 recombination sites ($r = 0.815$, d.f. = 14, $P < 0.01$) (experiment 1). Although the recombination sites were imprecise (except one), they showed strong association, yet the biological mechanism behind this strong statistical association remains unclear. No significant associations were found between junction sites for co-transfections involving pRLucWT + CVA21BKSabin1P1 or pT7Rep3-L + CVA21BKPV3P1 RNA partners in L929 cells (experiments 2 and 3). A perfectly negative association was found between PV3 and CVA21BKPV3P1 RNA partners in BsrT7 cells ($r = -1.0$, d.f. = 2, $P < 0.01$). These recombination sites were imprecise and were strongly negatively associated, meaning that the donor and template sites of recombination were linked by similar distances between both points.

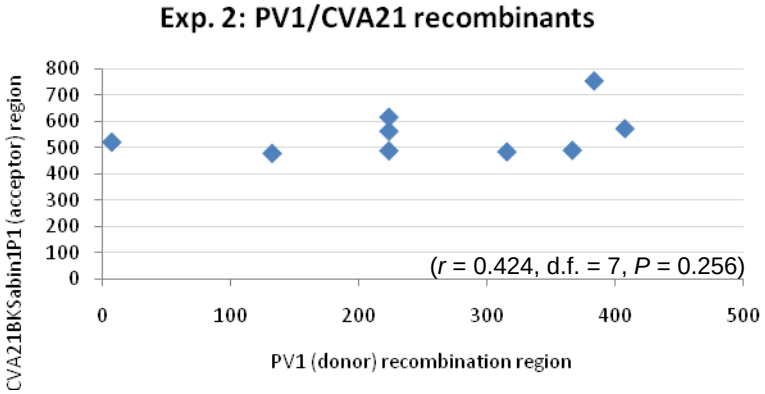
To conclude from this analysis, there could be an association between the donor and acceptor template sites, but this could only be confirmed and expanded at a later date with larger datasets.

The regions of template switching were also of interest and warranted further investigation. It was observed that once the RdRp removed itself (and the nascent RNA strand) from the donor template, its 'reattachment' to the acceptor strand did not extend beyond a certain distance, particularly from the 3' end of donor PV1 2A (figures 5.4 and 5.17). This suggested there were restrictions imposed on template switching distances between templates. In theory, the RdRp could switch from the 3' end of region 2A and reattach in region 2C, but this was not observed. One possible restriction would be that a packaging limit is almost certainly applied to an

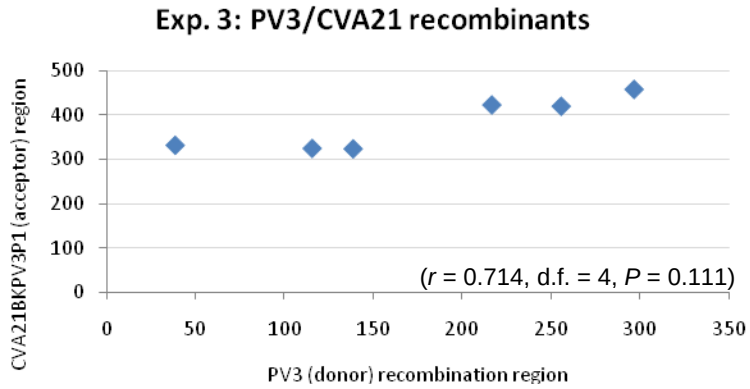
Experiment 1: PV3/PV1 recombinants from pRLucWT and pT7FLC/SL3 (table 5.1). Spearman rank-order correlation indicated a *significant positive association* between PV3 and PV1 recombination sites.



Experiment 2: Recombinants from pRLucWT and CVA21BKSabin1P1 cotransfections in BsrT7 cells (table 5.4). Spearman rank-order indicated *no significant positive association* between PV1 and CVA21 recombination sites.



Experiment 3: Recombinants from pT7Rep3-L and CVA21BKPV3P1 cotransfections in L929 cells (table 5.4). Spearman rank-order indicated *no significant positive association* between PV3 and CVA21 recombination sites.



Continued over page.

Experiment 4: Recombinants from pT7Rep3-L and CVA21BKPV3P1 cotransfections in BsrT7 cells (table 5.4). Spearman rank-order indicated a *significant negative association* between PV3 and CVA21 recombination sites.

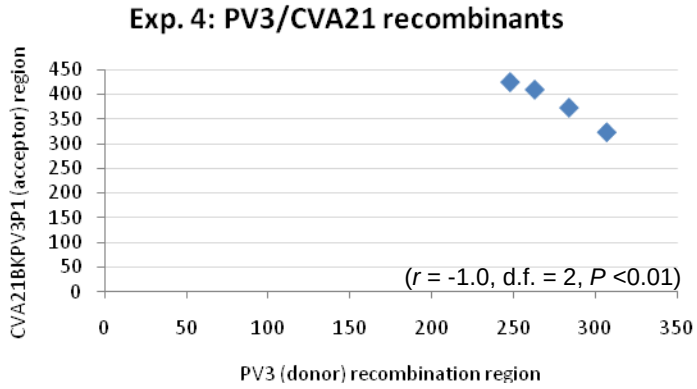


Figure 5.28 Correlation coefficients for locations of junction sites in recombinants recovered. Locations of donor and acceptor sites were plotted along the targeted region of recombination only (being the same sized fragment for both templates). The corresponding genome fragments for donor and template are plotted on the X- and Y- axis, but exact nt positions in the genome are not provided (fragment locations of recombination occur between ~ 3404 to 4465 nts in aligned PV1 and PVl3 sequences). This was applied to the 16 distinct crossover junctions identified from one co-transfection with PV1 and PV3 partners (pRLucWT and pT7FLC/SL3) (table 5.1), and the three co-transfection experiments involving CVA21 containing RNA partners (table 5.4). Spearman rank-order correlation indicates a significant positive (or negative) association between X and Y (r = value, degrees of freedom = number of pairs – 2, P <0.05). Values of -1 or 1 are perfectly negatively or positively correlated.

RNA genome with inserted sequence, and this could be approximately 513 nt (the largest insertion size observed in this study). This is discussed in chapter seven. Preliminary analysis was performed to determine whether the length of inserted sequence was related to the location of the RdRp switching from the donor template. Figure 5.29 includes graphs that plotted the nt difference from donor to acceptor template locations, versus the donor location in the recombination region where the RdRp removed itself.

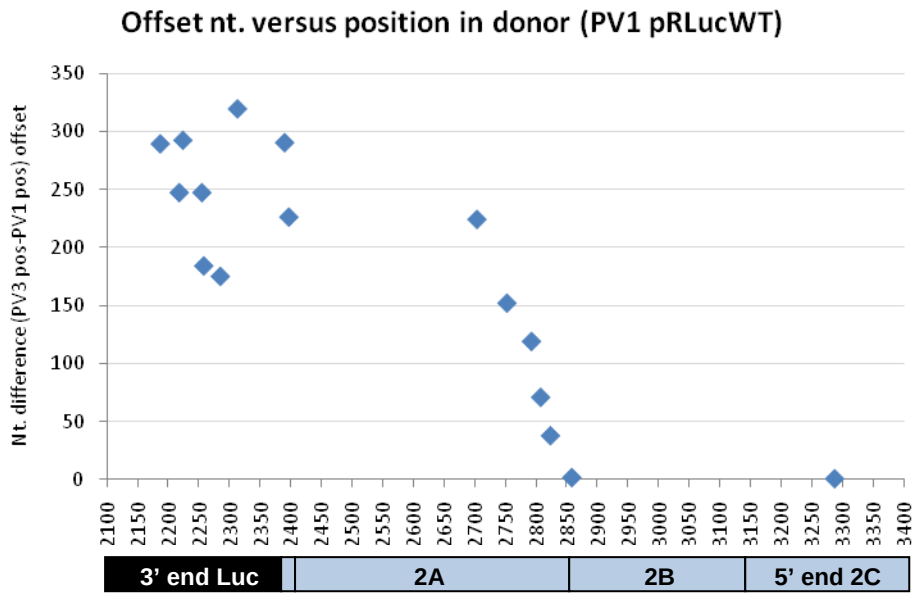
In mapping the distance the RdRp travelled from donor to acceptor templates, it was noted that the RdRp reattached on the acceptor template at indiscriminate distances from the luciferase sequence on the donor stand (clear in experiment 1, figure 5.29). Results from all other experiments suggested that as more donor strand was incorporated at the crossover site, this was offset by the addition of less acceptor template at the junction. In other words, the RdRp travelled a shorter distance (measured in nts) from donor to acceptor templates.

5.5 Discussion

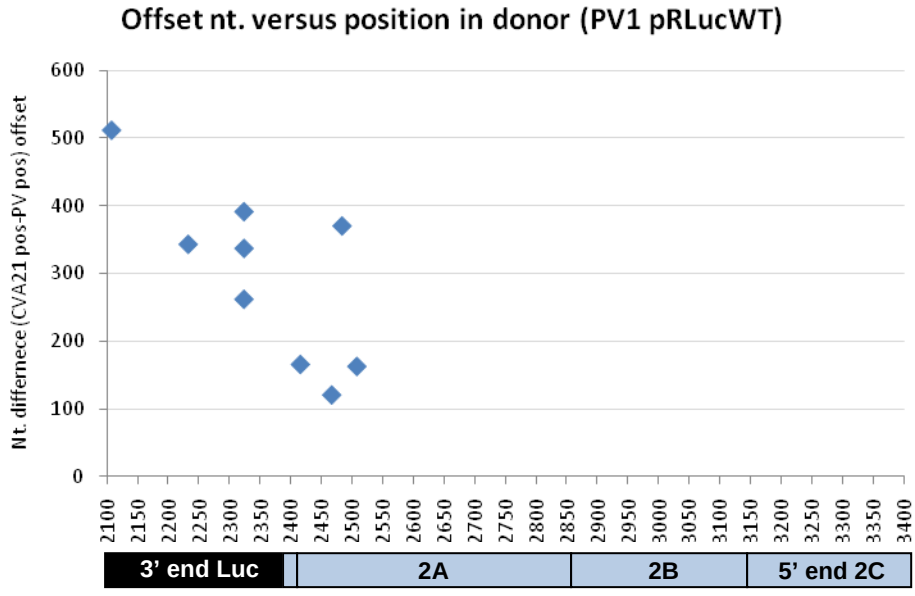
All viruses recovered from the system that was developed during this project underwent recombination in the region targeted during the experimental design. In all cases, recombination occurred predominantly in the non-structural 2A region between replication competent virus RNA partners. Upon generating a population of natural PV3/PV1 recombinants *in vitro*, it remained to be determined if template switching correlated with specific events at the genome level. Substantial lengths of extra sequence were incorporated at junction sites, and preliminary analysis indicated that these additional sequences were lost from such recombinants during subsequent rounds of replication. This method showed that viable virus was generated by imprecise recombination between two complementary RNA sequences, indicating that close interactions across templates occurred, but not necessarily at corresponding sequences.

The unanticipated nature of recombinants signified very early events in the virus life cycle which are rarely available for study because of experimental design. Recombinants identified from isolates associated with enterovirus outbreaks have

Experiment 1: PV3/PV1 recombinants from pRLucWT and pT7FLC/SL3 (table 5.1).

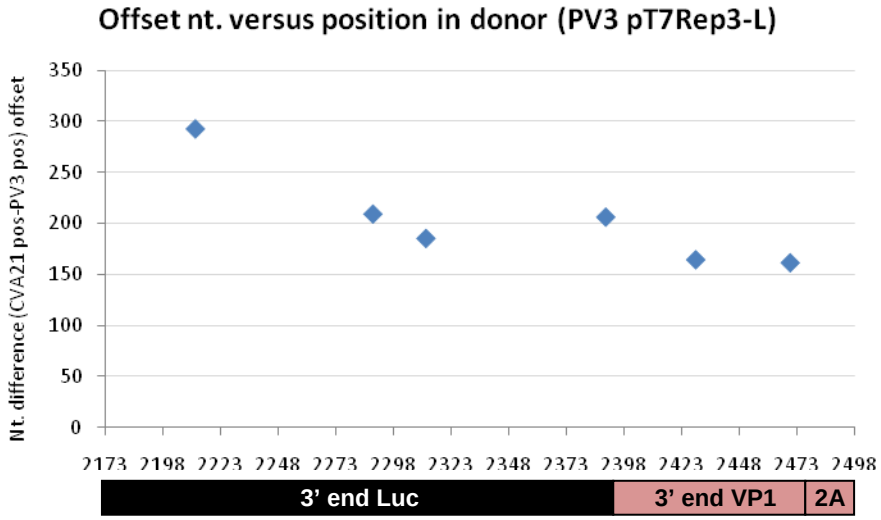


Experiment 2: Recombinants from pRLucWT and CVA21BKSabin1P1 cotransfections in BsrT7 cells (table 5.4).



Continued over page.

Experiment 3: Recombinants from pT7Rep3-L and CVA21BKPV3P1cotransfections in L929 cells (table 5.4).



Experiment 4: Recombinants from pT7Rep3-L and CVA21BKPV3P1cotransfections in BsrT7 cells (table 5.4).

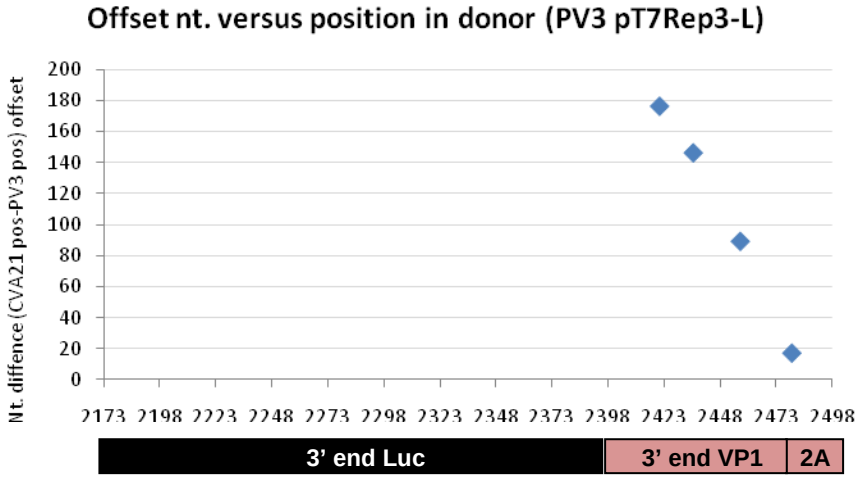


Figure 5.29 Distance of template switching from original displacement from donor template. X-axis is the genome position where the RdRp left the donor template. The distance from donor to acceptor template was calculated from Clustal alignments containing full sequence data of each RNA template.

been exposed to numerous selective pressures *in vivo* (i.e. several levels/layers of complex host systems and host resistance mechanisms) and may therefore reflect highly adapted viruses. Although studying recombination using RNA fragments has been reported previously whether in picornaviruses or plant viruses, no results have shown the frequency and variability of insertions at crossover junctions on this scale (Cascone *et al.*, 1990; Chetverin *et al.*, 1997; Gmyl *et al.*, 1999; Gmyl *et al.*, 2003; Raju *et al.*, 1995). One study does describe a deletion at the crossover site, and this phenomenon was reported in recombinants generated from Sindbis virus RNA transcripts, where in one instance, the crossover site contained a 76 nt deletion (Raju *et al.*, 1995).

The overall data suggests that recombination is at least a two stage process that initially produced random chimeras combining two RNA partners in frame before further rounds of replication. The nature of the recombinants recovered suggests that template switching occurred in a 'promiscuous' and unpredictable manner, particularly with the unexpected insertion of non-virus sequence in the final virus genome. Such recombinants were eventually corrected so that the junctions no longer contained any insertions – this was aided by rapid passages at high MOI and selection through susceptible cells. To demonstrate the rapid evolutionary properties of RNA viruses, further passages of original clonal virus populations produced new heterogeneous progeny of recombinants containing several variants (five different crossovers identified from one clone). The generation of quasispecies and continuous recombination served to stabilise genomes and remove unfavourable mutations while creating newly defined genomes. Ideally, it would have been beneficial to perform single-step growth curves to compare growth characteristics of passaged recombinants to their original seed and wild-type parents.

Tolskaya *et al* had previously found that several non-random interserotypic PV recombination events took place within RNA secondary structure elements (Tolskaya *et al.*, 1987). Crossover sites were mapped on a model of the secondary structure of a 793 nt RNA region of PV1 (between 3823 and 4616 nts) which corresponded to the region of interest in the current project. Crossover sites fell within hairpin-like structures, with the majority showing unpaired nucleotide sequences which would seem favourable for acceptor templates. Unfortunately, most analysis was carried out

on a large segment of RNA sequence with extensive secondary structure, suggesting that all crossover sites were located amid this secondary structure. This analysis predated more accurate and current RNA folding methods and would be better served with smaller fragments for analysis. The only conserved secondary RNA structure in this region is the CRE, which was not evident within the potential folding figure in this publication.

There was no evidence for the occurrence of multiple crossover events within the regions, although this could not be discounted for the remainder of the unsequenced genome. Recombination seemed to occur in preferred locations (5' end of 2A or 2B) as the crossover sites were not evenly distributed throughout the region, but it was difficult to identify factors relating to nucleotide or amino acid sequences that provided evidence for preferences to those sites. The region of recombination was relatively unstructured as a whole, but contained several potential local RNA structures and sequence motifs (such as stem loops and homopolymers) that did not correlate strongly with junction sites.

CHAPTER SIX: Construction of RNA Partners for Co-transfections

6.1 Introduction

Chapter four demonstrated proof of concept by using two RNA molecules to create recombinant species C viruses comprising sequences from PV1 and PV3 *in vitro*. A number of recombinant viruses were recovered in the absence of any background full-length parental PV1 or PV3 viruses. These isolated, recombinant species C viruses contained sequence similarities to recombinant clinical isolates obtained from humans (Cuervo *et al.*, 2001; Georgescu *et al.*, 1994) and suggested that this approach could be used to simulate intraspecies recombination events that may arise *in vivo*. The new model described in chapter four may also be expanded to study virus recombination between different species of enterovirus (i.e. interspecies recombination). A review of the literature suggests that interspecies recombination is limited to viruses of the same host species and tropism, those that share close phylogenetic relationships, and happen to be circulating concurrently (Bolanaki *et al.*, 2007; Yozwiak *et al.*, 2010). These mostly host related restrictions make it difficult to determine if virus recombination between species is possible at the genomic level. The method described in chapter four used naked RNA transfected *in vitro* in order to bypass cell entry processes which would otherwise restrict viral tropism (such as the need for specific host cell receptors).

The method relied on two defective virus sequences (herein described as "partners") that could recombine within a single host cell to create viable progeny virus. Defective virus genomes were derived from plasmids encoding several forms of enterovirus cDNA, such as subgenomic replicons and full-length virus cDNA containing a mutated CRE sequence. Mutating the CRE would limit virus RNA replication to the negative-strand, meaning it would be semi-replication competent. The construction of full-length virus cDNA clones and their derivatives are discussed in chapters one and four.

Echovirus 30 (E30) belongs to species B enteroviruses and is the major cause of meningitis in both children and adults (Khetsuriani *et al.*, 2006; Rice *et al.*, 1995). Surges of E30 transmission correlate with distinct new genomic lineages which

usually replace older circulating ones. High frequencies of E30 detection and isolation occur over large geographical areas and phylogenetic analysis of strains suggests that E30 sequence diversity differs between the structural and non-structural regions of the genome. Recombination with unidentified species B viruses, evident in the non-structural regions (McWilliam Leitch *et al.*, 2009; Mirand *et al.*, 2007) may play an important role during outbreaks of E30. The outbreak patterns of circulation and continuous dynamic changes to genomic lineage make it an ideal virus to study further by incorporation into recombination experiments. To date there have been no publications identified describing the construction of a full-length infectious clone or subgenomic replicon for E30.

Chapter four described the use of defective virus RNA partners in co-transfection experiments for generating recombinant viruses for further characterisation. The aim here was to incorporate E30 RNA into such experiments by constructing either a defective full-length clone with mutated CRE, or a subgenomic replicon for E30. Both forms of virus RNA would fit the requirements to generate infectious virus during co-transfections with other virus RNA partners.

Three E30 isolates were provided by Professor Peter Simmonds (Centre for Infectious Diseases, University of Edinburgh) and had been obtained as part of an ongoing geographical and evolutionary epidemiological study of E30 transmission worldwide over an eight year period (McWilliam Leitch *et al.*, 2009). Because of the nature of the molecular study, only small regions of the E30 genome (VP1 and 3Dpol) had been sequenced. As the complete E30 virus sequence was unknown, the complete genome could be amplified by long distance PCR using oligonucleotides targeted to the conserved 5' and 3' ends of enterovirus genomes (Lindberg *et al.*, 1997). A general strategy for the construction of an E30 infectious clone into pRiboI vector using restriction enzymes for ligations is outlined in figure 6.1.

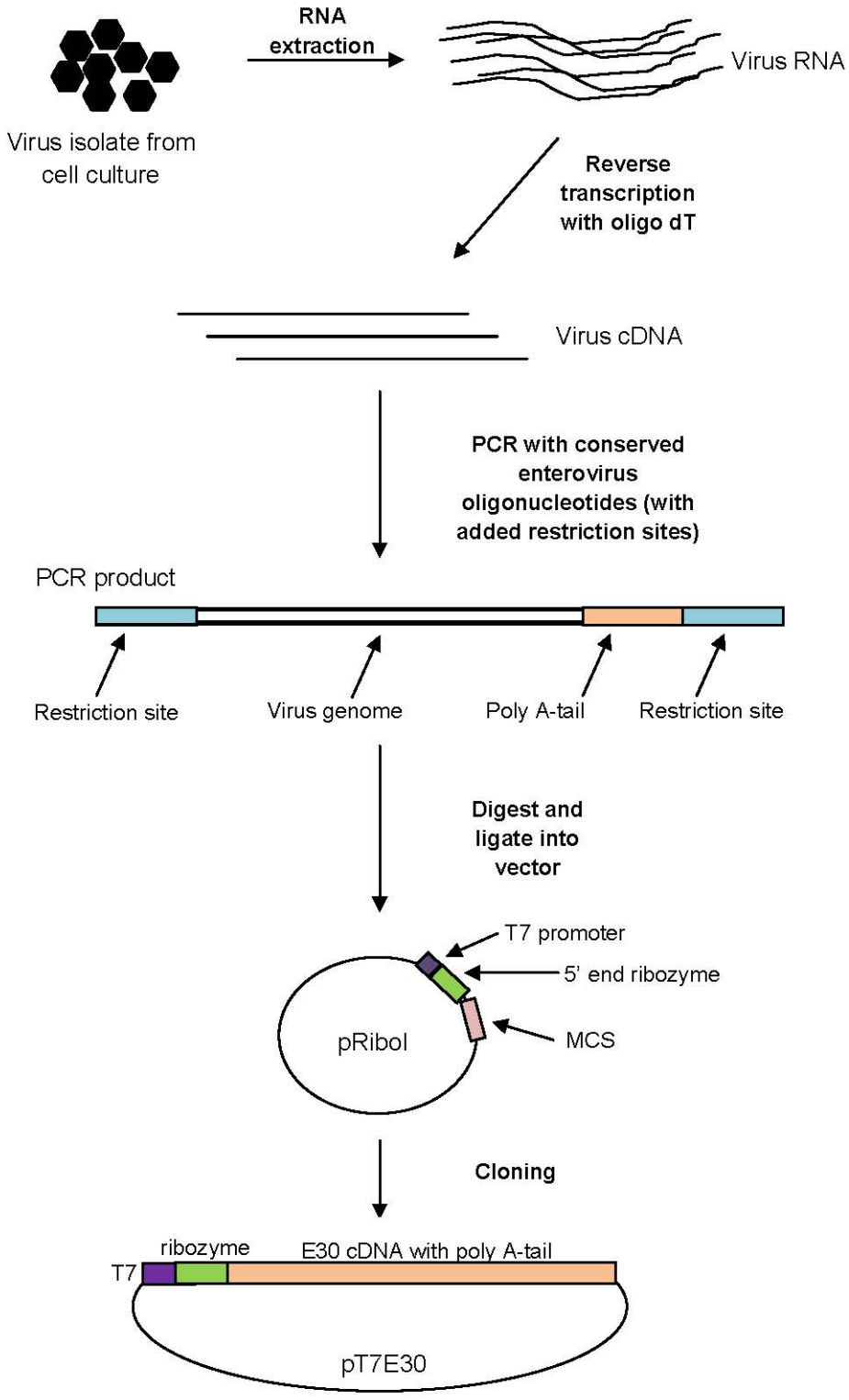


Figure 6.1 Overview of proposed E30 full-length infectious clone construction. Virus RNA is extracted from virus cultured in RD cells. Virus cDNA is synthesised from RNA and complete genome amplified by PCR using conserved enterovirus oligonucleotides with inserted restriction sites. Virus amplicon and pRiboI are digested and ligated together to form a full-length clone for E30. This will be tested for infectivity at a later date.

Plasmids containing insert of the correct size would be further characterised for infectivity. Ribonucleic acid transcribed *in vitro* from cloned virus cDNA would be transfected into susceptible cells and assessed for the generation of viable virus particles. A full-length E30 cDNA clone capable of generating virus would be sequenced in its entirety and used for additional cloning strategies.

Aims

The aim was to construct species B RNA partners for co-transfections to make populations of recombinant viruses in cell culture. This would allow the study of interspecies recombination as species C and D RNA partners were already available. The intent was two-fold. Firstly, in order to design E30 RNA partners, this necessitated the initial construction of a full-length infectious E30 cDNA. With the aid of a full-length E30 cDNA clone, it would be easier to construct a subgenomic replicon or mutate the CRE by modifying the plasmid by molecular cloning techniques. Secondly, to mutate the CRE in the full-length pT7E7 infectious clone in order to inhibit virus production yet allow synthesis of negative RNA strands, as previously demonstrated by Goodfellow *et al*., 2000 for the pT7FLC/SL3 construct. These two aims were necessary in order to complete a larger objective which was performed and discussed in chapter four; to integrate species B partners with species C and D partners as part of interspecies co-transfection experiments to determine if it was technically possible to generate chimeric interspecies viruses by bypassing host cell receptors and cell entry *in vitro*.

6.2 Generation of a Full-length Infectious E30 cDNA

Amplification of E30 genome by RT-PCR

Isolates Fi13 (Finland), GB27, and GB37 (both Great Britain) were amplified by passage in RD cells. Ribonucleic acid was extracted from tissue culture using the RNAeasy Mini kit (Qiagen) according to the manufacturer's instructions. Reverse transcription was performed using Superscript II Reverse Transcriptase (Invitrogen) and the oligonucleotide MJREV containing an extended T region (table 2.3), with a longer extension step of 50 min at 46°C. This was necessary to ensure complete transcription to the extreme 5' end of the genome. The E30 genome was amplified by PCR using conditions optimised during supervision of an undergraduate student

project (Megan Jackson). Briefly, the forward oligonucleotide (E30-1FClaI) (table 2.3) contained the first 27 bases corresponding to those of the start of E30 (as determined by sequences in GenBank) as well as a *Cla*I restriction site to aid in subsequent cloning into the pRiboI vector (Herold & Andino, 2000). The reverse oligonucleotide used was the same as that used for RT (MJREV), and included a run of 29 thymine bases in order to incorporate a longer and more stable poly(A)-tail at the 3' end of the virus transcript. The KOD polymerase (Novagen) was effective in amplifying the entire E30 genome with cDNA produced, and was applied using 30 cycles with an annealing temperature of 61°C and a 6 min extension incubation. The higher annealing temperature was required for priming the longer oligonucleotides to template and was optimised by gradient PCR (data not shown). A 5-6 min extension incubation was recommended by the polymerase manufacturer. The outcome was an amplicon of approximately 7.5 kbp (figure 6.2A), with smaller non-specific bands evident at 3 kbp and 1.7 kbp.

Restriction digestions to determine restriction enzyme sites in virus sequence

As these virus isolates had not been previously sequenced, it was important to determine which restriction enzymes could digest the virus sequence for future cloning procedures. A *Sal*I digest produced a similar banding pattern to the original PCR product, indicating the absence of *Sal*I sites within the virus genome. This did not exclude *Sal*I sites from the genome altogether, as one might exist near the ends of the genome which makes visualisation of products difficult on an agarose gel. Single (not shown) and double digests with *Sma*I and *Sal*I showed the existence of at least three *Sma*I sites limiting the use of the pRiboI vector (figure 6.2B). Digestions were incomplete, as full-length product (~7.5 kbp) remained. Strains Fi13 and GB37 indicated the presence of five gel bands (~7.5 kbp, faint ~5.75kbp doublet, ~4.5 kbp, ~2.5 kbp, and ~0.85 kbp) meaning that three possible *Sma*I sites existed if undigested band fragments were taken into account. The strain GB27 digest contained five gel bands, but the extra band at approximately 1.7 kbp was evident in the original PCR product. With regards to pRiboI, the 5' terminus of virus cDNA could only be cloned immediately following the *cis*-acting hammerhead ribozyme using the single *Sma*I site contained within the ribozyme cDNA (Herold & Andino, 2000).

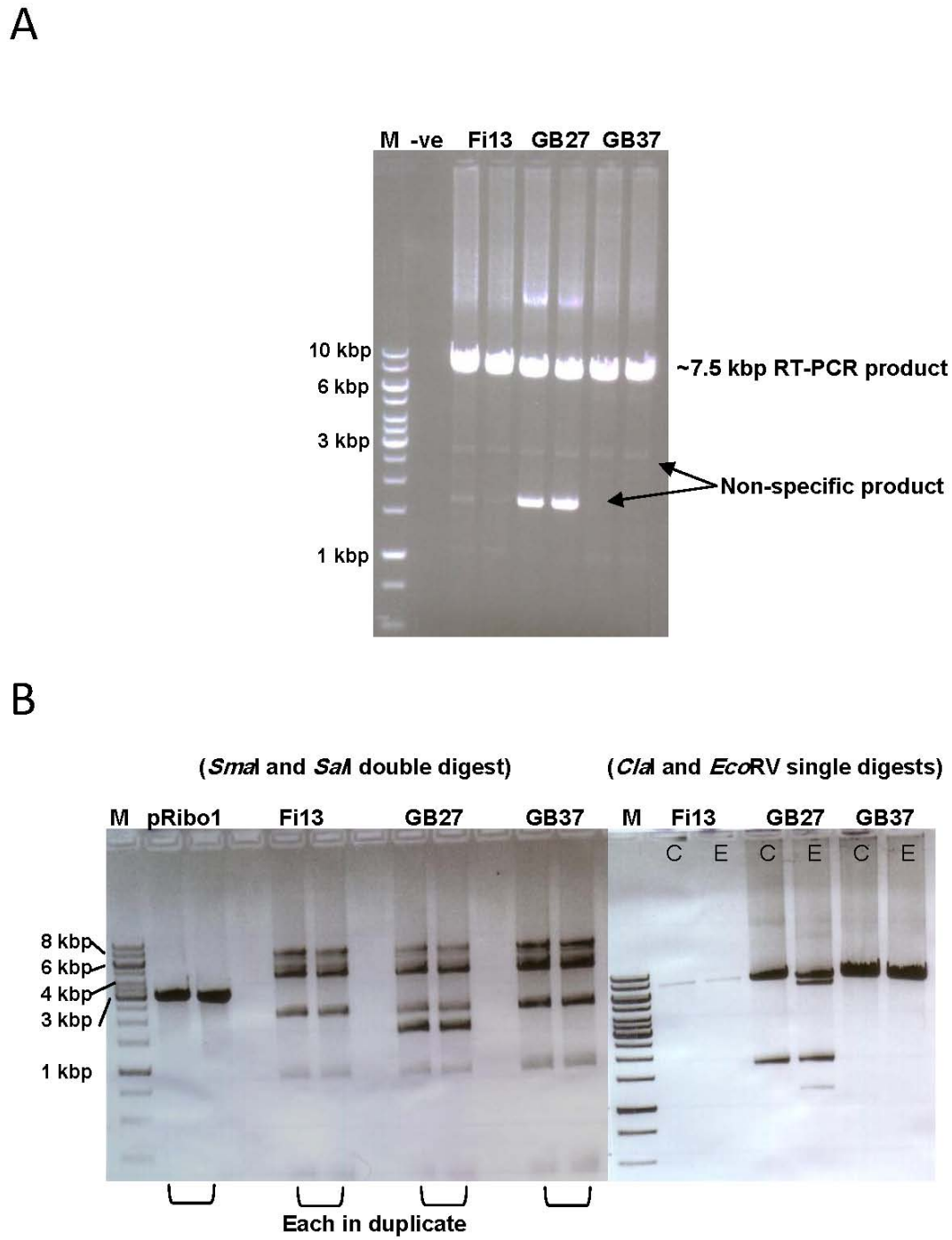


Figure 6.2 Amplification and restriction digests of E30 cDNA. (A) Agarose gel of full-length E30 RT-PCR product, stained with ethidium bromide. PCR was performed in duplicate per strain; -ve indicates PCR control with no template (H_2O). (B) Restriction enzyme digests to confirm cut sites present in the virus sample. Left gel indicates the presence of several *Sma*I or *Sal*I sites within the virus sequence, while the right gel shows that one *Eco*RV site (E) is present in virus strain GB27. The three strains of E30 (Fi13, GB27, and GB37) do not contain *Cla*I restriction sites (C) (see text).

Virus cDNA containing several *Sma*I sites would make this difficult to achieve. Single digests with *Cla*I and *Eco*RV showed that the viruses most likely did not contain a *Cla*I site (figure 6.2B), but indicated the existence of an *Eco*RV site within E30 virus strain GB27 (two extra bands appear in lane at ~6.5 kbp and 1.25 kbp). Multiple *Eco*RI restriction sites were detected in all strains (data not shown).

To conclude, cloning could not easily be performed using available restriction sites in pRiboI because such sites were present in the virus genomes and would increase the complexity of cloning required.

Cloning of full-length RT-PCR product into pGEM®-T Easy vector

As the sequence (and therefore restriction sites) was unknown for the virus isolates, it was apparent that blunt-end cloning would be a useful method for ligating full-length virus PCR product into a single vector without the need for restriction digests. Full-length amplicon could be cloned directly into linearised vectors if the DNA polymerase used during PCR generated single A overhangs attached to the 3' end of amplified fragments. An amplicon with 3'-A overhangs can be ligated directly into linear vectors containing single 3' terminal T overhangs at insertion sites. These A and T overhangs flanking amplicon and vector respectively, improve the efficiency of ligation by providing compatible bases for attachment.

A full-length virus amplicon was therefore purified from 1% agarose gel in order to remove non-specific PCR product using the DNA Extraction Kit (Fermentas). To increase ligation efficiency, the purified amplicon was further modified to incorporate 3'A residues using the *Taq* polymerase kit (Fermentas) and cleaned using a High Pure PCR Clean Up Micro Kit (Roche) in order to remove excess dATPs that could interfere with cloning. Full-length PCR product (with 3' A overhangs) for all three E30 strains were cloned into linearised pGEM®-T Easy vector using the kit supplied by Promega. The 3015 bp pGEM®-T easy vector contains a single 3' T extension at each end to aid in efficient ligation, and includes T7 and SP6 RNA polymerase promoters that flank the multiple cloning site (MCS). A diagrammatic representation of this cloning method is shown in (figure 6.3A).

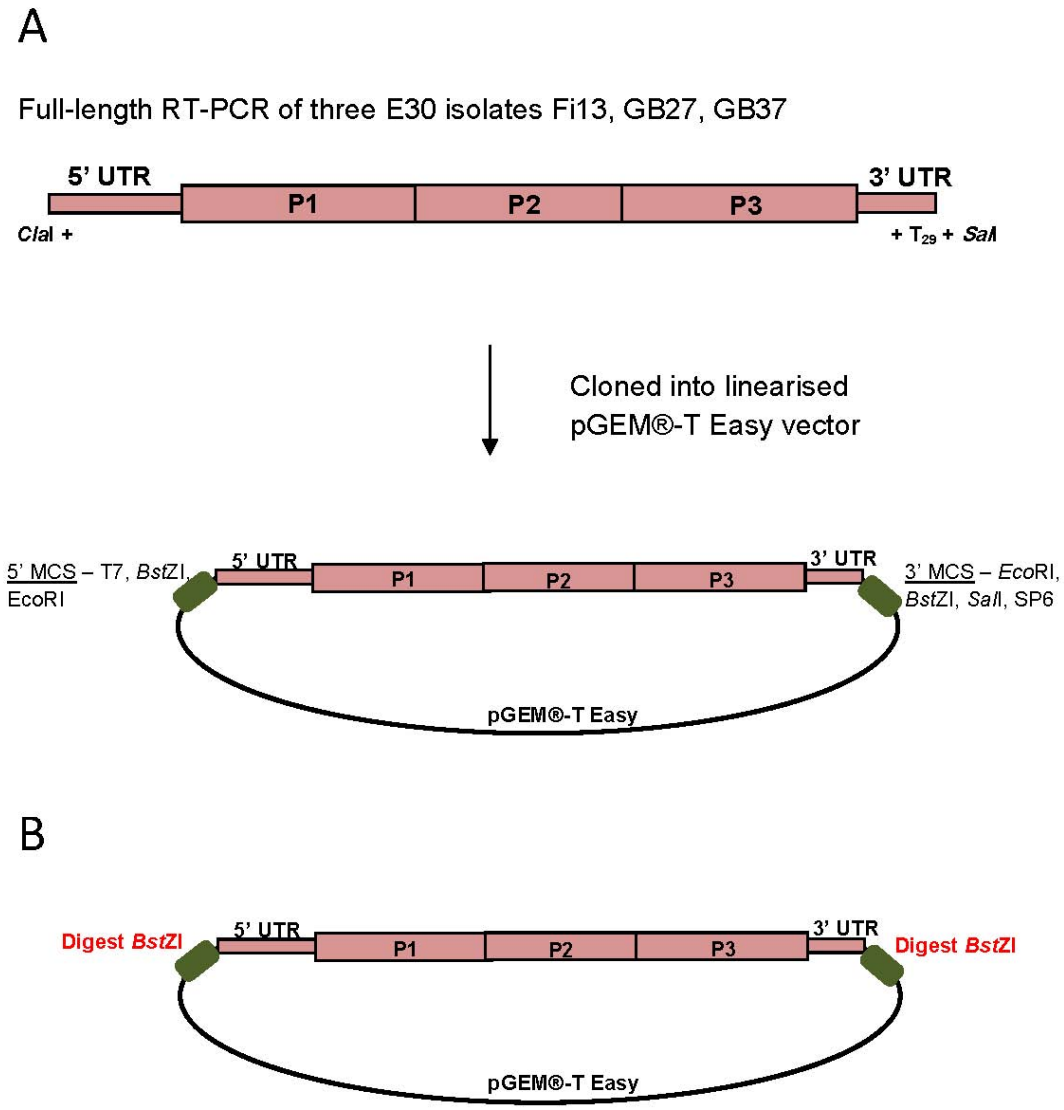


Figure 6.3 Construction of pT7E30 full-length infection clone for E30. (A) The entire amplified genome of E30 (strains Fi13, GB27, and GB37) was separately cloned into linearised commercial vector pGEM®-T Easy. The 5' MCS begins with a T7 promoter, and the 3' MCS ends with an SP6 promoter in reverse orientation. (B) As this was blunt-ended cloning, virus cDNA could be inserted in either orientation. To reorientate the genome in order to use T7 promoter, the plasmid was digested with *Bst*ZI and re-ligated with all components present, therefore removing the genome and small flanking MCS. Resulting clones were analysed for correct T7 orientation by sequence analysis.

This cloning strategy proved to be inefficient, with only six clones containing the appropriately sized DNA recovered from several hundred clones for all three strains. One clone was recovered for strain Fi13, three clones for strain GB27, and two full-length clones for strain GB37. Single and double digests were performed with *Cla*I and *Sal*I on all six plasmids to confirm the presence of virus sequence of correct length and of these, only three plasmids contained the entire virus sequence, one for each strain. The double digest displayed two DNA bands on the agarose gel, the smaller (3015 bp) representing the vector backbone, and larger band (~7500 bp) being virus cDNA. As this was essentially achieved by blunt-end ligation, the orientation of virus sequence in the plasmid was also determined by restriction digestion using *Sal* I alone. A single *Sal* I site was present in the vector 3' end of the MCS (when linear) and this digest would determine whether the virus was in reverse (a *Sal*I site at each end of virus) or in the correct orientation with both *Sal*I sites at the 3' end of the virus (figure 6.3A).

The orientation of virus cDNA in pGEM®-T Easy was determined in order to establish which promoter (T7 or SP6) would be required to transcribe RNA from cDNA. The insertion site in pGEM®-T Easy was flanked by either a T7 or SP6 promoter. As such, all three strains of E30 were cloned in reverse with regard to the T7 promoter (3' to 5') and would require the SP6 promoter for transcription. The use of the standardised T7 promoter was preferred and therefore, the virus sequence was digested out of the vector with *BstZ*I to remove the virus sequence in its entirety from the backbone before re-ligating it back in the opposite direction (figure 6.3B). A *BstZ*I restriction site was located on either side of the inserted virus cDNA in the MCS and was not present in the virus (data not shown). The virus cDNA and vector backbone was gel extracted from 1% agarose and religated together once the vector was dephosphorylated with alkaline phosphatase in order to remove 5' phosphate groups to prevent religation to itself. The cloning process was repeated and plasmids were isolated with the desired orientation of insert using the T7 promoter for RNA transcription. These new clones were designated pT7E30 Fi13 (15.6, 15.7, 15.16), pT7E30 GB27 (27.29), and pT7E30 GB37 (2.44). The numbers in parenthesis indicated clone numbers belonging to each strain.

Sequence analysis of E30 GB37 cDNA

As one clone from those above proved to be infectious (details in the following section), it was sequenced in its entirety. The pT7E30 GB37 (2.44) cDNA sequence, which was 7428 nt long excluding the A tail, contained an open reading frame of 2194 amino acids, had a 5' NCR of 743 nt, and was followed by a shorter 3' NCR containing 103 nt and poly$(A)_{29}$ tract. The nucleotide identity of E30 GB37 was compared to four full-length E30 genomes available in GenBank (figure 6.4A). It shared between 81.2-84.4% nucleotide identity with the database strains, including two separate submissions for prototype strain Bastianni (accession numbers AF162711 and AF311938). The E30 GB37 strain appeared as a more recently evolved strain from the prototype, as is evident from the phylogenetic tree calculated using ClustalW in MegAlign (DNASTAR Lasergene®)(figure 6.4B).

Full genome sequencing of non-infectious clones pT7E30 GB27 (27.29) and pT7E30 Fi13 (15.6) was performed to characterise the particular strains of E30 and identify potential mutations that could be corrected at a later date. Clone pT7E30 GB27 (27.29) contained a T deletion at nt position 6463 (AATCT[T]AGGCT) that resulted in the introduction of a stop codon at amino acid 2166. Virus sequence in pT7E30 Fi13 (15.6) contained three mutations to render this non-infectious. It contained an inserted A at nt position 4363 (AGCTG[A]AAGCA), an encoded stop codon at nt position 4908 (TTTCTAACAAT), and a G deletion at nt position 6926 (CCTAC[G]GAGAC). The mutations listed above could be corrected by single rounds of site-directed mutagenesis in the future.

RNA transfections and recovery of infectious E30 progeny

In order to determine whether the full-length E30 clones were capable of generating infectious virus, clones pT7E30 Fi13 (15.6, 15.7, 15.16), pT7E30 GB27 (27.29), and pT7E30 GB37 (2.44) were linearised with *Sal*I (a unique site located downstream of the poly(A)-tail), ethanol precipitated to remove excess enzyme, and transcribed using T7 RNA Polymerase Kit (Fermentas)(figure 6.5A). Rhabdomyosarcoma cells were transfected with 2 µg RNA using Lipofectamine 2000 (Invitrogen) and incubated for 48 hrs. Transfection supernatant was removed and the viable virus identified and quantified using plaque assays. Supernatant was serially diluted in

A

Percent Identity

	1	2	3	4	5
1	-	81.1	81.2	83.1	84.4
2	21.4	-	99.7	82.1	82.1
3	21.2	0.3	-	82.2	82.2
4	18.7	20.7	20.5	-	85.9
5	16.9	20.7	20.6	16.0	-

Divergence

1 pT7E30GB37
2 E30 (AF162711) Bastianni prototype 1999
3 E30 (AF311938) Bastianni prototype 2000
4 E30 (DQ246620) China 2003 (aseptic meningitis)
5 E30 (DQ534205) Netherlands 1987 (diabetes)

B

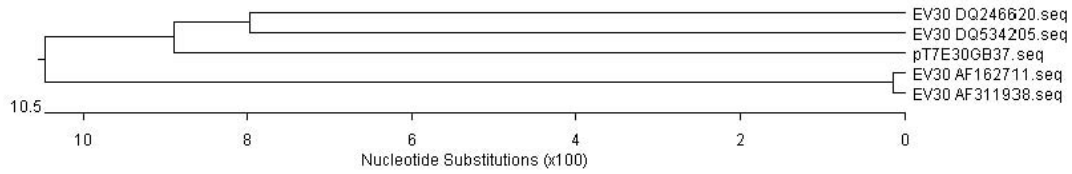
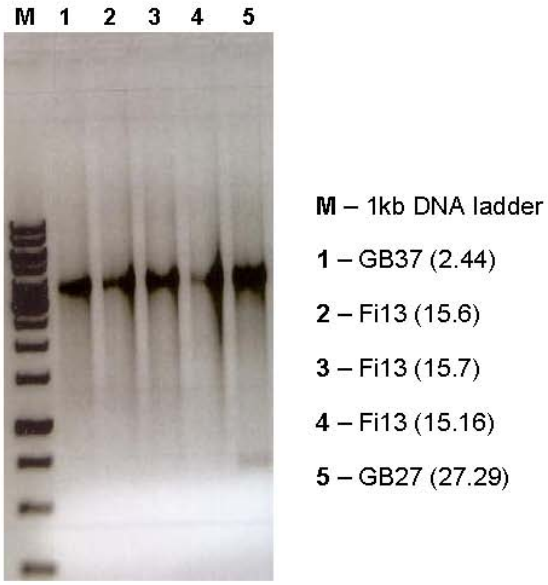


Figure 6.4 Phylogenetic analysis of cloned E30 strain GB37. Nucleotide sequences of E30 strains from GenBank were aligned with pT7E30 GB37 sequence using ClustalW in MegAlign software (DNASTAR Lasergene®). This produced (A) a table outlining percent identity and divergence for all five sequences, and (B) a phylogenetic tree to describe evolutionary distances.

A



B

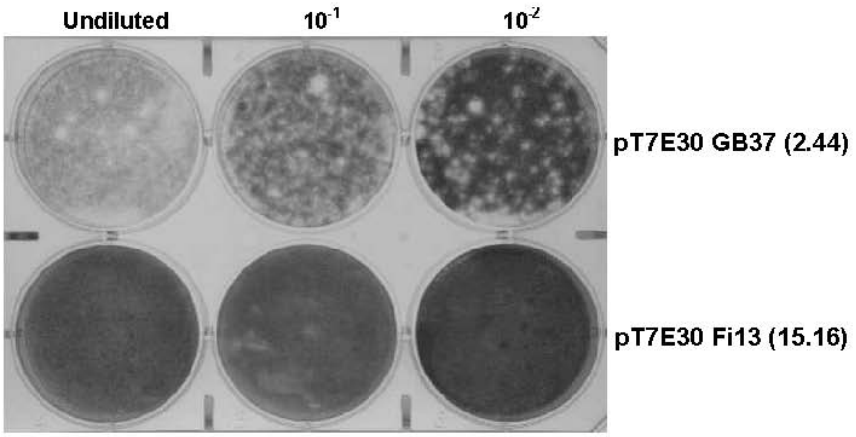


Figure 6.5 Recovery of infectious E30 from RNA transfections. (A) Ethidium bromide stained 1% agarose gel showing RNA transcription product of pT7E30 clones from T7 promoter. (B) Plaque assays of serially diluted transfection supernatant on RD cells showing infectious virus generated from pT7E30 GB37 (2.44) RNA, but not pT7E30 Fi13 (15.16) clone. The remaining clones were negative (plates not shown).

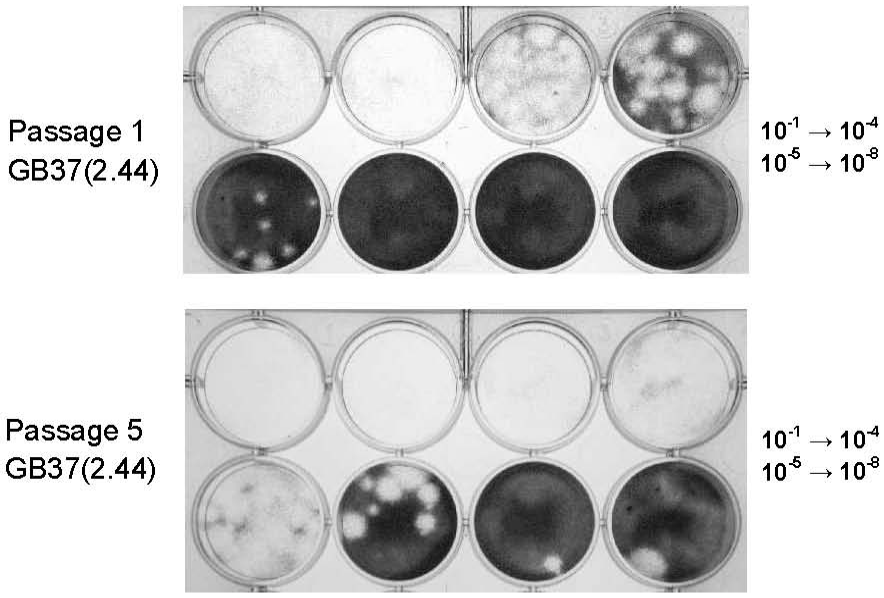
medium before using it to inoculate fresh RD monolayers. Once overlay plaque media was added to each well, plates were incubated at 37°C until CPE was evident in the cells. Only one clone demonstrated CPE after 48 hrs (pT7E30 GB37 (2.44)) (figure 6.5B). Infectious virus was recovered after transfections with clone pT7E30 GB37 (2.44), but not for the remaining clones (figure 6.5B).

Growth characteristics of cloned E30 GB37

To determine the stability of clone pT7E30 GB37(2.44), (designated as pT7E30GB37 henceforth), infectious virus generated from the full-length clone was serially passaged five times through RD cells. A fresh monolayer of RD cells was infected with RNA transfection supernatant containing E30 GB37 virus from the infectious clone and cultures were harvested after 24 hrs or when cells displayed ≥50% CPE. A quarter of the previously harvested supernatant was used for each subsequent inoculum. There were no observable changes in plaque phenotype between passages 1 to 5, with the initial passage yielding 7.2 x 10^5 PFU/ml and the fifth passage yielding 3.6 x 10^7 PFU/ml (figure 6.6A).

Replication kinetics of the original isolate and virus recovered from RNA transfection was compared in single-step growth experiments performed in 12-well plates to determine the kinetics of virus growth (PFU over time) as a measure of viral fitness. Briefly, RD cells were infected at an MOI of 5 for 30-60 mins before cells were washed twice with sterile PBS and incubated with fresh medium. Supernatant was harvested at various time points over a 24 hr period and virus titre was determined by plaque assay. Despite irregularities in wild-type E30 GB37 titres at 11 and 24 hrs (possibly due to aliquoting errors at the time of infection), virus recovered after 1 and 5 passages displayed similar growth kinetics to the original isolate (figure 6.6B). This experiment indicated that clone-derived virus (p1) had a slightly lower rate of growth, (lower by approximately one $\log_{10}$) from 4 to 9 hrs, but at later time points was at least as good as p5 E30 GB37. The early 'lag' in growth of p1 virus may have been the result of inserted nts between the T7 promoter and beginning of the 5' NCR, which could affect replication and packaging efficiency until insertions are gradually removed through repeated passage. Any nt changes introduced in virus sequence during subsequent passages were not investigated further due to time restraints.

A



B

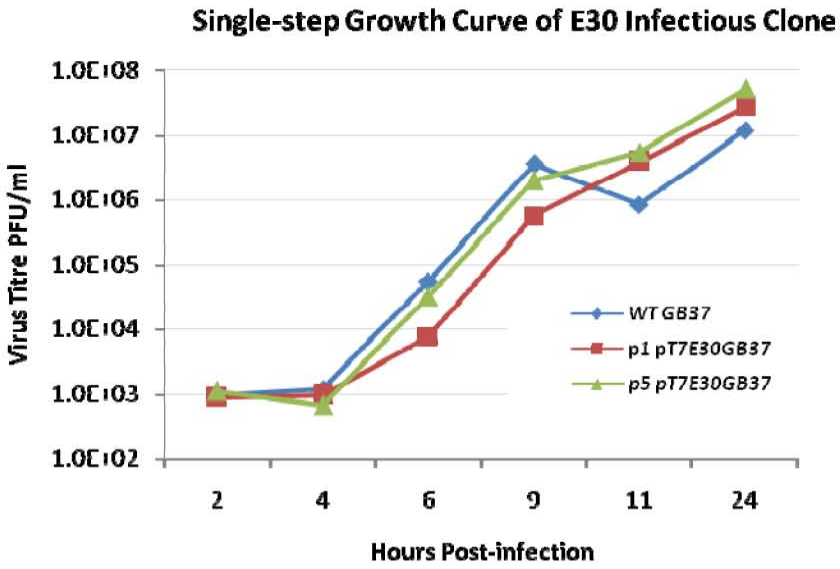


Figure 6.6 Growth characteristics of infectious full-length pT7E30 GB37 clone. (A) Plaque phenotypes of clone-derived E30 GB37 after 1 or 5 passages through RD cells. Cells were infected with 10-fold dilutions of virus and plated with plaque assay overlay medium. Cells were stained and photographed 72 hrs post-infection. (B) Growth kinetics comparing clone-derived E30 GB37 (passages 1 and 5) to original isolate. RD cells were infected at an MOI of 5; inoculum was removed after 30-60 mins and replacement medium removed at various time points from 2 to 24 hrs.

In summary, virus generated from the pT7E30 GB37 infectious clone had similar phenotypic and growth kinetics to wild-type GB37 from which it was derived. This meant that clone-derived virus clearly represented the wild-type virus isolated from a patient. The infectious clone was not characterised further (e.g. specific infectivity of clone-derived RNA) due to time constraints.

Construction of an E30 subgenomic virus replicon

Attempts were made to construct a luciferase encoded subgenomic replicon in a pRiboI vector backbone using full-length pT7E30 GB37 virus cDNA as template. The strategy is outlined in figure 6.7. Briefly, cloning was performed in two major steps: insertion of E30 5' NCR and the luciferase gene into pRiboI vector using the unique *Sma*I site in the vector ribozyme sequence, followed by insertion of the 3' end of the E30 genome downstream of the luciferase gene. The E30 5' NCR was amplified from pT7E30 GB37 plasmid using oligonucleotides RiboSmaIE30F and E305UTRSacIR (table 2.3). The luciferase gene was amplified from pT7Rep3-L (PV3 subgenomic replicon encoding luciferase) using oligonucleotides SacILucF and BstZ17ISmaILucR. Both PCR products were double digested with *Sma*I and *Sac*I and ligated into the unique *Sma*I site in pRiboI in a triple ligation (figure 6.7). In the final stage, the 3' end of E30 (non-structurals and 3' NCR) was amplified using oligonucleotides E30-2551F and MJREV to produce a PCR product of approximately 5 kbp. This and a correct pRiboI/5' NCR/luciferase backbone from the previous step was double digested with *BstZ17*I and *Sal*I and subsequently ligated together. This step exploited the unique *BstZ17*I site in pT7E30 GB37 at nt position 3150, and the unique *Sal*I sites in the pRiboI vector and *Sal*I site incorporated at the 3' terminus of the PCR product. Although apparently complete constructs were recovered and confirmed by restriction digests, all clones tested gave low luciferase readings, comparable to controls in the presence of guanidine hydrochloride. Unfortunately, these results were not explored further due to time limitations.

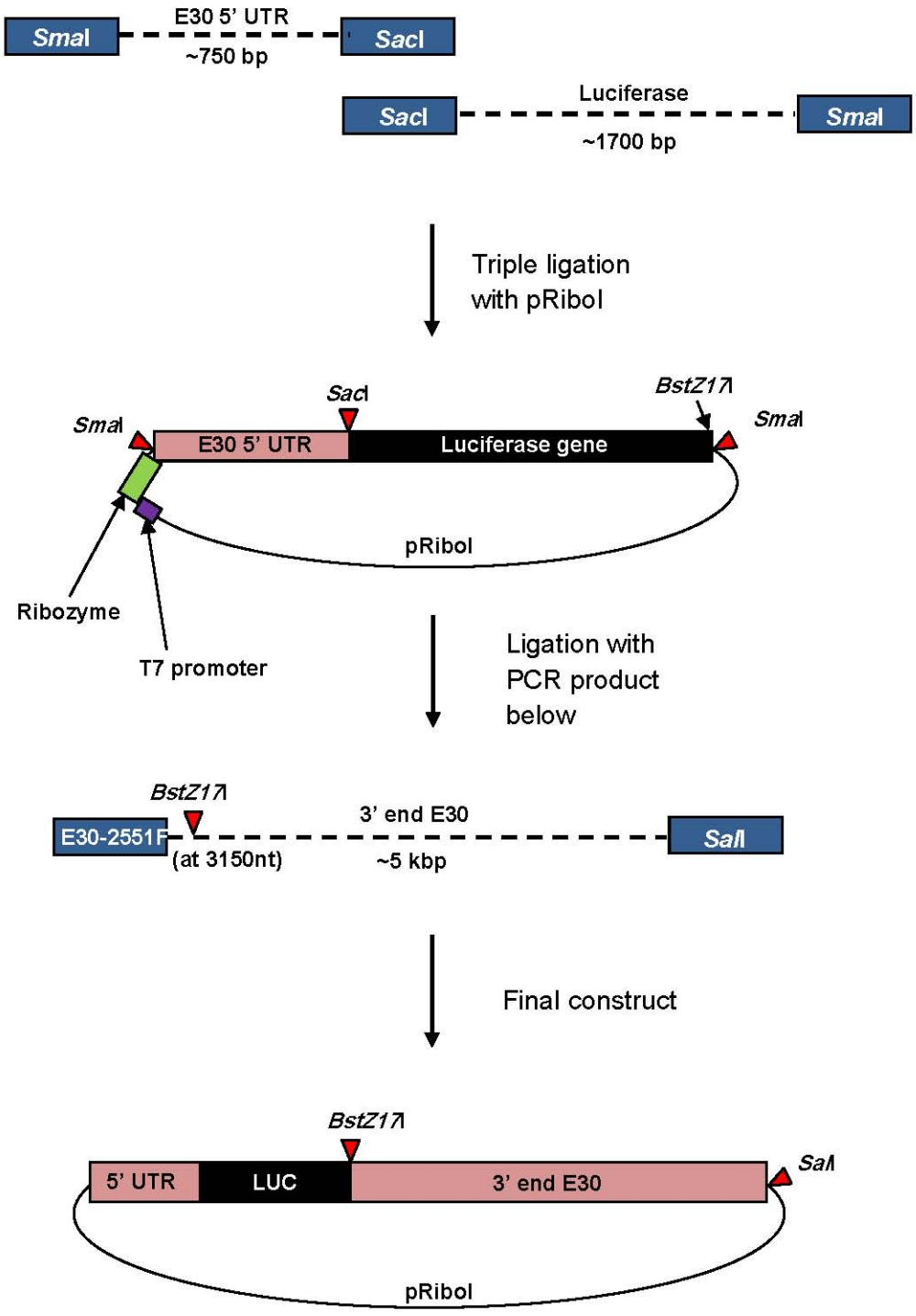


Figure 6.7 Overview of construction of E30 subgenomic replicon encoding luciferase. All three cDNA fragments were amplified from plasmids by PCR using oligonucleotides discussed in the text. Restriction sites used for cloning were incorporated into PCR product via oligonucleotides encoding these sites.

6.3 Generation of a Full-length E7 CRE Mutant cDNA

Mutating CRE region in full-length infectious E7 plasmid by site-directed mutagenesis

The second aim was to construct a species B enterovirus RNA partner for use in interspecies recombination experiments. Echovirus type 7 was chosen to represent a species B partner because both a full-length infectious clone and subgenomic replicon was available at the time. The CRE is present and highly conserved in the enteroviruses and is located in the E7 strain Wallace at position 4390 to 4450 according to complete genome sequence in GenBank (AY302559). The species B E7 partner was modified by mutating the CRE in full-length infectious clone pT7E7 in order to disrupt the ability of this clone to generate E7 virus particles while still allowing for negative-strand RNA synthesis. The aim was to disrupt the conserved secondary RNA structure of the CRE stem loop with point mutations in the nt sequence. The wild-type RNA structure of E7 CRE is presented in figure 6.8A and was predicted *in silico* using MFOLD (Zuker, 1989). Seven synonymous mutations were incorporated into the CRE sequence and were designed to disrupt the structure from forming. Most point mutations occurred at the third-base position of codons to avoid changing the translated 2C product and mimicked those used in early CRE studies (Goodfellow *et al.*, 2000). Figure 6.8B shows substitutions in the stem loop at the following nt positions: 4402 (C→T), 4405 (A→C), 4408 (G→A), 4412 (C→T), 4414 (G→A), 4419 (A→G), and 4429 (T→C), and further illustrates how these changes were predicted to affect the overall shape of the CRE to prevent RNA synthesis.

Initially, a portion of the 2C coding region (spanning the CRE) was removed from pT7E7 by digestion using restriction enzyme *Pst*I (4332 and 6434 nts) and was subcloned into the unique *Pst*I site in pUC19. Attempts were made to mutate the CRE by using a site-directed mutagenesis kit (Stratagene) following the manufacturer's instructions. Unfortunately this was unsuccessful as no inserts with mutated CREs were recovered. One interpretation of this result suggested that this method was ineffective due to the length of oligonucleotides and the large number of mutations contained within them. An alternative approach was therefore used where a pT7E7ΔCRE mutant was constructed by overlapping PCR mutagenesis, as outlined

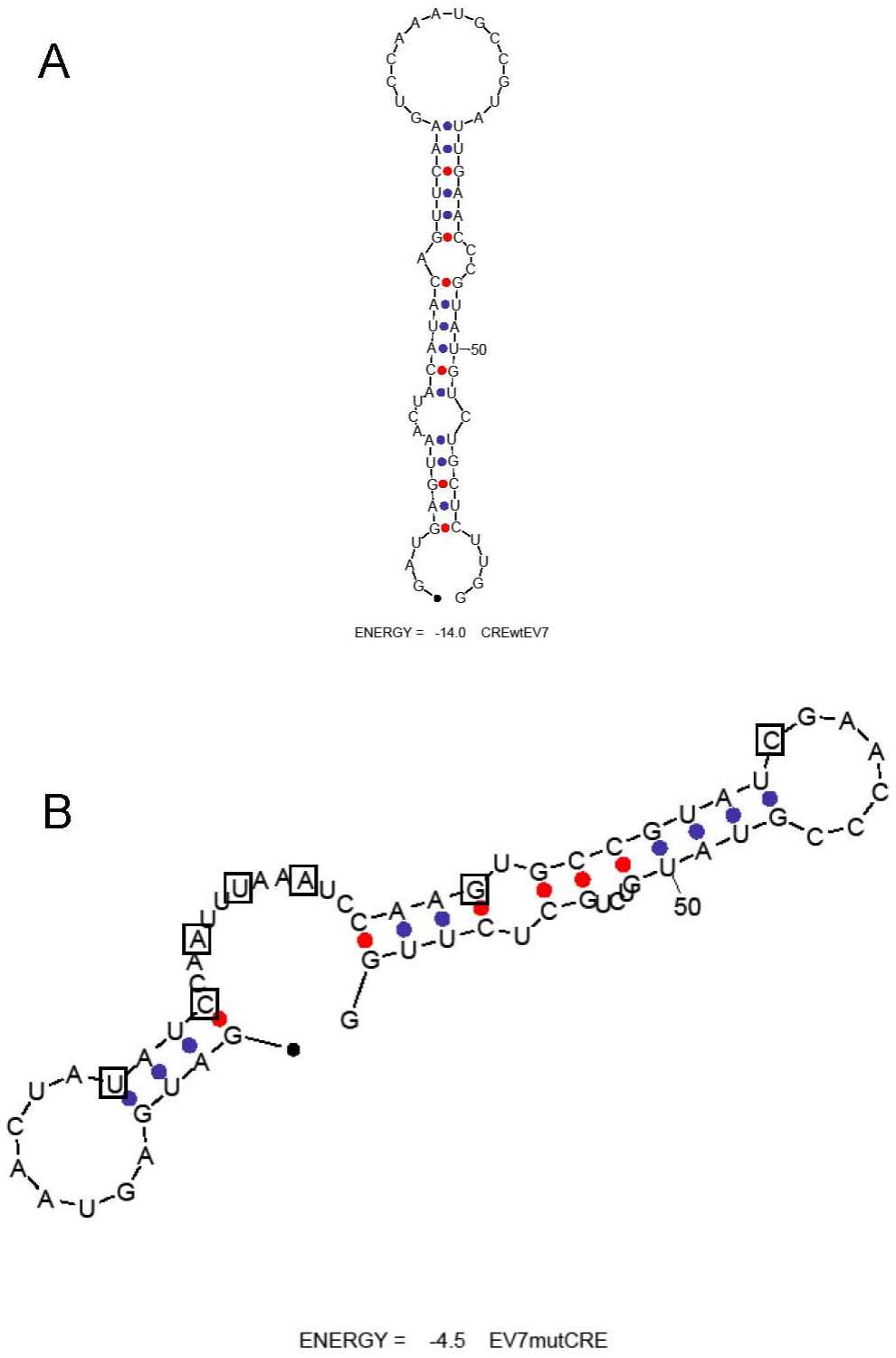


Figure 6.8 Predicted secondary structure of wild-type and mutagenised E7 CRE stem loop. (A) Wild-type CRE stem loop generated by an MFOLD program run by Burnet Institute website (software originally created by Zuker and Stiegler, 1981). (B) Predicted secondary structure of mutagenised CRE. Seven mutations introduced are boxed.

in figure 6.9. The same oligonucleotides were applied to incorporate the aforementioned nt mutations.

Briefly, the CRE was amplified in two fragments in order to incorporate the necessary mutations into the oligonucleotides. Oligonucleotides E7-3782F and E7RcreMUT were used to amplify one fragment from plasmid pT7E7 while E7FcreMUT and M13R amplified the remaining fragment from the initial subclone in pUC19 described above. The fragments contained overlapping sequence (mutatedCRE) and acted as template to construct the larger amplicon. Both fragments were purified from agarose gel, the DNA quantified, and incorporated into an initial '5 cycle' PCR performed without oligonucleotides in order to complete amplicons with precise ds ends. This PCR product was used as template and subject to final assembly amplification using *Taq* DNA Polymerase with standard conditions (figure 6.10A). Column-purified product was digested with *Pst*I, gel-purified and ligated into phosphatase treated *Pst*I digested pGEM®-T Easy vector. This vector was used to aid in the sequencing of cloned fragments with the use of T7 and SP6 oligonucleotide regions flanking the insertion site, see below. Mutations introduced to the CRE also incorporated a unique *Swa*I site which was used to screen plasmid DNA minipreps (figure 6.10B). Sequencing was used to confirm integrity of the fragment before it was rebuilt into the full-length infectious pT7E7 plasmid to generate pT7E7ΔCRE. Once again, the unique *Swa*I site in mutated CRE was used to confirm its presence in pT7E7ΔCRE, and orientation of the reinserted fragment was later determined by sequencing.

Sequence analysis of CRE overlapping PCR fragment in pGEM®-T Easy

The following seven mutations to the CRE were confirmed in the full-length genome, at nt positions: 4402 (C→T), 4405 (A→C), 4408 (G→A), 4412 (C→T), 4414 (G→A), 4419 (A→G), and 4429 (T→C). Two further mutations were located in the fragment, and these occurred at nt positions: 4590 (C→T) and 5993 (G→A). Both were responsible for amino acid changes from threonine to isoleucine and from alanine to threonine, respectively. As the initially cloned E7 fragment had not been

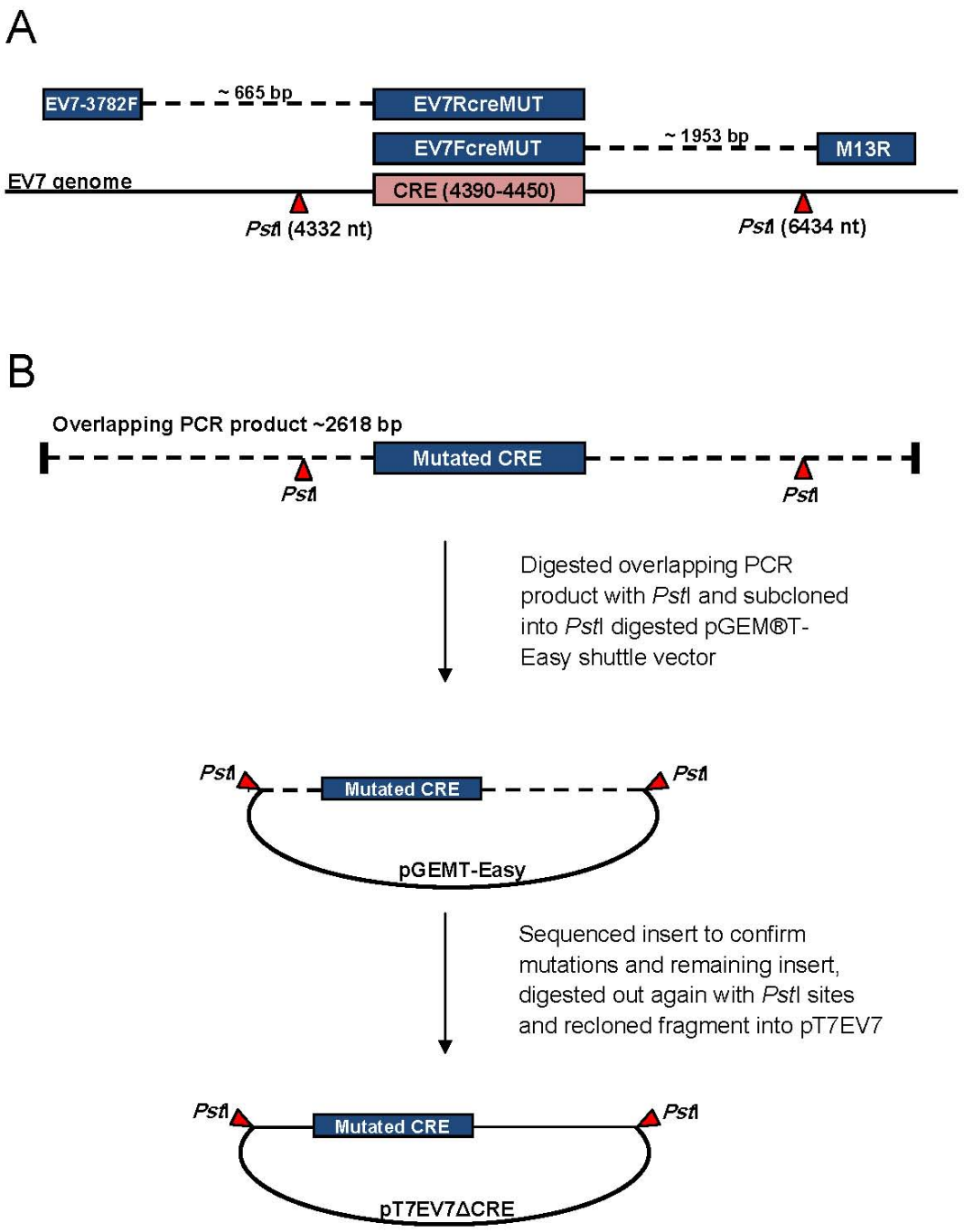


Figure 6.9 Strategy for mutagenising CRE in full-length infectious clone pT7E7. (A) The CRE region was amplified in two fragments using oligonucleotides containing the seven nt changes required to alter the CRE sequence. The second fragment (E7FcreMUT to M13R) was amplified from the original subclone constructed in pUC19 (see text). (B) Overlapping PCR performed to join both CRE fragments generated in (A). The ~ 2618 bp fragment was digested with *Pst*I and subcloned into pGEM®T-Easy to confirm mutagenesis by sequence analysis. The CRE fragment was digested from the shuttle vector and rebuilt in the pT7E7 backbone using the *Pst*I sites again.

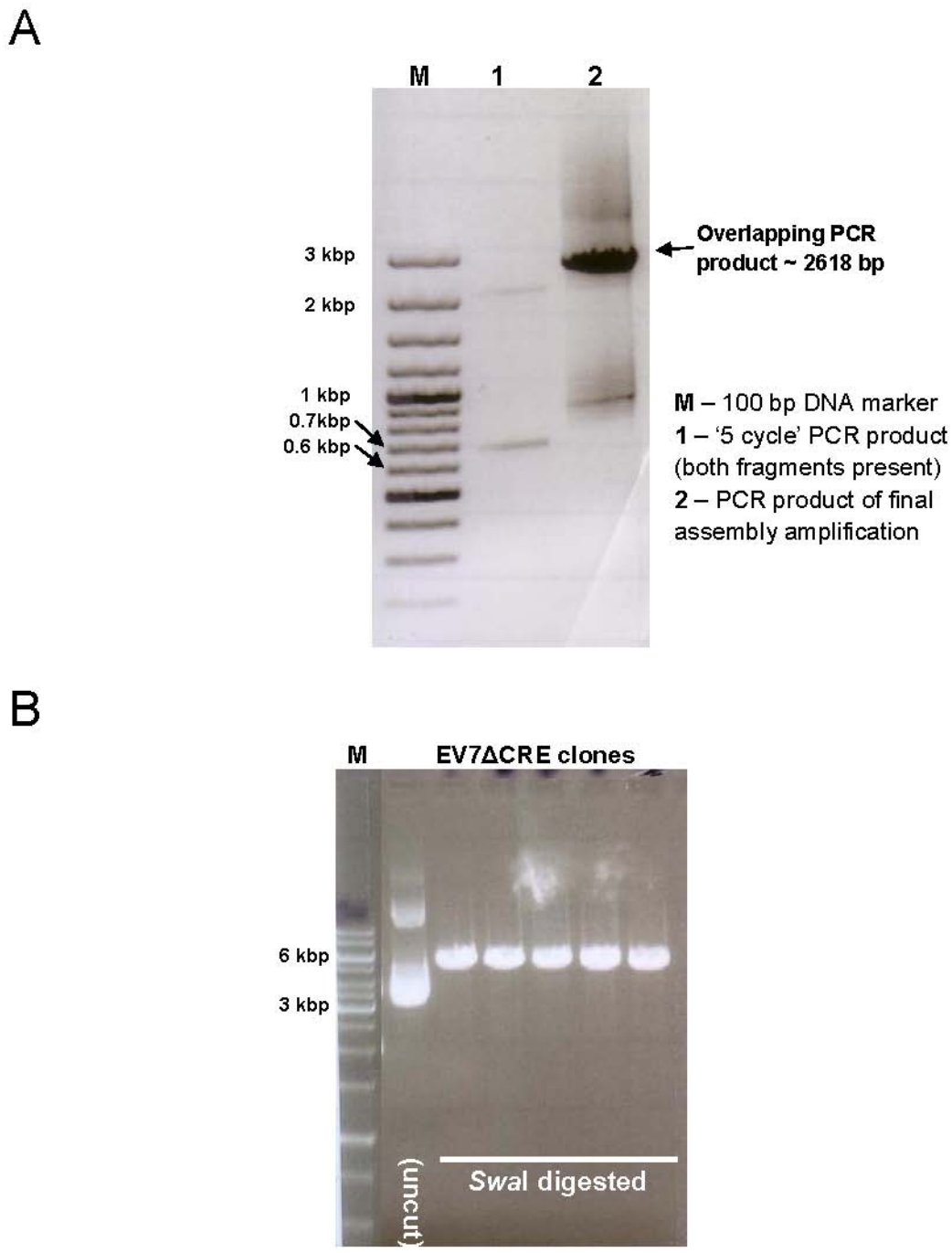


Figure 6.10 Overlapping PCR of E7 CRE fragments. (A) Agarose gel photo showing '5 cycle' PCR product (lane 1) complete with both fragments amplified separately and added to the reaction to create precise ds amplicons. Lane 2 shows the result of the final assembly PCR (overlapping PCR) which used lane 1 product as template. The large band highlighted in lane 2 is the combined overlapping product of both fragments. (B) Digestion of E7ΔCRE/pGEM®-T Easy clones to confirm presence of introduced *Swa*I site. All clones are positive for *Swa*I and therefore have a mutated CRE sequence.

sequenced prior to mutagenesis, it was not apparent whether the two mutations were already present in the E7 genome.

RNA transfections and recovery of revertant virus to confirm stability of CRE mutations

It was important that the introduced mutations could not revert to their original nts by continuous passage resulting in restored infectivity to clone-derived RNA. This would no longer be a viable partner in the interspecies recombination experiments. Previous studies of PV3 CRE have shown that a treble mutant in the stem loop (2C*mut*6) destroyed replication of full-length genomes and function was only restored after three repeated blind passages, which resulted in reversion of at least one of the three mutations (Goodfellow *et al.*, 2000). There was no reason to believe that the conserved CRE would behave differently for E7, and there should be no detectable revertants, as too many would be required simultaneously to restore infectivity. Integrity of the seven mutations in the 2C section of the CRE was tested by serial passage of the non-viable RNA through RD cells.

Several pT7E7ΔCRE clones, together with wild-type pT7E7, were linearised with *Not*I and transcribed into RNA as described in Materials and Methods. Briefly, 2 μg of clone-derived RNA was transfected into cells and supernatant harvested 48 hrs later. A portion of supernatant was inoculated onto fresh RD monolayers for 30 mins, removed and cells washed twice with sterile PBS, before serum-free medium was added to the cells. Supernatant was again removed 48 hrs post-transfection and the process repeated five times. Wild-type pT7E7 supernatant was not passaged in this way as it provided a positive control for RNA transfections. Supernatant harvested from transfections (p0) and p5 infections was diluted in serial 10-fold increments and plaqued onto fresh RD monolayers. It should be noted that no CPE was observed during these stages. Plaques were not detected in supernatant at passages 0 or 5 in the pT7E7ΔCRE clones, indicating that they remained non-infectious (data not shown). It has been reported that the 2C loop can accommodate a small number of variations in the region, but the non-viable clone would have to overcome the majority of the seven introduced mutations to restore infectivity (Goodfellow *et al*., 2000).

6.4 Discussion

A full-length infectious clone of E30 strain GB37 was constructed (pT7E30 GB37) and its complete nucleotide sequence determined. This was the first reported infectious clone for E30. Although this plasmid does not contain a ribozyme for more efficient RNA processing from plasmid, the clone-derived RNA produced infectious virus with similar growth characteristics to wild-type strain GB37 in a single-step growth curve and was therefore deemed suitable for the project's requirements. Although the specific infectivity for clone-derived virus RNA in RD cells was not calculated, it was estimated to be lower than that calculated for naked PV RNA in HeLa cells of around 10^6 PFU/μg RNA (Crotty *et al.*, 2001) based on the cloning strategy used. There was approximately 70 nt of extraneous sequence located between the T7 promoter and the start of the virus, which would reduce early virus infectivity as has been shown in studies with PV (Herold & Andino., 2000). Herold and Andino showed that replication of PV was more efficient when the 5' end of the RNA was cleaved precisely by a *cis*-acting hammerhead ribozyme (Herold & Andino, 2000). This could be overcome by subcloning the E30 virus sequence into a pRiboI vector (a pBluescript II derived cloning vector with a hammerhead ribozyme situated adjacent to the T7 promoter).

Cloning of full-length E30 RT-PCR product proved to be an inefficient process which is evident with only one viable infectious clone recovered. A DNA polymerase was used with rapid amplification and proof reading ability, and this provided adequate quantities of PCR product. The KOD polymerase possesses strong 3'-5' exonuclease activity and is reported to have an error rate of approximately 1 in 10^6 (Nishioka *et al.*, 2001). The polymerase forms a mix of blunt and 3' A ends on PCR product which should be ideal for cloning in T-vectors such as pGEM®-T Easy.

Only six full-length E30 clones were obtained over several repeated cloning experiments. It is likely that several factors contributed to this low efficiency, including the low rate of 3' A additions to amplicon during and after PCR, and decreased efficiencies in further cloning steps. Transformation efficiencies are reported to decrease as the size of plasmid increases and this is a probable reason for

the low number of clones obtained[4]. Promega literature suggests that an empty vector the same size as that used in the present study (pGEM®-T Easy at 3015 bp) yields approximately 60 times more colonies per microgram of transformed DNA than 13.6 kbp plasmids, and this is even lower for transformations using chemically competent cells. Competition was therefore created between empty vector (3015 bp) and vector containing virus cDNA (3015 + 7515 = 10530 bp) for insertion into *E. coli*, and this would have been compounded by the lower proportion of constructs in the ligation mix. As mentioned, this could have been caused by the inefficient addition of A nts to 3' ends of amplicon by the polymerases used, thereby reducing ligated fragments prior to transformation.

The non-infectious E30 clones (strains Fi13 and GB27) contained several insertions/deletions upon sequencing and these are likely to have destroyed the genomes' function. The E30 clones could be repaired at a later date using site-directed mutagenesis or overlapping PCR to generate repaired infectious clones representing the remaining E30 strains. These would provide a useful model to study recombination between E30 strains in the future.

Due to time restraints, attempts to construct a luciferase-encoding subgenomic replicon for E30 were not successful. Attempts were also made to mutagenise the CRE to inhibit the generation of viable virus from clone-derived virus RNA, but this was not completed. The latter project was performed in the same manner as that described for constructing pT7E7ΔCRE. An E30 subgenomic replicon and full-length CRE mutant would have made valuable species B partners for *in vitro* recombination experiments using co-transfections.

The construction of a full-length semi-replicating E7 clone was successful and could constitute one RNA partner in co-transfection experiments to generate recombinant viruses in cell culture. The modification of the replication element was stable during continuous RNA replication cycles and is unlikely to restore its infectivity in the long term based on data from 5 serial passages. The secondary structure of the CRE was completely disrupted from forming by mutagenesis so that it no longer presented as a

[4] (http://www.promega.com/enotes/faqspeak/0107/fq0030.htm)

stem loop. The CRE is involved in templating the uridylylation of VPg and is therefore required for the initiation of virus RNA synthesis. Disrupting the CRE structure meant that the two adenosines located at the CRE terminal loop could not function as a template for the addition of two uridylate residues to VPg, hence leading to inhibition of the formation of VPg-pUpU. As this modified VPg-pUpU is required as a primer for virus $3D^{pol}$, and is linked to the 5'-end of the positive- and negative-strand genomes. Any disruptions to this process would therefore negatively affect replication. As negative-strand RNA synthesis was not affected, this provided an ideal species B partner for recombination experiments presented in this thesis.

Due to time restraints, the replication ability (negative-strand synthesis) of pT7E7ΔCRE was not investigated. Instead, this new RNA partner was incorporated into a co-transfection assay with homologous pT7E7luc subgenomic replicon (refer to chapter four). Viable recombinant E7 was recovered, demonstrating semi-replication of pT7E7ΔCRE in cell culture and the use of this model in the study of intraserotypic recombination events in both species B and species C picornaviruses.

CHAPTER SEVEN: General Discussion

7.1 Current Understanding

There are wide ranging implications of enterovirus evolution, particularly when recombination events contribute to the emergence of modified viruses associated with disease outbreaks. Outbreaks of non-wild-type PV are usually caused by recombination in low OPV vaccinated populations, and it is not clear which circulating enteroviruses will contribute to future outbreaks.

Although it has been thoroughly investigated in several virus systems, little is known about the underlying molecular mechanisms of recombination between positive-sense single stranded RNA viruses. A better understanding of these processes has been advanced in this study by the development of an experimental method for the generation of recombinant PVs. This method is beneficial for studying recombinant viruses because there are no competing parental-type viruses in the progeny, and the assay is not restricted by certain aspects of natural viral tropism, such as susceptible cell type and infection into the cell post-attachment. These characteristics allow the method to capture initial recombination events prior to further virus propagation.

During this project, an assay was developed that produces populations of natural PV recombinants that leads to replication competent viruses. The system only allows for recombination events to occur in the genome region following VP1 and before the CRE, which is the location of commonly detected sites 'in the wild'.
The work presented in this thesis established general rules for recombination as observed in individual recombinant PVs:

- The majority of junction sites are imprecise with insertions up to ~ 500 nts, in the 2A and 2B coding regions.
- A minority of junction sites are precise, but usually located in the 2C coding region.
- Template switching appears to be a sequence independent event, since inserts contain non-virus sequence derived from the luciferase reporter gene.
- Template switching occurs independently of local RNA motifs, such as strings of As or AUs, potential stem loop structures, and palindromes.

- Recombinants contain a complete 2A and/or 2B coding region from one or other parent. The generation of chimeric 2A and 2B coding regions are rare in initial recombinants recovered, and only appear after subsequent serial passage.
- Template switching is localised to regions, and not evenly distributed throughout the entire targeted region. As a consequence of the above.
- Imprecise junction sites do not show deletions of sequence, only insertions.
- Recombination is at least a two stage process, with initial template switching occurring in a 'promiscuous' and unpredictable manner (conforming to the general 'rules' above), while further rounds of replication correct imprecise crossover junctions to reflect wild-type virus genome length.

The question of whether PV RdRp switches templates at preferred sites in the viral genome was explored by King in 1988, who analysed the crossover sites of approximately 50 reported interserotypic recombinants for poliovirus (King, 1988). King suggested that secondary structure was not involved in polymerase pausing, by stating that "most ssRNAs are extensively folded" naturally, which is in agreement with data presented in this project. Although previous reports had concluded that poliovirus recombination occurred in more general regions and not by a site-specific mechanism (Kirkegaard & Baltimore, 1986), King found a strong bias for the occurrence of precise crossover sites immediately after AA dinucleotides (two-thirds of those analysed), indicating a site-specific selectivity. It was not clear from the report whether this pertained to donor, acceptor, or both templates. Of the nine precise crossover sites identified in this study (figure 5.13), only one virus, PV3(3640-54)PV1, contained an adjacent 3' AA dinucleotide on the donor template. It is difficult to conclude from this result without applying a larger dataset to this analysis. It was striking to discover, however, that 67% of crossover sites in this project occurred next to dipurine sequences (AA, AG, GA, or GG) on the donor strand. The significance of this finding could be investigated further with more experimental data.

The nature of imprecise recombinant PVs may reflect restrictions imposed on the viruses during packaging. Imprecise recombinants contained inserted sequence up to

513 nt in length, and these genomes (now ~ 8 kb) could be approaching the theoretical packaging limit for capsid protein shells. Of current picornaviruses, FMDV has the largest genome at ~8.2 kb, and this is contained within similar sized particles to PV (~30 nm) (Fox *et al.*, 1987). Teterina and colleagues successfully inserted a ~ 700 nt sequence encoding a fluorescent protein tag into $2A^{pro}$ of PV, and this increased the genome length to just over 8.1 kb (Teterina *et al.*, 2010). But like the recombinants generated in this project, the fluorescently tagged virus was only stable for four passages before deletions appeared in the non-virus sequence. It is possible that longer genomes can only be packaged in larger particles, and longer recombinant PVs may be approaching the theoretical packaging limit for the conserved particle size, therefore imposing a selection pressure on the size of insert and final genome length.

There are many steps in the virus life cycle that lead to the generation of a viable recombinant virus, and there are equally numerous constraints imposed by the host and virus itself that could impact on this occurring. The possible constraints at each level of the virus life cycle are described in figure 7.1. It is not known what proportion of cells in the host are co-infected, or the proportion of virus RNA partners that reside within the same replication complex. This relates directly to the MOI and would also be limited to the number of sites in which replication complexes can be established within an infected cell. When looking at a subset of partners that have made it through early selection as just mentioned, a rough probability of successful template switches can be calculated. Those recombinants with forward crossovers will not survive since they contain deletions, and crossovers that are too far back (> 513 nt, the maximum observed in this study) may be too long for successful packaging. If the location of template switching is evenly distributed throughout the genome, each nt position would have 1/7440 chance of being the location of a junction site (PV1 Mahoney genome contains 7440 nt), but this is not the case. This is also influenced by the location in the genome where recombination occurs, as its clear that it rarely occurs in the capsid region, which is approximately 1/3 of the genome. This means that ~4800 nt would be accessible in the donor template, and ~5300 nt in the acceptor template (including up to ~500 nt insertion). If a virus can only tolerate a 513 nt insert, then there are theoretically 513/4800 possible junction sites in the genome. Subsequently, only 1/3 will be translated in

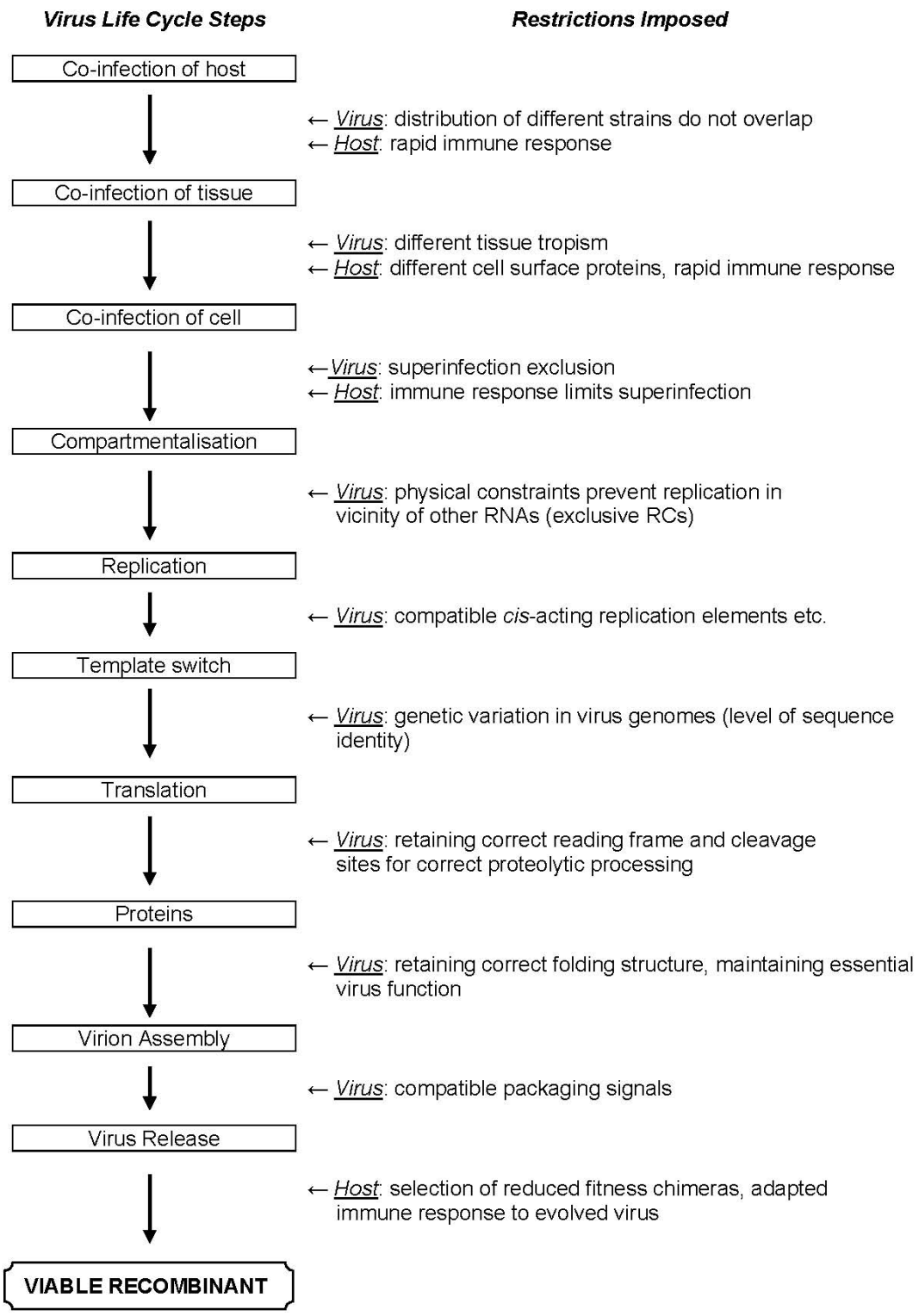


Figure 7.1 Production (and restrictions) of a viable recombinant RNA virus. The steps necessary to generate a recombinant and viable RNA virus in a host, and the constraints imposed at each stage by a combination of virus and host. (This model is adapted from Worobey & Holmes, 1999).

frame, meaning that ~2% (513/5300 x 4800/7440 x 1/3 x 100%) of recombinant genomes will fulfil all the 'rules'. This small subset of successful recombinants must then prove viable at all subsequent steps of the life cycle. In a simplified way, this demonstrates the low frequency of recombination events that must occur in the host.

Many restrictions can be imposed on the creation of a viable replication-competent recombinant virus, from both the host and virus itself. Considering the number of restrictions that could impact on the generation of recombinant viruses, it is not surprising that intra- and interspecies recombinants are rarely identified, and why interspecies viruses were not recovered during this project. At the amino acid level, chimeric virus proteins containing both CVA21 and PV sequence were not observed in isolated recombinants. Such junctions may have occurred during replication, but were later selected against during protein processing or morphogenesis. Similarly, since the construction of the recombinant CVA21/PV full-length cDNAs by Dr. Claire Blanchard, it is now known that direct interactions between proteins VP3 and $2C^{ATPase}$ are vital for virus morphogenesis (refer to chapter four).

This method has highlighted the promiscuous nature of early recombination events between two closely related viruses. Recombinant genomes may have been generated between input RNA species within the cell, but might not have been functional at later stages of the life cycle. Only a few interspecies recombinants have been identified from natural isolates (Bolanaki *et al.*, 2007; Yozwiak *et al.*, 2010), and only small regions of genome are replaced (e.g. 5' NCR or 3D) compared to significant portions of genome aimed for in these experiments (Smura *et al.*, 2007). Chimeric PV genomes with an HRV2 $2A^{pro}$ did not generate viable virus, but those with a CVB4 $2A^{pro}$ did, suggesting that replacements with more closely related virus sequences will be more likely to conserve function (Lu *et al.*, 1995). In another approach, Bell and colleagues replaced the $3D^{pol}$ coding sequence with that from CVB3 in a PV1 infectious clone. No infectious virus was produced from the chimeric RNA and no RNA synthesis occurred in transfected HeLa cells, indicating that these chimeric genome regions were not compatible for essential interactions (Bell *et al.*, 1999). It seems that interspecies recombinants, with a junction site in the centre of the genome, have either remained undiscovered, or are not viable due to incompatibilities in RNA-RNA or RNA-protein interactions during replication. Such

sequence motifs required for these interactions are poorly defined and are difficult to identify, although they must be localised to the coding region, as NCR functions and the CRE are similar between species. This could explain the lack of chimeric viruses with different structural and non-structural coding regions during intra- and interspecies co-transfection experiments.

7.2 Future experiments

As this method for generating recombinant viruses has proven efficient and reproducible, it can be expanded upon to gain further insight into the influences of recombination at the virus and host levels.

- If crossover junctions identified in this project are not dependent on sequence identity between RNA partners or local sequence/structure motifs, then it would be of interest to perform the same co-transfection, with PV1 replicon and PV3 CRE mutant, several times independently to determine whether recombination events follow the 'general rules' mentioned above. The observations above were obtained from a limited number of experiments, and it is inconclusive whether only a single host cell produces a virus with a distinct junction site that replicates well (reflected in the number of clones identified), or whether the same recombinant can be generated by more than one host cell in the population. Recombination junction sites from replicate assays can be compared to investigate this.
- Resistance markers in mutagen-induced viruses have previously been applied in order to calculate recombination frequencies (Kirkegaard & Baltimore, 1986; McCahon *et al.*, 1977). Similar controlled experiments can be applied to the method in this project in order to accurately calculate recombination frequencies in relation to distance of RNA partner relatedness. Any influences of the distance between lesions in both RNA partners on recombination frequencies can also be determined.
- Due to time restraints, it was not possible to incorporate different species B RNA partners into this method. The availability of a full-length infectious E30 clone makes it possible to construct a CRE mutant using either site-directed mutagenesis or by overlapping PCR (as was the case for mutating CRE in pT7E7). Similarly, a portion of the capsid coding region can be removed by restriction digest followed by religation of flanking 5' and 3'

termini, to construct a replication competent E30 DI particle or a luciferase (or similar reporter) coding replicon. Co-transfections with E7 and E30 RNA partners can be performed to validate this method further and to determine whether both viruses will recombine *in vitro*.

- Mechanistic studies of PV RdRp have uncovered an amino acid substitution resulting from a single nt change (G64S) in the $3D^{pol}$ encoding region that improves fidelity during replication (Pfeiffer & Kirkegaard, 2003) and reduces the elongation rate (Arnold *et al.*, 2005). Both of these events may impact on the frequency of template switching by PV RdRp and can be determined experimentally by incorporating the G64S substitution into RNA partners prior to co-transfections. A less error-prone RdRp may reduce the frequency of template switching between RNA templates, and the opposite phenomenon can be investigated by decreasing RdRp fidelity. This could be achieved by performing recombination experiments in the presence of ribavirin.
- As briefly mentioned above, the distance between engineered "lesions" contained within RNA partners can be altered by relocating the CRE to other regions of the genome with little impact on function (Goodfellow et al., 2003). For example, a mutated CRE at opposing ends of each RNA partner genome will allow an even greater region for targeted recombination and may establish 'hot spots' outside the scope of this project.
- There has been much interest in host and environmental factors contributing to RNA recombination, particularly involving plant viruses. Genome-wide screens representing ~95% of host genes have been performed with Tomato bushy stunt virus (TBSV) in a yeast model, to identify host genes affecting RNA recombination (Serviene *et al.*, 2006; Serviene *et al.*, 2005). Interestingly, at least two important pathways have been identified in yeast that directly affect virus replication and recombination in the presence of external chemical factors. In one study, it was reported that inactivation of yeast host Pmr1p, an ion pump that controls Ca^{2+}/Mn^{2+} movement into the Golgi from the cytosol, led to an ~160-fold increase in TBSV recombination in yeast (Jaag *et al.*, 2010). This involved deleting non-overlapping fragments of *PMR1* gene (highly conserved gene that encodes Pmr1p) following by the initiation of TBSV RNA replication from expression plasmids. Analysis of

northern blots revealed the more efficient accumulation of RNA recombinants in deletion mutants compared than in wild-type yeast. In another study, it was demonstrated that the *MET22/Xrn1* pathway (nucleotidases/ ribonucleases) regulates the frequency of TBSV RNA recombination and efficiency of virus replication in a yeast system (Jaag & Nagy, 2010). Not only did deletions of *MET22* gene increase TBSV RNA replication and recombination, but the addition of either LiCl or NaCl was also shown to inhibit MET22p activity to produce similar outcomes. To conclude, it would be meaningful to identify similar genes/pathways in humans and determine whether such applied host and external factors could affect PV replication and recombination in the same manner. Preliminary searching has revealed that a human equivalent of the conserved *PMR1* gene exists with a similar function, and is known as calcium-transporting ATPase type 2C member 1 (ATP2C1) (NCBI reference sequence NM_001199185). Likewise, human equivalents of *MET22* gene encoding bisphosphate-3'-nucleotidase can be identified and subsequently altered experimentally (with siRNA) to determine affects on PV replication and recombination efficiencies.

- Observations in *in vitro* assays using cell lines do not always represent mechanisms in *in vivo* models. The experimental method presented in this thesis is an artificial system, as naked RNA is introduced directly into cells (bypassing natural virus infection stages), and RNA partners contain non-virus sequence. It would be beneficial to apply this method to an animal model using transgenic mice that express PVR. A proposed experiment is to inject a solution containing two PV RNA partners directly into the spine (intrathecally) of mice. After a short period of time, tissue would be removed from different regions of the mouse, including from sites surrounding the injection location. Any viable recombinant PVs in mouse tissue can be initially identified by plaque assay, following by isolation of individual viruses for characterisation. If recombinant PV is recovered, crossover sites can be characterised according to the location of tissue they are derived from. Are recombinant PVs similar to those generated *in vitro*, and do they follow the same 'general rules'? Do crossover sites change according to the distance from the initial injection site? The prediction would be that the further away

from the site of inoculation, the more precise the recombinants would be as those viruses can spread more efficiently. Such recombinants will have to overcome multiple host restrictions that are not present in an *in vitro* system; this may influence the frequency and nature of recombination.

The work in this thesis contributes to our understanding of the underlying mechanisms involved in RNA recombination. Viruses must survive various constraints imposed by the infected host and by the virus itself, and they achieve this through rapid adaptation of their large population sizes and with the aid of recombination.

Bibliography:

Agol, V. I. (1997). Recombination and Other Genomic Rearrangements in Picornaviruses. *Seminars in Virology* **8**, 77-84.

Ahlquist, P. (2002). RNA-dependent RNA polymerases, viruses, and RNA silencing. *Science* **296**, 1270-3.

Almond, J. W. (1987). The attenuation of poliovirus neurovirulence. *Annu Rev Microbiol* **41**, 153-80.

Ambros, V., Pettersson, R. F. & Baltimore, D. (1978). An enzymatic activity in uninfected cells that cleaves the linkage between poliovirion RNA and the 5' terminal protein. *Cell* **15**, 1439-46.

Andino, R., Rieckhof, G. E., Achacoso, P. L. & Baltimore, D. (1993). Poliovirus RNA synthesis utilizes an RNP complex formed around the 5'-end of viral RNA. *EMBO J* **12**, 3587-98.

Ansardi, D., Porter, D., Anderson, M. & Morrow, C. (1996). Poliovirus Assembly and Encapsidation of Genomic RNA. *Advances in Virus Research* **46**, 1-68.

Arden, K. E. & Mackay, I. M. (2009). Human rhinoviruses: coming in from the cold. *Genome Med* **1**, 44.

Arnold, J., Vignuzzi, M., Stone, J., Andino, R. & Cameron, C. (2005). Remote site control of an active site fidelity checkpoint in a viral RNA-dependent RNA polymerase. *J Biol Chem* **280**, 25706-16.

Bachrach, H. L. (1968). Foot-and-mouth disease. *Annu Rev Microbiol* **22**, 201-44.

Balvay, L., Soto Rifo, R., Ricci, E., Decimo, D. & Ohlmann, T. (2009). Structural and functional diversity of viral IRESes. *Biochim Biophys Acta* **1789**, 542-57.

Banerjee, R. & Dasgupta, A. (2001). Interaction of picornavirus 2C polypeptide with the viral negative-strand RNA. *J Gen Virol* **82**, 2621-7.

Barclay, W., Li, Q., Hutchinson, G., Moon, D., Richardson, A., Percy, N., Almond, J. W. & Evans, D. J. (1998). Encapsidation studies of poliovirus subgenomic replicons. *Journal of General Virology* **79**, 1725-1734.

Barton, D. & Flanegan, J. (1997). Synchronous replication of poliovirus RNA: initiation of negative-strand RNA synthesis requires the guanidine-inhibited activity of protein 2C. *J Virol* **71**, 8482-9.

Basavappa, R., Syed, R., Flore, O., Icenogle, J. P., Filman, D. J. & Hogle, J. M. (1994). Role and mechanism of the maturation cleavage of VP0 in poliovirus assembly: structure of the empty capsid assembly intermediate at 2.9 A resolution. *Protein Sci* **3**, 1651-69.

Baxter, N. J., Roetzer, A., Liebig, H. D., Sedelnikova, S. E., Hounslow, A. M., Skern, T. & Waltho, J. P. (2006). Structure and dynamics of coxsackievirus B4 2A proteinase, an enyzme involved in the etiology of heart disease. *J Virol* **80**, 1451-62.

Bedard, K. M. & Semler, B. L. (2004). Regulation of picornavirus gene expression. *Microbes and Infection* **6**, 702-713.

Bell, Y. C., Semler, B. L. & Ehrenfeld, E. (1999). Requirements for RNA replication of a poliovirus replicon by coxsackievirus B3 RNA polymerase. *J Virol* **73**, 9413-21.

Bienz, K., Egger, D. & Pasamontes, L. (1987). Association of polioviral proteins of the P2 genomic region with the viral replication complex and virus-induced membrane synthesis as visualized by electron microscopic immunocytochemistry and autoradiography. *Virology* **160**, 220-226.

Bienz, K., Egger, D., Rasser, Y. & Bossart, W. (1983). Intracellular distribution of poliovirus proteins and the induction of virus-specific cytoplasmic structures. *Virology* **131**, 39-48.

Bienz, K., Egger, D., Troxler, M. & Pasamontes, L. (1990). Structural organization of poliovirus RNA replication is mediated by viral proteins of the P2 genomic region. *Journal of Virology* **64**, 1156-1163.

Bolanaki, E., Kottaridi, C., Markoulatos, P., Kyriakopoulou, Z., Margaritis, L. & Katsorchis, T. (2007). Partial 3D gene sequences of Coxsackie viruses reveal interspecies exchanges. *Virus Genes* **35**, 129-140.

Bouslama, L., Nasri, D., Chollet, L., Belguith, K., Bourlet, T., Aouni, M., Pozzetto, B. & Pillet, S. (2007). Natural Recombination Event within the Capsid Genomic Region Leading to a Chimeric Strain of Human Enterovirus B. *Journal of Virology* **81**, 8944-8952.

Buchholz, U., Finke, S. & Conzelmann, K. (1999). Generation of bovine respiratory syncytial virus (BRSV) from cDNA: BRSV NS2 is not essential for virus replication in tissue culture, and the human RSV leader region acts as a functional BRSV genome promoter. *J Virol* **73**, 251-9.

Bushati, N. & Cohen, S. (2007). microRNA functions. *Annu Rev Cell Dev Biol* **23**, 175-205.

Caliguiri, L. A. & Tamm, I. (1968). Action of guanidine on the replication of poliovirus RNA. *Virology* **35**, 408-17.

CDC (2006). Progress Toward Interruption of Wild Poliovirus Transmission --- Worldwide, January 2005-March 2006. *Morbidity and Mortality Weekly Report* **55**, 458-462.

Chetverin, A. B., Chetverina, H. V., Demidenko, A. A. & Ugarov, V. I. (1997). Nonhomologous RNA recombination in a cell-free system: evidence for a transesterification mechanism guided by secondary structure. *Cell* **88**, 503-13.

Colbere-Garapin, F., Blondel, B., Saulnier, A., Pelletier, I. & Labadie, K. (2005). Silencing viruses by RNA interference. *Microbes and Infection* **7**, 767-775.

Copper, P. D., Steiner-Pryor, A., Scotti, P. D. & Delong, D. (1974). On the nature of poliovirus genetic recombinants. *J Gen Virol* **23**, 41-9.

Crotty, S., Cameron, C. & Andino, R. (2001). RNA virus error catastrophe: direct molecular test by using ribavirin. *Proc Natl Acad Sci U S A* **98**, 6895-900.

Dahourou, G., Guillot, S., Le Gall, O. & Crainic, R. (2002). Genetic recombination in wild-type poliovirus. *Journal of General Virology* **83**, 3103-3110.

Dayhoff, M. O., Schwartz, R. & Orcutt, B. C. (1978). "A model of Evolutionary Change in Proteins". In *Atlas of protein sequence and structure*, 3 edn: Nat. Biomed. Res. Found.

Domingo, E. & Holland, J. J. (1997). RNA virus mutations and fitness for survival. *Annual Review of Microbiology* **51**, 151-178.

Dougherty, J. D., White, J. P. & Lloyd, R. E. (2011). Poliovirus-mediated disruption of cytoplasmic processing bodies. *J Virol* **85**, 64-75.

Duggal, R. & Wimmer, E. (1999). Genetic recombination of poliovirus in vitro and in vivo: temperature-dependent alteration of crossover sites. *Journal of Virology* **258**, 30-41.

Dytham, C. (2003). Choosing and Using Statistics: A Biologist's Guide, Second edn, pp. 248. Oxford: Blackwell Publishing.

Dytham, C. (2009). Evolved dispersal strategies at range margins. *Proc Biol Sci* **276**, 1407-13.

Egger, D. & Bienz, K. (2002). Recombination of poliovirus RNA proceeds in mixed replication complexes originating from distinct replication start sites. *J Virol* **76**, 10960-71.

Emini, E. A., Leibowitz, J., Diamond, D. C., Bonin, J. & Wimmer, E. (1984). Recombinants of Mahoney and Sabin strain poliovirus type 1: analysis of in vitro phenotypic markers and evidence that resistance to guanidine maps in the nonstructural proteins. *Virology* **137**, 74-85.

Evans, D. J. & Almond, J. W. (1998). Cell receptors for picornaviruses as determinants of cell tropism and pathogenesis. *Trends Microbiol* **6**, 198-202.

Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E. & Mello, C. C. (1998). Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature* **391**, 806-811.

Flint, S. J., Enquist, L. W., Racaniello, V. R. & Skalka, A. M. (2004). Principles of Virology, 2 edn, pp. 918. Washington: ASM Press.

Fox, G., Stuart, D., Acharya, K. R., Fry, E., Rowlands, D. & Brown, F. (1987). Crystallization and preliminary X-ray diffraction analysis of foot-and-mouth disease virus. *J Mol Biol* **196**, 591-7.

Freistadt, M. S., Vaccaro, J. A. & Eberle, K. E. (2007). Biochemical characterization of the fidelity of poliovirus RNA-dependent RNA polymerase. *Virol J* **4**, 44.

Gallei, A., Pankraz, A., Thiel, H. J. & Becher, P. (2004). RNA recombination in vivo in the absence of viral replication. *J Virol* **78**, 6271-81.

Gao, Y., Abreha, M., Nelson, K. N., Baird, H., Dudley, D. M., Abraha, A. & Arts, E. J. (2011). Enrichment of intersubtype HIV-1 recombinants in a dual infection system using HIV-1 strain-specific siRNAs. *Retrovirology* **8**, 5.

Gerber, K., Wimmer, E. & Paul, A. V. (2001). Biochemical and genetic studies of the initiation of human rhinovirus 2 RNA replication: identification of a cis-replicating element in the coding sequence of 2A(pro). *J Virol* **75**, 10979-90.

Gitlin, L., Stone, J. K. & Andino, R. (2005). Poliovirus Escape from RNA Interference: Short Interfering RNA-Target Recognition and Implications for Therapeutic Approaches. *Journal of Virology* **79**, 1027-1035.

Goodfellow, I., Chaudhry, Y., Richardson, A., Meredith, J., Almond, J. W., Barclay, W. & Evans, D. J. (2000). Identification of a cis-acting replication element within the poliovirus coding region. *Journal of Virology* **74**, 4590-4600.

Goodfellow, I. G., Kerrigan, D. & Evans, D. J. (2003). Structure and function analysis of the poliovirus cis-acting replication element (CRE). *RNA* **9**, 124-137.

Goodfellow, I. G., Polacek, C., Andino, R. & Evans, D. J. (2003). The poliovirus 2C cis-acting replication element-mediated uridylylation of VPg is not required for synthesis of negative-sense genomes. *Journal of General Virology* **84**, 2359-2363.

Greve, J. M., Davis, G., Meyer, A. M., Forte, C. P., Yost, S. C., Marlor, C. W., Kamarck, M. E. & McClelland, A. (1989). The major human rhinovirus receptor is ICAM-1. *Cell* **56**, 839-47.

Hahn, C. S., Lustig, S., Strauss, E. G. & Strauss, J. H. (1988). Western Equine Encephalitis Virus is a Recombinant Virus. *Proceedings of the National Academy of Sciences* **85**, 5997-6001.

Herold, J. & Andino, R. (2000). Poliovirus requires a precise 5' end for efficient positive-strand RNA synthesis. *J Virol* **74**, 6394-400.

Herrewegh, A. A., Vennema, H., Horzinek, M. C., Rottier, P. J. & de Groot, R. J. (1995). The molecular genetics of feline coronaviruses: comparative

sequence analysis of the ORF7a/7b transcription unit of different biotypes. *Virology* **212**, 622-31.
Higgins, D. G. & Sharp, P. M. (1988). CLUSTAL: a package for performing multiple sequence alignment on a microcomputer. *Gene* **73**, 237-44.
Hughes, P. J., North, C., Minor, P. D. & Stanway, G. (1989). The complete nucleotide sequence of coxsackievirus A21. *J Gen Virol* **70 (Pt 11)**, 2943-52.
Hutvágner, G., McLachlan, J., Pasquinelli, A., Bálint, E., Tuschl, T. & Zamore, P. (2001). A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. *Science* **293**, 834-838.
Hyypiä, T., Hovi, T., Knowles, N. J. & Stanway, G. (1997). Classification of enteroviruses based on molecular and biological properties. *J Gen Virol* **78 (Pt 1)**, 1-11.
Ishihama, A., Mizumoto, K., Kawakami, K., Kato, A. & Honda, A. (1986). Proofreading function associated with the RNA-dependent RNA polymerase from influenza virus. *J Biol Chem* **261**, 10417-21.
Jaag, H. M. & Nagy, P. D. (2010). The combined effect of environmental and host factors on the emergence of viral RNA recombinants. *PLoS Pathog* **6**, e1001156.
Jaag, H. M., Pogany, J. & Nagy, P. D. (2010). A host Ca2+/Mn2+ ion pump is a factor in the emergence of viral RNA recombinants. *Cell Host Microbe* **7**, 74-81.
Jarvis, T. C. & Kirkegaard, K. (1992). Poliovirus RNA recombination: mechanistic studies in the absence of selection. *EMBO Journal* **11**, 3135-3145.
Jiang, P., Faase, J. A. J., Toyoda, H., Paul, A., Wimmer, E. & Gorbalenya, A. E. (2007). Evidence for emergence of diverse polioviruses from C-cluster coxsackie A viruses and implications for global poliovirus eradication. *Proceedings of the National Academy of Sciences* **104**, 9457-9462.
Jopling, C. L., Yi, M., Lancaster, A. M., Lemon, S. M. & Sarnow, P. (2005). Modulation of hepatitis C virus RNA abundance by a liver-specific MicroRNA. *Science* **309**, 1577-81.
Kafasla, P., Morgner, N., Robinson, C. V. & Jackson, R. J. (2010). Polypyrimidine tract-binding protein stimulates the poliovirus IRES by modulating eIF4G binding. *EMBO J* **29**, 3710-22.
Kaplan, G. & Racaniello, V. R. (1988). Construction and characterization of poliovirus subgenomic replicons. *J Virol* **62**, 1687-96.
Kapuscinski, J. (1995). DAPI: a DNA-specific fluorescent probe. *Biotech Histochem* **70**, 220-33.
Kawamura, N., Kohara, M., Abe, S., Komatsu, T., Tago, K., Arita, M. & Nomoto, A. (1989). Determinants in the 5' noncoding region of poliovirus Sabin 1 RNA that influence the attenuation phenotype. *J Virol* **63**, 1302-9.
Kew, O., Morris-Glasgow, V., Landaverde, M., Burns, C., Shaw, J., Garib, Z., André, J., Blackman, E., Freeman, C., Jorba, J., Sutter, R., Tambini, G., Venczel, L., Pedreira, C., Laender, F., Shimizu, H., Yoneyama, T., Miyamura, T., van Der Avoort, H., Oberste, M., Kilpatrick, D., Cochi, S., Pallansch, M. & de Quadros, C. (2002). Outbreak of poliomyelitis in Hispaniola associated with circulating type 1 vaccine-derived poliovirus. *Science* **296**, 356-359.
Kew, O. M., Sutter, R. W., de Gourville, E. M., Dowdle, W. R. & Pallansch, M. A. (2005). Vaccine-derived Poliovirus and the Endgame Strategy for Global Polio Eradication. *Annual Review of Microbiology* **59**, 587-635.

Khetsuriani, N., Lamonte-Fowlkes, A., Oberst, S., Pallansch, M. A. & Prevention, C. f. D. C. a. (2006). Enterovirus surveillance--United States, 1970-2005. *MMWR Surveill Summ* **55**, 1-20.

King, A. M. (1988). Preferred sites of recombination in poliovirus RNA: an analysis of 40 intertypic cross-over sequences. *Nucleic Acids Research* **16**, 11705-11723.

Kirkegaard, K. & Baltimore, D. (1986). The mechanisms of RNA recombination in Poliovirus. *Cell* **47**, 433-443.

Kräusslich, H. G., Nicklin, M. J., Toyoda, H., Etchison, D. & Wimmer, E. (1987). Poliovirus proteinase 2A induces cleavage of eucaryotic initiation factor 4F polypeptide p220. *J Virol* **61**, 2711-8.

Kuge, S., Saito, I. & Nomoto, A. (1986). Primary structure of poliovirus defective-interfering particle genomes and possible generation mechanisms of the particles. *J Mol Biol* **192**, 473-87.

Lai, M. M. (1992). RNA recombination in animal and plant viruses. *Microbiol. Mol. Biol. Rev* **56**, 61-79.

Lauring, A. S. & Andino, R. (2010). Quasispecies theory and the behavior of RNA viruses. *PLoS Pathog* **6**, e1001005.

Lazzarini, R. A., Keene, J. D. & Schubert, M. (1981). The origins of defective interfering particles of the negative-strand RNA viruses. *Cell* **26**, 145-54.

Lee, H. S., Ahn, J., Jee, Y., Seo, I. S., Jeon, E. J., Jeon, E.-S., Joo, C. H., Kim, Y. K. & Lee, H. (2007). Universal and mutation-resistant anti-enteroviral activity: potency of small interfering RNA complementary to the conserved cis-acting replication element within the enterovirus coding region. *Journal of General Virology* **88**, 2003-2012.

Levy, H. C., Bostina, M., Filman, D. J. & Hogle, J. M. (2010). Catching a virus in the act of RNA release: a novel poliovirus uncoating intermediate characterized by cryo-electron microscopy. *J Virol* **84**, 4426-41.

Lindberg, A. M., Polacek, C. & Johansson, S. (1997). Amplification and cloning of complete enterovirus genomes by long distance PCR. *J Virol Methods* **65**, 191-9.

Lu, H. H., Li, X., Cuconati, A. & Wimmer, E. (1995). Analysis of picornavirus 2A(pro) proteins: separation of proteinase from translation and replication functions. *J Virol* **69**, 7445-52.

Macadam, A. J., Pollard, S. R., Ferguson, G., Skuce, R., Wood, D., Almond, J. W. & Minor, P. D. (1993). Genetic basis of attenuation of the Sabin type 2 vaccine strain of poliovirus in primates. *Virology* **192**, 18-26.

Mason, P. W., Bezborodova, S. V. & Henry, T. M. (2002). Identification and characterization of a cis-acting replication element (cre) adjacent to the internal ribosome entry site of foot-and-mouth disease virus. *J Virol* **76**, 9686-94.

McGoldrick, A., Macadam, A. J., Dunn, G., Rowe, A., Burlison, J., Minor, P. D., Meredith, J., Evans, D. J. & Almond, J. W. (1995). Role of mutations G-480 and C-6203 in the attenuation phenotype of Sabin type 1 poliovirus. *J Virol* **69**, 7601-5.

McKnight, K. L. & Lemon, S. M. (1998). The rhinovirus type 14 genome contains an internally located RNA structure that is required for viral replication. *RNA* **4**, 1569-84.

McLaren, L. C., Holland, J. J. & Syverton, J. T. (1959). The mammalian cell-virus relationship. I. Attachment of poliovirus to cultivated cells of primate and non-primate origin. *J Exp Med* **109**, 475-85.

Mendelsohn, C. L., Wimmer, E. & Racaniello, V. R. (1989). Cellular receptor for poliovirus: molecular cloning, nucleotide sequence, and expression of a new member of the immunoglobulin superfamily. *Cell* **56**, 855-65.

Merino, E. J., Wilkinson, K. A., Coughlan, J. L. & Weeks, K. M. (2005). RNA structure analysis at single nucleotide resolution by selective 2'-hydroxyl acylation and primer extension (SHAPE). *J Am Chem Soc* **127**, 4223-31.

Meyers, G., Rümenapf, T. & Thiel, H. J. (1989). Ubiquitin in a togavirus. *Nature* **341**, 491.

Minor, P. D. (1985). Growth, Assay and Purification of Picornaviruses. In *Virology: a practical approach*, pp. 25-41. Edited by B. W. J. Mahy. Oxford: IRL Press.

Minor, P. D. (2004). Polio eradication, cessation of vaccination and re-emergence of disease. *Nat Rev Microbiol* **2**, 473-82.

Mizutani, S. & Colonno, R. J. (1985). In vitro synthesis of an infectious RNA from cDNA clones of human rhinovirus type 14. *J Virol* **56**, 628-32.

Mueller, S., Wimmer, E. & Cello, J. (2005). Poliovirus and poliomyelitis: a tale of guts, brains, and an accidental event. *Virus Res* **111**, 175-93.

Nagy, P. D. & Simon, A. E. (1997). New insights into the mechanisms of RNA recombination. *Virology* **235**, 1-9.

Nathanson, N. & Kew, O. M. (2010). From emergence to eradication: the epidemiology of poliomyelitis deconstructed. *Am J Epidemiol* **172**, 1213-29.

Nishimura, Y., Shimojima, M., Tano, Y., Miyamura, T., Wakita, T. & Shimizu, H. (2009). Human P-selectin glycoprotein ligand-1 is a functional receptor for enterovirus 71. *Nat Med* **15**, 794-7.

Nishioka, M., Mizuguchi, H., Fujiwara, S., Komatsubara, S., Kitabayashi, M., Uemura, H., Takagi, M. & Imanaka, T. (2001). Long and accurate PCR with a mixture of KOD DNA polymerase and its exonuclease deficient mutant enzyme. *J Biotechnol* **88**, 141-9.

Novak, J. & Kirkegaard, K. (1991). Improved method for detecting poliovirus negative strands used to demonstrate specificity of positive-strand encapsidation and the ratio of positive to negative strands in infected cells. *J Virol* **65**, 3384-7.

Oberste, M., Maher, K., Nix, W., Michele, S., Uddin, M., Schnurr, D., al-Busaidy, S., Akoua-Koffi, C. & Pallansch, M. (2007). Molecular identification of 13 new enterovirus types, EV79–88, EV97, and EV100–101, members of the species Human Enterovirus B. *Journal of Virus Research* **128**, 34-42.

Oberste, M. S., Maher, K., Kennett, M. L., Campbell, J. J., Carpenter, M. S., Schnurr, D. & Pallansch, M. A. (1999). Molecular epidemiology and genetic diversity of echovirus type 30 (E30): genotypes correlate with temporal dynamics of E30 isolation. *J Clin Microbiol* **37**, 3928-33.

Oberste, M. S., Maher, K., Kilpatrick, D. R., Flemister, M. R., Brown, B. A. & Pallansch, M. A. (1999). Typing of human enteroviruses by partial sequencing of VP1. *J Clin Microbiol* **37**, 1288-93.

Pääbo, S., Irwin, D. M. & Wilson, A. C. (1990). DNA damage promotes jumping between templates during enzymatic amplification. *J Biol Chem* **265**, 4718-21.

Paul, A. V., van Boom, J. H., Filippov, D. & Wimmer, E. (1998). Protein-primed RNA synthesis by purified poliovirus RNA polymerase. *Nature* **393**, 280-4.

Paximadi, E., Karakasiliotis, I., Mamuris, Z., Stathopoulos, C., Krikelis, V. & Markoulatos, P. (2006). Genomic analysis of recombinant Sabin clinical isolates. *Virus Genes* **32**, 203-210.

Pelletier, J. & Sonenberg, N. (1988). Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA. *Nature* **334**, 320-5.

Percy, N., Barclay, W., Sullivan, M. & Almond, J. (1992). A poliovirus replicon containing the chloramphenicol acetyltransferase gene can be used to study the replication and encapsidation of poliovirus RNA. *J Virol* **66**, 5040-6.

Pfeiffer, J. K. & Kirkegaard, K. (2003). A single mutation in poliovirus RNA-dependent RNA polymerase confers resistance to mutagenic nucleotide analogs via increased fidelity. *Proc Natl Acad Sci U S A* **100**, 7289-94.

Pilipenko, E. V., Gmyl, A. P. & Agol, V. I. (1995). A model for rearrangements in RNA genomes. *Nucleic Acids Research* **23**, 1870-1875.

Pilipenko, E. V., Poperechny, K. V., Maslova, S. V., Melchers, W. J., Slot, H. J. & Agol, V. I. (1996). Cis-element, oriR, involved in the initiation of (-) strand poliovirus RNA: a quasi-globular multi-domain RNA structure maintained by tertiary ('kissing') interactions. *EMBO J* **15**, 5428-36.

Powell, R. M., Schmitt, V., Ward, T., Goodfellow, I., Evans, D. J. & Almond, J. W. (1998). Characterization of echoviruses that bind decay accelerating factor (CD55): evidence that some haemagglutinating strains use more than one cellular receptor. *J Gen Virol* **79 (Pt 7)**, 1707-13.

Putnak, J. R. & Phillips, B. A. (1981). Differences between poliovirus empty capsids formed in vivo and those formed in vitro: a role for the morphopoietic factor. *J Virol* **40**, 173-83.

Raju, R., Subramaniam, S. V. & Hajjou, M. (1995). Genesis of Sindbis virus by in vivo recombination of nonreplicative RNA precursors. *J Virol* **69**, 7391-401.

Reed, L. & Muench, H. (1938). A simple method of estimating fifty percent end points. *American Journal of Hygeine* **27**, 493-497.

Roivainen, M., Piirainen, L., Hovi, T., Virtanen, I., Riikonen, T., Heino, J. & Hyypiä, T. (1994). Entry of coxsackievirus A9 into host cells: specific interactions with alpha v beta 3 integrin, the vitronectin receptor. *Virology* **203**, 357-65.

Romanova, L., Tolskaya, E., Kolesnikova, M. & Agol, V. (1980). Biochemical evidence for intertypic genetic recombination of polioviruses. *FEBS Letters* **118**, 109-112.

Sabin, A. & Boulger, L. (1973). History of Sabin attenuated poliovirus oral live vaccine strains. *Journal of Biological Standardisation* **1**, 115-118.

Saitou, N. & Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* **4**, 406-25.

Salk, J. E. (1953). Studies in human subjects on active immunization against poliomyelitis. I. A preliminary report of experiments in progress. *J Am Med Assoc* **151**, 1081-98.

Sambrook, J., Russell, D. W., Maniatis, T. & Fritsch, E. F. (2000). Molecular Cloning: A Laboratory Manual, 3rd edn, pp. 999. New York: Cold Spring Harbor Laboratory Press.

Santti, J., Hyypiä, T., Kinnunen, L. & Salminen, M. (1999). Evidence of recombination among enteroviruses. *J Virol* **73**, 8741-9.

Savolainen-Kopra, C., Samoilovich, E., Kahelin, H., Hiekka, A., Hovi, T. & Roivainen, M. (2009). Comparison of poliovirus recombinants: accumulation of point mutations provides further advantages. *JOURNAL OF GENERAL VIROLOGY* **90**, 1859-1868.

Schwarz, D. S., Hutvágner, G., Du, T., Xu, Z., Aronin, N. & Zamore, P. D. (2003). Asymmetry in the assembly of the RNAi enzyme complex. *Cell* **115**, 199-208.

Shafren, D., Dorahy, D., Ingham, R., Burns, G. & Barry, R. (1997). Coxsackievirus A21 binds to decay-accelerating factor but requires intercellular adhesion molecule 1 for cell entry. *J Virol* **71**, 4736-43.

Shafren, D. R., Bates, R. C., Agrez, M. V., Herd, R. L., Burns, G. F. & Barry, R. D. (1995). Coxsackieviruses B1, B3, and B5 use decay accelerating factor as a receptor for cell attachment. *J Virol* **69**, 3873-7.

Shimizu, H., Thorley, B., Paladin, F. J., Brussen, K. A., Stambos, V., Yuen, L., Utama, A., Tano, Y., Arita, M., Yoshida, H., Yoneyama, T., Benegas, A., Roesel, S., Pallansch, M., Kew, O. & Miyamura, T. (2004). Circulation of type 1 vaccine-derived poliovirus in the Philippines in 2001. *J Virol* **78**, 13512-21.

Simmonds, P. (2006). Recombination and selection in the evolution of picornaviruses and other Mammalian positive-stranded RNA viruses. *J Virol* **80**, 11124-40.

Simmonds, P. & Welch, J. (2006). Frequency and dynamics of recombination within different species of human enteroviruses. *J Virol* **80**, 483-93.

Smura, T., Blomqvist, S., Paananen, A., Vuorinen, T., Sobotová, Z., Bubovica, V., Ivanova, O., Hovi, T. & Roivainen, M. (2007). Enterovirus surveillance reveals proposed new serotypes and provides new insight into enterovirus 5'-untranslated region evolution. *J Gen Virol* **88**, 2520-6.

Stanway, G., Hughes, P. J., Mountford, R. C., Reeve, P., Minor, P. D., Schild, G. C. & Almond, J. W. (1984). Comparison of the complete nucleotide sequences of the genomes of the neurovirulent poliovirus P3/Leon/37 and its attenuated Sabin vaccine derivative P3/Leon 12a1b. *Proc Natl Acad Sci U S A* **81**, 1539-43.

Steinhauer, D., Domingo, E. & Holland, J. (1992). Lack of evidence for proofreading mechanisms associated with an RNA virus polymerase. *Gene* **122**, 281-288.

Sutter, R. W., Kew, O. M. & Cochi, S. L. (2008). Poliovirus vaccine-live. In *Vaccines*, 5 edn, pp. 1748. Edited by S. A. Plotkin, W. Orenstein & P. A. Offit: Saunders.

Taniguchi, T., Palmieri, M. & Weissmann, C. (1978). QB DNA-containing hybrid plasmids giving rise to QB phage formation in the bacterial host. *Nature* **274**, 223-8.

Teterina, N., Levenson, E. & Ehrenfeld, E. (2010). Viable polioviruses that encode 2A proteins with fluorescent protein tags. *J Virol* **84**, 1477-88.

Tolskaya, E. A., Romanova, L. I., Blinov, V. M., Viktorova, E. G., Sinyakov, A. N., Kolesnikova, M. S. & Agol, V. I. (1987). Studies on the recombination between RNA genomes of poliovirus: the primary structure and nonrandom distribution of crossover regions in the genomes of intertypic poliovirus recombinants. *Virology* **161**, 54-61.

van Rij, R. (2008). Virus meets RNAi. Symposium on antiviral applications of RNA interference. *EMBO Rep* **9**, 725-9.

Vennema, H., Poland, A., Foley, J. & Pedersen, N. C. (1998). Feline infectious peritonitis viruses arise by mutation from endemic feline enteric coronaviruses. *Virology* **243**, 150-7.

Wang, Z., Day, N., Trifillis, P. & Kiledjian, M. (1999). An mRNA stability complex functions with poly(A)-binding protein to stabilize mRNA in vitro. *Mol Cell Biol* **19**, 4552-60.

Wang, Z., Ren, L., Zhao, X., Hung, T., Meng, A., Wang, J. & Chen, Y. (2004). Inhibition of severe acute respiratory syndrome virus replication by small interfering RNAs in mammalian cells. *J Virol* **78**, 7523-7.

Ward, T., Pipkin, P. A., Clarkson, N. A., Stone, D. M., Minor, P. D. & Almond, J. W. (1994). Decay-accelerating factor CD55 is identified as the receptor for echovirus 7 using CELICS, a rapid immuno-focal cloning method. *EMBO J* **13**, 5070-4.

Weidman, M. K., Yalamanchili, P., Ng, B., Tsai, W. & Dasgupta, A. (2001). Poliovirus 3C protease-mediated degradation of transcriptional activator p53 requires a cellular activity. *Virology* **291**, 260-71.

Westrop, G. D., Wareham, K. A., Evans, D. M., Dunn, G., Minor, P. D., Magrath, D. I., Taffs, F., Marsden, S., Skinner, M. A., Schild, G. C. & Almond, J. W. (1989). Genetic basis of attenuation of the Sabin type 3 oral poliovirus vaccine. *Journal of Virology* **63**, 1338-1344.

White, J. P., Cardenas, A. M., Marissen, W. E. & Lloyd, R. E. (2007). Inhibition of cytoplasmic mRNA stress granule formation by a viral proteinase. *Cell Host Microbe* **2**, 295-305.

Whitton, J. L., Cornell, C. T. & Feuer, R. (2005). Host and virus determinants of picornavirus pathogenesis and tropism. *Nat Rev Microbiol* **3**, 765-76.

WHO (1988). *Polio eradication by the year 2000. Rep. Resolution 41.28.* Geneva: World Health Organisation.

Wimmer, E., Hellen, C. U. & Cao, X. (1993). Genetics of poliovirus. *Annu Rev Genet* **27**, 353-436.

Worobey, M. & Holmes, E. C. (1999). Evolutionary aspects of recombination in RNA viruses. *Journal of General Virology* **80**, 2535-2543.

Wu, W., Blumberg, B. M., Fay, P. J. & Bambara, R. A. (1995). Strand transfer mediated by human immunodeficiency virus reverse transcriptase in vitro is promoted by pausing and results in misincorporation. *J Biol Chem* **270**, 325-32.

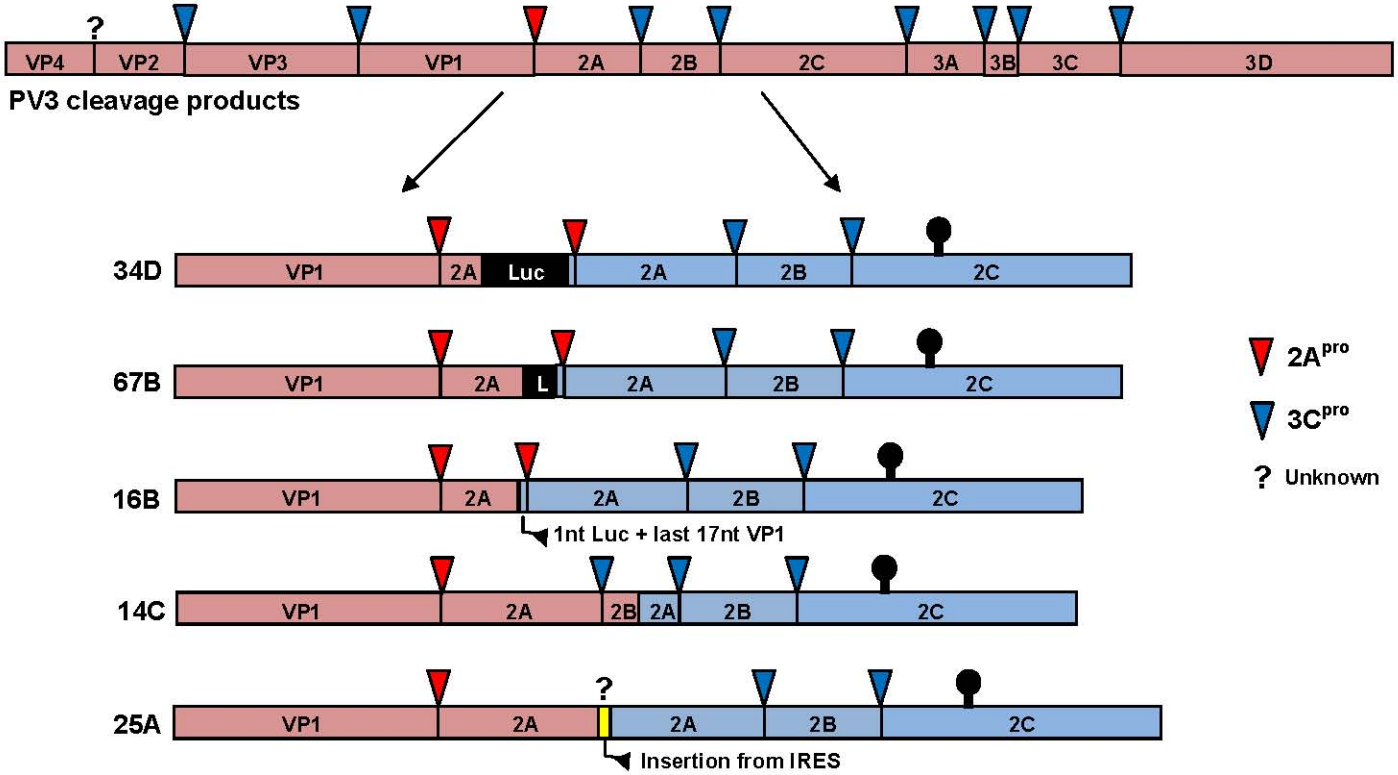
Xiao, C., Bator, C. M., Bowman, V. D., Rieder, E., He, Y., Hébert, B., Bella, J., Baker, T. S., Wimmer, E., Kuhn, R. J. & Rossmann, M. G. (2001). Interaction of coxsackievirus A21 with its cellular receptor, ICAM-1. *J Virol* **75**, 2444-51.

Yamayoshi, S., Yamashita, Y., Li, J., Hanagata, N., Minowa, T., Takemura, T. & Koike, S. (2009). Scavenger receptor B2 is a cellular receptor for enterovirus 71. *Nat Med* **15**, 798-801.

Yi, R., Qin, Y., Macara, I. G. & Cullen, B. R. (2003). Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes Dev* **17**, 3011-6.

Zoll, J., Galama, J. M. & van Kuppeveld, F. J. (2009). Identification of potential recombination breakpoints in human parechoviruses. *J Virol* **83**, 3379-83.

Zuker, M. (1989). On finding all suboptimal foldings of an RNA molecule. *Science* **244**, 48-52.



Appendix 1 Potential cleavage products of selected imprecise recombinants. The PV3 polyprotein and cleavage pattern is outlined at the top of the page. The central polyprotein region has been expanded to further depict the genome structure of several imprecise recombinants (outlined in table 5.1), together with known cleavage sites for virus proteases $2A^{pro}$ and $3C^{pro}$. Approximately 17 nt of the 3' end of VP1 coding region remains after the luciferase gene in the pRLucWT (blue shading), in order to maintain $2A^{pro}$ cleavage site.