INVESTIGATION INTO CELLULAR GENES INVOLVED IN NIPAH VIRUS PATHOGENESIS

BY

NINA J. L. KUZENKO

A Thesis
Submitted to the Faculty of Graduate Studies
In Partial Fulfillment of the Requirements for the Degree of

MASTER OF SCIENCE

Department of Medical Microbiology and Infectious Diseases
University of Manitoba
Winnipeg, Manitoba

THE UNIVERSITY OF MANITOBA

FACULTY OF GRADUATE STUDIES

*****

COPYRIGHT PERMISSION PAGE

Investigation into Cellular Genes Involved in Nipah Virus Pathogenesis

BY

Nina J.L. Kuzenko

A Thesis/Practicum submitted to the Faculty of Graduate Studies of The University of Manitoba in partial fulfillment of the requirements of the degree of

MASTER OF SCIENCE

Table of Contents

Acknowledgements

I thank my supervisor, Dr. Markus Czub, for being so generous with his time and being so eager to help me succeed. Without his help and confidence, I would never have come this far. I also thank him for all the hours he spent completing the level 4 portion of my work, and for helping me with radioactive experiments.

I thank the members of my committee, Dr. Joanne Embree and Dr. R. Daniel Gietz for serving on my committee and for providing valuable suggestions and support during my studies.

I thank my friends and family for their understanding, love and support during my late nights and weekends at the lab, my frustration over failed PCR, and my grumpy days of writing. I thank my sister, Natasha, for her expertise and help with designing my figures.

I owe thanks to each member, both past and present, of the Special Pathogens Program for their incredible willingness to listen and help. In particular, I thank Allen "the human thermometer" Grolla, for devoting so many hours to my work in level 4, and for patience with a seemingly endless stream of questions. You are truly a fountain of knowledge and a great asset to the entire lab. I also thank Bevan Sawatsky for sharing his hard-earned constructs, and for volunteering his time for level 4 work on my behalf. Thanks to the Nipah group for everything, for understanding when I forgot to empty the lab waste and refill the glassware, for listening to my bad renditions of 80s music, and for making the road seem less long.

I thank the staff and students of the Department of Medical Microbiology as well as the NML staff for all of their help. In particular I would like to thank Dr. Michael Carpenter for offering his helpful technical advice and suggestions and all of the members of the DNA Core Facility for helpful discussion and excellent sequencing and primer services.

I also would like to thank everyone who helped me out immensely by proof-reading my thesis: Dennis, Charlene, Bevan, Allen, Markus, and Lisa. Lastly, I would like to thank all of my fellow students who made being a grad student in Med Micro so much fun and an experience I will remember fondly.

...the most informative experiments are frequently those which met most difficulties and had, in principle, less chance to be successful...

- excerpt from Meselsohn, Stahl, and the Replication of DNA, 2001, Yale University Press

List of Figures

List of Tables

List of Abbreviations

α	alpha
A	ampere
ABI	Applied Biosystems
BHK-21	Baby hamster kidney 21
BLAST	basic local alignment search tool
bp	base pair
°C	degrees Celcius
cDNA	complimentary deoxyribonucleic acid
Ci	curie
CL	containment level
CMV	Cytomegalovirus
CO_2	carbon dioxide
CPE	cytopathic effect
CRFK	Crandell feline kidney
dATP	deoxy adenosine triphosphate
DEPC	diethyl pyrocarbonate
dH_2O	distilled water
dNTP	deoxynucleotide triphosphate
DMEM	Dulbecco's modified Eagle's medium
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
E. coli	*Escherichia coli*
EDTA	ethylenediamine tetraacetic acid
EGFP	enhanced green fluorescent protein
ELISA	enzyme linked immunosorbant assay
ES	embryonic stem
F	fusion protein
FA	formaldehyde agarose
FBS	fetal bovine serum
g	gram
G	glycoprotein
GFP	green fluorescent protein
GMEM	Glasgow's minimum essential medium
HeV	Hendra virus
HRP	horseradish peroxidase
Ig	immunoglobulin
∞	infinity
iPCR	inverse polymerase chain reaction
IU	infectious unit
J	joule
JE	Japanese encephalitis
kb	kilobases
L	litre; viral polymerase protein
LB	Luria-Bertani
LTR	long terminal repeat

m	metre
M	matrix protein; molar
MEM	Minimum Essential medium
$MgCl_2$	magnesium chloride
MoMuLV	Moloney Murine Leukemia virus
MOI	multiplicity of infection
MOPS	3-(N-mopholino) propane-sulfonic acid
mRNA	messenger ribonucleic acid
MV	Measles virus
N	nucleoprotein
NaCl	sodium chloride
NaOH	sodium hydroxide
NCBI	National Center for Biotechnology Information
Neo^R	neomycin resistance gene
NiV	Nipah virus
NTR	nitroreductase
OAS	2'-5'-oligoadenylate synthetase
OD	optical density
Opti-MEM	Opti-modified Eagle's medium
^{32}P	radioactive isotope of phosphorus
pA	polyadenylation signal
PBS	phosphate buffered saline
PBS-T	phosphate buffered saline with 0.1% Tween-20
PCR	polymerase chain reaction
PGK-1	phosphoglycerate kinase
PKR	protein kinase receptor
ψ	psi
PVC	phosphoprotein, C protein, V protein
RACE	rapid amplification of complimentary ends
rcf	relative centrifugal force
RNA	ribonucleic acid
RT-PCR	reverse-transcriptase polymerase chain reaction
SA	splice acceptor
SAP-PCR	single strand-specific second-strand arbitrary polymerase chain reaction
SDS	sodium dodecyl sulfate
SLIP	solo long terminal repeat inverse polymerase chain reaction
SSC	salt sodium citrate
TAE	tris acetate ethylene diamine tetraacetic acid
TE	tris ethylene diamine tetraacetic acid
TEF-1	transcription enhancer factor 1
U	unit
UTP	uridine triphosphate
UV	ultraviolet
V	volt
VSV	Vesicular Stomatitis virus

ABSTRACT

Nipah virus (NiV) is a recently emergent zoonotic containment level 4 agent. It is highly detrimental to mammalian hosts and tissue culture cells. As a paramyxovirus, many aspects of the viral life cycle have been inferred by analogy to more extensively studied paramyxoviruses, although certain specific characteristics have yet to be deduced. Several cellular genes have the potential to play a role in NiV pathogenesis, for instance, the (as yet unknown) gene or genes encoding the cellular receptor may code for a protein that is integral to the susceptibility of the cell to NiV. Without expression of such genes, the normally susceptible cell may become resistant to infection by NiV.

In order to investigate cellular genes that potentially have a role in viral pathogenesis, a gene trapping strategy using a retroviral vector was developed. A retrovirus functioning as a gene trap was created and used to infect cells. As in any retroviral infection, the genome of the retrovirus, in this case the gene trap, was integrated at random into the host genome, causing an insertional mutation, possibly leading to knockout of the gene at the location of integration. In a population of cells infected with the retroviral vector, each cell likely contained a single integration at a different location in the genome. This population of cells was exposed to NiV for selection of cellular clones that were made resistant to infection or cytopathic effects by NiV through the mutation of a gene by integration of the retrovirus. Cell clones that survived exposure to NiV were analysed using polymerase chain reaction-based strategies to investigate a strategy to identify the location of integration and the gene that was knocked out leading to the creation of a NiV resistant cellular clone.

1. INTRODUCTION

1.1 Nipah virus

Nipah virus (NiV), a recently emergent paramyxovirus, was first identified in 1999 following an outbreak of severe febrile encephalitis of humans, and to a lesser extent in pigs, in the states of Perak, Negeri Sembilan, and Selangor within Peninsular Malaysia between September 1998 and April 1999 (Field, *et al.*, 2001; Middleton, *et al.*, 2002). The virus was named after Sungai Nipah in Negeri Sembilan, the home village of the worker from whom the virus was first isolated (Nor, 1999; Nor, *et al.*, 1999).

The disease associated with NiV infection was first characterised as a pronounced respiratory illness in pigs, at times with the sudden death of sows and boars (Nor, *et al.*, 1999). Initial investigations into the causative agent focused on Japanese encephalitis (JE) virus, which is endemic to the region. Vaccination and insect control attempts that would have controlled JE failed to make an impact on containing the new disease, and thus JE was discounted. Further investigations confirmed a new "Hendra-like" paramyxovirus (Nor, 1999), NiV, as the etiologic agent (Chua, *et al.*, 2000). The outbreak of disease in the pig population appeared to be associated with the concurrent epidemic of viral meningitis and encephalitis (Nipah encephalitis) in pig farm workers in the region. Molecular characterisation later confirmed NiV to be the cause of both diseases (Nor, *et al.*, 1999). Whereas porcine disease has been observed to be respiratory or neurological, human disease is predominantly neurological (Nor, 1999; Nor, *et al.*, 1999). During the outbreak, 265 human cases of NiV infection were reported which led to 105 associated fatalities, indicating a case fatality rate of almost 40% (Nor, *et al.*, 1999; Bossart, *et al.*, 2001). Within months of the start of the outbreak, NiV spread

rapidly to other parts of Malaysia, as well as into Singapore. As a result of the outbreak, over 1 million pigs were culled in an effort to contain the disease, devastating the Malaysian swine industry.

It is suspected that the probable reservoir of NiV and Hendra virus (HeV), a very close relative, are fruit bats of the genus *Pteropus* as animals in many species have tested as seropositive (Nor, *et al.*, 1999; Field, *et al.*, 2001). Wildlife surveillance in Malaysia has shown that some fruit bats contain neutralising antibodies to NiV (Field, *et al.*, 2001) as well as show positive results for the virus neutralisation test (Nor, 1999). NiV has displayed the capability to cause disease in a wide range of species including humans, pigs, cats, dogs, horses, bats, and goats (Nor, *et al.*, 1999; Chua, *et al.*, 2000, Middleton, *et al.*, 2002).

Although the mode of transmission between animals is not certain, experimental infections of pigs with a human isolate of NiV have provided suggestions and insight into this question. In the laboratory, detection of NiV in respiratory and oropharyngeal secretions of animals penned with infected animals has suggested that infection may have occurred via exposure to secretions and excretions of penmates. In the field, direct contact of pigs housed together, along with coughing and sneezing which produce aerosolised droplets, have been suggested to have roles in transmission (Middleton, *et al.*, 2002).

Human to human transmission of NiV has yet to be documented and this mode of transmission is not believed to occur. A recent study in Malaysia employing the detection of anti-NiV immunoglobulin M (IgM) and IgG in serum as a measure of NiV transmission from infected patients to health care workers found no evidence for

nosocomial transmission despite extensive contact between infected patients and health care workers. Furthermore, even though health care workers are exposed to patients shedding virus in urine and respiratory secretions, the data nonetheless support the conclusion that the risk of nosocomial transmission of NiV is very low (Mounts, *et al.*, 2001).

1.2 Virus Characteristics

Electron microscopy has revealed that NiV particles are pleomorphic and can be filamentous or spherical, ranging in diameter from 40-600 nm. These particles consist of a nucleocapsid surrounded by an envelope (*The Springer Index of Viruses*, 2001). Within the nucleocapsid lies the viral genome (Figure 1) that is a linear, non-segmented, single-stranded, negative sense RNA of 18.2 kilobases (kb) (Wang, *et al.*, 2001).

NiV is very closely related to HeV, another recently emergent containment level 4 (CL4) pathogen. In Canada, NiV is handled as a CL4 agent on the basis of its high risk to individuals and the community, and the lack of a vaccine or effective treatment for at risk or exposed individuals (as per in-house Laboratory Centre for Disease Control Material Safety Data Sheet). Both NiV and HeV are members of the family *Paramyxoviridae*, subfamily *Paramyxovirinae*. NiV and HeV are sufficiently different from the other paramyxoviruses, with respect to their unusually large genomes, long untranslated regions, and their ability to infect a wide range of host species (Bossart, *et al.*, 2001; Wang, *et al.*, 2001), that creation of a new genus was warranted. They have been classified into their own genus, *Henipavirus*, which was recently approved as an official genus by the International Committee on Taxonomy of Viruses.

1.3 Viral Life Cycle

The life cycle of NiV is described on the basis of a typical paramyxovirus life cycle. It includes several steps that require or have the potential for interaction with the host cell. The first step is host cell recognition. The cellular receptor, as yet unidentified, is recognised by the virus. This is followed by fusion of the viral membrane with the cell membrane that is mediated by the viral glycoproteins. The NiV G (glycoprotein) protein is thought to recognise the host receptor while the NiV F (fusion) protein is believed to mediate fusion of the viral envelope with the cellular membrane (Chua, *et al.*, 2000). Cytoplasmic entry is followed by uncoating of the virus and release of the negative sense viral RNA genome, that is subsequently transcribed into individual monocistronic messenger RNAs (mRNAs). Translation of the mRNA uses the cell's ribosomal machinery, thus providing another point of interaction between the virus and the host cell.

In order for the virus to replicate its full-length negative sense genome (Figure 1), it is postulated that a sufficient level of N (nucleoprotein) protein is required (Soria, *et al.*, 1974; Banerjee, 1987). Once sufficient viral N protein is present, it binds the newly synthesised positive sense RNA thus preventing synthesis from terminating after the first gene as would be the case for mRNA synthesis. Rather, it continues through to the end of the genome making a full-length positive sense RNA. This generates a template for synthesis of the full-length negative sense RNA genome. Another point of virus-host cell interaction occurs when the genome is transported to the cellular membrane for assembly with the viral proteins, which have undergone glycosylation and other modifications in

the endoplasmic reticulum and Golgi, for virus assembly. The virus is believed to bud from the cellular membrane, producing an enveloped particle. Viral replication likely occurs exclusively in the cytoplasm (*Fields Virology*, 2001).

As alluded to above, several steps in the life cycle of NiV involve interaction of the virus with the cell such as binding, fusion, entry, replication, and budding. Cellular gene products potentially involved in NiV pathogenesis may include various factors used by the virus in its life cycle such as the cell surface receptor for viral binding and entry or a cellular cofactor for the viral polymerase. Additionally, another group of genes whose products lead to cytopathic effects (CPE) may be induced. The cell may respond to virus infection by expressing genes that cause cellular damage that is observed as CPE in an effort to halt the spread of the virus by eliminating its host cell. Furthermore, the cell may also express antiviral genes in response to infection, such as protein kinase receptor (PKR), 2'-5'-oligoadenylate synthetase (OAS), or Mx genes. Specifically, Measles virus (MV), another paramyxovirus, has been shown to induce apoptotic genes (Esolen, *et al.*, 1995; Fugier-Vivier, *et al.*, 1997) and the production of interferons and other cytokines (Volckaert-Vervliet and Billiau, 1977; Schneider-Schaulies, *et al.*, 1993). Therefore, given all of the possible interactions between the virus and host cell, it is evident that an array of cellular genes may contribute to the overall pathogenesis of NiV.

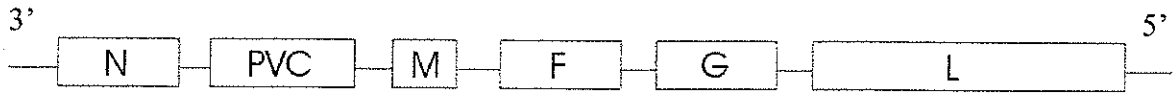

Figure 1. NiV genome organisation (adapted from *The Springer Index of Viruses*, 2001). N – nucleoprotein; PVC – phosphoprotein, V protein, C protein; M – matrix; F – fusion; G – glycoprotéine; L – polymerase.

1.4 Studying Virus Pathogenesis

The aim of this study was to develop a genetic approach to better understand the interactions between the host cell and the invading virus. Specifically, the goal was to identify cellular genes intrinsic to the viral life cycle or involved in defence against virus infection. One approach to study the function of a gene potentially involved in virus pathogenesis is to alter its expression. Mutation of an endogenous gene is likely to lead to its altered expression.

Although life cycles of closely related viruses such as MV have been described, many aspects related to the life cycle of NiV remain unknown. As a result, specific genes involved in pathogenesis, such as the cell surface receptor, have yet to be identified. As such, directed mutagenesis of these unknown genes is impossible. Instead, a mutational vector that works at random must be employed, so that a population of cells, each with its own random mutation, is created. This system allows screening of the entire genome in an attempt to identify candidate genes.

1.5 Gene Trapping

Gene trapping, a relatively new technique, has been used successfully for obtaining flanking gene sequence information (Friedrich and Soriano, 1991; Friedrich and Soriano, 1993; Chen, *et al.*, 1994; Skarnes, *et al.*, 1995; Wurst, *et al.*, 1995; Gogos, *et al.*, 1996; Gogos, *et al.*, 1997; Medico, *et al.*, 2001; Jin, *et al.*, 2002). This method allows for the global investigation of genes from which the entire genome can be screened in a single pass, while being cost-effective and non-labour intensive.

Gene trapping works by taking advantage of certain properties of retroviruses, namely the capability to integrate into host sequences and, consequently, the ability to cause a mutation in the host genome. A gene trap vector is a plasmid that, when transcribed, acts like the retroviral genome so that it is packaged into the retrovirus. Upon infection, the gene trap will integrate into the host genome, at which time a selectable marker in the trap is expressed, thus providing an easily detected confirmation of integration. As the integration is permanent, it will be passed on to all daughter cells that would also express the selectable marker.

The gene trap vector contains additional features that can be used for subcellular localisation and functional analysis experiments (Gogos, *et al.*, 1996). Also, gene trapping, in contrast to other genome screening methods, is an unbiased method of gene identification since it will knock out genes randomly, thus allowing for the study of genes independent of the abundance of mRNA in the cell.

Gene trap vectors were originally used for mutational analysis in mice (Friedrich and Soriano, 1991; Friedrich and Soriano, 1993; DeGregori, *et al.*, 1994; Skarnes, *et al.*, 1995; Wurst, *et al.*, 1995) by selecting for insertional mutations into developmentally regulated genes. Gogos and colleagues (1996, 1997) used gene trapping for the identification of several genes involved in the development of myoblasts. Their work demonstrated that the lysosomal cysteine protease cathepsin B is induced in myoblasts, and that its deficiency can lead to myopathy of I-cell disease (Gogos, 1996). In subsequent studies, this group investigated the effect of MyoD protein on myotube formation from myoblasts (Gogos, *et al.*, 1997). Similarly, Chen and colleagues (1994) have used a retroviral gene trap to disrupt the transcription enhancer factor 1 (TEF-1)

gene in embryonic stem (ES) cells in order to study tissue-specific regulation of gene expression. Despite many studies using gene trapping to study mammalian gene regulation, this technique has yet to be employed in the study of an infectious disease and the interaction between the host and the pathogen.

A gene trap vector contains retroviral long terminal repeats (LTRs), a psi (ψ) packaging signal, and a reporter gene or a selectable marker, all within a plasmid backbone. The integrase retroviral protein brought into the cell inside the virion is responsible for catalysing the integration of the provirus into the host genome. The LTRs and the sequence situated between them become integrated permanently into the host sequence (*Fields Virology*, 2001). The gene trap vector can thus be used to infect tissue culture cells to produce cell clones that will be referred to in this thesis as trapped clones or trapped cells. As this is a random event, the integration may occur at any locus, which will be different in each cell, giving rise to a mixed population of trapped cellular clones. When this integration occurs, the host sequence at that locus is mutated as the gene or regulatory element in that location is knocked out and rendered non-functional.

It is possible that a gene involved in NiV pathogenesis, such as the cell surface receptor, can be knocked out, thus producing a cell with one non-functional allele of this gene. Although the other allele of the gene is unaltered, the net effect may be downregulation of the expression of this protein. If, for instance, the receptor protein for NiV is absent, it is possible that the virus can no longer recognise and infect its target cell. Thus, a cellular gene that is involved in NiV pathogenesis is knocked out, cells resistant to either infection or CPE by the virus may be created.

It is important to note that retroviral vector infection is performed at a low multiplicity of infection (MOI) so that a single integration per cell is obtained and only a single gene per cell is knocked out (Gogos, *et al.*, 1996). This will facilitate analysis of the phenotype of a cell caused by a single gene mutation. Only one integration per cell is desired in order to ensure that a single gene knockout is being observed, however, it follows that only one allele can possibly be mutated. The other functional alleles can possibly take over expression of the gene, meaning the mutation is recessive. Unless a hypodiploid or functionally haploid cell line is used, a complete knockout eliminating gene expression is difficult. It is also possible that another gene altogether may take over expression of the protein or a similar protein with the same function.

1.6 pROSA-GFNR

The gene trap vector used in this work, pROSA-GFNR (Figure 2), was constructed by Medico and colleagues (2001) based on a vector constructed by Soriano and colleagues (1991) and was kindly provided for our group. This vector was constructed from a backbone of the plasmid pBR322 that was modified and named pGen$^-$ (Soriano, *et al.*, 1991). The gene trap contains two LTRs from Moloney Murine Leukemia virus (MoMuLV) that, along with the sequence in between the LTRs, constitute the provirus. This is the portion of the gene trap that integrates into the host genome upon retrovirus infection. The enhancer in the 3'LTR has been deleted, hence the name pGen$^-$, for enhancer minus. This deletion aids in avoiding possible interaction with a genomic promoter (Friedrich and Soriano, 1993).

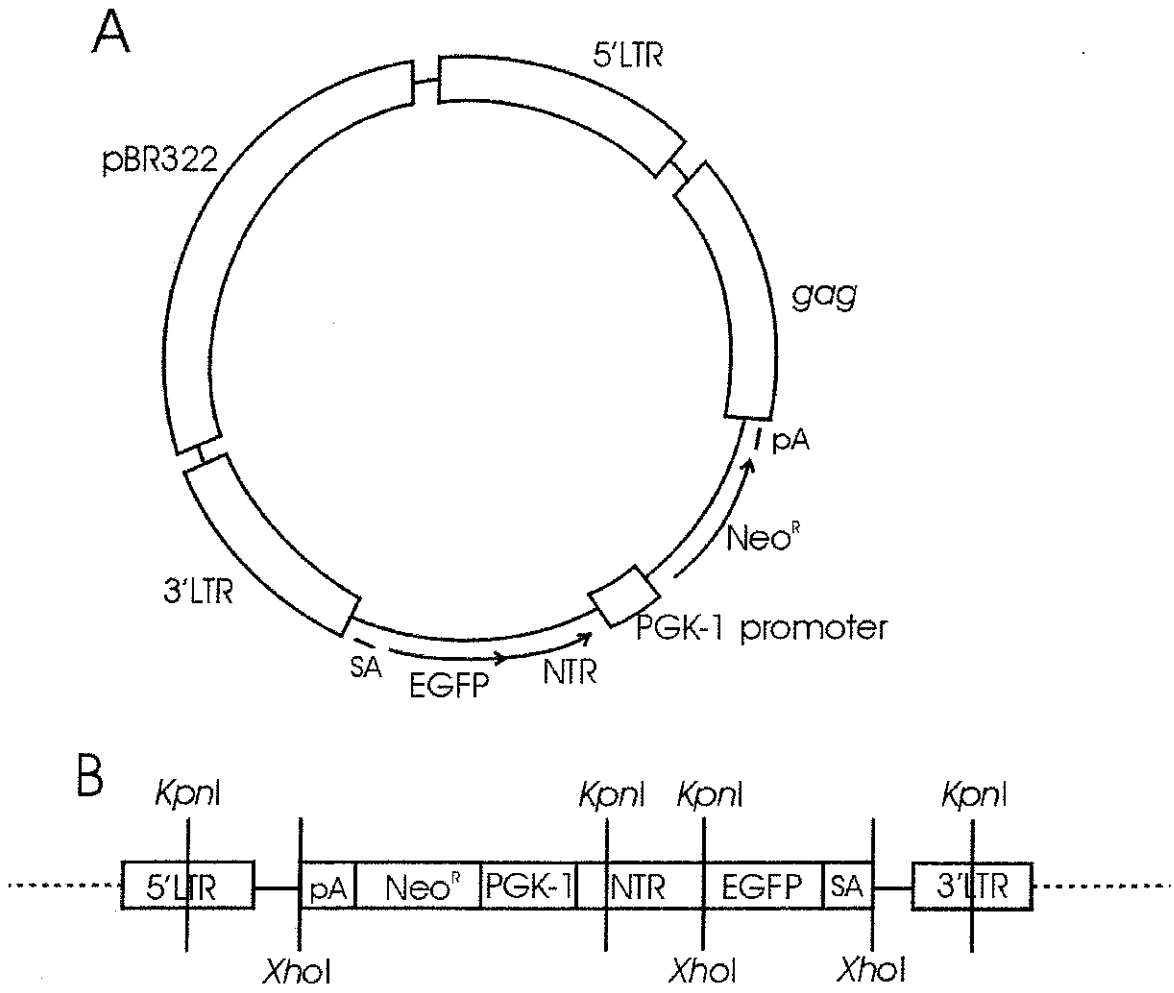

Figure 2a. pROSA-GFNR plasmid map. Circular representation of the approximate relative sizes of the regions of the plasmid (total size 8 kb).
b. Linear representation of pROSA-GFNR. The portion integrated as a provirus into the endogenous host sequence (dotted line). *Xho*I digest sites produce fragments of the following sizes: 4740 base pairs (bp), 2400 bp, and 860 bp. *Kpn*I digest sites produce fragments of 3119 bp, 2915 bp, 1554 bp, and 487 bp.

5'LTR – 5' long terminal repeat, pA – polyadenylation signal, Neo^R – neomycin resistance gene, PGK-1 promoter – phosphoglycerate kinase-1 promoter, NTR – nitroreductase gene, EGFP – enhanced green fluorescent protein gene, SA – splice acceptor, 3'LTR – 3' long terminal repeat.

Internal to the LTRs is a cassette containing several features, such as the selectable marker used in this study. The cassette is positioned in the opposite orientation relative to the LTRs to accommodate additional features not used in this study (Friedrich and Soriano, 1993). The neomycin phosphotransferase (resistance) gene is under the control of the PGK-1 gene promoter and confers resistance to the drug G418, also known as Geneticin. Expression from the PGK-1 promoter and consequently of the neomycin resistance gene is regulated by the cell, and only occurs when the provirus is integrated, that is, when the cell is trapped.

1.7 Identifying Pertinent Genes

Following trapping, cells are exposed to NiV to select for clones resistant to NiV infection (under normal conditions NiV kills all cells), based on the premise that a population of clones will contain knockouts in different cellular genes potentially involved in NiV pathogenesis. When several clones are present together in a population, the number of different clones in the sample can be identified by detecting the provirus using a Southern blot (Friedrich and Soriano, 1993; Chen, *et al.*, 1994; Gogos, *et al.*, 1997). Individual clones resistant to NiV can be analysed in more detail to identify which gene or regulatory element was mutated by the gene trap. Given the sequence of the trap, sequencing the flanking host sequence can be accomplished several ways. Among the previously described strategies are inverse polymerase chain reaction (iPCR) (Silver and Keerikatte, 1989; Gogos, *et al.*, 1996), solo LTR iPCR (SLIP) (Jin, *et al.*, 2002), 5' rapid amplification of complementary ends (RACE) (Gogos, *et al.*, 1996; Medico, *et al.*, 2001), and adapter ligation PCR (Gogos, *et al.*, 1997). The techniques

explored in this study were Southern blot, adapter ligation PCR, iPCR, and single strand specific second strand arbitrary PCR (SAP-PCR, Dr. Michael Carpenter, Division of Bloodborne Pathogens and Hepatitis, Health Canada, personal communications, 2003).

1.8 Hypothesis

Knockout of a cellular gene can enable normally susceptible cells to become resistant to infection or CPE by Nipah virus.

1.9 Thesis Objectives

The overall goal of this project was to establish a protocol for identifying host genes that play a role in viral pathogenesis in order to better understand these interactions. Specifically, NiV was chosen for establishing this method as it is recently emergent and provides a rapid and definitive read out system. To meet this goal, specific objectives were outlined:

- To create a retroviral gene trap vector.
- To create mutant, trapped cell lines by insertional mutagenesis using the retroviral gene trap vector.
- To select for trapped cells resistant to fusion induced by expression of the NiV glycoproteins.
- To expose the mutated cells to NiV to establish mutant cell lines resistant to infection or CPE.
- To investigate methods for determining the sequences flanking the site of insertion of the gene trap for identification of the interrupted gene.

2. MATERIALS AND METHODS

2.1 Cells and Viruses

Crandell feline kidney (CRFK) cells were grown in Minimum Essential medium (MEM, Invitrogen) supplemented with 10% heat inactivated fetal bovine serum (FBS, Wisent, Inc.), 1% non-essential amino acids (Invitrogen), 1% sodium bicarbonate (Invitrogen), and 1% sodium pyruvate (Invitrogen); 1% penicillin-streptomycin (Invitrogen) was added to the MEM.

Baby hamster kidney-21 (BHK-21) cells were grown in Glasgow's Minimum Essential medium (GMEM, Invitrogen) supplemented with 10% FBS; 1% penicillin-streptomycin was added to the GMEM, unless cells were to be used for transfection.

293T (human embryonic kidney) cells were grown in Dulbecco's Modified Eagle's medium (DMEM, Sigma), supplemented with 10% FBS.

All cultures were grown in a 5% CO_2, H_2O-saturated atmosphere at 37°C, and passaged by treatment with trypsin [Trypsin-EDTA (ethylenediamine tetraacetic acid), Invitrogen] and dilution in culture medium every two to three days.

Tissue culture cells that were to be frozen for storage were grown to confluency, treated with trypsin, and resuspended in 5 ml of medium. Cells were centrifuged at a speed of 300 relative centrifugal force (rcf) for three minutes to pellet the cells, the supernatant was discarded. Cells were resuspended in a few ml (typically 3 ml) of solution A (Appendix 2), followed by addition of an equal volume of solution B (Appendix 2). Cells were aliquoted into pre-cooled sterile cryovials. In order to cool the cells slowly, the tubes were placed in a Cryo 1°C™ container (Nalgene) at -80°C overnight, and transferred to a liquid nitrogen freezer the next day.

Escherichia coli (*E. coli*) XL1-Blue cells were made competent for transfomation in-house. Cells were grown overnight in a 5 ml starter culture in Luria-Bertani (LB) broth. The next day, 50 ml of LB broth was inoculated at a concentration of 1/500 with the overnight culture, and grown until the optical density (OD) of 0.5-0.8 absorbance units at 600 nm (OD_{600}) wavelength was reached. The cells were placed on ice for 20 minutes, then centrifuged at 900 rcf for 10 minutes at 4°C. Cells were resuspended in 5 ml TSS buffer (Appendix 2) then aliquoted prior to freezing at -80°C.

NiV and HeV were obtained from the Centers for Disease Control and Prevention (Atlanta, GA). Virus stocks were prepared by infecting an 80% subconfluent monolayer of Vero E6 or CRFK cells in a 162 cm^2 flask with NiV or HeV at an MOI=0.1. After two days, the supernatant was harvested, and cell debris was pelleted. The infectivity titre of 5 x 10^6 infectious units (IU)/ml was determined by end point titration using CPE and reverse-transcriptase PCR (RT-PCR) assays. Virus stocks were aliquoted then frozen in liquid nitrogen. Handling of NiV and HeV was done under CL4 conditions as outlined in an in-house Laboratory Centre for Disease Control Material Safety Data Sheet. Additional CL4 handling procedures are outlined in the Health Canada Laboratory Biosafety Guidelines (http://www.hc-sc.gc.ca/pphb-dgspsp/publicat/lbg-ldmbl-96/index.html).

2.2 Transfection

In general, transfections of plasmid DNA into tissue culture cells were performed as follows. Cells were seeded one day in advance in order to provide a confluency of 80% on the day of transfection. DNA for each well to be transfected was prepared in a

microcentrifuge tube as follows. Four μg of plasmid was combined with 250 μl of serum-free Opti-Modified Eagle's Reduced Serum medium (Opti-MEM, Invitrogen). As a positive control, 4 μg of $pGFP_{CMV}$ [Cytomegalovirus (CMV) promoter-driven green fluorescent protein (GFP) plasmid construct] was mixed with 250 μl of Opti-MEM. As a mock control, cells were given only 250 μl of Opti-MEM, without DNA. For each transfection, incubation of combined DNA and Opti-MEM was carried out at room temperature for five minutes. Five μl of a cationic lipid transfection reagent, Lipofectamine 2000™ (Invitrogen), was then added to each tube, gently tapped for mixing and allowed to incubate at room temperature for 15 minutes to allow DNA-Lipofectamine complexes to form. Finally, the volume in each tube was adjusted to 1 ml with Opti-MEM, and pipetted gently onto the cells after removal of spent culture medium and a single wash with sterile phosphate-buffered saline (PBS). For most transfections, medium was changed the next day.

2.3 Transformation

In general, transformations of plasmid DNA into competent cells were performed as follows. For each transformation, a 50 μl aliquot of competent cells was thawed on ice for 10 minutes. DNA was added to the cells, tapped to mix, and incubated on ice for 30 minutes. Cells were heat-shocked at 42°C for 60 seconds, returned to ice for two minutes, followed by addition of 500 μl of SOC medium (Appendix 2). Cells were incubated at 37°C with shaking for one hour, and plated on selective media (LB agar with either 100 μg/ml ampicillin or 35 μg/ml kanamycin). The following day, colonies were

picked, inoculated into 5 ml selective LB broth medium, grown overnight, and inoculated into a large-scale culture in selective LB broth medium.

2.4 Plasmids

The plasmid $pVSVG_{CMV}$ was created from a construct obtained from the Institute for Virology (Marburg, Germany). This vector was constructed by subcloning the Vesicular Stomatitis virus (VSV) G gene from the pBluescript II SK (+) phagemid™ (Stratagene) into the pBK-CMV phagemid™ (Stratagene). The vector M174 was obtained from the Institute of Virology (Wurzburg, Germany). This vector contains retroviral *gag* and *pol* genes from MoMuLV, with an ampicillin resistance gene for selection. $pGFP_{CMV}$ was obtained from the Institute for Virology (Marburg, Germany). This vector contains the GFP gene cloned into the pcDNA3.1™ vector (Invitrogen) whose expression is driven by the CMV promoter and which contains an ampicillin resistance marker. pROSA-GFNR (section 1.6, Figure 2) was kindly provided by Dr. Enzo Medico (University of Torino Medical School, Italy).

All plasmids were amplified by transformation of competent cells, as described in section 2.3, except pROSA-GFNR, for which the protocol was modified as follows. pROSA-GFNR was amplified by transformation into STBL2™ *E. coli* (high-efficiency transformation cells, Invitrogen). STBL2™ cells were used instead of the usual XL1-Blue competent cells, as attempts to amplify the plasmid using XL1-Blue cells were unsuccessful. Growth was carrried out in ampicillin selective media at 30°C, the required growth temperature for STBL2™ cells.

For all plasmids, DNA was extracted using a QIAGEN Plasmid Maxi Kit™ or a QIAGEN Plasmid Midi Kit™ according to manufacturer's instructions.

2.4.1 pROSA-GFNR Sequencing

pROSA-GFNR (Figure 2) was obtained for use by our group, however, the sequence of the construct was not provided. In order to conduct experiments including primer design and restriction digests, it was necessary to sequence the entire plasmid. This was begun using primers based on the neomycin resistance gene with new primers designed after each sequencing result was obtained. As the vector contains two LTRs that flank the cassette, it was considerably difficult to design primers in the vicinity of the LTRs while avoiding priming at both LTRs, leading to a double sequence result. Through background research, it was discovered that pROSA-GFNR was constructed from pBR322 (Soriano, *et al.*, 1991). Primers were designed based on pBR322 sequence in order to attempt to sequence through the LTRs from both directions, and eventually have the two sequences meet up and overlap. After a complete plasmid sequence was constructed, a final sequencing run was performed over the entire plasmid to verify the sequence and identify unknown nucleotides. All primers used to sequence the plasmid, as well as the complete sequence, are listed in Appendix 3.

DNA sequencing was performed by the in-house DNA Core Facility using Applied Biosystems (ABI) 3100 and ABI PRISM® 377 DNA sequencers with Applied Biosystems Big Dye Terminator chemistry, Version 3.0 reaction mix. Data was collected on Macintosh G4 computers using ABI Data Collection Software, Version 1.1. Data was

analysed using ABI Sequence Analysis Software, Version 3.7, Chromas and Sci Ed Central for Windows 95.

2.5 PCR, Amplicon Analysis

PCR reactions described in this manuscript were performed according to the following general guidelines using a HotStarTaq DNA Polymerase™ kit (Qiagen) and a Whatman Biometra TGradient thermocycler. A 50 µl reaction was prepared containing: 5 µl 10x PCR buffer, 1 µl 25 mM $MgCl_2$ solution, 1 µl 10 mM (each) deoxynucleotide triphosphate (dNTP) solution (Invitrogen), 0.25 µl [5 units (U)/µl] HotStarTaq™ polymerase, 0.5 µl each (0.1 µM final concentration) forward and reverse primers, a varying amount of DNA, and sterile dH_2O up to 50 µl. The amount of DNA added to the reaction depended upon the specific conditions necessary for that reaction. Cycling conditions, in general, are outlined in Table 1.

Table 1. General PCR cycling parameters

Number of Cycles	Time	Temperature	Cycle Type
1x	15 min	95°C	activation of HotStarTaq™ polymerase
30-50x	30 sec 30 sec 30 sec – 5 min	94°C 50-55°C 68-72°C	denaturation annealing elongation
1x	5 min	68-72°C	additional elongation
1x	∞	4°C	final cooling

RT-PCR reactions were performed, in general as follows using a QIAGEN One-Step RT-PCR™ kit. A 50 µl reaction volume was prepared: 10 µl 5x PCR buffer, 2 µl

400 µM (each) dNTP solution, 0.5 µl (0.6 µM final concentration) each forward and reverse primers, 2 µl One-Step RT-PCR Enzyme Mix™, a varying amount of RNA, and sterile, RNase-free dH_2O (provided in kit) up to 50 µl. Cycling conditions, in general, are outlined in Table 2.

Table 2. General RT-PCR cycling parameters

Number of Cycles	Time	Temperature	Cycle Type
1x	30 min	50°C	reverse transcription
1x	15 min	95°C	activation of HotStarTaq™ polymerase
30-50x	30 sec 30 sec 30 sec – 5 min	94°C 50-55°C 72°C	denaturation annealing elongation
1x	5 min	72°C	additional elongation
1x	∞	4°C	final cooling

2.5.1 Agarose Gel Electrophoresis

Visualisation of amplified PCR products was typically performed by 1% agarose gel electrophoresis where gels were made with and run in 1x tris acetate EDTA (TAE) buffer, and contained ethidium bromide (0.002% ethidium bromide in 1% agarose, using 10 mg/mL stock solution) for DNA staining. One tenth of a PCR reaction (typically 5 µl) was combined with 6x 30% glycerol loading buffer (Appendix 2). A molecular weight marker, 2-Log DNA Ladder™ (New England Biolabs, Appendix 4), was also run on the gel for approximation of PCR amplicon size. Electrophoresis was carried out for approximately 40 minutes at 100 volts (V) and 400 milliamps (mA).

2.5.2 Basic Local Alignment Search Tool (BLAST) Analysis

Sequencing results obtained by the ABI software were analysed using the internet program BLAST accessible through the National Center for Biotechnology Information (NCBI) website (http://www.ncbi.nlm.nih.gov). Comparison of sequences was conducted using the BLAST program for homology to those sequences already entered into the bank.

2.6 Gene Trap Vector Production

2.6.1 Concept

We wished to use a retroviral gene trap vector to produce gene knockout cellular clones. The retrovirus was created by transfecting each of the following plasmids into eukaryotic tissue culture cells for expression of retroviral proteins and transcription of the vector genome to produce retroviral vector particles: M174, which contains retroviral *gag-pol* genes, $pVSVG_{CMV}$, containing the VSV glycoprotein gene, and the gene trap plasmid, pROSA-GFNR, which contains the *cis*-acting sequences necessary for the retrovirus (Vile and Russell, 1995; Boris-Lawrie and Temin, 1994). The envelope glycoprotein gene, G, from VSV was used because VSV has the ability to infect a wide range of hosts, giving the resulting retrovirus the same property (*Fields Virology*, 2001). As well, the VSV glycoprotein is not readily shed from the surface of the retrovirus (Vile and Russell, 1995).

Transfected cells expressed the *gag-pol* and envelope genes that were required for production of viral proteins such as capsid proteins (*gag*), reverse transcriptase (*pol*) and envelope glycoproteins (G), which then assembled into viral particles. The gene trap

transcript was selectively packaged as the genome of the vector due to its ψ packaging signal, which is lacking in the other two plasmids, thus allowing recognition by the retroviral proteins for inclusion as the genome (Boris-Lawrie and Temin, 1994; Markowitz, *et al.*, 1988). Since the gene trap plasmid contained no information for viral proteins, the resulting retrovirus was only capable of a single infection and therefore not able to replicate, a feature that makes the vector safer to use (Vile and Russell, 1995). As an enveloped virus, it budded into the supernatant, from which it was harvested (Boris-Lawrie and Temin, 1994).

2.6.2 Cellular Transfection

Before seeding 293T cells, which do not adhere well to tissue culture dishes, the wells were incubated for one hour at 37°C with a 0.10 mg/ml solution of poly-D-lysine hydrobromide (Sigma), followed by two washes with sterile dH_2O. 293T cells were seeded into the pre-treated six-well dish with antibiotic-free DMEM at a density to provide approximately 80% subconfluency the next day. 293T cells were chosen for production of the retroviral particles as these cells are efficiently transfected.

Co-transfection of 4 μg each of $pVSVG_{CMV}$, M174, and pROSA-GFNR was performed as outlined in section 2.2. A positive control ($pGFP_{CMV}$) and a mock control were also performed. After 24 hours, the positive control cells were observed using a Zeiss Axiovert 100 microscope with a fluorescent light and an ultraviolet (UV) filter to assess the efficiency of transfection based on the percentage of GFP-expressing cells in the positive control well. The presence and proportion of green cells was used as an indicator for the efficiency of transfection in other wells, and thus for production of

infectious retrovirus in the other wells. The medium was not changed until the first harvest at 24 hours, when Opti-MEM was replaced by the same volume of DMEM. Supernatants from virus-containing wells were harvested at 24 and 48 hours, pooled between wells if applicable, centrifuged at 400 rcf for five minutes to pellet cells and debris, aliquoted, and frozen at -80°C.

2.7 Determination of Retroviral Stock Infectivity

2.7.1 Cell Infection

CRFK cells were seeded into a twelve-well dish at a density to yield a confluent monolayer after three days of growth, and were allowed to adhere overnight.

One aliquot (700 µl) of retroviral vector stock was thawed at 37°C. Spent medium was removed and cells were washed once with sterile PBS. The virus stock was mixed with an equal volume of medium and seven serial 1:10 dilutions were prepared giving a total of eight serial dilutions of virus, starting with undiluted and ending with 10^{-7} dilution. The total volume of virus dilution, 0.5 ml, was added to each of eight wells of CRFK cells in sequence, with one negative control well receiving only a medium change. Cells were incubated as usual and the virus was allowed to adsorb for a few hours after which an additional 0.5 ml of medium was added to each of the wells. The cells were allowed to grow without a medium change for three days.

2.7.2 Cell Selection, Calculation of Infectivity Titre

In order to select for cells with an integrated retroviral vector, cells were exposed to Geneticin (Invitrogen) at increasing concentrations. The gene trap cassette contains a

neomycin resistance gene under the control of the PGK-1 promoter, meaning that when a cell contains the integrated retroviral vector, the neomycin resistance gene will be expressed and the cell will be resistant to Geneticin, as detailed in section 1.6. The well in which no cells survive the selection process can be interpreted as having been exposed to no infectious retrovirus, and is considered the end point of the dilution, from which an approximate titre can be estimated.

Starting three days post-infection, cells were selected with 400 μg/ml Geneticin for two days. The drug concentration was increased gradually to avoid shocking the cells with a high concentration at the start. Cells were then selected with 600 μg/ml Geneticin for eleven days, at which time the drug concentration was increased to 800 μg/ml. The concentration of 800 μg/ml was chosen as it was similar to values seen in the literature (Gogos, *et al.*, 1996, Gogos, *et al.*, 1997, Weidhaas, *et al.*, 2000). Cells in the wells which received a low dilution (high amount of virus) grew to confluency soon after the start of selection, but were left untouched as these wells were not of interest for the calculation of infectivity titre. Selection was continued until all cells in a well died, similar to the negative control, thus representing the end point of the dilution. Infectivity titre was estimated as in the following example: if all cells in the 10^{-5} virus dilution well died, then the virus titre was approximately 1-10 x 10^4 IU/0.5 ml, i.e. ~ 10^4 IU/ml.

2.8 Gene Trapping of Tissue Culture Cells

Tissue culture cells (CRFK or BHK-21) were seeded into six-well tissue culture dishes at approximately 10^5 cells/well, and allowed to adhere overnight. This concentration of cells was found to be suitable as the cells grew to confluency in two

days. This period of growth post-exposure to the retroviral vector allowed time for cellular expression of the neomycin resistance gene from the provirus before drug selection began.

In order to be certain that only one insertional mutation occured per cell, an MOI=0.1, calculated based on the retroviral infectivity titre (10^4 IU/ml) and the number of cells seeded (10^5 cells/ml), was used as discussed in section 1.5. It was important for only a single integration to occur per cell so that only a single gene would be knocked out, and the effect of the single mutation could be studied.

The retroviral vector was thawed at 37°C, and diluted with culture medium to a volume of 1 ml. Spent medium was removed, cells were washed once with sterile PBS, and the virus solution was added to the cells.

2.9 Cell Selection

2.9.1 Integration Selection

Following exposure to the retroviral gene trap, cells containing an integrated gene trap expressed the neomycin resistance gene contained in the cassette. Cells containing the integrated provirus were selected with Geneticin. Cells grew to confluency in two days in the six-well dish, and were passaged to a 75 cm^2 flask where antibiotic selection was begun. Cells were grown in tissue culture medium containing 600 μg/ml Geneticin for two days, and were maintained in medium containing 800 μg/ml Geneticin. Throughout these experiments, all of the cells were passaged into fresh dishes with none of the cells discarded in order to ensure that all clones originally generated by the retroviral infection remained in the population. After approximately five days, the cells

in the flask reached confluency, and were transferred into six-well dishes for subsequent fusion experiments or into 162 cm^2 flasks for expansion and freezing for storage (section 2.1). Throughout the selection procedure, medium changes to remove excess cell debris were performed as needed.

2.9.2 Fusion Resistance Selection

2.9.2.1 Concept

The first steps in NiV infection include attachment to and fusion of virus with the host cell, utilising the NiV G and NiV F glycoproteins, respectively. When cells become infected with NiV, expression of the viral glycoproteins on the cell surface causes neighbouring cells to fuse, leading to the formation of large syncytia, which die. Cells not capable of fusion may contain a mutation in a gene related to virus entry. Cells unable to fuse upon exposure to NiV F and NiV G were positively selected using an *in vitro* tissue culture system employing the NiV glycoproteins, F and G, cloned individually into the vector pczCFG5 IEGZ (provided by Dr. Dirk Lindemann, Institute of Virology, Wurzburg, Germany) by Bevan Sawatsky (unpublished work, 2001) and transfected into tissue culture cells for expression. Expression of NiV F or NiV G alone in a cell will not lead to cell fusion. A single cell must express both of the glycoproteins in order to fuse with a neighboring cell (Bevan Sawatsky, unpublished work, 2001).

It is hypothesised that cells resistant to fusion contain the integrated retroviral vector in the gene for the cellular receptor, thus rendering them resistant to fusion. Fusion resistant cells are therefore of great interest with respect to identification of the cellular receptor used by NiV. In order to avoid unnecessary modification to trapped

cells, subconfluent wild type CRFK cells were transfected with both pczCFG5 IEGZ-NiV F and pczCFG5 IEGZ-NiV G, and trapped cells were transferred to the same dish for fusion with the transfected wild type cells by adherance to the dish in the empty spaces. Since wild type cells are fully susceptible to fusion leading to syncytia formation and cell death, these cells were eliminated at this step, along with any trapped cells susceptible to fusion. Following fusion selection, Geneticin selection further ensured that only trapped cells remained. Two rounds of fusion selection ensured that the trapped cells remaining had all been exposed to the fusion glycoproteins and had been selected as resistant to fusion.

2.9.2.2 Cell Transfection and Overlay

CRFK or BHK-21 wild type cells to be transfected were seeded in a six-well dish at a density able to give 50% subconfluency the next day. The trapped cells were seeded in a separate six-well dish such that the cells would also be approximately 50% subconfluent on the same day that the wild type cells were transfected.

For each well to be transfected, one tube of plasmid DNA and reagent mix was prepared as described in section 2.2. pczCFG5 IEGZ-NiV F and pczCFG5 IEGZ-NiV G were co-transfected into the wild type CRFK cells, and positive ($pGFP_{CMV}$) and mock control experiments were also performed. The transfected cells were allowed to take up the plasmid(s) for 3-4 hours. The trapped cells were then treated with trypsin and resuspended in 1 ml of medium, and transferred to the dish containing the transfected cells, from which spent medium was removed prior to addition of the trapped cells. Through overlay of the trapped cells, the transfected cells, at approximately 50%

subconfluency, became surrounded by the trapped cells, which adhered to the dish in the empty spaces.

2.9.2.3 Cell Fusion Selection

After 24 hours, cells were checked by microscopy using a UV filter for expression of the GFP positive control, and for fusion of cells, indicating expression of the NiV F and NiV G glycoproteins by the transfected cells. The formation of syncytia was monitored over several days, with medium changes to remove debris as needed, until no new syncytia formed (about one week). At this time, elimination of remaining wild type cells was performed by Geneticin selection (800 μg/ml) for 3 days, which is a sufficient length of time to select against and kill wild type cells. After the first round of fusion selection was complete, the procedure was repeated, once again using wild type transfected cells overlaid with the trapped cells followed by Geneticin selection. The fusion resistant selected cells were passaged; some were seeded into fresh six-well dishes for future experiments under CL4 conditions and some were expanded in a 162 cm^2 flask for expansion and freezing for storage (section 2.1).

2.10 NiV Resistance Selection

Cells that were selected for neomycin resistance and for fusion resistance were seeded into six-well dishes at approximately 80% subconfluency for exposure to NiV under CL4 conditions. An aliquot of NiV was thawed at 37°C, and diluted with culture medium to 100 IU/ml (MOI=0.0001). Spent culture medium was removed, cells were washed once with sterile PBS and the virus solution was added to the cells. Cells were

incubated as usual, and observed for CPE and for surviving cells. Medium changes to remove cell debris were performed as needed.

When cell clones reached a size of a few hundred cells per clone, all cells in the well were treated with trypsin and passaged into four wells in a 24-well dish for expansion. Subsequent passages were performed to expand cells for freezing and for extraction of DNA for molecular analysis. Cell supernatants were collected and frozen for future CPE assays for the presence of NiV persistent infection.

After transfer to the 24 well dishes, the cell clones were assigned names based on their well, for example, 1A or AA1. Clones with a name beginning with a number and containing one letter were CRFK cells and clone names containing two letters followed by a number were BHK-21 cells.

2.11 CL4 Samples

DNA from cells exposed to NiV under CL4 conditions was extracted using either DNAzol™ (Invitrogen) or a QiaAmp DNA Mini Kit™ (Qiagen), according to manufacturer's instructions. DNAzol™ is a guanidine-detergent lysing solution, which denatures all proteins and inactivates any NiV infectivity in the samples. Once DNAzol™ was added to the cells, the tubes were brought out of high containment via a dunk tank of quaternary ammonium compound detergent/disinfectant solution (5% Micro-Chem Plus™ solution, National Chemical Laboratories, Inc., Philadelphia, PA) for molecular work under CL2 conditions. Samples extracted using the Qiagen kit required additional inactivation steps in the protocol, where after addition of the first buffer (AL) and proteinase K, samples were heated to 56°C for 15 minutes and boiled for 30 minutes

to inactivate any remaining NiV, transferred to a fresh tube, and transferred out of high containment via a dunk tank.

RNA was extracted from NiV exposed cells using TRIzol LS™ reagent (Invitrogen), which is a mono-phasic solution of phenol and guanidine isothiocyanate. RNA was extracted from NiV using a QIAamp Viral RNA Mini Kit™ (Qiagen) where the first buffer, AVL, is a viral lysis buffer to inactivate any NiV. Samples were dunked out from high containment in the same manner as the DNA samples.

2.12 Analysis of Cellular Clone Properties

2.12.1 Persistent Infection CPE Test

A persistently infected cell supports a low level of NiV replication and may show CPE but is not killed by the virus infection. Persistent infection indicated that a cell was susceptible to initial infection by NiV. The first method to test for persistent infection in cells that survived exposure to NiV relied on observation of CPE such as syncytia formation, which is typical of NiV infection. All cell clones were tested for persistent infection by applying their cell culture supernatants to fresh wild type cells of the same cell line. If NiV was persistent in the cells, replication and budding into the supernatant would cause CPE in the wild type cells.

Wild type cells were seeded into a 24-well dish to provide a nearly confluent monolayer. Under CL4 conditions, supernatants were either transferred directly from NiV exposed trapped cells to wild type cells, or stored frozen supernatants were thawed at 37°C and applied to wild type cells. Cells were monitored by microscopy for CPE for several days.

2.12.2 Persistent Infection Viral RNA Detection Test

RT-PCR was the second method used for detection of NiV indicating persistent infection of cells. RNA was extracted from NiV exposed cells using TRIzol LS™ as per manufacturer's instructions (see also section 2.11). A typical RT-PCR reaction (section 2.5) was performed using 1 µl (~200 ng) of RNA from cells. A positive control was performed using 1 µl (~40 ng) of NiV RNA. A negative control water reaction was also included. Primers targeting NiV M and NiV NP (M5 and M3, NP5 and NP3, respectively, Table 4, Appendix 1) were used, with annealing at 50°C and 30 second elongation at 72°C for 35 cycles. One tenth of the reaction product was electrophoresed in a 2% agarose gel and photographed.

2.12.3 HeV Exposure Test

Selected clones resistant to infection by NiV were exposed to HeV. NiV exposed cellular clones were seeded into six-well dishes and exposed to HeV as described for NiV exposure in section 2.10. RNA was extracted from the cells using TRIzol LS™ and from HeV using the QIAamp Viral RNA Mini Kit™. Two sets of primers (HeV4912f, HeV5113r and HeV11266f, HeV11870r, Table 4, Appendix 1) that specifically detected HeV and not NiV were used. RT-PCR was performed using a QIAGEN One-Step RT-PCR™ kit using 1 µl (~200 ng) of the cellular RNA extraction, with 1 µl (~40 ng) of HeV RNA in a positive control reaction, and 1 µl (~40 ng) of NiV RNA in a HeV-specificity control reaction, with annealing at 55°C, and elongation at 72°C for 1 minute, for 30 cycles.

2.13 Southern Blot Assay - Biotinylated Probe

2.13.1 Riboprobe Production

Using pROSA-GFNR as a template, primers within the neomycin resistance gene (neoT7 and neoSP6, Appendix 1) were designed in order to produce a 500 bp amplicon (Figure 3). The forward primer, neoT7, contained a T7 RNA polymerase promoter recognition site. The reverse primer, neoSP6, contained an SP6 RNA polymerase promoter recognition site. Both primers were designed with at least 5 bases upstream of the promoter to aid the polymerase in recognition of the site.

PCR was carried out as outlined in section 2.5 using 10 ng of pROSA-GFNR plasmid DNA as the template with neoT7 and neoSP6 primers. The annealing temperature was 60°C and elongation was done at 72°C for 1.5 minutes for 40 cycles. The entire PCR reaction product was gel purified (1% agarose in 1x TAE), and the 500 bp amplicon was extracted and purified using the QIAquick Gel Extraction Kit™ (Qiagen). The purified fragment was quantified by gel electrophoresis using a standardised marker (2 Log Ladder™) for comparison.

The purified fragment was used as the template for synthesis of a biotinylated RNA probe, produced using the Biotin RNA Labelling Mix™ (Roche) with T7 RNA polymerase or SP6 RNA polymerase (both New England Biolabs) according to manufacturer's instructions. The reaction incorporates biotin-labelled uridine triphosphate (UTP), using either T7 or SP6 RNA polymerase. To verify the production and size of the RNA probe, half of the reaction product (10 µl), and an RNA marker (0.24-9.5 Kb RNA Ladder™, Invitrogen) were loaded on a 1% formaldehyde agarose

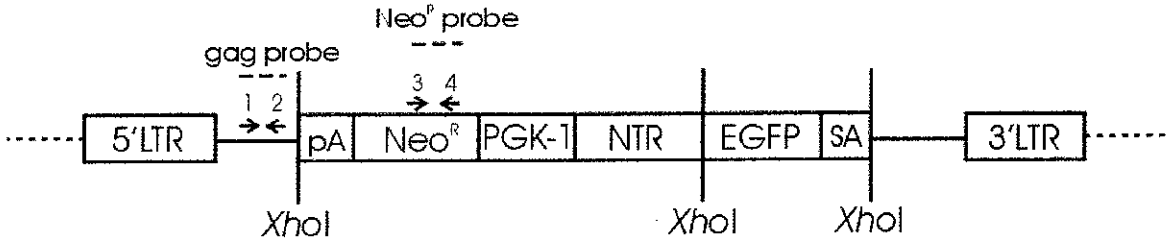

Figure 3. Design of probes for Southern blots. Primers were used to amplify regions of the plasmid pROSA-GFNR to produce templates for the production of probes for Southern blots. Primers 1 (ipcr20) and 2 (ipcr19) generated a 400 bp amplicon for the probe based on the *gag* region. Primers 3 (neoT7) and 4 (neoSP6) generated a 500 bp amplicon for the probe based on the neomycin resistance gene. (dotted line – endogenous sequence)

(FA) gel made with and run in 1x FA buffer (Appendix 2) with 5x RNA loading buffer (Appendix 2), and visualised by ethidium bromide staining with UV illumination.

2.13.2 Riboprobe Biotinylation Verification

To verify that the RNA produced was biotinylated, the above-mentioned FA gel was blotted onto a nylon membrane in a similar fashion to a Northern blot. After photographing the gel following electrophoresis, the gel was washed in dH_2O twice for 15 minutes at room temperature, then twice for 15 minutes in 10x salt sodium citrate buffer (SSC, Appendix 2) at room temperature. A capillary blotting stack was assembled, with all pieces of filter paper cut to the size of the gel and air bubbles removed between layers, from bottom to top as follows: wick GB002 filter paper (Turboblotter™ Blotter Pack, Schleicher & Schuell, Keene, NH), gel with wells facing up, Hybond N^+ ™ membrane (Amersham), a "frame" of plastic wrap on the edges of the gel, 3 sheets of GB002 filter paper soaked in 2x SSC, several sheets of GB004 filter paper, paper towels, weight. 10x SSC was used as the transfer buffer and transfer was allowed to proceed overnight at room temperature.

The membrane was removed from the stack and the RNA was cross-linked to the membrane by UV light at 70 000 joules (J)/cm^2 for 1 minute using a Hoefer UVC 500 UV Crosslinker, and then rinsed in 2x SSC. The membrane was incubated in 5% skim milk made with PBS with 0.1% Tween-20 (PBS-T) for 1 hour at room temperature, then incubated for 1 hour in 1% skim milk PBS-T with a streptavidin-horseradish peroxidase (HRP) conjugate (1:500 dilution, Amersham Life Science). The membrane was washed twice for 15 minutes in PBS-T at room temperature. The HRP was then detected using

an ECL^{+} plus^{TM} kit (Amersham) according to manufacturer's instructions, and the chemiluminescent reaction was detected on Hyperfilm^{TM} (Amersham) and developed using a Feline 14 Automatic X-Ray Film Processor (Fischer Industries Inc).

2.13.3 Southern Blot Control - pROSA-GFNR

In order to verify the specificity and sensitivity of detection of the biotinylated probe, decreasing amounts (10 ng, 1 ng, 100 pg, and 10 pg) of pROSA-GFNR plasmid DNA were restriction digested with *Xho*I (20 000 U/ml, New England Biolabs) at 37°C for 2 hours. The samples were electrophoresed in a 1% agarose gel as usual with a DNA marker, and photographed. The gel was prepared for blotting as follows: 30 minute incubation in depurination solution, two 20 minute incubations in denaturation solution, two 20 minute incubations in neutralisation solution (all recipes in Appendix 2), with rinsing in dH_2O between buffers and again at the end. The capillary transfer stack, as described in section 2.13.2, was assembled in the same way, with the exception that the membrane was Hybond N^{TM} (Amersham) and the transfer buffer was 20x SSC. Following the transfer procedure, the membrane was removed from the stack and cross-linked for 1 minute.

Pre-hybridisation of the membrane was performed at 68°C with rotation (VWR Scientific Products hybridisation oven) in a 50 ml tube containing the membrane and 3 ml of QuikHyb^{TM} solution (Stratagene) for 15 minutes. One aliquot (2 µl) of the riboprobe, and 100 µl sonicated salmon sperm DNA (blocking agent, 100 mg/ml, Stratagene) were thawed on ice and mixed. After prehybridisation, 1 ml of the pre-hybridisation solution was removed from the tube, added to the probe mixture, and the

entire volume returned to the hybridisation tube for hybridisation at 68°C with rotation for 1 hour. Two 15 minute stringency washes were performed at room temperature in each of the following solutions: 2x SSC with 0.1% sodium dodecyl sulfate (SDS), 1x SSC with 0.1% SDS, 0.1x SSC with 0.1% SDS.

The membrane was incubated in 5% skim milk in PBS-T for 1 hour at room temperature, then incubated for 1 hour in 1% skim milk in PBS-T with a 1:500 dilution of the streptavidin-HRP conjugate. Two 15 minute and two 5 minute washes of the membrane in PBS-T at room temperature were performed. The HRP was detected using ECL^{+} plus^{TM} reagent and Hyperfilm^{TM} as described in section 2.13.2.

2.13.4 Southern Blot Control - Genomic DNA Samples

A control blot with a probe for the neomycin resistance gene (500 bp) was performed in order to verify that the detection system could detect the provirus in genomic DNA. The fragment being detected was the 2.4 kb *Xho*I restriction digest product (Figure 2). The expected size in the genomic DNA was known and was the same as in the plasmid positive control.

DNA from cellular clones (3D and 4A) and wild type (CRFK) cells was prepared for Southern blot analysis as follows. Ten µg of genomic DNA (each of NiV exposed trapped cells and wild type cells) was digested with 1 µl *Xho*I for 1 hour, boosted with 1 µl *Xho*I, and digested overnight at 37°C. A positive control of 10 ng of pROSA-GFNR was digested with 0.5 µl *Xho*I, also overnight at 37°C. Samples were electrophoresed, photographed, and Southern blotted followed by probing the membrane and exposing to film as described in section 2.13.3.

2.14 Southern Blot Assay - Radioactively Labeled Probe

2.14.1 Labeled Probe Production

PCR to produce a 500 bp amplicon of the neomycin resistance gene, gel purification and quantitation were performed as described in section 2.13.1. The purified amplicon (~50 ng) was used as template for synthesis of a α^{32}P deoxy adenosine triphosphate (dATP) labelled probe using a Random Primed DNA Labeling Kit™ (Roche). Redivue™ α^{32}P dATP (3000 curie (Ci)/mmol, Amersham) was incorporated into a DNA probe using Klenow enzyme (provided in kit), as per kit instructions. This probe was always made fresh and used immediately.

2.14.2 Southern Blot Control - pROSA-GFNR

A dilution series (10 ng, 1 ng, 100 pg, 10 pg) of *Xho*I-digested pROSA-GFNR plasmid DNA was prepared, electrophoresed in a 1% agarose gel, blotted to Hybond N and cross-linked as described in section 2.13.3. Pre-hybridisation was carried out at 64°C with rotation in a Techne Hybridiser HB-1D hybridisation oven with 10 ml QuikHyb™ for approximately 30 minutes. The freshly prepared probe was mixed with 100 µl sonicated salmon sperm DNA and boiled for 2 minutes prior to addition to the hybridisation tube. Hybridisation was carried out at 61°C overnight. Stringency washes were performed: one rinse and 15 minute incubation each in 2x SSC with 0.1% SDS and 0.1x SSC with 0.1% SDS.

Blots were sealed in plastic pouches and exposed to a phosphor screen (Molecular Dynamics) for varying lengths of time. After exposure, the screen was scanned using a Typhoon 9410 Variable Mode Imager phosphorimager (Molecular Dynamics) using

Typhoon Scanner Control software, and data was analysed with ImageQuant 5.2 software (Molecular Dynamics). The phosphor screen was recharged by a 5 minute exposure to bright light.

2.14.3 Southern Blot Control - Genomic DNA Samples

As in section 2.13.4, the neomycin resistance gene was detected in genomic DNA as a control for the detection system with plasmid DNA as the positive control. Ten μg of DNA from 4A and from wild type CRFK cells, as well as 10 ng of pROSA-GFNR plasmid DNA, were digested, electrophoresed and Southern-blotted to Hybond N^{TM} as described in sections 2.13.3 and 2.13.4. Freshly made α^{32}P dATP labelled probe (500 bp neomycin resistance gene probe) was hybridised to the blots and detected on a phosphorimager as described in section 2.14.2.

2.14.4 Southern Blot Assay - *gag* Region Detection

A Southern blot for detection of the *gag* region of the provirus was performed in order to approximate the size of the restriction fragment that would later be the target in PCR assays. Knowing the size of the fragment would have been useful in choosing the elongation time, and for predicting the size of the amplicon.

Primers (ipcr19 and ipcr20, Appendix 1) were designed based on the *gag* region of pROSA-GFNR (Figure 3) to produce a 400 bp amplicon to use as a template for the production of a Southern blot probe. This probe was designed to detect an *Xho*I digest fragment containing both retroviral vector and endogenous sequences (Figure 3). The

length of this digest fragment was unknown due to the unknown location of an *Xho*I site in the endogenous sequence.

PCR to produce the template for the probe was performed using 10 ng of pROSA-GFNR plasmid DNA with 55°C annealing temperature and 30 second elongation at 72°C for 40 cycles. The entire reaction was gel purified as described in section 2.13.1. As in section 2.14.1, the amplicon was used as a template to produce a α^{32}P dATP labelled probe. A pROSA-GFNR dilution blot (as in section 2.14.2) was prepared and blots of genomic DNA (1B, 3C, 3D, 4A, wild type CRFK, AA6, AD6, BD3, wild type BHK-21) with a plasmid positive control were prepared (as in section 2.13.4). The probe was hybridised to pROSA-GFNR and genomic DNA samples and detected as described in section 2.14.2.

2.15 Control PCR Reaction

In order to ensure that the cellular DNA to be used as template in all future PCR reactions contained the integrated gene trap as expected, it was necessary to perform PCR to detect the integrated trap within the genomic DNA. A region unique to the gene trap not present elsewhere in the genome was the neomycin resistance gene. Primers were designed based on the 5' and 3' ends of this gene (neo5f, neo3r, Appendix 1), to amplify an 800 bp PCR product. PCR was performed using 100 ng of genomic DNA, with annealing at 55°C, and elongation at 72°C for 1 minute, for 45 cycles. A positive control of 10 ng of pROSA-GFNR plasmid DNA was used. Negative controls of wild type CRFK and BHK-21 DNA and water reactions were performed. PCR products were visualised on a 1% agarose gel as usual.

2.16 Adapter-ligation PCR

2.16.1 Concept

The first strategy employed to address the task of identifying the sequences flanking the site of retroviral integration was adapter-ligation PCR (Figure 5). Since the sequence flanking the integration site was unknown, designing a primer based on the endogenous sequence was impossible. Although the location of restriction sites in the flanking sequence could not be predicted, it was reasonable to assume that a site would be located in the relative vicinity of the integration site. With the aim of producing known sequence ends for a specific PCR reaction, the genomic DNA was restriction digested with *Kpn*I to produce overhangs that were then ligated to adapters. The adapters generated an end of known sequence from which a primer could be designed and paired with a primer based on the provirus. This would selectively amplify the fragment containing parts of both the proviral and endogenous sequences.

2.16.2 Adapter Preparation

Two oligonucleotides (linker and Kpnlink, Figure 4a, Appendix 1) were designed with complimentary sequences. One oligonucleotide was longer, having a restriction site hangover on the 3' end to hybridise to the genomic DNA digested with the same restriction enzyme (Figure 4b). Five µl of linker and 20 µl of Kpnlink, with 10 µl 10x hybridisation buffer (Appendix 2) and 65 µl dH_2O were combined and annealing was

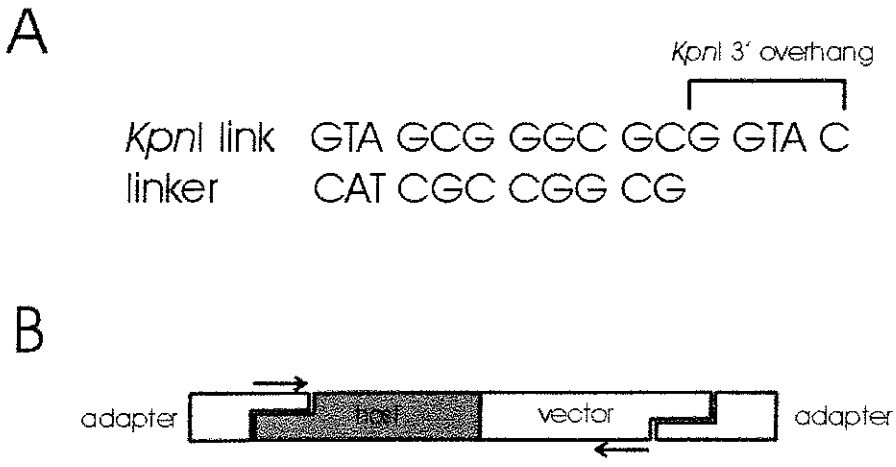

Figure 4a. Oligonucleotides for production of the adapter. Complimentary oligonucleotides were designed to hybridise and produce an adapter with a 3' *Kpn*I digest overhang.

b. Primer binding sites for adapter-ligation PCR. Schematic representation of primer binding sites onto the digested target template containing part proviral vector and part endogenous host sequence. One primer binds the adapter sequence and the other binds to the proviral vector sequence.

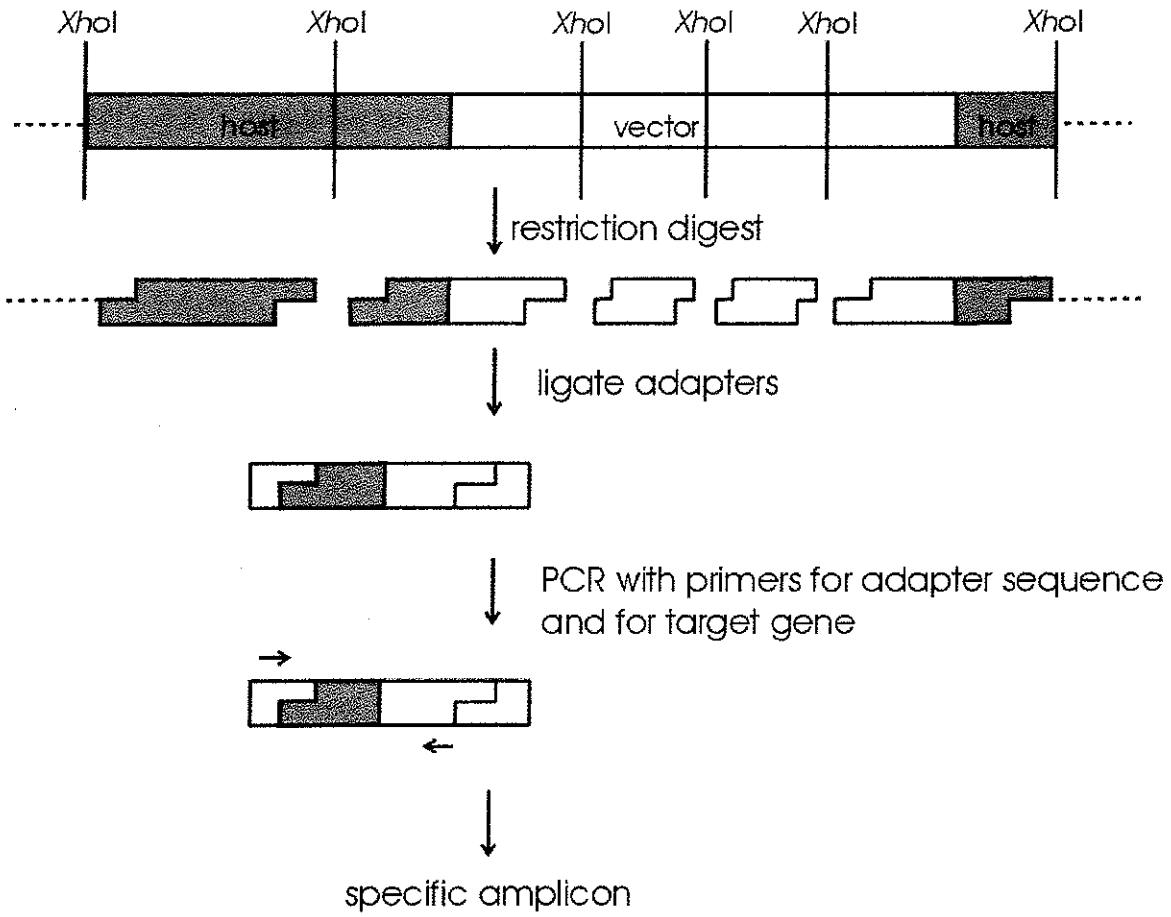

Figure 5. Adapter-ligation PCR schematic representation. The genomic DNA is digested with a restriction enzyme, adapters are ligated to the overhangs, and PCR is carried out using two primers. One primer is specific for the adapter sequence, the other is specific for the provirus sequence.

conducted at 72°C for 10 minutes, 32°C for 2 hours, followed by incubation overnight at 4°C. The hybridised adapter was stored at -20°C.

In order to approximate the ratio of genomic DNA to adapter, the predicted number of digest sites (New England Biolabs 2002-03 Catalogue & Technical Reference, p. 263, 277) for murine genomic DNA was calculated as follows. Six µg of DNA contains 10^6 genomes, and one genome contains 3 x 10^9 bases. *Kpn*I digest sites occur every 3000 bases, making 10^6 digests and 2 x 10^6 ends. In 10^6 genomes, this would be 2 x 10^{12} ends and therefore, 2 x 10^{12} adapter molecules would be needed. One µl of adapter, based on measured concentration and length of the oligonucleotides was calculated to contain 2 x 10^{13} molecules. The adapter should be added in 10x excess to ensure all ends receive an adapter, so 1 µl was used.

2.16.3 Adapter-ligation PCR

One µg of genomic DNA, and 1 µg of pROSA-GFNR plasmid DNA for a positive control, were digested individually with *Kpn*I (10 U/µl, Roche) at 37°C for several hours followed by clean up using a QIAquick PCR Purification™ kit (Qiagen) to remove the restriction enzyme (*Kpn*I cannot be heat inactivated). To make the restriction site overhangs available for ligation to the digested DNA, the adapter was melted at 37°C for five minutes and snap cooled on ice. One µl of adapter was ligated to the PCR purified DNA using 1 µl of T4 DNA ligase (high concentration ligase, 5 U/µl, Roche) at 16°C overnight, in a total volume of 60 µl.

As a control for successful digest and annealing of the adapters to the ends, PCR to detect the 2.9 kb *Kpn*I digest fragment (Figure 2) containing only proviral sequence

was performed. This control was necessary as the size of the digest fragment containing part endogenous and part proviral sequence was unknown, so a reaction where the amplicon size was known was performed first. In the control, pROSA-GFNR as well as trapped genomic DNA were both used. Since pROSA-GFNR was used as the template and was at high copy number compared to the ratio in trapped DNA, the plasmid acted as a good positive control for the PCR.

PCR was performed according to the general guidelines in section 2.5. One of the primers used was the Kpnlink primer, the same oligonucleotide used to make the adapter. The other primer was neo3f (Appendix 1), based on the neomycin resistance gene in the provirus. Digested and ligated genomic DNA or pROSA-GFNR plasmid DNA (0.1 μg) was used as the template. Annealing was performed at 50°C, and elongation at 72°C for 1.5 minutes for 45 cycles. PCR products were visualised by typical agarose gel electrophoresis.

2.17 iPCR

2.17.1 Concept

Another strategy used to attempt identification of the endogenous sequence flanking the site of retroviral integration was iPCR. This strategy, like adapter-ligation PCR, also addressed the obstacle of designing a primer based on unknown endogenous sequence. For iPCR, genomic DNA was digested and the fragments were self-ligated. This produced a fragment that could be amplified using two primers based on the provirus (Figure 6).

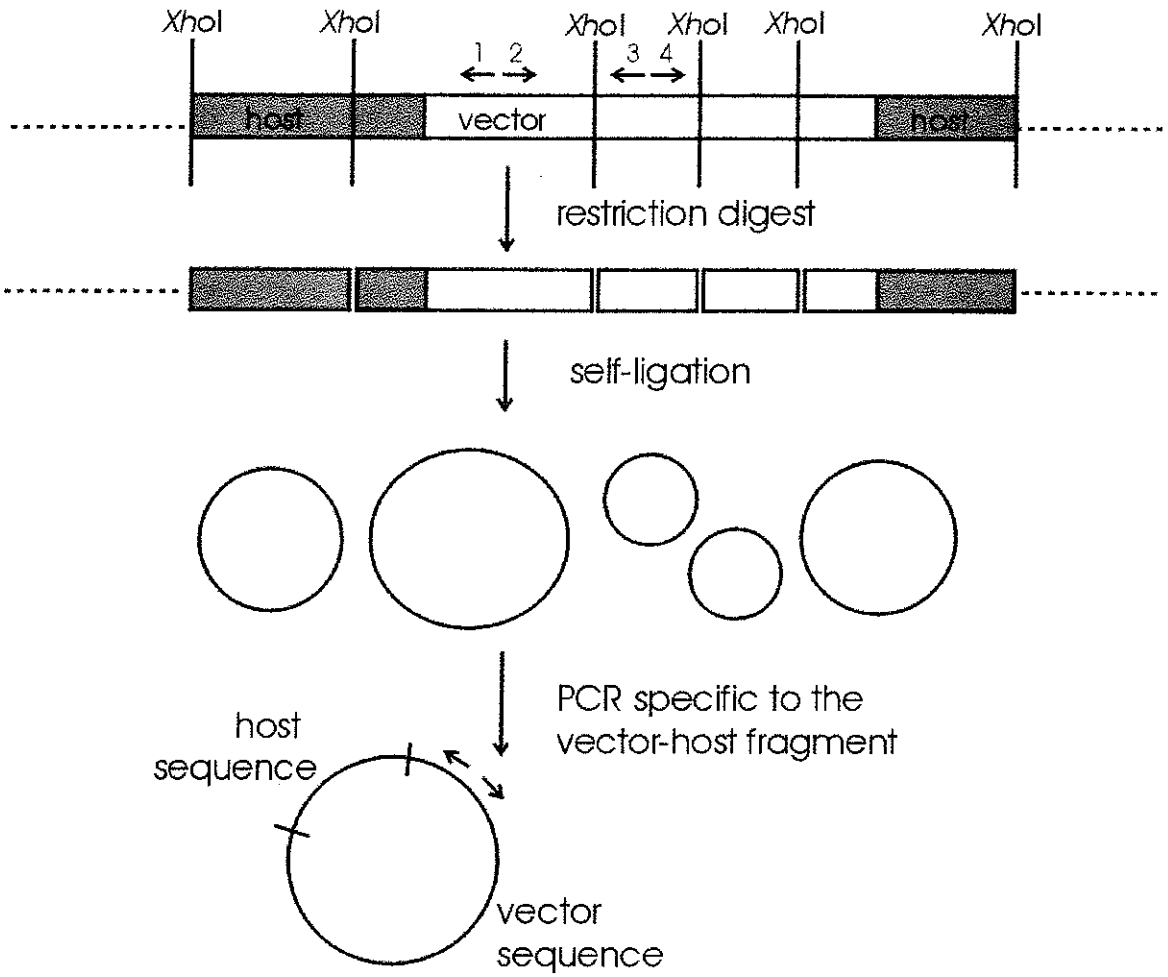

Figure 6. iPCR schematic representation. The genomic DNA was restriction digested, the digest fragments were self-ligated to produce circular fragments. A primer pair (1, 2) was designed based on the provirus sequence such that only the circular fragment containing part proviral and part endogenous sequence would be amplified. As a control for digest and ligation, primers 3 and 4 were used in a PCR reaction where the resulting amplicon size was known.

2.17.2 iPCR control

Within the provirus, there are three *Xho*I sites (Figure 2b). When DNA from trapped cells is digested with this enzyme, followed by ligation to produce self-ligated fragments, there are two products that will always be formed. These are the self-ligated fragments of 860 bp and 2.4 kb in size. Whether plasmid DNA or trapped genomic DNA is being used, these two products will always be present, thus the plasmid can be used as a positive control for digestion, ligation, and PCR. One of these products, the 2.4 kb size product, was chosen for use in a control iPCR reaction to provide an indication that digestion and self-ligation were successful. Only if the DNA is digested and ligated onto itself will the PCR amplicon be produced.

In order to increase the probability of producing self-ligated fragments, the DNA was diluted. A constant amount of genomic DNA, 200 ng, was digested with *Xho*I at 37°C overnight in a volume of 20 µl, after which the enzyme was heat inactivated at 65°C for 20 minutes. Several ligations were performed in increasing volumes: 20 µl, 50 µl, 100 µl, 200 µl, 500 µl to produce the effect of diluting the DNA. The T4 DNA ligase buffer was supplemented with additional ATP (Appendix 2) to a final concentration of 1 mM. Ligations were performed using T4 DNA ligase (high concentration, 5 U/µl) at 16°C overnight. The ligase was then heat inactivated at 65°C for 10 minutes. In order to decrease the volume to be able to use all 200 ng of DNA in a single PCR reaction, the DNA was precipitated. One tenth of a volume of 3 M sodium acetate and 2 volumes of anhydrous ethanol were added and tapped to mix. Samples were chilled at -20°C for 30 minutes and centrifuged at 20 000 rcf for 10 minutes. The supernatant was aspirated and 75% ethanol was added and tapped to wash the DNA. The samples were again chilled

and centrifuged. The DNA pellet was resuspended in 20 µl of tris EDTA (TE) buffer. A positive control, pROSA-GFNR plasmid DNA (0.1 pg/µl) and negative controls of wild type CRFK and BHK-21 DNA were prepared in the same way.

PCR was performed using nested primer sets in order to detect very small amounts of DNA. For the control reaction to amplify part of the 2.4 kb digest fragment (primers 3 and 4, Figure 6), the primers pRseq11 and pRseq12, followed by neo5f and pRseq02 (Appendix 1) were used to produce a 1.8 kb amplicon. Cycling conditions were set as described for general PCR in section 2.5, with 55°C annealing temperature, 1.5 minute elongation at 72°C, and 45 cycles. PCR products were visualised on 1% agarose as usual. Based on the results of the PCR using the varied ligation volumes, 200 µl was used as the ligation volume for future PCR reactions. Sample 3D was prepared along with wild type CRFK DNA and pROSA-GFNR plasmid DNA in order to detect the 2.4 kb fragment, and the amplicon was visualised by agarose gel electrophoresis as usual.

2.17.3 iPCR

iPCR was performed on digested and ligated pROSA-GFNR plasmid DNA to test the primers in a positive control reaction. Primer pairs (ipcr21, ipcr22, and ipcr23, ipcr24, Appendix 1) were designed as nested sets based on the *gag* region (which is a portion of the retroviral *gag* gene) of the gene trap vector (primers 1 and 2, Figure 6), a region between the 5'LTR and the cassette of the gene trap, as opposed to the LTR region. The 5'LTR sequence is repeated as the 3'LTR, therefore primers based on this region would produce double amplicons and the results would be inconclusive.

pROSA-GFNR plasmid DNA was prepared by digestion and ligation as described above. PCR was performed with annealing at 55°C and elongation at 68°C for 4.5 minutes with 45 cycles, both for the first and nested reactions. The first primer set used was ipcr23 with ipcr24, followed by ipcr21 with ipcr22 for nested PCR. PCR products were visualised as usual by agarose gel electrophoresis.

2.18 Single Strand Specific Second Strand Arbitrary (SAP-) PCR

2.18.1 Concept

The third strategy used for identification of sequences flanking the site of retroviral integration was SAP-PCR. This strategy, unlike adapter-ligation PCR and iPCR, uses a primer (P3) that binds in the endogenous host sequence, however, this primer is not specific. The principle is that the 3' end of primer 3 (also referred to as P3, Figure 7) will bind a few bases in the host sequence, and with successive rounds of amplification, this error will eventually produce amplicons containing the entire primer binding site. Through nested PCR (with P2 followed by P1, Figure 7), the fragment containing part endogenous and part proviral sequence will be amplified. Since it is possible that the samples of genomic DNA contain more than one population of clones, the PCR may produce more than one specific amplicon.

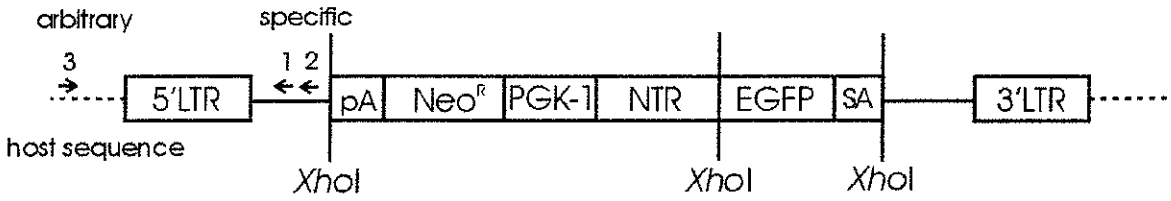

Figure 7. SAP-PCR schematic representation. Primers 1 (sappcr04) and 2 (sappcr03) bind to the *gag* region of the integrated provirus and are used in nested PCR reactions where primer 1 is used in the first round of amplification and primer 2 is used in the nested round. Primer 3 (Table 2) is not specific to the target sequence and binds arbitrarily upstream of primers 1 and 2 in the endogenous host sequence. (dotted line – endogenous sequence)

All of the P3 primers chosen for use were specific for viruses not related to NiV, HeV or retroviruses. The primers chosen were specific for viruses so as to minimise the chance that they would contain sequence homologous to eukaryotic cellular DNA. The viruses to which the primers are specific are listed in Table 5. Primers were obtained from common stocks and diagnostic stocks within the Special Pathogens Program, and the Hepatitis C virus primers were kindly provided by Julie Boutilier (Division of Bloodborne Pathogens and Hepatitis, Health Canada).

2.18.2 SAP-PCR

A control reaction was performed first where only the P3 primer (not specific) was included in the reaction mix (P3/P3 reaction). It was possible that even though a provirus primer and an arbitrary primer were both in the reaction mix, the arbitrary primer alone might prime in both forward and reverse directions producing an amplicon solely using the P3 arbitrary primer. This type of amplicon would not contain proviral sequence and hence would not aid in the identification of genes flanking the site of integration. It was necessary to test whether these types of amplicons would be produced using a given primer so that they could be excluded from use for future reactions.

Unmodified genomic DNA (3D, 4A, wild type CRFK, 100 ng per reaction) was used. The P3/P3 reactions were performed with a single P3 primer (Table 5) using the following conditions: 55°C annealing, 72°C elongation for 3 minutes, for 40 cycles. A nested reaction identical to the first round, except with only 35 cycles, containing 1 µl of the first reaction as template was performed, followed by typical agarose gel electrophoresis to screen for suitable P3 primers.

After the P3/P3 reactions were complete, the P1/P3 and P2/P3 nested reactions were performed. The first round of the nested reaction was performed with sappcr04 and one of the suitable P3 primers at the same conditions as above, followed by nested PCR with sappcr03 and the P3 primer at the same conditions as above.

2.18.3 Cloning and Sequencing of Amplicons

Since it was possible that the samples could contain multiple clones, it was expected that several bands would be produced by the nested PCR reaction. In order to facilitate sequencing of individual amplicons, these were cloned into the pCR4-TOPO™ vector of the TOPO TA Cloning Kit for Sequencing™ (Invitrogen). Two µl of the nested PCR reaction was added to the cloning reaction according to kit instructions and transformed cells were plated onto LB agar with 100 µg/ml ampicillin. Colonies were picked into 5 ml of LB broth with 100 µg/ml ampicillin, and cultures were extracted for plasmid DNA using the QIAprep Spin Miniprep Kit™ (Qiagen). Miniprep plasmid DNA was digested with *Eco*RI (New England Biolabs) to screen for clones containing amplicons. Selected miniprep DNA samples containing cloned inserts were sequenced (section 2.4.1) using T3 and T7 primers (Appendix 1, primer sites contained in the pCR4-TOPO™ vector) and analysed for identity of the sequence by alignment to pROSA-GFNR plasmid sequence and using BLAST (section 2.5.2).

3. RESULTS

3.1 Retroviral Vector Production, Infectivity Determination

In order to produce the retroviral vector used as a gene trap, tissue culture cells (293T) were transfected with the gene trap plasmid and two other plasmids that provided retroviral proteins *in trans*. Mock transfected cells (no DNA) appeared healthy, an indication that no contamination was introduced and that the Lipofectamine 2000 was not toxic to the cells. Approximately 80% (estimated visually) of positive control cells expressed GFP. This positive result in the positive control was extrapolated to infer that plasmids transfected in the other wells were also expressed successfully. The expression of these other plasmids led to retroviral particle production.

Retroviral particles were titred for infectivity using CRFK tissue culture cells and neomycin resistance selection. Geneticin (up to 800 μg/ml) treatment of mock infected CRFK cells (no retrovirus) resulted in CPE after 3-4 days and led to cell death. Accordingly, cells in wells that received a more diluted suspension of virus also began to show CPE and cell death whereas cells in wells that received a less diluted suspension of virus showed little to no CPE or cell death upon exposure to Geneticin.

After 6 days, all cells in the mock infected well (no retrovirus) had died. The other wells were examined against the mock control. Wells in which all cells had died were those wells that had received 10^{-5}, 10^{-6}, and 10^{-7} dilutions of the stock solution of retrovirus. The 10^{-4} well contained a few surviving cellular clones, and increasingly more clones were observed where the virus applied to the wells had been less dilute. It was therefore concluded that since a small amount of virus (1-10 IU/ml) still remained at 10^{-4} dilution, the retroviral gene trap titre was 10^4 IU/ml.

3.2 Transduction of Cells With Retrovirus

In order for a single integration to occur in the genome of a cell, an MOI=0.1 was used for transduction of cells with retroviral particles. The MOI=0.1 would result in one out of ten cells infected with a retrovirus, hence leading to one out of ten cells surviving Geneticin selection. The majority (>90%) of cells (both CRFK and BHK-21 cells) in the wells died upon selection with Geneticin, which is consistent with the low MOI used. Cells that survived Geneticin selection appeared completely healthy and grew at the same rate as untreated wild type cells would, indicating that the neomycin resistance gene was being expressed from the integrated gene trap.

3.3 Fusion Resistance Selection

With the intention of focusing the investigation of genes involved in pathogenesis to those for the cellular receptor or its regulation of expression, cells were selected for resistance to fusion induced by the NiV glycoproteins. Expression of the NiV F and NiV G glycoproteins in cell culture causes fusion of neighboring cells as would occur during NiV infection. Trapped cells (CRFK and BHK-21) were selected for resistance to fusion by two rounds of exposure to NiV glycoprotein-expressing wild type cells. After 24 hours post-transfection of pczCFG5 IEGZ-NiV F and pczCFG5 IEGZ-NiV G in wild type cells, glycoprotein expression was evident by the formation of syncytia. Over the next few days, large syncytia formed between the majority of cells (Figure 8a). Clones of cells resistant to syncytia formation, that is, resistant to fusion which was caused by the interaction between the NiV glycoproteins and the cellular receptor, were seen. Resistance to fusion by this *in vitro* method was likened to an increased probability of

selecting for clones with an integration in the gene for the cellular receptor causing decreased expression of the protein.

Approximately one week post-transfection and overlay, new syncytia formation ceased, and the cells were then selected with Geneticin. At this point, all surviving cells were resistant to fusion and resistant to Geneticin, meaning that all surviving cells were trapped cells. The selected cells grew in clonal populations at the normal rate and appeared healthy. All killed cells were therefore either trapped cells susceptible to fusion or wild type cells susceptible to Geneticin or fusion.

3.4 NiV Selection

Cells that were trapped and selected for resistance to fusion were exposed to NiV for selection of clones resistant to infection or CPE induced by NiV. Upon exposure to NiV, both CRFK and BHK-21 cells showed very rapid CPE, visible within 24-48 hours post-exposure. Several forms of CPE were observed including necrotic-appearing cell death, apoptotic-appearing cell death, and syncytia formation (Figure 9). Daily observation revealed widespread cell death in the wells, with much cell debris floating in the supernatant for approximately seven days post-exposure. During this stage, the debris made observation of surviving cells difficult. When most of the cells had died and this debris was removed, it was possible to see some cells peeling from the bottom of the dish, some syncytia, and some cells that looked healthy and comparable to the mock infected cells. After approximately two weeks post-exposure, we were able to discern clones that could possibly be described as resistant to the effects of NiV, and their growth

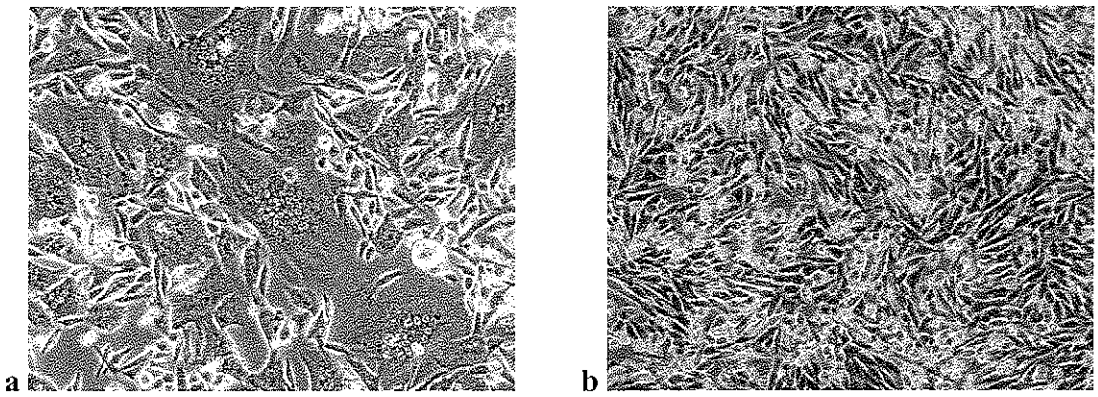

Figure 8a. Syncytia in BHK-21 cells. Wild type cells were transfected with plasmids and expressed the NiV F and NiV G glycoproteins causing fusion and syncytia formation with neighbouring wild type and trapped cells.
b. Wild type BHK-21 cells. Cells were mock transfected (incubated with Lipofectamine 2000 and Opti-MEM, without DNA).

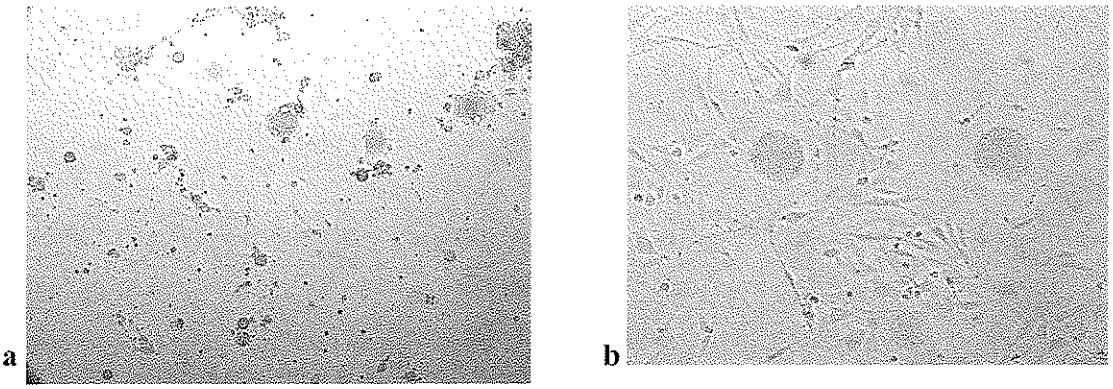

Figure 9. CPE in tissue culture cells exposed to NiV. Several forms of CPE are demonstrated in tissue culture cells upon exposure to NiV including necrotic-appearing cell death (**a**), apoptotic-appearing cell death (**a**), and syncytia formation (**b**).

was monitored over several weeks. It appeared that several clonal populations were present per well.

3.5 Cloning of Cells

After exposure of the CRFK or BHK-21 cells to NiV, the single surviving cells needed to be expanded into larger populations so that they could be passaged, used in additional experiments, and eventually their DNA extracted. Although both CRFK and BHK-21 cells were exposed to NiV on the same day with the same amount of infectivity, after exposure, it became quickly evident that BHK-21 cells grew far more rapidly than CRFK cells, for reasons unknown. It seemed that the CRFK cells took much longer to recover from NiV exposure and exposure to possibly toxic cellular debris than did the BHK-21 cells, although both cell lines were equally susceptible and showed an equivalent amount and type of CPE at equivalent time post-exposure.

Cells that survived exposure to NiV were left to grow until clones reached an average size of a few hundred cells, although the clones grew at different rates. At this stage of growth, remaining cells were passaged from the original six well dishes into several wells of a 24 well dish where they continued to grow. The BHK-21 cells were passaged within three weeks post-exposure, however only after several weeks of slow growth were CRFK cells passaged into fresh dishes.

Growth and signs of CPE in the cell clones were monitored. Many cell clones while having survived exposure to NiV still showed some CPE. Although most clones continued to grow after passage, some clones died. It is unclear why some cells were able to survive initial exposure to NiV and yet be unable to continue to grow.

Furthermore, several clones were believed to be persistently infected due to the observation of CPE (any initial virus would have been eliminated with medium changes) and yet the cells grew normally and appeared healthy.

3.6 Properties of Cellular Clones

After passage and growth of the clones in 24 well dishes, supernatants were tested for the presence of infectivity to test whether the cells were persistently infected. For persistent infection to occur, the cells must be infected by NiV, but not killed, and continue to support NiV replication.

As there were numerous BHK-21 clones that survived exposure, only some of the clones were chosen for further experiments. All BHK-21 supernatants that were tested produced CPE in the wild type cells, were considered positive for the presence of infectivity, and therefore, these were classified as persistently infected with NiV (Table 3). In contrast, all CRFK clones tested for infectivity did not cause CPE in wild type cells, and hence were classified as not persistently infected (Table 3). In both experiments, wild type cells were susceptible to NiV as a positive control, while mock infected cells (no virus) remained healthy.

The cell clones that survived initial exposure to NiV were re-exposed to NiV. This experiment was intended to investigate whether clones could survive re-exposure to NiV. Several clones of both BHK-21 and CRFK cells were re-exposed to NiV. Clone AD6 (BHK-21) and all of the CRFK cell clones grew at the normal rate after re-exposure. In contrast, most of the BHK-21 clones (some results not shown) showed CPE and cell death (Table 3).

It is hypothesised that NiV and HeV are very closely related viruses, however, little is known about the genes involved in the pathogenesis of either virus. Certain cellular clones, the CRFK clones, were shown to be resistant to infection with NiV, and so infection with HeV was performed to investigate whether these cells would also be resistant to HeV infection, which would further demonstrate the relatedness of these two viruses.

Selected CRFK clones (1B, 3C, 3D, 4A) that were previously exposed to NiV were exposed to HeV and their RNA was extracted. RT-PCR was performed to detect the presence of either NiV or HeV RNA. The primers for NiV detection (M5, M3, NP5, NP3) could amplify both NiV and HeV RNA. In contrast, the HeV primers (HeV4912f, HeV5113r, HeV11266f, HeV11870r) were specific only to HeV. The positive controls produced the expected amplicons for each primer pair (Table 4). None of the CRFK samples were found to contain either NiV or HeV RNA (Table 3, Figure 10). This result supports the negative supernatant CPE result and also supports the hypothesis that NiV and HeV have very similar properties as cells resistant to NiV were also resistant to HeV. As all of the supernatants from the BHK-21 cells were clearly positive for CPE, RT-PCR was not done.

In summary, all of the CRFK cellular clones were resistant to NiV and HeV infection whereas all of the BHK-21 cellular clones were susceptible to NiV infection and clone AD6 was the only BHK-21 clone to survive NiV re-exposure.

Table 3. Properties of selected cellular clones

Clone	Persistently infected (CPE)[1]	Growth after re-exposure to NiV	NiV RNA detection by RT-PCR[2]	HeV RNA detection by RT-PCR[2]
AA6	yes	no growth	ND	ND
AC6	yes	no growth	ND	ND
AD6	yes	growth	ND	ND
BC3	yes	no growth	ND	ND
1B	no	growth	not detected	not detected
3C	no	growth	not detected	not detected
3D	no	growth	not detected	not detected
4A	no	growth	not detected	not detected

[1]Persistent infection was determined by CPE in cells exposed to supernatants.
[2]Cells exposed to NiV or to HeV were tested for infectability by specific RT-PCR for the viral RNA genome.

ND – not done

Table 4. Expected amplicon size for detection of NiV or HeV RNA by RT-PCR

Primer Pair[1]	Amplicon Size[2]
M5, M3	360 bp
NP5, NP3	130 bp
HeV4912f, HeV5113r	200 bp
HeV11266f, HeV11870r	600 bp

[1]Primers used to detect NiV and/or HeV RNA in whole RNA extracted from virus-exposed cells.
[2]The size of amplicon expected with the presence of NiV and/or HeV RNA.

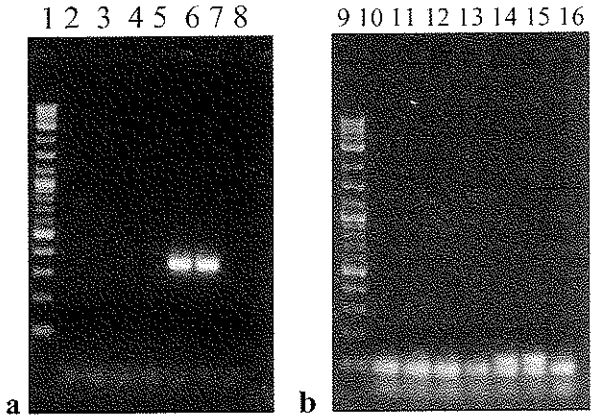

Figure 10a, b. RT-PCR for NiV RNA. RT-PCR specific to both NiV and HeV was performed using RNA extracted from cells exposed to NiV and later to HeV, with HeV RNA, and NiV RNA for positive controls. Primers M3 and M5 (lanes 2-8, 360 bp amplicon) and NP3 and NP5 (lanes 10-16, 130 bp amplicon) were used. Four CRFK cellular RNA samples were tested (1B, 3C, 3D, 4A, lanes 2-5 and 10-13). Both HeV (lanes 6, 14) and NiV RNA (lanes 7, 15) were used as positive controls. A water reaction (lanes 8, 16) was also performed. (lanes 1, 9 are 2 Log Ladder™; 2% agarose)

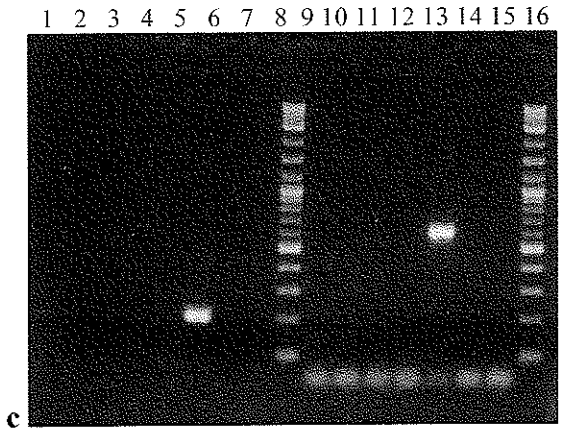

Figure 10c. RT-PCR for HeV RNA. RT-PCR specific to only HeV was performed using RNA extracted from cells exposed to NiV and later HeV, HeV RNA (positive control), and NiV (specificity control). Two sets of primers were used (4912f with 5113r, lanes 1-7, 200 bp amplicon; 11266f with 11870r, lanes 9-15, 600 bp amplicon). Four CRFK cellular RNA samples were tested (1B, 3C, 3D, 4A, lanes 1-4 and 9-12). The positive control, HeV (lanes 5, 13), was the only sample to produce an amplicon. NiV RNA (lanes 6, 14) was used as a control to demonstrate specificity to HeV. A water reaction (lanes 7, 15) was also performed. (lanes 8, 16 are 2 Log Ladder™; 2% agarose)

3.7 DNA Extraction for Molecular Analysis

Cellular clones resistant to infection or to CPE effects by NiV were created. Several of the clones were further tested for growth after re-exposure to NiV, for persistent infection with NiV, and for susceptibility to HeV. These cells survived exposure to NiV due to an integrational mutation that caused a change in the phenotype of these cells compared to the wild type, however, the location of the mutation or the gene that was affected was not known. DNA extraction for molecular analysis to confirm the presence of an integration and determine its location in the genome was necessary.

DNA was extracted from cellular clone samples from CL4 using DNAzol™. Approximately 2-5% of the total extraction sample was electrophoresed and photographed, from which RNA and protein contaminations were observed to be minimal. Additionally, DNA concentration was estimated by comparison to the DNA marker run on the same gel. DNA concentration was also measured by spectrophotometry and found to give approximately equal values to those estimated by the gel method. The spectrophotometer also served to support that little protein contamination was present, giving an absorbance ratio (260 nm:280 nm) close to 1.8. On average, a total of 50 μg of DNA was extracted per sample from a 75 cm^2 flask (approximately 10^7 cells). This DNA was considered of good quality for use in Southern blot and PCR assays.

3.8 Southern Blot

Molecular analysis of the DNA from the cellular clones was begun with Southern blot assays. The goal of the Southern blot assays was to confirm the presence of the gene

trap in the genome and to predict cycling conditions and expected amplicon sizes in future PCR assays. Two types of detection were used: biotinylation of a riboprobe, and α^{32}P dATP labelling of a DNA probe.

3.8.1 Biotinylated Probe

Before its use, the biotinylated riboprobe was verified for expected size in base pairs and for biotinylation signal by gel electrophoresis and blotting to a nylon membrane followed by streptavidin-HRP conjugation. The probe was observed to be approximately 500 bp in size, as expected, and when blotted to a membrane and conjugated to streptavidin-HRP was observed to produce a high level of chemiluminescence during a very short exposure time (10-30 seconds) to film (Figure 11a). This result was taken to mean that the probe was correctly produced to the proper length and efficiently biotinylated, and could therefore be used for detection of target DNA sequences by Southern blots. This probe was designed to detect the neomycin resistance gene in the *Xho*I digest fragment of 2.4 kb (Figure 2).

The biotinylated probe successfully detected the positive control, pROSA-GFNR at amounts of 100 ng, 10 ng and 1 ng, which was observed to be the limit of sensitivity for this probe (Figure 11b). The probe was unable to detect the 2.4 kb digest fragment (neomycin resistance gene) in the genomic DNA samples tested (1B, 3D, 4A, Figure 11c). As the results for the detection of the 2.4 kb pROSA-GFNR fragment were unsuccessful for the genomic DNA samples, detection of the fragment containing the vector and endogenous DNA (size unknown) was not attempted.

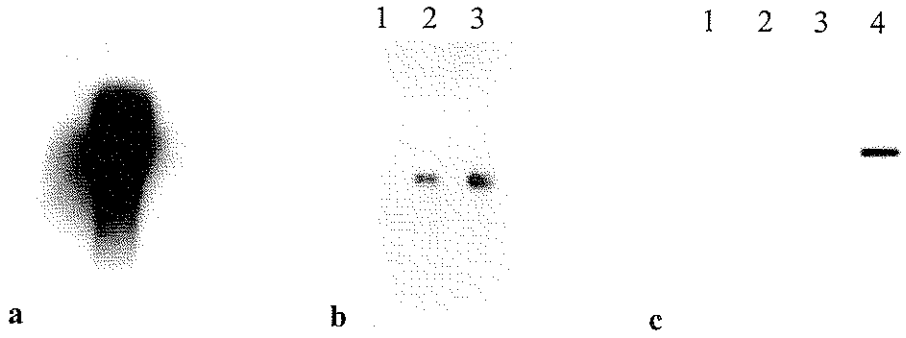

Figure 11a. Chemiluminescent detection (10 second exposure) of the biotinylated probe. Verification of biotinylation was performed on the probe (500 bp) used for detection of the neomycin resistance gene.

b. pROSA-GFNR plasmid DNA, detected with a biotinylated probe. pROSA-GFNR plasmid DNA was digested with *Xho*I, diluted 1:10 (1 ng, 10 ng, 100 ng, lanes 1 to 3, respectively) and Southern blotted for a sensitivity test of the biotinylated probe for the neomycin resistance gene (2.4 kb fragment).

c. Detection of trapped CRFK cellular DNA and pROSA-GFNR with a biotinylated probe. Genomic DNA samples from cellular clones that survived NiV exposure (lane 1 – clone 1B, lane 2 - 3D, lane 3 - 4A) and a pROSA-GFNR positive control (lane 4, 10 ng) were Southern blotted and detected using the biotinylated probe for the neomycin resistance gene (2.4 kb fragment).

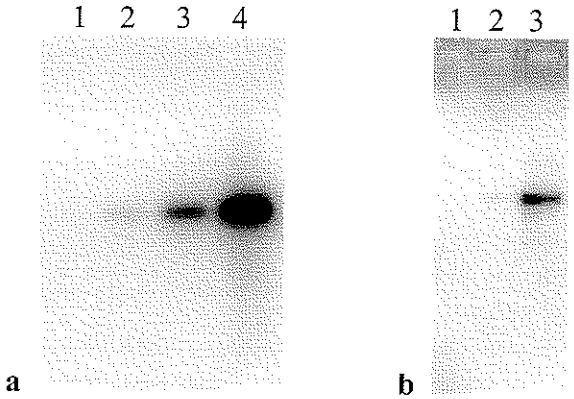

Figure 12a. Detection of pROSA-GFNR with a α^{32}P dATP labeled neomycin resistance gene probe. pROSA-GFNR plasmid DNA was digested, diluted 1:10 (10 pg, 100 pg, 1 ng, 10 ng, lanes 1 to 4, respectively) and Southern blotted. Detection of the 2.4 kb band was performed using an α^{32}P dATP labeled DNA probe for the neomycin resistance gene.

b. Detection of trapped cellular CRFK DNA and pROSA-GFNR using a α^{32}P dATP labeled neomycin resistance gene probe. Detection of the neomycin resistance gene (2.4 kb fragment) was performed on genomic DNA from cells that survived NiV exposure (lane 1 - wild type control, lane 2 - 4A) and on pROSA-GFNR plasmid DNA (lane 3, 10 ng) by Southern blot with an α^{32}P dATP labeled DNA probe.

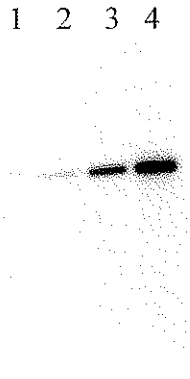

Figure 13a. Detection of pROSA-GFNR using an α^{32}P dATP labeled *gag* probe. Dilutions of pROSA-GFNR plasmid DNA (10 pg, 100 pg, 1 ng, 10 ng, lanes 1 to 4, respectively) were digested with *Xho*I and probed with the α^{32}P dATP labeled *gag* gene probe to detect the 4.7 kb band.

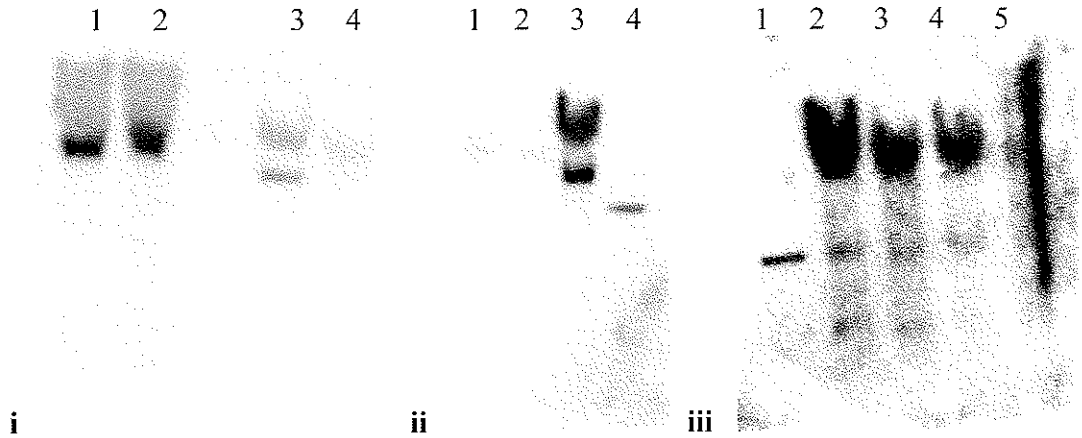

Figure 13b. Detection of trapped, genomic CRFK DNA using an α^{32}P dATP labeled *gag* gene probe.
i. Lane 1 – clone 1B, lane 2 – 3C, lane 3 – 4A, lane 4 – wild type CRFK control.
ii. Lane 1 – wild type CRFK control, lane 2 – 4A, lane 3 – 3D, lane 4 – pROSA-GFNR plasmid DNA positive control (100 pg, 4.7 kb).
iii. Lane 1 – pROSA-GFNR plasmid DNA positive control (100 pg, 4.7 kb), lane 2 – AA6, lane 3 – AD6, lane 4 – BD3, lane 5 – wild type BHK-21 control.

3.8.2 α^{32}P dATP Labeled Probe

An α^{32}P dATP labeled DNA probe with the same target fragment as the biotinylated riboprobe (neomycin resistance gene in the 2.4 kb fragment) was produced. The probe was tested for sensitivity on a dilution series of *Xho*I digested pROSA-GFNR. The probe detected pROSA-GFNR in amounts of 10 ng, 1 ng, 100 pg, and 10 pg, which was determined to be the limit of detection of this probe (Figure 12a). The 2.4 kb fragment was also detectable in trapped genomic DNA (clone 4A, Figure 12b) when at least 10 µg of DNA was digested and blotted while no hybridisation was detectable in wild type DNA. The capability of the probe to detect the neomycin resistance gene in the genomic DNA sample indicated that the sensitivity of this probe and the procedure used were suitable for detecting the provirus in the trapped DNA, and that this system could potentially detect the fragment containing part proviral and part endogenous sequence.

The next probe produced was targeted to the portion of the *gag* gene in the *Xho*I digest fragment containing part proviral and part endogenous sequence (*gag* probe in Figure 3). Future PCR assays would use this digest fragment as template for amplification, however, in the case of the genomic DNA, the size of this fragment was unknown. The Southern blot targeting this segment was useful in predicting the size of this fragment and in planning the PCR experiments.

The probe was tested against pROSA-GFNR where the target fragment was 4.7 kb. The probe was again tested for sensitivity using a dilution series of pROSA-GFNR, and the sensitivity was the same as with the neomycin resistance gene probe, 10 pg (Figure 13a). The probe successfully detected the target fragment in genomic DNA samples (1B, 3C, 3D, 4A, AA6, AD6, BD3, Figure 13b). In several instances, there were

two sizes of DNA that were detected by the probe, usually a band, as well as a smear (>20 kb). The sizes of the detected fragments were approximated based on the relative location of the positive control hybridisation signal (4.7 kb). This result indicates that more than one target for the probe existed in the sample, possibly due to the presence of more than one cellular clone, and hence more than one site of integration represented within the sample.

The *gag* gene probe also detected a target sequence in the wild type DNA, which was not trapped and did not contain a retroviral integration (Figure 13b ii, lane1). This indicates that the band being detected was not a proviral sequence, and was likely an endogenous retroviral sequence.

Overall, the Southern blots specifically confirmed the presence of the gene trap in the genomic DNA samples (both CRFK and BHK-21). The first type of detection, biotinylation, showed poor results with respect to sensitivity whereas the α^{32}P dATP probe showed 100x greater sensitivity in detection of the plasmid DNA control. This led to an approximation as to the size of the restriction digested fragment to be targeted for amplification in future PCR studies.

3.9 PCR Amplification of Flanking Sequences

Several PCR assays were investigated to develop a method for identification of the sequences flanking the site of retroviral integration, that is, the interrupted endogenous sequence. As the site of integration was completely unknown, methods had to be devised to amplify the flanking region based on the known sequence of the

provirus. The strategies investigated were adapter-ligation PCR, inverse PCR (iPCR), and single strand specific second strand arbitrary (SAP-) PCR.

3.9.1 Adapter-ligation PCR

The target DNA for PCR amplification was in a region where the endogenous and proviral sequences met. By restriction digest, this region could be cut out from the genomic DNA producing a fragment with both endogenous and proviral DNA. In order to amplify this fragment containing part proviral and part endogenous host sequence (Figures 4, 5), adapters were ligated to the ends of the fragments to create known ends from which a primer for PCR could be designed. PCR amplification using one primer based on the adapter and the other based on the provirus was successful when using pROSA-GFNR as the template. The control reaction, amplification of a *Kpn*I digest fragment containing only proviral sequence produced the expected 1.4 kb amplicon (results not shown, Figure 2).

In order to control for sensitivity of detection of a single copy gene, the digested and ligated DNA was tested for amplification of the neomycin resistance gene as described in section 2.15. The amplicon was detected, indicating that the amount of DNA was sufficient and that the DNA was trapped. Nonetheless, the 1.4 kb amplicon using an adapter-specific primer was not obtained using trapped genomic DNA samples despite attempts to optimise the reaction by changing annealing temperatures, elongation times and $MgCl_2$ concentrations added to the buffer. As this control reaction was unsuccessful, amplification of the target fragment containing part proviral and part endogenous sequence was not attempted.

3.9.2 iPCR

iPCR was performed by digest and self-ligation of DNA, followed by nested PCR. When pROSA-GFNR was used, the ligation produced a 2.4 kb product. This was used in PCR to produce a 1.8 kb amplicon (using primers 3 and 4, Figure 6) as a control for successful digest and ligation, and PCR sensitivity. A second control reaction used primers targeting the *gag* region. In the pROSA-GFNR plasmid positive control, a 4.7 kb amplicon would be obtained whereas with genomic DNA, the amplicon size would be unknown (primers 1 and 2, Figure 6).

As with adapter-ligation PCR, it was necessary to ensure that the genomic DNA was fully digested and so overnight digests were performed. The next critical step was the self-ligation. In order to obtain self-ligated fragments, the concentration of DNA needed to be fairly low so that the fragments would preferentially ligate to their own ends rather than to another fragment. This was done by ligating a small amount of DNA in a large volume, precipitating the DNA, and resuspending it into a small volume for use in PCR. The ligation was difficult to optimise at first as T4 DNA ligase is extremely dependant upon a sufficient concentration of ATP in the buffer. Repeated freeze-thawing of the buffer was avoided by storage in small aliquots, and the buffer was boosted with additional ATP.

The control for digest and ligation, the PCR to amplify 1.8 kb of the 2.4 kb circular fragment, was successful using pROSA-GFNR plasmid DNA (Figure 14). Successful amplification was used as an indication that the plasmid had been properly digested and ligated. The digest and ligations were also performed on genomic DNA (wild type CRFK and 3D, 200 ng) where amplification of the 1.8 kb fragment from

trapped DNA was also successful, indicating that the proviral DNA was digested and ligated to produce the 2.4 kb circular fragment (Figure 14) in the same way as the plasmid.

Unfortunately, detection of the *gag* region was not successful. The 4.7 kb fragment containing the 5'LTR and the *gag* region could not be amplified from pROSA-GFNR. Although attempts to optimise the cycling conditions through changes in the elongation time, annealing temperature or $MgCl_2$ concentrations were done, it is not clear why the reaction produced only smaller amplicons, not the desired 4.7 kb amplicon. The plasmid was used as a control to test the primers and their specificity. As the plasmid PCR was not successful, amplification of this region was not attempted with genomic DNA samples.

3.9.3 SAP-PCR

The principle of SAP-PCR was to use one primer specific to the provirus and one primer not specific to any sequence (similar to random priming) to produce an amplicon containing part endogenous and part proviral sequence (Figure 7). To ensure that the amplicon was truly produced using both the specific and arbitrary primer, the arbitrary primer (P3) alone was used in a control reaction to test whether it would produce random amplicons that might have been mistaken for provirus-specific amplicons.

These control reactions using only the P3 primers were performed first as any amplicons obtained in these reactions were clearly not products of amplification from the provirus. Several primers were screened for the production of non-provirus-specific amplicons (Table 5).

The only P3 primer able to produce amplicons when used in the nested reaction with P1 and P2 was EBO 30F. This reaction produced several amplicons up to 1.2 kb in size. Unfortunately, most of these amplicons were too short in length to be amplicons containing sequence from both the provirus and flanking DNA as the P1 and P2 primers were designed 1.2 kb from the 5' end of the provirus (see schematic diagram, Figure 7). The amplicon 1.2 kb in size was investigated further as the size was possibly long enough to contain flanking endogenous sequence. The sequencing results of the amplicon cloned into the pCR4-TOPO™ vector, section 2.18.3, showed that the amplicon matched the sequence of the 5'LTR of pROSA-GFNR, meaning that the P1 and P2 primers were specifically binding to the *gag* region as intended. Unfortunately, the P3 primer did not bind in the endogenous sequence, but within the provirus near the 5' end of the 5' LTR, thus the amplicon contained only proviral sequence, and no flanking sequence.

Several changes to the PCR conditions were performed in an attempt to optimise the reaction, such as lowering the annealing temperature, using a step-up type of program where the annealing temperature was gradually increased, as well as the addition of more amplification cycles. In the end, despite optimisation efforts, the PCR was not successful. These results thus provide no sequence data for regions flanking the site of retroviral integration, although specificity of the P1 and P2 primers to the provirus was demonstrated.

Overall, difficulties in optimisation of the PCR assays prevented identification of sequences flanking the retroviral integration sites within the trapped genomic DNA samples. In several instances, the use of the pROSA-GFNR plasmid as a positive control proved useful in demonstrating that primers were properly designed and that cycling

conditions were suitable. Although these conditions could not necessarily be extended for use in genomic DNA amplifications, they provided a starting point from which to continue optimisation.

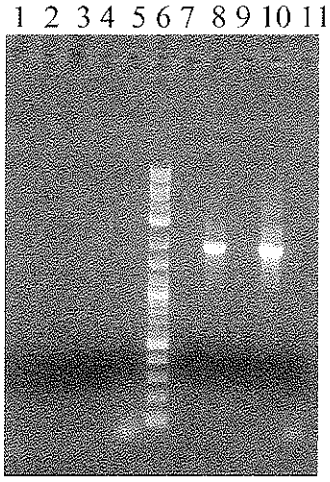

Figure 14. iPCR control reaction. Primers pRseq11 and pRseq12 (lanes 1, 2, 4, and 5) followed by neo5f and pRseq02 in a nested reaction (lanes 7, 8, 10, and 11; 1.8 kb) were used for a control iPCR reaction. Wild type genomic DNA (lanes 1, 7) and trapped 3D DNA (lanes 2, 8) as well as a positive control with pROSA-GFNR (lanes 4, 10) and a water control (lanes 5, 11) were tested. (lane 6, 2 Log Ladder™; lanes 3, 9 are empty)

Table 5. List of primers used in SAP-PCR

Primer	P3/P3 reaction amplicon production[1]	Nested reaction amplicon production[2]	Virus for which the primer is specific
EBO 30F	-	+	Ebola virus
EBO 30R	-	-	Ebola virus
HCV 215R	-	-	Hepatitis C virus
HCV 216F	+	ND	Hepatitis C virus
HCV 213R	+	ND	Hepatitis C virus
HCV 227R	-	-	Hepatitis C virus
LassaGP 188f	+	ND	Lassa virus
LassaGP 649f	-	-	Lassa virus
Var crmB3	+	ND	Variola virus
Var crmB5	+	ND	Variola virus
Var HA3	+	ND	Variola virus
Var HA5	+	ND	Variola virus
HerBV polF	-	-	Herpes B virus
HerBV 1-22 Ir-2	+	ND	Herpes B virus
HerBV GS4	+	ND	Herpes B virus
HerBV GAS4	+	ND	Herpes B virus
HSV P1	+	ND	Herpes Simplex virus
HSV P2	+	ND	Herpes Simplex virus
Lassa 1010C	+	ND	Lassa virus
CCHF SM19	+	ND	Crimean Congo Hemorrhagic Fever virus
CCHF SM20	-	-	Crimean Congo Hemorrhagic Fever virus

[1]All primers were tested for P3/P3 amplification products.
[2]Primers that did not produce amplicons in the P3/P3 reaction were then tested for amplicon production in a nested reaction with P1 followed by P2.

Legend: + – amplicon produced
- – no amplicon produced
ND – not done, as P3/P3 reaction produced inappropriate amplicon(s)

4. DISCUSSION

The general concept addressed in this project was a method to identify cellular genes involved in NiV pathogenesis. Many studies with respect to other viruses have been devoted to this area of research, such as the identification of genes required for virus entry or replication (Vazquez and Esteban, 1999; Qing, *et al.*, 1999; Nadal, *et al.*, 2002), or the identification of genes that are expressed as a result of infection (Fugier-Vivier, *et al.*, 1997; Johansen, *et al.*, 2003). Following identification, several techniques can be employed for analysis of such cellular genes. One example that has recently gained widespread use is gene arrays (Bolt, *et al.*, 2002; DeBiasi, *et al.*, 2003) where specific genes are probed. This assay can provide information on upregulation or downregulation of several genes. A yeast two-hybrid system (*Yeast Hybrid Technologies*, 2000; Gietz and Woods, 2002) is a method that could be used for identification through the investigation of protein interactions. A new method for confirmation of action of a protein in the pathogenesis of a virus is small interfering RNA (siRNA) silencing (Elbashir, *et al.*, 2001a; Elbashir, *et al.*, 2001b) whereby interaction of mRNA and homologous RNA abrogates protein expression, and the effect can be studied.

The technique used in this study, gene trapping (Medico, *et al.*, 2001), was chosen for several reasons. Gene trapping provides a global approach to identification of pathogenesis genes. In this technique, a vector integrates randomly, and any location in the genome could be mutated leading to knockout or downregulation of the gene at that location. Furthermore, the relative expression level of the gene and hence the mRNA levels do not affect the ability to identify the interrupted gene as the method is DNA-

based. Lastly, the method, once established can be a quick and non-labour intensive method of gene identification.

Generation and Analysis of Trapped Cells

The purpose of creating a population of knockout cellular clones was to create clones with a mutation in a gene important to NiV pathogenesis and therefore select for clones resistant to the effects of exposure to NiV. In theory, any location in the genome could be a target for integration of the retroviral genome. Whether in a regulatory sequence, in a coding region, or in a non-coding region, the integration produced an insertional mutation at a random locus in the genome. Each cell in the population contained a knockout at a random location, hence leading to populations of cells containing knockouts at different locations.

NiV is an extremely quick and deadly virus in tissue culture. Wild type cells of many cell lines such as CRFK and BHK-21 cells show CPE within 24 hours and no cell survives exposure. Given the profound effect of NiV on wild type cells, we expected that very few trapped cells would survive exposure.

The first step in this gene identification strategy was the creation of a retroviral gene trap vector by plasmid DNA transfection. Successful production of the retrovirus was inferred by a positive transfection control, and later shown by selection of trapped cells with Geneticin. The titration experiment as well as the transduction experiments both resulted in cells resistant to Geneticin, where wild type cells were susceptible and died under the same selection conditions. Since trapped cells survived Geneticin selection where wild type cells did not, this was proof that the neomycin resistance gene

was expressed in these cells, meaning that they were successfully trapped with the retroviral gene trap vector. The presence of the neomycin resistance gene was further shown by PCR. Using the plasmid as a positive amplification control, the neomycin resistance gene was amplified in trapped genomic DNA samples and not in wild type genomic DNA.

The creation of a single knockout per cell was very important to the study of the resulting phenotype. With only a single integration in a cell, only one mutation would be created, and the resulting change in phenotype from the wild type could be studied. More than one mutation would make detection of the location of integration difficult and would make study of the effect of the mutation on the phenotype more complicated. In order to ensure knockout of only one gene per cell, an MOI=0.1 was used so that if only one retrovirus per ten cells was present, the chances of two viruses infecting the same cell would be low. The ensuing cell death in approximately 90% of the cells during Geneticin selection was consistent with the MOI=0.1. This procedure produced cells with a single knockout that each were subsequently expanded into a clonal population of cells.

After selecting for cells containing the retroviral integration, cells were tested for resistance to fusion induced by the NiV glycoproteins. Fusion (syncytia formation) is a form of CPE typical of NiV infection in tissue culture cells and occurs due to the expression of the NiV glycoproteins during replication on the cell surface which is necessary for virus budding. The glycoproteins bind to the cellular receptor on neighboring cells leading to syncytia and thus, the cellular receptor is an integral component of syncytia formation. The aim of the fusion resistance selection experiment was to focus the identification of genes involved in pathogenesis to the gene for the

cellular receptor, or a related gene such as a gene involved in the synthesis pathway or regulation of expression of the cellular receptor. The implications of identification of such genes can be appreciated with respect to virus entry, CPE production, and tissue or organism tropism.

Wild type cells were made to express the glycoproteins in tissue culture by transfection of plasmids coding for the glycoproteins (NiV F and NiV G). Fusion induced by transfection of these plasmids produced an environment where all cells expressing the cellular receptor were killed by CPE in the form of syncytia formation. At this point, the trapped cells were mixed with the glycoprotein-expressing (non-trapped) cells. A cell with a knockout in the cellular receptor leading to a reduced expression level would be likely to survive this selection. The majority of cells were killed by fusion, meaning that the surviving cells likely contained a knockout in the gene for the cellular receptor or a related gene. In contrast, the fusion-susceptible cells must have contained an integration in another gene not involved in virus entry or fusion that left the cell susceptible to fusion. The fusion resistance selection was done in two rounds, and after each round, the number of surviving fusion resistant cells increased. The selection was therefore successfully eliminating cells that were less likely to contain an integration in a gene for the cellular receptor in an increasingly efficient manner with each round.

The cells selected for resistance to Geneticin and resistance to fusion were subsequently selected for resistance to infection or CPE by NiV. Upon exposure to NiV, approximately 99% of these cells were killed within a few days, and all of the wild type control cells were killed upon exposure to NiV. The cells surviving exposure to NiV were not necessarily resistant to infection, but perhaps only resistant to CPE. Another

possibility is that the virus was able to infect the cells but was not able to replicate due to mutation of a gene required for replication. It is interesting to note that although both cell lines (CRFK and BHK-21) were equally susceptible to NiV, the CRFK cells took much longer to recover from NiV exposure than did the BHK-21 cells. Although the reason for this was not investigated, it may be that since BHK-21 cells are diploid and CRFK cells are polyploid, under these stress conditions, the faster growth rate of the BHK-21 cells was remarked more than under normal conditions. Both types of wild type cells grew at about the same rate. Clonal populations of cells surviving NiV exposure were expanded, passaged, and frozen for storage.

At this point, the cells had been exposed to Geneticin, exposed to *in vitro* expression of the NiV glycoproteins, and exposed to NiV itself. The cells remaining were therefore trapped, likely to contain an integration in a gene (or related gene) for the cellular receptor, and due to the effect of the integrational mutation, resistant to the pathogenic effects related to exposure to NiV.

Several assays were performed to characterise the properties of the cellular clones surviving exposure to NiV. It was unclear at this stage whether surviving cellular clones were resistant to infection by NiV or only resistant to CPE. If the latter, then cells would likely be persistently infected with NiV, which could easily be detected by transferring the supernatant from these cells to fresh wild type cells and observing CPE production or by RT-PCR for detection of the NiV genome.

All of the CRFK supernatants that were tested were negative for the production of CPE, meaning that these cell clones were not infected with NiV. It may be that these surviving clones contained a knockout in the gene for the cellular receptor, or a related

gene as the virus was unable to gain entry at all, and hence was unable to establish a persistent infection. It is possible that the fusion resistance selection experiment may have enriched for cells with a knockout of this phenotype. Alternatively, these clones could have contained a knockout in a gene important to virus replication that prevented NiV infection. In contrast, all of the BHK-21 clones tested positive for the presence of NiV persistent infection by the CPE test. It is unclear if there is a reason or what the reason may be as to why only the CRFK clones were resistant to NiV infection.

RT-PCR for the presence of NiV in cellular RNA extractions was performed on the CRFK clones. The results of this assay were intended to support the results from the supernatant CPE test. As only the CRFK clones were found to be resistant to NiV, only these were tested. By RT-PCR, no NiV RNA was found to be present in the cells, a result that lends support to the supernatant test results that these cells were not infected with NiV. These CRFK cells were also exposed to HeV. NiV and HeV are very closely related, and so clones that were resistant to NiV were expected to also be resistant to HeV. After exposure to HeV, RT-PCR for HeV RNA was performed, and these cells were found to be uninfected with HeV. This result supports existing information indicating that NiV and HeV are very closely related (Chua, *et al.*, 2000; Harcourt, *et al.*, 2000) and likely utilise the same cellular receptor, same replication cofactor, or other genes involved in pathogenesis.

All cellular clones that survived the first exposure to NiV were re-exposed to NiV to observe the outcome of re-exposure to the virus. All of the CRFK clones survived re-exposure to NiV, and grew at their normal rate. Based on the data from the CPE and RT-PCR tests that indicated that the CRFK clones were resistant to NiV infection, the

survival after re-exposure was expected. These cells probably contained a mutation in the cellular receptor (or related) gene, or in a gene involved in virus replication, thus preventing either virus entry or virus replication which would cause CPE and cell death.

Most of the BHK-21 cell clones that were re-exposed to NiV showed CPE and cell death. Based on survival during the first exposure to NiV, it was expected that most cell clones would exhibit the same response as the first time and survive re-exposure, but this was not the case and the BHK-21 re-exposed cells died. This supports the hypothesis that gene trapping only mutates a single allele. If only one allele is mutated while the other is functional, most likely only decreased expression of genes involved in pathogenesis would occur. Hence, expression of a gene such as the cellular receptor or a replication-related gene would still occur, allowing initial infection or replication, but at a decreased level. With decreased NiV entry, perhaps the cell might not have been overwhelmed at first exposure with a large influx of virus particles due to less available entry sites. Furthermore, if less entry sites were available, it follows, then, that less receptors were available for fusion with neighboring cells, which could explain why a typical strong CPE was not observed, and why the cells did not die at the first exposure.

While most BHK-21 cells died upon re-exposure to NiV, AD6 cells did not. AD6 cells survived initial exposure to NiV, however, in contrast to all other BHK-21 clones, AD6 cells also grew after re-exposure to NiV as the CRFK clones did. The results of supernatant testing, however, showed that in contrast to all of the CRFK clones, the AD6 supernatant was infectious. This meant that the cells were persistently infected and virus was replicating and being shed from the cells. For this to be the case, these cells must have been infectable in the first place in order for NiV to enter the cells and replicate.

For the AD6 cells to have been susceptible to NiV entry, expression of the cellular receptor was required. The receptor is required for virus attachment and fusion, and when in contact with NiV-infected cells, causes fusion leading to syncytia. When the cells were exposed to NiV F and NiV G-expressing cells (transfected), the AD6 cells were resistant to fusion while the wild type cells fused. This result would tend to imply that the AD6 cells lacked expression of the cellular receptor and thus were incapable of fusion, however, the cells were undoubtedly expressing the receptor as NiV was able to enter the cells and establish a persistent infection.

One possible explanation for this apparent contradiction is that having only a single allele knocked out, expression of the cellular receptor was simply downregulated and not completely abolished in AD6 cells after initial infection. BHK-21 cells are a diploid cell line, meaning that two copies of each allele exist. If only a single integration occurs per cell, as desired, only one allele is mutated, leaving the other intact and able to be expressed. A low amount of receptor production still enabled the virus to enter the cell, but may have left the cells resistant to *in vitro* fusion. Once the cell was infected, it could not clear the infection and it shed virus into the supernatant as seen by the CPE test.

The difference between the AD6 cells and the other BHK-21 clones is that these cells did not show fusion and syncytia formation leading to cell death when re-exposed to NiV. Keeping in mind that the AD6 cells were persistently infected and must therefore have had some expression of the cellular receptor, a possible explanation exists. Once NiV began replication, the glycoproteins NiV F and NiV G were produced within the cell and were transported to the cellular membrane to await virus budding. This was a requirement for virus replication, which occurred in these cells because they were

persistently infected. The glycoprotein expression was also shown by the observation that when wild type cells were mixed with the AD6 cells, fusion occurred due to the glycoproteins on the AD6 cells and the receptor on the wild type cells. The question then remained as to why the AD6 cells did not fuse with each other.

When the NiV G (attachment) protein is expressed in an infected cell, it exists inside the cell before transport to the cellular membrane. It is possible that NiV G became bound to the cellular receptor protein internally before it was transported to the cell surface. If this occurred, the receptor would have been already occupied with NiV G, and fusion with neighboring NiV F and NiV G expressing cells could not have occurred. This phenomenon can be described as viral interference (Delwart and Panganiban, 1989; Czub, *et al.*, 1995). This could explain why AD6 cells did not fuse with each other, and may also explain why the cells grew normally after re-exposure, as the virus may not have been able to enter the cells with the receptor blocked and expressed at a decreased level.

Southern Blotting

Southern blotting was performed on the DNA extracted from cells that survived NiV exposure and from control wild type cells. Probes targeting different areas of the provirus were used, each with a purpose. The overall goal of this assay was to detect the provirus within the genomic DNA digested with a particular restriction enzyme to be able to predict the size of the restriction fragment produced, and hence the elongation time necessary in a PCR reaction.

Experiments with the first probe for detection of the neomycin resistance gene using 4A, wild type CRFK and plasmid DNA served three purposes. First, it was determined that the assay was sensitive enough to detect the provirus in a genomic DNA sample. Second, it was demonstrated that the correct size in the genomic DNA samples could be detected, meaning that the digest conditions were suitable to cut the DNA efficiently to produce the desired sizes. Third, it was demonstrated that the assay could specifically detect a target sequence in a trapped DNA sample and not in a wild type DNA sample, meaning that hybridisation conditions were suitable to be specific and produce little background. Having obtained very good results, the next probe, for detection of the *gag* region, was tested.

The *gag* gene probe was designed based on the locations of the *Xho*I digest sites (Figure 2b) to detect the fragment containing part proviral and part host endogenous sequence. The purpose of this assay was to first support the statement that these cells contained the retroviral integration, and second to provide an approximation for the sizes of this digest fragment as it was a target for amplification in PCR reactions.

Although the probe for detection of the fragment containing part proviral and part endogenous sequence was designed to be specific for the *gag* region of pROSA-GFNR, evidently, the probe was able to hybridise to endogenous retroviral sequences. In Figure 13, detection of target sequences by the probe was seen in the wild type DNA control lanes. Although it is unfortunate that the probe was not as specific as desired, this result supports the statement that endogenous retroviral sequences present in the genome had sequences similar to the provirus, as will later be discussed with respect to the PCR. Nevertheless, the probe successfully detected proviral sequences in trapped DNA

samples (Figure 13). The differences between the target sequences sizes detected in the different samples and the wild type DNA demonstrated that these clones contained trapped DNA and that the integrations in the clones were different from each other.

In some samples, such as 3C (Figure 13b i), 4A (Figure 13b i), and 3D (Figure 13b ii) more than one band was detected. This may have been due to the presence of more than one clonal population within a sample, as was suspected based on observations of clonal growth of NiV expressing cells under CL4 conditions.

Flanking Sequence Identification Methods

The creation of cell clones containing a retroviral integration that alters the expression of a gene is the first step in identification of cellular genes involved in pathogenesis. The observed phenotype must now be correlated to a molecular explanation for the survival of these cells upon exposure to NiV using molecular biology tools such as PCR and DNA sequencing.

Previously, gene trapping accompanied by a PCR-based method of identification of flanking endogenous sequences has been used. Methods that have proven successful include adapter-ligation PCR (Gogos, *et al.*, 1997), iPCR (Silver and Keerikatte, 1989; Gogos, *et al.*, 1996), SLIP (Jin, *et al.*, 2002), and 5' RACE (Chen, *et al.*, 1994; Gogos, *et al.*, 1996; Medico, *et al.*, 2001). Based on these previous successes, some of these methods were used for this project.

The methods investigated for obtaining sequence information flanking the site of retroviral integration were adapter-ligation PCR, iPCR, and SAP-PCR. All of these methods shared similar drawbacks such as the need to design provirus-specific primers at

a large distance from the 5' terminus of the provirus. Furthermore, the distance to the restriction site was unknown, but was likely on the order of several kilobases, necessitating a long (and unknown) elongation time in the PCR reaction. Another disadvantage was using a retroviral gene trap, as the terminal sequences of the provirus are LTRs. Primer binding sites at the terminus would have been ideal as the amplicon would have started directly into the flanking endogenous sequence, however, the LTR sequences were too similar to endogenous retroviral sequences, which would have led to non-provirus specific amplifications. Also, the LTRs were repeated at the 5' and 3' termini, and so primer biding would have occurred at both termini, leading to a double sequencing result.

Table 6. Advantages and disadvantages of the PCR methods used in this study

Method	Advantages	Disadvantages
Adapter-ligation PCR	· Creates known ends for PCR	· Correct adapter preparation difficult to prove · DNA ends can self-ligate instead of onto adapters · Ligation of adapter to DNA inefficient · Risk of amplification of endogenous retroviral sequences
iPCR	· Less modifications required than in adapter-ligation PCR	· Risk of amplification of endogenous retroviral sequences · Provirus primers far from 5' LTR terminus
SAP-PCR	· No template DNA modifications required · Arbitrary primers increase the probability of priming from the endogenous sequence	· Provirus primers far from 5' LTR terminus · Arbitrary primers do not bind as randomly as expected, and may bind in the provirus itself · Risk of amplification of endogenous retroviral sequences

In general, the main obstacle to obtaining flanking sequence information was to produce a PCR amplicon without having sequence information on which to base a primer for the endogenous sequence. The first of the strategies was adapter-ligation PCR. The idea behind this method was that if a known sequence could be added onto a digested end, this could be used to design a primer for PCR. In order to create an end of known sequence, however, several steps were required, leaving room for potential problem areas.

It was necessary for the restriction enzyme to completely digest the DNA at all sites; this was likely achieved. The ligation of the adapters to the restricted ends, however, was probably not as efficient, leading to the absence of an adapter at some of the restricted ends.

The reasons for poor annealing efficiency are not known, but several possibilities exist. First, it is possible that the digested genomic DNA ends did not remain free for annealing to an adapter. As there was so much DNA and so many available sticky ends, it is probable that the ends re-annealed to each other. Second, it was not known whether the adapter, made from two hybridised oligonucleotides, was made correctly. We attempted to visualise a change in molecular size between the individual oligonucleotides versus the annealed product using a polyacrylamide gel (results not shown), however, it was difficult to judge whether there was a change in size and the results were deemed inconclusive. Additionally, it was not known whether the melting step performed before each use was effective in melting sticky ends of the adapter to make them available for annealing.

The last step in the procedure, the PCR, also had several potential problem areas. First, the sensitivity of the PCR may have been questionable. As the target DNA (provirus) was in such low ratio to all other DNA present in the sample, it was unlikely that PCR would have been sensitive enough to produce an amplicon without the extra amplification step of a nested reaction. Unfortunately, a nested reaction could not be performed as the adapter that was made was only long enough for a single primer binding site. Perhaps a longer adapter with space for two primer binding sites for nested PCR might have produced better results. Second, the primer based on the provirus was designed to bind in the neomycin resistance gene, which is quite a distance from the 5' terminus of the provirus, so in addition to sensitivity, the PCR reaction was also required to amplify a large region of the provirus, without even taking into account the length of the endogenous sequence which was, of course, unknown. This was likely a major contributing factor to the poor results as optimisation of such a long PCR program is quite difficult, especially when using genomic DNA of unknown amplicon length.

iPCR was the next strategy used to try to identify flanking endogenous sequences. This method addressed the problem of primer design in a different way. Here, the DNA was digested to produce a fragment containing part proviral and part endogenous sequence which would ligate on itself. Then, using the known proviral sequence, a pair of primers were designed to amplify the endogenous portion of the circular fragment. The digest and self-ligation with the plasmid positive control produced good results, however, when the genomic DNA was used, only the control reaction worked well.

Once again, complete digestion of the genomic DNA was likely achieved. The ligation in this method was probably more efficient than in the previous method, leading

to a higher number of template DNA copies available for PCR, as demonstrated by the results presented in Figure 14.

Due to the location of the digest site of the chosen restriction enzyme (*Xho*I) which was located between the *gag* region and the cassette, the only reasonable location in which to design primers was the *gag* region. It would have been extremely difficult to design primers in the LTR region and ensure that amplification from any of the endogenous retroviral sequences would not occur. Nonetheless, the primers designed in the *gag* region still produced many non-specific amplicons rather than the desired 4.7 kb product in a control reaction using digested and ligated plasmid DNA. The Southern blot results showed that the target fragment would be quite large (>10 kb, exact size unknown) necessitating a long elongation time in the PCR program. With the adapter-ligation PCR, optimisation of a long PCR program was difficult, and could not be achieved for amplifications using genomic DNA. It is not known why the primers chosen for this reaction did not amplify the desired region as expected. As the region in which primers could be designed was quite limited (*gag* region), changing the primers within this narrow region was not likely to change the outcome.

SAP-PCR was the third method examined for identification of the flanking sequences. This strategy was different from the others in that a primer binding within the endogenous sequence was used. Here, in contrast to the other strategies, an arbitrary (non-specific) primer was used, meaning it was necessary for an error in primer binding to occur in order to begin amplification from within the endogenous sequence. In this way, similar to a random priming strategy, amplicons of varied lengths could be produced and neither their lengths nor their start sites could be predicted. Similar

obstacles to the design of the provirus-specific primer were encountered with this strategy as well. It was necessary to design the primer far enough from the terminus out of the LTR region so that non-specific binding to endogenous sequence would not occur. Unfortunately, the farther the primer was from the 5' terminus, the greater the chance that the arbitrary primer would bind in the provirus and amplify from there instead of from the endogenous sequence.

After several control reactions to test the suitability of the candidate P3 primers, it was surprising that many P3 primers did not prime arbitrarily at all, producing no amplicons. The reason for the low priming efficiency is not known, and although several modifications to the PCR program were attempted, the outcomes were not successful.

Future Directions

At the outset of this study, the goals included production of cellular clones resistant to infection or CPE by NiV and analysis of the mutations created in these cells. Numerous cell clones were generated, however, molecular analysis proved to be rather technically difficult. Future PCR optimisation efforts will be required in order to take advantage of these unique cell lines that were created. The ideal result would be identification of knocked out genes and determination of their roles in NiV pathogenesis.

Further work in the future would require knock-in (versus knockout) experiments of any identified genes. Many cell clones have been generated in which there is a retroviral vector integration, however, it remains to be shown that the gene knockout is the key to the resistance of the cell to CPE or infection by NiV. This gene knockout must be complemented by a knock-in to ensure that restoration of the expression of the gene

leads to the normal, susceptible phenotype. By causing the cell clone to become susceptible to NiV infection, this will provide proof that the gene, when knocked out, confers resistance to CPE or infection by NiV, and that this gene plays a role in virus pathogenesis.

Previously, Gogos and colleagues (1997) have carried out knock-in experiments by introducing wild-type cathepsin B cDNA into the knockout cellular clones and observing the resulting phenotype. A similar strategy may be employed for future studies with the cellular clones created during this work.

Research Impact

The cells created in this project have several potential uses. Some of these cell clones are persistently infected and as such could be used for the development of an enzyme-linked immunosorbant assay (ELISA) for NiV. Persistently infected cells would provide a better detection material than acutely infected cells that die immediately. Also, some of the cell clones respond to NiV re-exposure with varied forms of CPE not typically seen during NiV exposure. These cells provide a means to study CPE responses to NiV without the drawback of killing the cells quickly, as is normally the case.

The concept of viral interference was discussed with respect to the BHK-21 clone AD6, where it was hypothesised that the reason for resistance to re-exposure to NiV was due to occupation of the cellular receptor by NiV G. This cellular clone could be exploited to further study this phenomenon.

In addition, the CRFK cells that were exposed to HeV after NiV exposure, and remained free of HeV, support the suggestion that NiV and HeV are very closely related.

Although specific genes involved in NiV pathogenesis are unclear, when these genes are identified using the cellular clones created here, these studies can be expanded to studying HeV pathogenesis genes. This will provide further investigations into the predicted similarities between NiV and HeV pathogenesis, and their similarities to and differences from other paramyxoviruses.

The techniques described in this manuscript have the potential to be applied to the study of other viruses not related to NiV, for instance, emerging viruses that require study in order to characterise and understand their pathogenesis. Gene trapping provides a global approach to identifying potential interactions between the virus and host cell, and once optimised, could provide a fast, easy tool for virus pathogenesis study.

Conversely, genes that confer susceptibility to a virus infection could also be identified using gene trapping. For instance, viruses that cannot be grown in tissue culture could be studied to investigate which cellular genes play a role in preventing the growth of a particular virus in tissue culture. Specifically, a major obstacle in Hepatitis C virus research is the lack of a susceptible cell line in which to study cell responses and to grow the virus. Using the principles of the techniques described here, one could study the genes involved in susceptibility, rather than resistance to infection.

Appendix 1

Table 7. List of primers used in PCR reactions

Section	Name	Sequence	Experiment
2.12.2	M5	ATT GC(TC) TTC AAT CTT CTA GT	NiV detection
2.12.2	M3	GAT CTC ACA ACT GTT GTT CC	NiV detection
2.12.2	NP5	ATC AAT CGT GGT TAT CTT GA	NiV detection
2.12.2	NP3	CAG C(GC)A GTT CTG CAA CTT GAT C	NiV detection
2.12.3	HeV4912f	GCC CTG CGC CAA CAA GGC	HeV detection
2.12.3	HeV5113r	ATC AAG GTT ATC ACT CAC AC	HeV detection
2.12.3	HeV11266f	CGT GCA TGA TTG AGA TAG TCA G	HeV detection
2.12.3	HeV11870r	TAT GTT ACT GAG TCT GTC CTG	HeV detection
2.13.1	neoT7	GAG GAC TAT AAT ACG ACT CAC TAT AGG GTG ATG CCG CCG CCG TTC CGG	Southern blot
2.13.1	neoSP6	GAC TAC ATA CGA TTT AGG TGA CAC TAT AGA CCA TGA TAT TCG GCA AGC	Southern blot
2.14.4	ipcr19	GAG AAG GAG TGA GGG CTG	Southern blot
2.14.4	ipcr20	GCC AGA CTG TTA CCA CTC C	Southern blot
2.15, 2.17.2	neo5f	TTG AAC AAG ATG GAT TGC ACG	PCR control, iPCR
2.15	neo3r	TCA GAA GAA CTC GTC AAG AAG	PCR control
2.16.2	linker	CAT CGC CGG CG	Adapter-ligation PCR
2.16.2	Kpnlink	GTA GCG GCC GCG GTA C	Adapter-ligation PCR
2.16.3	neo3f	GGT ATC GCC GCT CCC GAT TCG	Adapter-ligation PCR
2.17.2	pRseq11	GGC ATT CTG CAC GCT TCA AA	iPCR
2.17.2	pRseq12	GAA TTG ATC CTT ACA CTT CGG	iPCR
2.17.2	pRseq02	TCA ACG CTG TGA TGA CCT AC	iPCR
2.17.3	ipcr21	ACA TCG TGA CCT GGG AAG C	iPCR
2.17.3	ipcr22	GGA CCT GGT CTG GGT GTC	iPCR
2.17.3	ipcr23	ACC CTA AGC CTC CGC CTC	iPCR
2.17.3	ipcr24	GGT CTC GGT TAA AGG TGC C	iPCR
2.18.2	sappcr03	CAA ACT TAA GGG AGT GGT AAC	SAP-PCR
2.18.2	sappcr04	TCC GCT CGA CAT CTT TCC AG	SAP-PCR
2.18.2	EBO 30F	ATG GAA GCT TCA TAT GAG AGA GG	SAP-PCR
2.18.2	EBO 30R	TTA AGG GGT ACC CTC ATC AGA C	SAP-PCR
2.18.2	HCV 215R	TGA TAA GCT TCT ACG CGT CGA CGC CGG CA	SAP-PCR

2.18.2	HCV 216F	TGT TGA ATT CAT GGA AAC CCA CGT CAC CGG G	SAP-PCR
2.18.2	HCV 213R	TGA TGG ATC CTA GGC TGA AGC GGG CAC AGT	SAP-PCR
2.18.2	HCV 227R	TGA TGG ATC CTA GCA TGG AGT GTT ACA CTC CG	SAP-PCR
2.18.2	LassaGP 188f	GGA CTT ACG AGC TGC AAA CCC TTG AG	SAP-PCR
2.18.2	LassaGP 649f	CAA TAT CTA ATC ATT CAG AAT ACA ACC	SAP-PCR
2.18.2	Var crmB3	GAC GCT AGA TAG ACA GTC	SAP-PCR
2.18.2	Var crmB5	TAC CGG TCT CAG CGA ATC	SAP-PCR
2.18.2	Var HA3	TCT TGT CTG TTG TGG ATT CT	SAP-PCR
2.18.2	Var HA5	ATG CCG GTA CTT ATG TAT GTG C	SAP-PCR
2.18.2	HerBV polF	AAC TCG GTG TAC GGG TTC AC	SAP-PCR
2.18.2	HerBV 1-2 I-r2	CTC CAG GCG GTC CAG GGT CAT	SAP-PCR
2.18.2	HerBV GS4	CCG CGT ACG ACT ACG AGA TCC	SAP-PCR
2.18.2	HerBV GAS4	GTT CGC GGC CAC GAT CCA	SAP-PCR
2.18.2	HSV P1	GTG GTG GAC TTT GCC AGC CTG TAC CC	SAP-PCR
2.18.2	HSV P2	TAA ACA TGG AGT CCG TGT CGC CGT AGA TGA	SAP-PCR
2.18.2	CCHF SM19	GCA GGG TTG TTG GAT AGT C	SAP-PCR
2.18.2	CCHF SM20	CAT GTT GAA CAT GCC GGT AG	SAP-PCR
2.18.3	T3	ATT AAC CTC ACT AAA GGG A	SAP-PCR
2.18.3	T7	TAA TAC GAC TCA CTA TAG GG	SAP-PCR

Appendix 2 - Recipes

Section 2.1 - Cell Culture Freezing Solutions (in house recipes)

Solution A

80% FBS
20% cell culture medium
Filter sterilise, cool to 4°C.

Solution B

80% FBS
20% dimethyl sulfoxide (DMSO)
Filter sterilise, cool to 4°C.

Section 2.1 - TSS Buffer (in house recipe)

volume/mass	chemical	final concentration
85 ml	LB broth	85%
10 g	polyethylene glycol	10%
5 ml	DMSO	5%
1.00165 g	$MgCl_2$	50mM

Combine all components and adjust volume to 100 ml with dH_2O. Filter sterilise. Store at 4°C.

Section 2.3 – SOC Medium (Sambrook, 1989, A.2)

volume/mass	chemical
20 g	bacto-tryptone
5 g	bacto-yeast extract
0.5 g	NaCl

Combine in 900 ml of dH_2O, shake to dissolve. Add 10 ml of 250 mM solution of KCl. Adjust to pH 7.0 with 5 N NaOH. Adjust volume to 1 L with dH_2O. Autoclave. Add 20 ml of sterile 1 M glucose solution. Just before use, add 5 ml sterile 2 M $MgCl_2$ solution.

Section 2.5.1 - 6x gel loading buffer (Sambrook, 1989)
(modified buffer type III, B.24)

final concentration	chemical
30%	glycerol
0.25%	bromophenol blue

Section 2.13.1 - 10x FA gel buffer
(www.qiagen.com/literature/BenchGuide/pdf/1017778_BenchGuide_chap_3.pdf#45)
1 L total volume in 18 MΩ H_2O:

volume	chemical
41.9 g	3-(N-morpholino) propane-sulfonic acid (MOPS)
6.8 g	sodium acetate•H_2O
20 ml	0.5 M EDTA, pH 8.0

pH to 7.0 with NaOH.
Treat entire volume with 0.1% diethyl pyrocarbonate (DEPC) at 37°C overnight.
Autoclave for 15 minutes.
Dilute to 1x in DEPC-treated H_2O.

Section 2.13.1 - 5x RNA loading buffer
(www.qiagen.com/literature/BenchGuide/pdf/1017778_BenchGuide_chap_3.pdf#45)
10 ml total volume in DEPC-treated H_2O:

volume/mass	chemical
80 µl	0.5 M EDTA
750 µl	37% formaldehyde
2 ml	glycerol
3.084 ml	formamide
4 ml	10x FA gel buffer
to desired colour	bromophenol blue

Section 2.13.2 - 20x SSC
20x SSC was either purchased commercially from Invitrogen or made in house according to Sambrook, 1989 (B.13)

volume/mass	chemical
175.3 g	NaCl
88.2 g	sodium citrate

Adjust pH to 7.0 with 10 M NaOH. Adjust volume to 1 L with dH_2O. Autoclave.
10x, 2x, 1x, and 0.1x solutions were made by diluting the 20x stock solution with dH_2O.

Section 2.13.3 - Depurination Solution
(Current Protocols, http://www.mrw2.interscience.wiley.com/)

final concentration	chemical
0.25 M	HCl

Dilute 1 M HCl solution to 0.25 M working concentration.

Section 2.13.3 - Denaturation Solution (Current Protocols)

final concentration	chemical
1.5 M	NaCl
0.5 M	NaOH

Section 2.13.3 - Neutralisation Solution (Current Protocols)

final concentration	chemical
1.5 M	NaCl
0.5 M	Tris·Cl, pH 7.0

Section 2.17.2 – ATP Solution for Ligase Buffer (Sambrook, 1989, B.10)
Dissolve 60 mg of ATP in 0.8 ml of dH_2O. Adjust to pH 7.0 with 0.1 N NaOH. Adjust volume to 1 ml with dH_2O. Aliquot, store at –80°C.

Appendix 3a

Table 8. Primers used to sequence pROSA-GFNR

Name	Sequence	Binding Location (base number)
Neo5r	GAA TAG CCT CTC CAC CCA AGC	2686
Neo3r	TCA GAA GAA CTC GTC AAG AAG	1949
Neo5f	TTG AAC AAG ATG GAT TGC ACG	2741
Neo3f	GGT ATC GCC GCT CCC GAT TCG	2011
pBRPvur	GAC AAG CTG TGA CCG TCT	5871
pBREcof	ATT ATT ATC ATG ACA TTA ACC	8030
pRseq01	GTC GAA CGA GGA GGT TCA	1595
pRseq02	TCA ACG CTG TGA TGA CCT AC	3347
pRseq03	CCG TGT ATC CAA TAA ACC CTC	6506
pRseq04	GCG GTG GCG GCC GCT C	5789
pRseq05	CAG TTG CAT CCG ACT TGT GG	664
pRseq06	TCG GTG ATG ACG GTG AAA AC	5812
pRseq07	CTT TAA CAA ATT GGA CTA ATC GA	36
pRseq08	GAA CTT CTT ATT CTC AGT TAT G	5163
pRseq09	GTC TGT TGT GCC CAG TCA TA	2663
pRseq10	CTG GGC AAG GGT CTC CCG	785
pRseq11	GGC ATT CTG CAC GCT TCA AA	2838
pRseq12	GAA TTG ATC CTT ACA CTT CGG	3273
pRseq13	GGT CTT TGC TCA GGG CGG	4101
pRseq15	GTT TGT CCT CAA CCG CGA G	4793
pRseq16	GCG TCA TTT GAT ATG ATG CGC	5351
pRseq17	TTC TGG AGA CTA AAT AAA ATC TTT	135
pRseq18	AAC CAG CTG GGG CTC GAT CC	1671
pRseq19	CAA TAT TAT TGA AGC ATT TAT CAG	7898
pRseq20	TGC TCT GAT GCC GCA TAG TTA	6006
pRseq21	TTT CTC AAC GTA ACA CTT TAC AGC	5378
pRseq22	ATG GTC CCC AGA TGC GGT	418
pRseq23	TGG CCG CTC GGG AAC CC	5266
pRseq24	GAA GCT CCC TCG TGC GCT C	6369
pRseq25	ATT GTT GCC GGG AAG CTA GAG	7272
pRseq26	TGG GAA GCC TTG GCT TTT GAC	1480
pRseq27	CAC GCT TCC CGA AGG GAG AA	6448
pRseq28	CCT TGA TCG TTG GGA ACC GG	7408
pRseq29	TAA TGT CAT GAT AAT AAT GGT TTC TT	8047
pRseq30	AAA CTG CTG AGG GCT GGA CC	452
pRseq31	CGT TGG CTA CCC GTG ATA TTG C	2080
pRseq32	GCC ATC CTC GAT GCA GAA TTT G	3419
pRseq33	GGA CGG CAG CGT GCA GCT C	4216
pRseq34	CCC GGG CTG CAG ATC TGT AG	4894

Appendix 3b. Complete sequence of pROSA-GFNR. Base 1 immediately follows the end of the pBR322 sequence.

Legend: blue – 5' and 3' LTRs
green – *gag* region
red underline – neomycin resistance gene
red – PGK-1 promoter
blue underline – NTR, EGFP fusion gene
fuschia – pBR322

```
1     GAGGTCGACG GTATCGATTA GTCCAATTTG TTAAAGACAG GATATCAGTG
51    GTCCAGGCTC TAGTTTTGAC TCAACAATAT CACCAGCTGA AGCCTATAGA
101   GTACGAGCCA TAGATAAAAT AAAAGATTTT ATTTAGTCTC CAGAAAAAGG
151   GGGGAATGAA AGACCCCACC TGTAGGTTTG GCAAGCTAGC TTAAGTAACG
201   CCATTTTGCA AGGCATGGAA AAATACATAA CTGAGAATAG AGAAGTTCAG
251   ATCAAGGTCA GGAACAGATG GAACAGCTGA ATATGGGCCA AACAGGATAT
301   CTGTGGTAAG CAGTTCCTGC CCCGGCTCAG GGCCAAGAAC AGATGGAACA
351   GCTGAATATG GGCCAAACAG GATATCTGTG GTAAGCAGTT CCTGCCCCGG
401   CTCAGGGCCA AGAACAGATG GTCCCCAGAT GCGGTCCAGC CCTCAGCAGT
451   TTCTAGAGAA CCATCAGATG TTTCCAGGGT GCCCCAAGGA CCTGAAATGA
501   CCCTGTGCCT TATTTGAACT AACCAATCAG TTCGCTTCTC GCTTCTGTTC
551   GCGCGCTTCT GCTCCCCGAG CTCAATAAAA GAGCCCACAA CCCCTCACTC
601   GGCGCGCCAG TCCTCCGATT GACTGAGTCG CCCGGGTACC CGTGTATCCA
651   ATAAACCCTC TTGCAGTTGC ATCCGACTTG TGGTCTCGCT GTTCCTTGGG
701   AGGGTCTCCT CTGAGTGATT GACTACCCGT CAGCGGGGGT CTTTCATTTG
751   GGGGCTCGTC CGGGATCGGG AGACCCTTGC CCAGGGACCA CCGACCCACC
801   ACCGGGGAGG TAAGCTGGCC AGCAACTTAT CTGTGTCTGT CCGATTGTCT
851   AGTGTCTATG ACTGATTTTA TGCGCCTGCG TCGGTACTAG TTAGCTAACT
901   AGCTCTGTAT CTGGCGGACC CGTGGTGGAA CTGACGAGTT CGGAACACCC
951   GGCCGCAACC CTGGGAGACG TCCCAGGGAC TTCGGGGGCC GTTTTTGTGG
1001  CCCGACCTGA GTCCAAAAAT CCCGATCGTT TTGGACTCTT TGGTGCACCC
1051  CCCTTAGAGG AGGGATATGT GGTTCTGGTA GGAGACGAGA ACCTAAAACA
1101  GTTCCCGCCT CCGTCTGAAT TTTTGCTTTC GGTTTGGGAC CGAAGCCGCG
1151  CCGCGCGTCT TGTCTGCTGC AGCATCGTTC TGTGTTGTCT CTGTCTGACT
1201  GTGTTTCTGT ATTTGTCTGA AAATATGGGC CAGACTGTTA CCACTCCCTT
1251  AAGTTTGACC TTAGGTCACT GGAAAGATGT CGAGCGGATC GCTCACAACC
1301  AGTCGGTAGA TGTCAAGAAG AGACGTTGGG TTACCTTCTG CTCTGCAGAA
1351  TGGCCAACCT TTAACGTCGG ATGGCCGCGA GACGGCACCT TTAACCGAGA
1401  CCTCATCACC CAGGTTAAGA TCAAGGTCTT TTCACCTGGC CCGCATGGAC
1451  ACCCAGACCA GGTCCCCTAC ATCGTGACCT GGGAAGCCTT GGCTTTTGAC
1501  CCCCCTCCCT GGGTCAAGCC CTTTGTACAC CCTAAGCCTC CGCCTCCTCT
1551  TCCTCCATCC GCCCCGTCTC TCCCCCTTGA ACCTCCTCGT TCGACCCCGC
1601  CTCGATCCTC CCTTTATCCA GCCCTCACTC CTTCTCGAGG TCGACTCTAG
1651  AGGATCGAGC CCCAGCTGGT TCTTTCCGCC TCAGAAGCCA TAGAGCCCAC
1701  CGCATCCCCA GCATGCCTGC TATTGTCTTC CCAATCCTCC CCCTTGCTGT
1751  CCTGCCCCAC CCCACCCCCC AGAATAGAAT GACACCTACT CAGACAATGC
1801  GATGCAATTT CCTCATTTTA TTAGGAAAGG ACAGTGGGAG TGGCACCTTC
1851  CAGGGTCAAG GAAGGCACGG GGGAGGGGCA AACAACAGAT GGTGGCAACT
1901  AGAAGGCACA GTCGAGGCTG ATCAGCGAGC TCTAGAGAAT TGATCCCCTC
1951  AGAAGAACTC GTCAAGAAGG CGATAGAAGG CGATGCGCTG CGAATCGGGA
2001  GCGGCGATAC CGTAAAGCAC GAGGAAGCGG TCAGCCCATT CGCCGCCAAG
2051  CTCTTTCAGC AATATCACGG GTAGCCAACG CTATGTCCTG ATAGCGGTCC
```

2101 GCCACACCCA GCCGGCCACA GTCGATGAAT CCAGAAAAGC GGCCATTTTC
2151 CACCATGATA TTCGGCAAGC AGGCATCGCC ATGGGTCACG ACGAGATCAT
2201 CGCCGTCGGG CATGCGCGCC TTGAGCCTGG CGAACAGTTC GGCTGGCGCG
2251 AGCCCCTGAT GCTCTTCGTC CAGATCATCC TGATCGACAA GACCGGCTTC
2301 CATCCGAGTA CGTGCTCGCT CGATGCGATG TTTCGCTTGG TGGTCGAATG
2351 GGCAGGTAGC CGGATCAAGC GTATGCAGCC GCCGCATTGC ATCAGCCATG
2401 ATGGATACTT TCTCGGCAGG AGCAAGGTGA GATGACAGGA GATCCTGCCC
2451 CGGCACTTCG CCCAATAGCA GCCAGTCCCT TCCCGCTTCA GTGACAACGT
2501 CGAGCACAGC TGCGCAAGGA ACGCCCGTCG TGGCCAGCCA CGATAGCCGC
2551 GCTGCCTCGT CCTGCAGTTC ATTCAGGGCA CCGGACAGGT CGGTCTTGAC
2601 AAAAAGAACC GGGCGCCCCT GCGCTGACAG CCGGAACGGC GGCGGCATCA
2651 GAGCAGCCGA TTGTCTGTTG TGCCCAGTCA TAGCCGAATA GCCTCTCCAC
2701 CCAAGCGGCC GGAGAACCTG CGTGCAATCC ATCTTGTTCA AGTGGCCGAT
2751 CCCATATTGG CTGCAGGTCG AAAGGCCCGG AGATGAGGAA GAGGAGAACA
2801 GCGCGGCAGA CGTGCGCTTT TGAAGCGTGC AGAATGCCGG GCCTCCGGAG
2851 GACCTTCGGG CGCCCGCCCC GCCCCTGAGC CCGCCCCTGA GCCCGCCCCC
2901 GGACCCACCC CTTCCCAGCC TCTGAGCCCA GAAAGCGAAG GAGCAAAGCT
2951 GCTATTGGCC GCTGCCCCAA AGGCCTACCC GCTTCCATTG CTCAGCGGTG
3001 CTGTCCATCT GCACGAGACT AGTGAGACGT GCTACTTCCA TTTGTCACGT
3051 CCTGCACGAC GCGAGCTGCG GGGCGGGGGG GAACTTCCTG ACTAGGGGAG
3101 GAGTAGAAGG TGGCGCGAAG GGGCCACCAA AGAACGGAGC CGGTTGGCGC
3151 CTACCGGTGG ATGTGGAATG TGTGCGAGGC CAGAGGCCAC TTGTGTAGCG
3201 CCAAGTGCCC AGCGGGGCTG CTAAAGCGCA TGCTCCAGAC TGCCTTGGGA
3251 AAAGCGCCTC CCCTACCCGG TAGAATTGAT CCTTACACTT CGGTTAAGGT
3301 GATGTTTTGC GGCAGACGAG ATTTCGGCAG CGTAGCGTTA AAATCTTCAA
3351 CGCTGTGATG ACCTACCGGA ACAACCACCA GACTGGTGTA GCCTTTCTCT
3401 TTCAGACCAA ATTCTGCATC GAGGATGGCG GCGTCAAAAC CTTCGATGGG
3451 TACCGCGTCC AGACCCAGAG CCGCCACGCC GAGCAGGAAG TTACCGACGT
3501 TGAGATAAAC CTGTTTTGCC ATCCACTCTG CATCATCATG CAGATCTTTA
3551 CGGTGCATAT CAGCGAAGAA CTTGCGACCT TTATCGTTCG CGGCTTTCGC
3601 TTCCGGCGTG GCAAAGCGGC CATCGGCATC TTCCTGGTCA ACAACCAGCT
3651 TCAGCCAGAC ATCGTCCATC GCGGTTTTTG CACAGAACAC CACGACGTGC
3701 GAGGCATCAA GCATTTTACG CTCGTTGAAC ACGTAATTAC CGGCAGCGGA
3751 TTTGGCAACA CGCGCTTTAC CTTCTTCCGT GCTGGCAACA ATAAAATGCC
3801 ACGGCTGGGA GTTGGTGCTG GATGGGCTGT ATTGCAGTAG CGTTTTGATC
3851 TGCTCGGCCT GTTCCGGGGT AAGTTTTTTG CTGGCATCAA ATGCCTTAGT
3901 GGAATGACGC TTTAAGGCGA CAGAAATGAT ATCCATGGTA CCGGTGAATT
3951 CGATATCAAG CTTATCGATA CCGTCGACTG CAGAATTCGA AGCTTGAGCT
4001 CGAGATCTGA GTCCGGACTT GTACAGCTCG TCCATGCCGA GAGTGATCCC
4051 GGCGGCGGTC ACGAACTCCA GCAGGACCAT GTGATCGCGC TTCTCGTTGG
4101 GGTCTTTGCT CAGGGCGGAC TGGGTGCTCA GGTAGTGGTT GTCGGGCAGC
4151 AGCACGGGGC CGTCGCCGAT GGGGGTGTTC TGCTGGTAGT GGTCGGCGAG
4201 CTGCACGCTG CCGTCCTCGA TGTTGTGGCG GATCTTGAAG TTCACCCTTG
4251 ATGCCGTTCT TCTGCTTGTC GGCCATGATA TAGACGTTGT GGCTGTTGTA
4301 GTTGTACTCC AGCTTGTGCC CCAGGATGTT GCCGTCCTCC TTGAAGTCGA
4351 TGCCCTTCAG CTCGATGCGG TTCACCAGGG TGTCGCCCTC GAACTTCACC
4401 TCGGCGCGGG TCTTGTAGTT GCCGTCGTCC TTGAAGAAGA TGGTGCGCTC
4451 CTGGACGTAG CCTTCGGGCA TGGCGGACTT GAAGAAGTCG TGCTGCTTCA
4501 TGTGGTCGGG GTAGCGGCTG AAGCACTGCA CGCCGTAGGT CAGGGTGGTC
4551 ACGAGGGTGG GCCAGGGCAC GGGCAGCTTG CCGGTGGTGC AGATGAACTT
4601 CAGGGTCAGC TTGCCGTAGG TGGCATCGCC CTCGCCCTCG CCGGACACGC
4651 TGAACTTGTG GCCGTTTACG TCGCCGTCCA GCTCGACCAG GATGGGCACC
4701 ACCCCGGTGA ACAGCTCCTC GCCCTTGCTC ACCATGGTGG CGACCGGTAG
4751 CGCTAGCGAT ACCGTCGATC CCCACTGGAA AGACCGCGAA GAGTTTGTCC

4801 TCAACCGCGA GCTGTGGAAA AAAAAGGGAC AGGATAAGTA TGACATCATC
4851 AAGGAAACCC TGGACTACTG CGCCCTACAG ATCTGCAGCC CGGGGGATCC
4901 ACTAGTTCTA GCCTCGAGGT CGACGGTATC GATTAGTCCA ATTTGTTAAA
4951 GACAGGATAT CAGTGGTCCA GGCTCTAGTT TTGACTCAAC AATATCACCA
5001 GCTGAAGCCT ATAGAGTACG AGCCATAGAT AAAATAAAAG ATTTTATTTA
5051 GTCTCCAGAA AAAGGGGGGA ATGAAAGACC CCACCTGTAG GTTTGGCAAG
5101 CTAGCTTAAG TAACGCCATT TTGCAAGGCA TGGAAAAATA CATAACTGAG
5151 AATAGAGAAG TTCAGATCGG GATCCCAATT CTTTCGGACT TTTGAAAGTG
5201 ATGGTGGTGG GGGAAGGATT CGAACCTTCG AAGTCGATGA CGGCAGATTT
5251 AGAGTCTGCT CCCTTTGGCC GCTCGGGAAC CCCACCACAG GTAATGCTTT
5301 TACTGGCCTG CTCCCTTATC GGGAAGCGGG GCGCATCATA TCAAATGACG
5351 CGCCGCTGTA AAGTGTTACG TTGAGAAAGA ATTGGGATCC CGATCAAGGT
5401 CAGGAACAGA TGGAACAGCT AGAGAACCAT CAGATGTTTC CAGGGTGCCC
5451 CAAGGACCTG AAATGACCCT GTGCCTTATT TGAACTAACC AATCAGTTCG
5501 CTTCTCGCTT CTGTTCGCGC GCTTCTGCTC CCCGAGCTCA ATAAAAGAGC
5551 CCACAACCCC TCACTCGGCG CGCCAGTCCT CCGATTGACT GAGTCGCCCG
5601 GGTACCCGTG TATCCAATAA ACCCTCTTGC AGTTGCATCC GACTTGTGGT
5651 CTCGCTGTTC CTTGGGAGGG TCTCCTCTGA GTGATTGACT ACCCGTCAGC
5701 GGGGGTCTTT CACCCAGAGT TTGGAACTTA CTGTCTTCTT GGGACCTGCA
5751 GCCCGGGGGA TCCACTAGTT CTAGAGCGGC CGCCACCGCG GTGGATTCTG
5801 CCTCGCGCGT TTCGGTGATG ACGGTGAAAA CCTCTGACAC ATGCAGCTCC
5851 CGGAGACGGT CACAGCTTGT CTGTAAGCGG ATGCCGGGAG CAGACAAGCC
5901 CGTCAGGGCG CGTCAGCGGG TGTTGGCGGG TGTCGGGGCG CAGCCATGAC
5951 CCAGTCACGT AGCGATAGCG GAGTGTATAC TGGCTTAACT ATGCGGCATC
6001 AGAGCAGATT GTACTGAGAG TGCACCATAT GCGGTGTGAA ATACCGCACA
6051 GATGCGTAAG GAGAAAATAC CGCATCAGGC GCTCTTCCGC TTCCTCGCTC
6101 ACTGACTCGC TGCGCTCGGT CGTTCGGCTG CGGCGAGCGG TATCAGCTCA
6151 CTCAAAGGCG GTAATACGGT TATCCACAGA ATCAGGGGAT AACGCAGGAA
6201 AGAACATGTG AGCAAAAGGC CAGCAAAAGG CCAGGAACCG TAAAAAGGCC
6251 GCGTTGCTGG CGTTTTTCCA TAGGCTCCGC CCCCCTGACG AGCATCACAA
6301 AAATCGACGC TCAAGTCAGA GGTGGCGAAA CCCGACAGGA CTATAAAGAT
6351 ACCAGGCGTT TCCCCCTGGA AGCTCCCTCG TGCGCTCTCC TGTTCCGACC
6401 CTGCCGCTTA CCGGATACCT GTCCGCCTTT CTCCCTTCGG GAAGCGTGGC
6451 GCTTTCTCAT AGCTCACGCT GTAGGTATCT CAGTTCGGTG TAGGTCGTTC
6501 GCTCCAAGCT GGGCTGTGTG CACGAACCCC CCGTTCAGCC CGACCGCTGC
6551 GCCTTATCCG GTAACTATCG TCTTGAGTCC AACCCGGTAA GACACGACTT
6601 ATCGCCACTG GCAGCAGCCA CTGGTAACAG GATTAGCAGA GCGAGGTATG
6651 TAGGCGGTGC TACAGAGTTC TTGAAGTGGT GGCCTAACTA CGGCTACACT
6701 AGAAGGACAG TATTTGGTAT CTGCGCTCTG CTGAAGCCAG TTACCTTCGG
6751 AAAAAGAGTT GGTAGCTCTT GATCCGGCAA ACAAACCACC GCTGGTAGCG
6801 GTGGTTTTTT TGTTTGCAAG CAGCAGATTA CGCGCAGAAA AAAAGGATCT
6851 CAAGAAGATC CTTTGATCTT TTCTACGGGG TCTGACGCTC AGTGGAACGA
6901 AAACTCACGT TAAGGGATTT TGGTCATGAG ATTATCAAAA AGGATCTTCA
6951 CCTAGATCCT TTTAAATTAA AAATGAAGTT TTAAATCAAT CTAAAGTATA
7001 TATGAGTAAA CTTGGTCTGA CAGTTACCAA TGCTTAATCA GTGAGGCACC
7051 TATCTCAGCG ATCTGTCTAT TTCGTTCATC CATAGTTGCC TGACTCCCCG
7101 TCGTGTAGAT AACTACGATA CGGGAGGGCT TACCATCTGG CCCCAGTGCT
7151 GCAATGATAC CGCGAGACCC ACGCTCACCG GCTCCAGATT TATCAGCAAT
7201 AAACCAGCCA GCCGGAAGGG CCGAGCGCAG AAGTGGTCCT GCAACTTTAT
7251 CCGCCTCCAT CCAGTCTATT AATTGTTGCC GGGAAGCTAG AGTAAGTAGT
7301 TCGCCAGTTA ATAGTTTGCG CAACGTTGTT GCCATTGCTG CAGGCATCGT
7351 GGTGTCACGC TCGTCGTTTG GTATGGCTTC ATTCAGCTCC GGTTCCCAAC
7401 GATCAAGGCG AGTTACATGA TCCCCCATGT TGTGCAAAAA AGCGGTTAGC
7451 TCCTTCGGTC CTCCGATCGT TGTCAGAAGT AAGTTGGCCG CAGTGTTATC

```
7501 ACTCATGGTT ATGGCAGCAC TGCATAATTC TCTTACTGTC ATGCCATCCG
7551 TAAGATGCTT TTCTGTGACT GGTGAGTACT CAACCAAGTC ATTCTGAGAA
7601 TAGTGTATGC GGCGACCGAG TTGCTCTTGC CCGGCGTCAA CACGGGATAA
7651 TACCGCGCCA CATAGCAGAA CTTTAAAAGT GCTCATCATT GGAAAACGTT
7701 CTTCGGGGCG AAAACTCTCA AGGATCTTAC CGCTGTTGAG ATCCAGTTCG
7751 ATGTAACCCA CTCGTGCACC CAACTGATCT TCAGCATCTT TTACTTTCAC
7801 CAGCGTTTCT GGGTGAGCAA AAACAGGAAG GCAAAATGCC GCAAAAAAGG
7851 GAATAAGGGC GACACGGAAA TGTTGAATAC TCATACTCTT CCTTTTTCAA
7901 TATTATTGAA GCATTTATCA GGGTTATTGT CTCATGAGCG GATACATATT
7951 TGAATGTATT TAGAAAAATA AACAAATAGG GGTTCCGCGC ACATTTCCCC
8001 GAAAAGTGCC ACCTGACGTC TAAGAAACCA TTATTATCAT GACATTAACC
8051 TATAAAAATA GGCGTATCAC GAGGCCCTTT CGTCTTC
```

Appendix 4

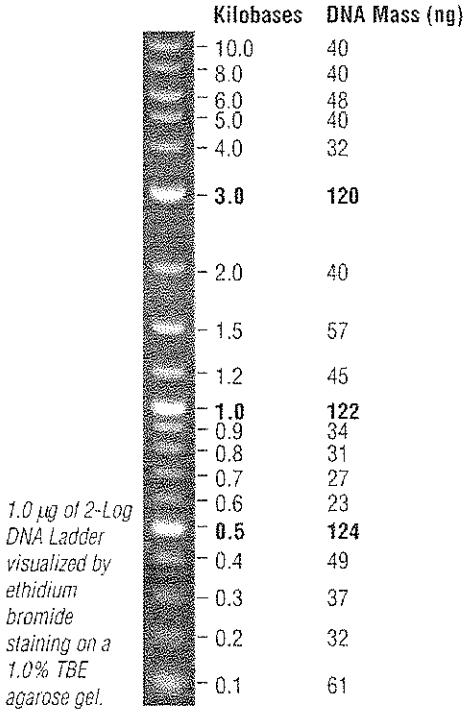

Figure 15. New England Biolabs 2 Log Ladder™

Reference List

1. Bench Guide (RNA). Qiagen Literature General Molecular Biology Protocols Online . 2003.
   Ref Type: Electronic Citation

2. 1999. Biosafety in Microbiological and Biomedical Laboratories. U. S. Government Printing Office, Washington.

3. Current Protocols in Molecular Biology, Chapter 2, Section IV, Unit 2.9A. Current Protocols in Molecular Biology Online . 2003. John Wiley and Sons Inc.
   Ref Type: Electronic Citation

4. 2001. Fields Virology. Lippincott Williams & Wilkins, New York.

5. New England Biolabs Inc. 2002-03 Catalog & Technical Reference. 2002. New England Biolabs Inc.
   Ref Type: Catalog

6. 2001. The Springer Index of Viruses. Springer, New York.

7. 2000. Yeast Hybrid Technologies. Eaton Publishing, Natick, MA.

8. **Banerjee, A. K.** 1987. Transcription and replication of rhabdoviruses. Microbiol. Rev. **51**:66-87.

9. **Bolt, G., K. Berg, and M. Blixenkrone-Moller**. 2002. Measles virus-induced modulation of host-cell gene expression. J. Gen. Virol. **83**:1157-1165.

10. **Boris-Lawrie, K. and H. M. Temin**. 1994. The retroviral vector. Replication cycle and safety considerations for retrovirus-mediated gene therapy. Ann. N. Y. Acad. Sci. **716**:59-70.

11. **Bossart, K. N., L. F. Wang, B. T. Eaton, and C. C. Broder**. 2001. Functional expression and membrane fusion tropism of the envelope glycoproteins of Hendra virus. Virology **290**:121-135.

12. **Chen, Z., G. A. Friedrich, and P. Soriano**. 1994. Transcriptional enhancer factor 1 disruption by a retroviral gene trap leads to heart defects and embryonic lethality in mice. Genes Dev. **8**:2293-2301.

13. **Chua, K. B., W. J. Bellini, P. A. Rota, B. H. Harcourt, A. Tamin, S. K. Lam, T. G. Ksiazek, P. E. Rollin, S. R. Zaki, W. Shieh, C. S. Goldsmith, D. J. Gubler, J. T. Roehrig, B. Eaton, A. R. Gould, J. Olson, H. Field, P. Daniels, A. E. Ling, C. J. Peters, L. J. Anderson, and B. W. Mahy**. 2000. Nipah virus: a recently emergent deadly paramyxovirus. Science **288**:1432-1435.

14. **Czub, M., F. J. McAtee, S. Czub, W. P. Lynch, and J. L. Portis**. 1995. Prevention of retrovirus-induced neurological disease by infection with a nonneuropathogenic retrovirus. Virology **206**:372-380.

15. **DeBiasi, R. L., P. Clarke, S. Meintzer, R. Jotte, B. K. Kleinschmidt-Demasters, G. L. Johnson, and K. L. Tyler**. 2003. Reovirus-induced alteration in expression of apoptosis and DNA repair genes with potential roles in viral pathogenesis. J. Virol. **77**:8934-8947.

16. **DeGregori, J., A. Russ, H. von Melchner, H. Rayburn, P. Priyaranjan, N. A. Jenkins, N. G. Copeland, and H. E. Ruley**. 1994. A murine homolog of the yeast RNA1 gene is required for postimplantation development. Genes Dev. **8**:265-276.

17. **Delwart, E. L. and A. T. Panganiban**. 1989. Role of reticuloendotheliosis virus envelope glycoprotein in superinfection interference. J. Virol. **63**:273-280.

18. **Elbashir, S. M., J. Harborth, W. Lendeckel, A. Yalcin, K. Weber, and T. Tuschl**. 2001. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. Nature **411**:494-498.

19. **Elbashir, S. M., W. Lendeckel, and T. Tuschl**. 2001. RNA interference is mediated by 21- and 22-nucleotide RNAs. Genes Dev. **15**:188-200.

20. **Esolen, L. M., S. W. Park, J. M. Hardwick, and D. E. Griffin**. 1995. Apoptosis as a cause of death in measles virus-infected cells. J. Virol. **69**:3955-3958.

21. **Field, H., P. Young, J. M. Yob, J. Mills, L. Hall, and J. Mackenzie**. 2001. The natural history of Hendra and Nipah viruses. Microbes. Infect. **3**:307-314.

22. **Friedrich, G. and P. Soriano**. 1993. Insertional mutagenesis by retroviruses and promoter traps in embryonic stem cells. Methods Enzymol. **225**:681-701.

23. **Friedrich, G. and P. Soriano**. 1991. Promoter traps in embryonic stem cells: a genetic screen to identify and mutate developmental genes in mice. Genes Dev. **5**:1513-1523.

24. **Fugier-Vivier, I., C. Servet-Delprat, P. Rivailler, M. C. Rissoan, Y. J. Liu, and C. Rabourdin-Combe**. 1997. Measles virus suppresses cell-mediated immunity by interfering with the survival and functions of dendritic and T cells. J. Exp. Med. **186**:813-823.

25. **Gietz, R. D. and R. A. Woods**. 2002. Screening for protein-protein interactions in the yeast two-hybrid system. Methods Mol. Biol. **185**:471-486.

26. **Gogos, J. A., R. Thompson, W. Lowry, B. F. Sloane, H. Weintraub, and M. Horwitz**. 1996. Gene trapping in differentiating cell lines: regulation of the

lysosomal protease cathepsin B in skeletal myoblast growth and fusion. J. Cell Biol. **134**:837-847.

27. **Gogos, J. A., W. Lowry, and M. Karayiorgou**. 1997. Selection for retroviral insertions into regulated genes. J. Virol. **71**:1644-1650.

28. **Harcourt, B. H., A. Tamin, T. G. Ksiazek, P. E. Rollin, L. J. Anderson, W. J. Bellini, and P. A. Rota**. 2000. Molecular characterization of Nipah virus, a newly emergent paramyxovirus. Virology **271**:334-349.

29. **Hyatt, A. D., S. R. Zaki, C. S. Goldsmith, T. G. Wise, and S. G. Hengstberger**. 2001. Ultrastructure of Hendra virus and Nipah virus within cultured cells and host animals. Microbes. Infect. **3**:297-306.

30. **Jin, Y. F., T. Ishibashi, A. Nomoto, and M. Masuda**. 2002. Isolation and analysis of retroviral integration targets by solo long terminal repeat inverse PCR. J. Virol. **76**:5540-5547.

31. **Johansen, L. M., C. D. Deppmann, K. D. Erickson, W. F. Coffin, III, T. M. Thornton, S. E. Humphrey, J. M. Martin, and E. J. Taparowsky**. 2003. EBNA2 and activated Notch induce expression of BATF. J. Virol. **77**:6029-6040.

32. **Markowitz, D., S. Goff, and A. Bank**. 1988. A safe packaging line for gene transfer: separating viral genes on two different plasmids. J. Virol. **62**:1120-1124.

33. **Medico, E., G. Gambarotta, A. Gentile, P. M. Comoglio, and P. Soriano**. 2001. A gene trap vector system for identifying transcriptionally responsive genes. Nat. Biotechnol. **19**:579-582.

34. **Middleton, D. J., H. A. Westbury, C. J. Morrissy, B. M. van der Heide, G. M. Russell, M. A. Braun, and A. D. Hyatt**. 2002. Experimental Nipah virus infection in pigs and cats. J. Comp Pathol. **126**:124-136.

35. **Mounts, A. W., H. Kaur, U. D. Parashar, T. G. Ksiazek, D. Cannon, J. T. Arokiasamy, L. J. Anderson, and M. S. Lye**. 2001. A cohort study of health care workers to assess nosocomial transmissibility of Nipah virus, Malaysia, 1999. J. Infect. Dis. **183**:810-813.

36. **Nadal, A., M. Martell, J. R. Lytle, A. J. Lyons, H. D. Robertson, B. Cabot, J. I. Esteban, R. Esteban, J. Guardia, and J. Gomez**. 2002. Specific cleavage of hepatitis C virus RNA genome by human RNase P. J. Biol. Chem. **277**:30606-30613.

37. **Nor, M. N. M.** 1999. Nipah Disease in Malaysia. O. I. E. Disease Information **12**:67-70.

38. Nor, M. N. M., Gan, C. H., and Ong, B. L. Nipah virus infection of pigs in Peninsular Malaysia. Ministry of Agriculture, Malaysia . 1999.
Ref Type: Generic

39. **Qing, K., C. Mah, J. Hansen, S. Zhou, V. Dwarki, and A. Srivastava**. 1999. Human fibroblast growth factor receptor 1 is a co-receptor for infection by adeno-associated virus 2. Nat. Med. **5**:71-77.

40. **Sambrook, J., E. F. Fritsch, and T. Maniatis**. 1989. Molecular Cloning: A Laboratory Manual 2nd Edition. Cold Spring Harbor Laboratory Press, Plainview, New York.

41. **Schneider-Schaulies, S., J. Schneider-Schaulies, M. Bayer, S. Loffler, and V. ter Meulen**. 1993. Spontaneous and differentiation-dependent regulation of measles virus gene expression in human glial cells. J. Virol. **67**:3375-3383.

42. **Silver, J. and V. Keerikatte**. 1989. Novel use of polymerase chain reaction to amplify cellular DNA adjacent to an integrated provirus. J. Virol. **63**:1924-1928.

43. **Skarnes, W. C., J. E. Moss, S. M. Hurtley, and R. S. Beddington**. 1995. Capturing genes encoding membrane and secreted proteins important for mouse development. Proc. Natl. Acad. Sci. U. S. A **92**:6592-6596.

44. **Soria, M., S. P. Little, and A. S. Huang**. 1974. Characterization of vesicular stomatitis virus nucleocapsids. I. Complementary 40 S RNA molecules in nucleocapsids. Virology **61**:270-280.

45. **Soriano, P., G. Friedrich, and P. Lawinger**. 1991. Promoter interactions in retrovirus vectors introduced into fibroblasts and embryonic stem cells. J. Virol. **65**:2314-2319.

46. **Vazquez, M. I. and M. Esteban**. 1999. Identification of functional domains in the 14-kilodalton envelope protein (A27L) of vaccinia virus. J. Virol. **73**:9098-9109.

47. **Vile, R. G. and S. J. Russell**. 1995. Retroviruses as vectors. Br. Med. Bull. **51**:12-30.

48. **Volckaert-Vervliet, G. and A. Billiau**. 1977. Induction of interferon in human lymphoblastoid cells by Sendai and measles viruses. J. Gen. Virol. **37**:199-203.

49. **Wang, L. F. and B. T. Eaton**. 2001. Henipavirus, p. 641-644. *In* C. A. Tidona and G. Darai (eds.), The Springer Index of Viruses. Springer, New York.

50. **Wang, L. F., M. Yu, E. Hansson, L. I. Pritchard, B. Shiell, W. P. Michalski, and B. T. Eaton**. 2000. The exceptionally large genome of Hendra virus: support

for creation of a new genus within the family Paramyxoviridae. J. Virol. **74**:9972-9979.

51. **Wurst, W., J. Rossant, V. Prideaux, M. Kownacka, A. Joyner, D. P. Hill, F. Guillemot, S. Gasca, D. Cado, A. Auerbach, and .** 1995. A large-scale gene-trap screen for insertional mutations in developmentally regulated genes in mice. Genetics **139**:889-899.