THE ROLE OF REVERSE GENETICS SYSTEMS IN THE STUDY OF FILOVIRUS PATHOGENICITY

BY

Steven Theriault
H.B.Sc., University of Manitoba 1999
M.Sc., University of Manitoba 2001

THESIS

Submitted to the University of Manitoba
In Partial Fulfillment of the
Requirements for the Degree of

Doctor of Philosophy

in

Medical Microbiology and Infectious Diseases

THE UNIVERSITY OF MANITOBA

FACULTY OF GRADUATE STUDIES

*****

COPYRIGHT PERMISSION

THE ROLE OF REVERSE GENETICS SYSTEMS IN THE STUDY OF FILOVIRUS PATHOGENICITY

BY

Steven Theriault

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

Of

DOCTOR OF PHILOSOPHY

Acknowledgements

I would to thank Dr. Heinz Feldmann for his valuable insight, constant patience, and willingness to teach. My thesis would have not been possible otherwise. Special thank to Dr. Strong, Dr. Kobasa, and Dr. Czub for their help during my studies, all of our late night talks paid off. Thanks to my parents and Charlotte who made this thesis possible with their encouragement and constant support. This thesis is dedicated to Chewy, may all the camping and ball throwing he missed be returned ten fold.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

The family *Filoviridae* is comprised of two genera: *Marburgvirus* and *Ebolavirus*. To date minigenome systems have been developed for two *Ebolavirus* species (*Reston ebolavirus* and *Zaire ebolavirus* [ZEBOV]) as well as for *Lake Victoria marburgvirus*, the sole member of the *Marburgvirus* genus. The use of these minigenome systems has helped characterize functions for many viral proteins in both genera as well as having provided valuable insight towards the development of an infectious clone system in the case of ZEBOV. The recent development of two such infectious clone systems, one for ZEBOV and MARV now allow effective strategies for experimental mutagenesis to study the biology and pathogenesis of one of the most lethal human pathogens.

In order to better understand and optimize the reverse genetic system, we studied the relatedness of VP35, VP30, NP and viral polymerase (L) for their role in transcription and replication. We expressed the above mentioned proteins derived from *Reston ebolavirus* and Marburg virus, strain Musoke, using the chicken beta actin promoter. After optimizing the reverse genetic system (nearly 100% rescue), we studied the capacity of heterologous support proteins in virus rescue of *Zaire ebolavirus*, strain Mayinga. This was done by determining the expression of the heterologous protein(s) in the cell compared to virus rescue as determined by the cytopathogenic effect and virus characterization.

We also utilized this genetic system to characterize the gene mutation seen within the ZEBOV guinea pig adapted virus. In characterizing the gene mutations the nuclear protein in combination with VP24 mutations produced a viral variant, which was 100% lethal in guinea pigs. We also demonstrated that signal mutations could not produce lethal

virus genotypes. We also generated a full length ZEBOV guinea pig adapted cDNA construct. This reverse genetic system is 100% lethal in guinea pig and is now a valuable tool in the study of filovirus pathogenicity.

In the development of diagnostic tools to help in the study of filoviruses, we generated a ZEBOV-GFP cDNA construct which expresses GFP in infected cells. Using this newly generated virus we determined the ability of ZEBOV-GFP to be detected *in vivo* and *in vitro*.

THE ROLE OF REVERSE GENETICS SYSTEMS IN THE STUDY OF FILOVIRUS PATHOGENICITY

BY

Steven Theriault
H.B.Sc., University of Manitoba 1999
M.Sc., University of Manitoba 2001

THESIS

Submitted to the University of Manitoba
In Partial Fulfillment of the
Requirements for the Degree of

Doctor of Philosophy

in

Medical Microbiology and Infectious Diseases

THE UNIVERSITY OF MANITOBA

FACULTY OF GRADUATE STUIES

COPYRIGHT PERMISSION

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

Doctor of Philosophy

1. LITERATURE REVIEW

1.1. INTRODUCTION

In early August 1967 the first observed case of filoviral disease emerged in simultaneous outbreaks in both Marburg and Frankfurt, Germany, with later infections coming in Belgrade, Yugoslavia. This virus Marburg (MARV) named for the town where illness was initially observed was linked to exposure to African green monkeys (*Cercopithecus aethiops*). These monkeys were imported form Uganda and a total of 31 patients were affected (103). Since MARV appearance in 1967 only sporadic outbreaks have occurred in Central Africa, with major outbreaks occurring among gold miners in the Democratic Republic of Congo, and in children within Uigie region of Angola in 2004 / 05.

In 1976 a new type of filovirus hemorrhagic fever emerged in two simultaneous outbreaks within the continent of Africa. The isolate viruses from these outbreaks would later be classified as the second species within the family *Filoviridae.* This virus was named ebola (EBOV) after a river within the Democratic Republic of Congo.

One of the first outbreaks caused by the most pathogenic species seen in humans occurred in Northern Zaire within the region of Yambuku. During this outbreak there were 318 cases with a mortality rate of 88 %. At approximately the same time but unrelated, another outbreak occurred in Southern Sudan in which there were 284 cases with a fatality rate of 56 %(1, 2). Following the general nomenclature pattern for naming filovirus subtypes, both viral isolates were named after the regions in which the outbreaks originated. Following the same pattern two other species of EBOV were isolated, one in the Ivory Coast in 1992 in which one human infection was noted (69) and one in Reston

Virginia were imported cynomolgus monkeys from the Philippines were associated with the only human non-pathogenic strain of EBOV viral hemorrhagic fever (84). EBOV viruses cause a severe and often fatal viral hemorrhagic fever in which variable degrees of hemorrhage, marked hepatic involvement, coagulation disorders, and widespread necrosis of internal organs is seen (123). A similar type of fever is characteristic in only four viral families: *Arenaviridae, Bunyaviridae, Flaviviridae*, and *Filoviridae*.

Filoviruses are classified as Risk Group 4 agents; the potential for human to human transmission and the lack of prophylaxis or therapeutic treatment make them a considerable public health concern in both endemic and non-endemic countries. Worldwide concern continues to rise with respect to their potential use as agents of bio-terrorism (15, 134).

Since the first isolation of the Ebola virus scientists have attempted to understand the replication, pathogenic, and evolutionary processes associated with this virus. Although great advances in the understanding of how the virus causes viral hemorrhagic fever have been seen in the past decade, there is still much to uncover.

1.2. TAXONOMY AND NOMENCLATURE

With initial findings using electron microscopy, MARV was proposed as a member of the family *Rhabdoviridae*. With the study of both EBOV and MARV this classification was challenged based on morphological, morphogenetic physiochemical, and biological features. With these data the family *Filoviridae* was established containing a single genus Filovirus (91) Filoviruses are enveloped, non-segmented, negative-sense RNA viruses and constitute a separate family within the order *Mononegavirales*. The family consists of the genera *Marburgvirus* (MARV) and *Ebolavirus* (EBOV). The genus *Ebolavirus* is

further subdivided into four distinct species: *Ivory Coast ebolavirus* (ICEBOV), *Reston ebolavirus* (REBOV), *Sudan ebolavirus* (SEBOV) and *Zaire ebolavirus* (ZEBOV) (42). Filoviral particles are bacilliform in shape, but can also appear as branched, circular, U-shaped, 6-shaped, and long filamentous forms (Figure. 1A). There are many features which distinguish the two genera, including limited antigenic cross reactivity, structural and genome size difference as well as protein expression strategy differences (48, 141)

1.3. Structure of Filovirus Particles

They display a uniform diameter of approximately 80 nm, but vary greatly in length. Negatively contrasted particles, regardless of serotype or host cell, contain an electron-dense central axis (19-25 nm in diameter) surrounded by an outer helical layer (45-50 nm in diameter) with cross-striations at 5 nm intervals. This central core is formed by the RNP complex, which is surrounded by a lipid envelope derived from the host cell plasma membrane. Spikes of approximately 7 nm in diameter and spaced at about 5-10 nm intervals are seen as globular structures on the surface of virions (Figure. 1A) (59, 112, 126). Virus particles are made of seven structural proteins with presumed identical functions for the different viruses. Four proteins make up the RNP complex [nucleoprotein (NP), virion protein (VP) 35, VP30 and RNA-dependent-RNA polymerase (L)] together with the viral RNA, while the remaining three proteins are membrane-associated [glycoprotein (GP), VP40, VP24].

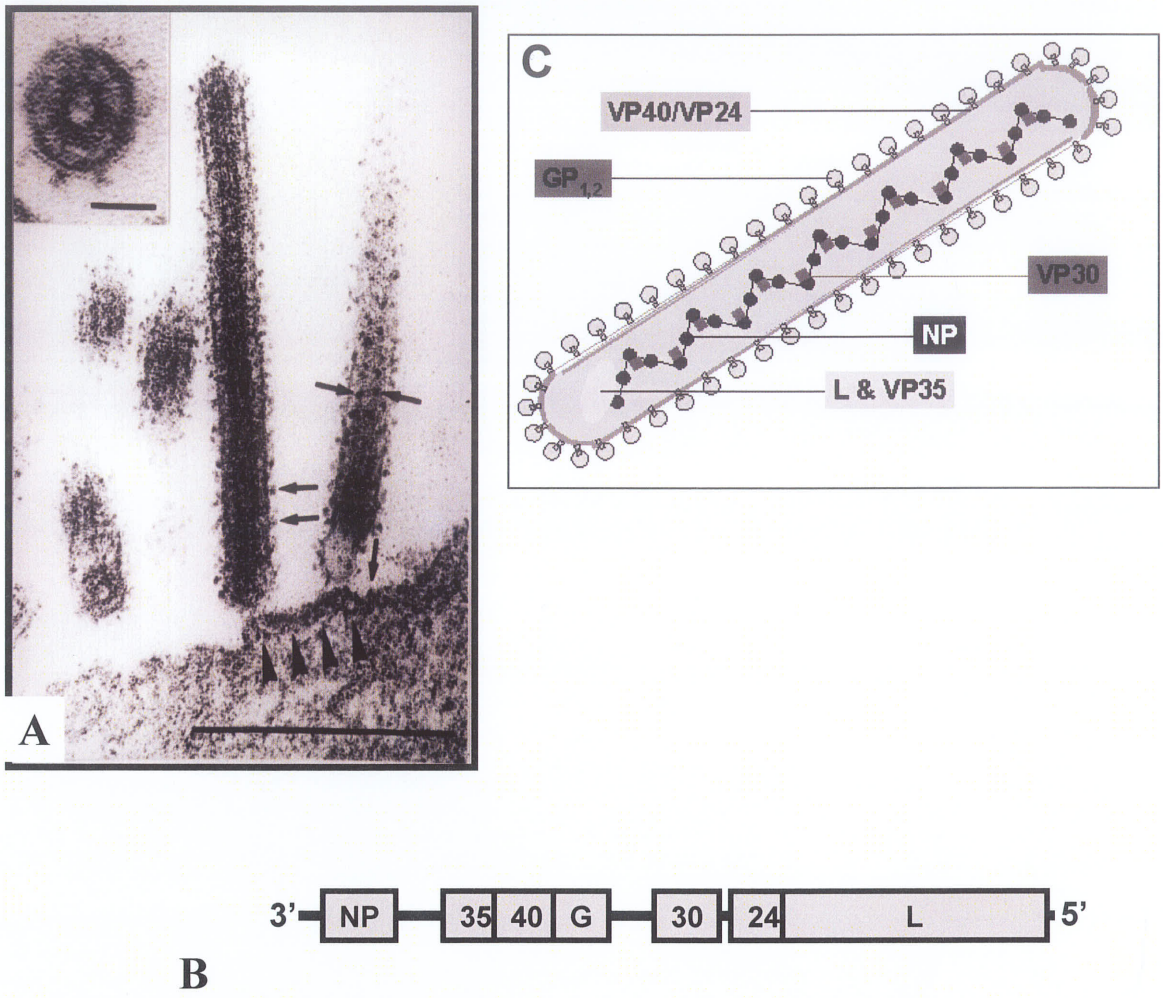


Figure. 1 Structure of filovirus particles. (A) Electron Micrograph. Marburg virus (MARV) particles shown here demonstrate a characteristic bacilliform shape. The electron-dense central axis, formed by the ribonucleoprotein complex and the surrounding lipid envelope are clearly visible. Additionally, the glycoprotein (GP) can be observed as projections on the surface of the particles. [altered from (43)] **(B) Genome organization.** The gene orders of fully sequenced filovirus genomes are presented. The intergenic regions are shown by black bars and the open reading frames in light gray boxes. Open reading frames joined together indicate the positions of the gene overlaps. Key: $GP_{1,2}$ = glycoprotein; L = RNA-dependent RNA polymerase; NP = nucleoprotein; 24, 30, 35, 40 = virion proteins (number indicates the molecular weight in kDa). [altered from (67)]. **(C) Schematic of a filoviral particle.** Representing all seven structural proteins VP40 and VP24 are both found in the outer enveloped membrane as well as the glycoprotein which protrudes out of the membrane. NP, VP30, 35, and L make up the ribonucleoprotein complex.

The single type I transmembrane glycoprotein (GP) is inserted in the envelope as a homotrimer and functions in receptor binding and fusion; VP40 has been identified and characterized as the matrix protein and a main contributor to viral budding. VP24 is still not very well characterized; however, reports have shown that VP24 has a minor role in budding and nucleocapsid assembly. EBOV expresses a nonstructural soluble glycoprotein (sGP) as the primary gene product of the glycoprotein gene, and the delta peptide, a cleavage product from the expression of sGP. sGP is efficiently secreted from infected cells and its functions remains unknown (43, 46, 109, 115, 141, 161, 173).

1.3.1 Filovirus proteins

Outlining the proteins involved in the Ebola viral life cycle (Figure. 1B, C), the first gene product is the nucleoprotein (NP), which is the major structural phosphoprotein in the virion and is required to form stable virion particles. It has been proposed that the highly conserved NH_2-termini may have a role in protein folding and/or RNA binding (7). The COOH-termini of Ebola's NP has been proposed to function in the assembly process by interacting with the matrix proteins (VP40, VP24) or the other nucleoprotein VP30 (37). NP has also been demonstrated to spontaneously form nucleocapsids in 293T cells in conjunction with VP35 and VP24. It was demonstrated that the O-glycosylation and sialation of NP were necessary for the association of all three proteins (79).

The viral structural protein 35 (VP35) appears to exist in both phosphorylated and non-phosphorylated forms. This protein has been shown to bind nucleic acids non-specifically which is consistent with its role as a polymerase cofactor and its localization around the ribonucleoprotein complex (37). Recently, major contributions illustrating a function of VP35 as a type I interferon antagonist has been proposed (12). VP35 was able

to block double-stranded RNA- and virus-mediated induction of an IFN-stimulated response element reporter gene and to block the IFN-beta promoter (12). Further analysis demonstrated that the blocking effect of VP35 was correlated to its ability to inhibit the activation of IRF-3, a cellular transcription factor responsible for the initiation of host cell IFN response (10). This may be an indication that VP35 likely plays a role in virulence by down regulating expression of host antiviral genes namely interferon beta genes.

The viral structural protein 40 (VP40) is not associated with the ribonucleoprotein complex and is located beneath the viral membrane. Recent reports demonstrating VP40s hydrophobic profile, abundance in virion particles, and genomic localization suggests that VP40 is a major contributor to the budding of progeny virions (37, 85). When expressed independently of other viral proteins, VP40 is sufficient to induce release of membrane-bound particles. This has been correlated to interactions with the PPXY motif of VP40 since loss of this motif results in a reduction in particle formation (85). VP40 has also been implicated in the filamentous formation of filoviruses (120) since expression of VP40 alone induces filamentous particles which are morphologically identical to wild-type virus.

The viral structural protein 30 (VP30) is associated with the ribonucleoprotein complex in which it has been proposed to work as a functional unit in encapsidation of the RNA genome (90). With its association with the ribonucleoprotein complex VP30 has been shown to be necessary for replication and transcription, and is considered a minor phosphoprotein with its main phosphorylation sites residing in amino acid region 40-51 (39, 40). It was shown that serines at position 40 and 42 are critical for interactions between NP and VP30 (105). VP30s critical role in EBOV-specific transcription was

demonstrated by mutational analysis of a cluster of four leucine residues, located between amino acids 94 – 112. It was demonstrated that when one of these residues is removed or changed oligomerization was no longer possible resulting in a molecule which was deficient in supporting EBOV-specific transcription (72).

Viral structural protein 24 (VP24) is not well understood although, it is known to be associated with the viral membrane and has a role as a second matrix protein which may bind to the cytoplasmic tail of GP and/or link the other membrane protein, VP40, to the ribonucleoprotein complex (37, 90). Results seen in the development of infectious virus like particles (iVLP) demonstrate this function by implicating VP24 in the formation of a functional ribonucleoprotein complex (77). VP24 has also been reported to be a major contributor to host cell adaptation, as viral mutation within both ZEBOV adapted small animal models carry one or two mutations in VP24 (36, 163).

The glycoprotein (GP) of MARV in contrast to EBOV, only produces $GP_{1,2}$, the predominant products for all EBOV species are the soluble secreted glycoproteins sGP and Δ-peptide, a small carboxyl-terminal peptide generated through cleavage by furin or a furin-like endoprotease from a precursor (pre-sGP) (168, 169). The transmembrane GP, found on the surface of mature EBOV particles, is produced through transcriptional RNA editing (139, 162), while that produced by MARV results from direct transcription and translation of the single open reading frame (ORF), and facilitates receptor binding and fusion with target cells and is associated with host cell cytotoxicity (56, 81, 82, 135, 155, 170, 175). GP is proteolytically processed by furin or a furin-like endoprotease into the cleavage fragments GP_1 and GP_2, which are disulfide linked and form the mature spike protein (140, 164). During processing, $GP_{1,2}$ becomes partially unstable and the non-

membrane-bound fragment GP_1 is released from infected cells (167). Recently, it was shown that another soluble product, $GP_{1,2\Delta TM}$, is produced through metalloprotease cleavage of the membrane-bound mature $GP_{1,2}$ (33). The mature $GP_{1,2}$ is known to form homotrimers on the surface of particles and it is speculated that trimerization is mediated through the GP_2 component of the protein (49, 140). The expression strategy of the glycoprotein gene and the roles of the different expression products have been summarized in detail in several review articles (46, 161).

Volchkov and colleagues (166) investigated the importance of the editing site within the glycoprotein of EBOV. The editing site, which consists of seven consecutive adenosine residues, is located within the GP gene at nucleotide positions 6918 – 6924 (GeneBank accession #AF 272001). Interestingly, approximately 80 % of the glycoprotein gene derived mRNA transcripts in infected cells are not edited and direct the synthesis of the nonstructural glycoproteins sGP and Δ-peptide (139, 162). Both proteins are secreted from EBOV infected cells and sGP has also been detected in blood of EBOV infected patients (139). Using site directed mutagenesis, the seven-adenosine residues (AAAAAAA) were interrupted by adding two guanidines (AAGAAGAA) and an additional adenosine to keep the defined open reading frame such that only $GP_{1,2}$ is produced (166). This construct was rescued and the effects of the deficient editing virus were assayed. It was demonstrated that without editing, effective replication and transcription were unhindered. However, the increased expression of full-length glycoprotein did not simultaneously increase viral release as one might have expected, but revealed a stronger cytopathic effect. It was demonstrated that glycoprotein synthesis was of an immature precursor with high-mannose type sugars, indicating that

glycoprotein transport was largely arrested in the endoplasmic reticulum or in an early Golgi compartment (166). Thus, over-expression of the glycoprotein might lead to cell death by exhausting the processing machinery of the cells. On the other hand, there is evidence that $GP_{1,2}$ displays cytotoxicity by itself which seems to be associated with the transmembrane subunit GP_2 and/or the mucin-like domain found in GP_1 (176). However, cytotoxicity depends on the level of glycoprotein expression and, thus, expression of sGP (non-edited transcripts) seems to control the cytotoxicity associated with the transmembrane glycoprotein $GP_{1,2}$ leading to enhanced virus load and spread in the infected organism.

Neumann and colleagues (114) have studied the importance of the proteolytic processing of the transmembrane glycoprotein precursor (pre-GP) for infectivity of virions. As mentioned above, it had been shown previously that EBOV transmembrane GP is cleaved by a subtilisin-like endoprotease such as furin (167). Interestingly, studies with murine leukemia virus (174) and VSV (82) pseudotyped with mutant ZEBOV GPs lacking a furin recognition site indicated that glycoprotein cleavage was not necessary for infectivity of the pseudotyped viruses. However, for many viruses, posttranslational cleavage of membrane glycoproteins by host proteolytic enzymes, including subtilisin-like proteases such as furin, is a prerequisite for fusion between the viral envelope and cellular membranes and, therefore, is an important step in pathogenesis (92). In the *Orthomyxoviridae* and *Paramyxoviridae* families, glycoprotein cleavage by furin and other host cell proteases is required for their infectivity and thus determines the extent of viral pathogenicity (92). As previously mentioned, MARV and EBOV are proteolytically processed by furin or furin-like proteases at a highly conserved sequence (R-X-K/R-R; X,

any amino acid) (82, 163, 164). Since the glycoprotein amino acid sequence and structural data of REBOV, the least pathogenic of all EBOV species in humans, deviates from the optimal furin recognition sequence (Figure. 2A), glycoprotein cleavage has been thought to be an important determinant of filovirus pathogenicity (46). Using the infectious clone technique, Neumann and colleagues (114) destroyed the furin recognition motif at the predictive cleavage site of ZEBOV by site directed mutagenesis. The resultant rescue of ZEBOV with uncleaved GP indicated that cleavage is not necessary for *in vitro* replication. The GP mutant virus was slightly attenuated *in vivo* showing decreased titers in tissue culture growth, indicating that cleavage may enhance infectivity. The EBOV fusion peptide has an unusual location 28 amino acids from the amino-terminal end of GP_2 (81, 140) (Figure. 2B). It also is flanked by two cysteine residues which are thought to form a disulfide bridge and, thus, expose the fusion peptide in form of a loop (46, 56). This unusual localization and structure might allow sufficient exposure of the fusion peptide to interact with the cellular membrane, even if GP remains uncleaved and thus, could explain the fact/observation that infectivity is largely independent of cleavage (Figure. 2C). However, the appearance and secretion of GP_1 (Figure. 3) is dependent on cleavage of preGP, which might therefore be essential for pathogenicity since soluble GP_1 has been postulated as a pathogenic determinant for filoviruses (46, 47).

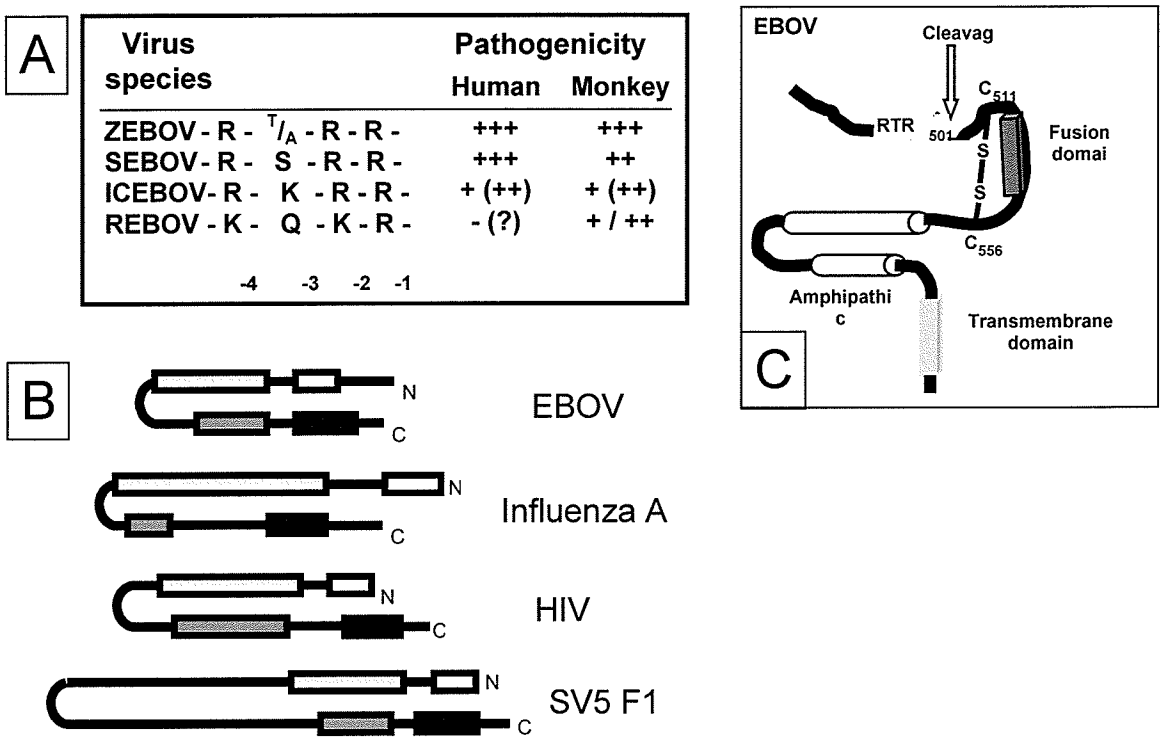


Figure. 2 Structure of Ebola virus GP2 (A) Proteolytic cleavage sites of filovirus glycoproteins. The amino acid sequences of the cleavage sites are presented from positions –1 to –4. Proteolytic cleavage occurs at the carboxy-terminus of the arginine residue at position –1. The relative pathogenicity in human and non-human primates is indicated. **(B) Structural features of fusogenic transmembrane glycoprotein domains.** Structural similarities between EBOV GP_2 and the transmembrane subunits HA_2 of the influenza virus hemagglutinin, gp41 of the HIV env protein and the F1 of the SV5 virus fusion protein are shown. Four domains can be distinguished in the fusion active state: the fusion peptide (a), an amino-terminal helix (b), a carboxy-terminal helix (c) and the membrane anchor (d). The transmembrane proteins assemble into trimers in which the large amino-terminal helices form an interior, parallel coiled-coil, while the smaller carboxy-terminal helices pack in an antiparallel fashion at the surface. Therefore, the fusion peptide and the membrane anchor are located at one end of the rod-like trimers. **(C) Proposed structure of GP2.** The ectodomain of GP_2 contains the fusion peptide followed by an amino-terminal helix, a peptide loop and a carboxy-terminal helix. Helices were proposed by the GARNIER program of PC/GENE (IntelliGenetics Inc.). The fusion peptide is predicted to be exposed on a loop formed by the disulphide linking of cysteine residues 511 and 556.Key: R= arginine, S= serine, K= lysine, Q= glutamine A= alanine T= threonine [altered from (46)]

The large protein, or L protein, is the RNA dependant RNA polymerase which has been shown to carry three common conserved boxes (A, B, and C) among filovirus, paramyxovirus and rhabdovirus L proteins. The highly conserved GDNQ motif located in the COOH-terminal is indicative of other RNA dependant RNA polymerases (147) and the high leucine and isoleucine content; as well as a net positive charge at neutral pH add to the similarities shared among non-segmented negative sense, single-stranded (NNS) RNA viruses (165).

The single negative-sense linear RNA genome of filoviruses does not contain a poly(A) tail and is noninfectious on its own. Upon entry into the cytoplasm of host cells it is transcribed by the viral polymerase to generate polyadenylated sub-genomic mRNA species. Filovirus genomes are approximately 19 kb in length and genes are organized in the following linear order: *3' leader – NP – VP35 - VP40 – GP - VP30 - VP24 – L - 5' trailer* (Figure. 1B). Genes are delineated by transcriptional signals at their 3' and 5' ends that have been identified by their conservation and by sequence analysis of mRNA species. Transcriptional start and stop signals are conserved among filoviruses, and the sequences *3'-CUNCNUNUAAUU-5'* and *3'-UNAUUCUUUUU-5'* represent the consensus start and stop motifs, respectively. Filoviral genes are usually separated from each other by intergenic regions that vary in length and nucleotide composition, but some gene overlap exists at characteristic positions (Figure. 1B).

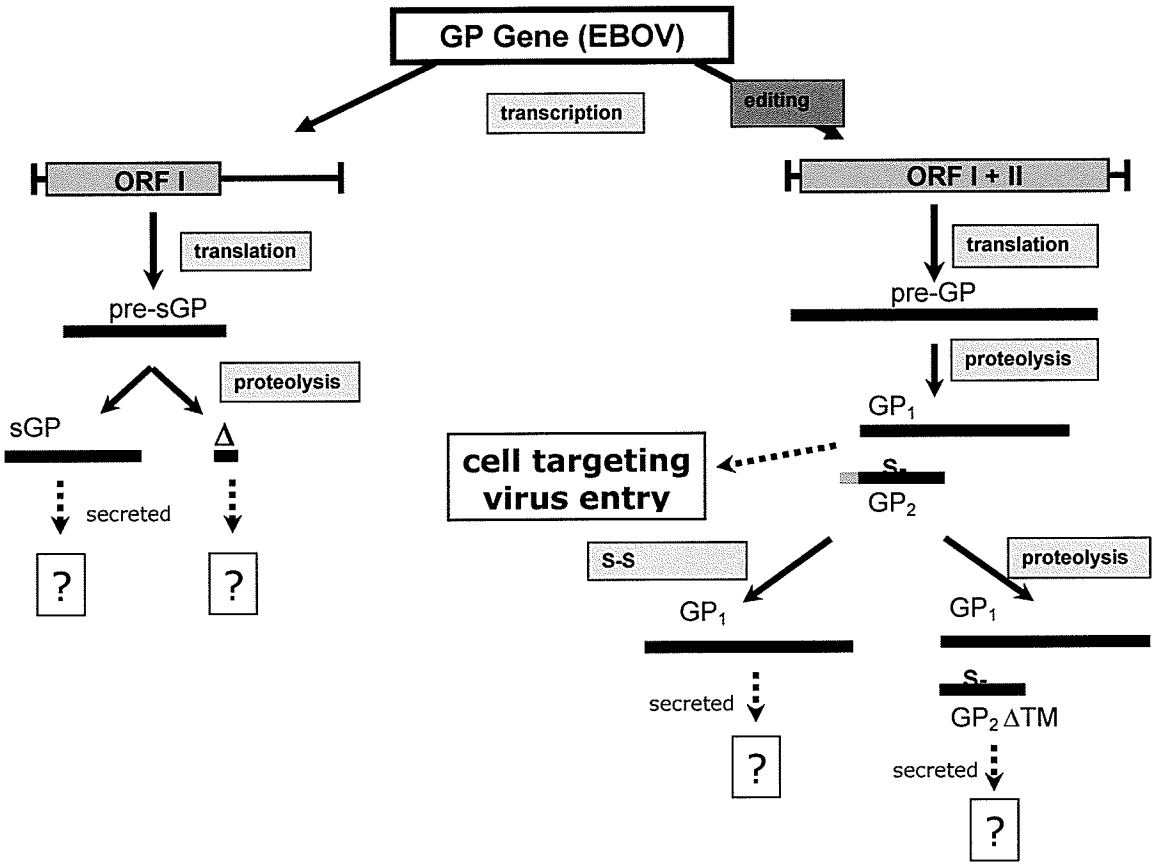


Figure. 3 Filovirus glycoprotein processing. Through a process of transcriptional editing at a series of 7 adenosine residues EBOV shifts its open reading frame and creates an mRNA transcript encoding the precursor of the full-length glycoprotein, pre-GP. This precursor protein is then proteolytically cleaved by furin or furin-like endoprotease into GP_1 and GP_2 which are linked by a disulfide bridge and expressed on the surface of the cell as $GP_{1,2}$. Destabilization of the disulfide bridge leads to release of a soluble form of GP_1. In addition, metalloprotease cleavage produces another soluble form of the glycoprotein, $GP_{1,2\Delta TM}$. Without editing a precursor of a soluble form of GP, pre-sGP, is produced and subsequently cleaved by furin or a furin-like endoprotease into two secreted products sGP and Δ peptide. The function of the secreted protein species is currently unknown, however, the membrane bound full-length $GP_{1,2}$ mediates cell targeting and virus entry. Key: EBOV = *Ebolavirus*; ORF = open reading frame; pre-sGP = precursor from which sGP and Δ peptide are produced by proteolytic cleavage; pre-GP = precursor from which GP_1 and GP_2 are produced by proteolytic cleavage. Note, MARV does not use RNA editing for the expression of preGP and, thus, does not express the soluble glycoproteins sGP and Δ-peptide. [altered from (46)]

The length of the overlaps is limited to five highly conserved nucleotides within the transcriptional signals (3'-*UAAUU*-5') that are found at the internal ends of the conserved sequences. Most genes tend to possess long non-coding sequences at their 3' and/or 5' ends which contribute to the increased length of the genome. Extragenic sequences are found at the 3'-leader and 5'-trailer ends of the genome. The leader and trailer sequences are complementary to each other at the extreme ends; a feature that is shared by many Negative non-segmented RNA viruses (43, 109, 141).

During viral replication within a host cell the viral negative sense genome must be transcribed in order to produce mRNA. Encapsidated RNA acts as a template for the generation of polyadenylated, monocistronic mRNA which is transcribed in a 3' to 5' direction (141). NP mRNA is detectable after 7 hours post infection with a transcription peak detected approximately 11 hours post infection (137). The host cell provides all the necessary components for viral transcription and translation leading to their build up within the cell. Subsequent to the translation of viral proteins there is a switch from transcription to replication that leads to the synthesis and encapsidation of full-length positive-sense RNA. This antigenome copy can serve as a template for the synthesis of full-length genomic RNA that is rapidly Encapsidated by the RNP complex (141). As newly synthesized negative-sense nucleocapsids in the cytoplasm as well as membrane bound proteins (VP24, VP40, and $GP_{1,2}$) accumulate, they amalgamate at the plasma membrane where viral assembly can occur (44, 141) (Figure. 4).

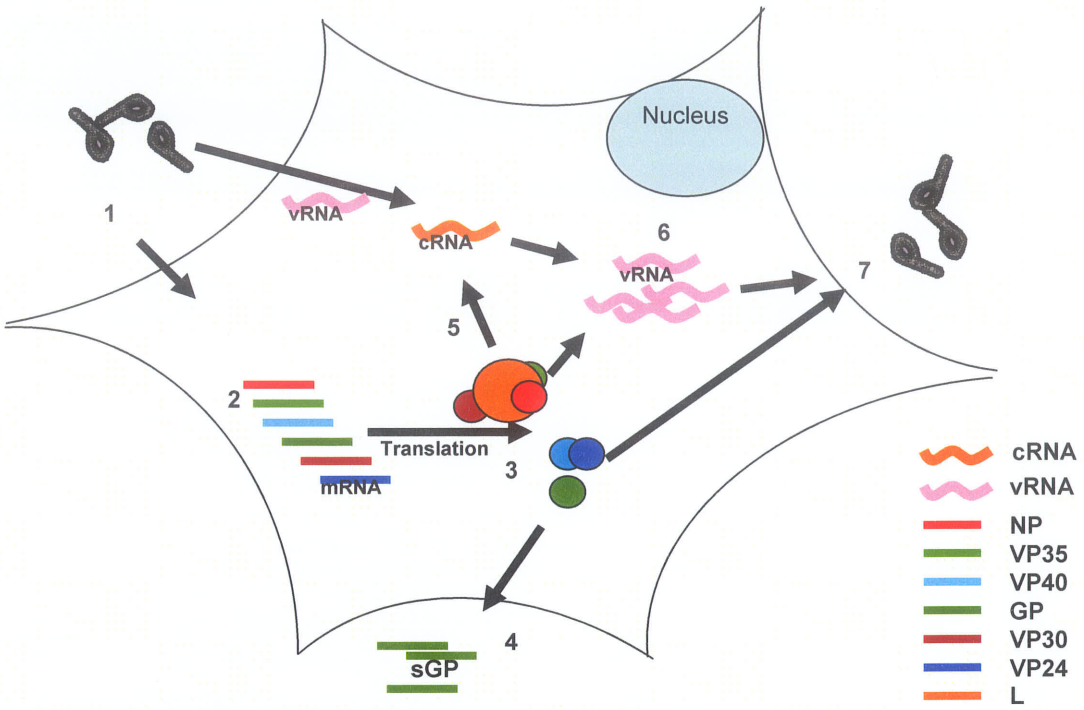


Figure. 4 Filovirus replication cycle in a susceptible cell line. **1.** Viral replication begins with attachment to a host cell receptor. **2.** Viral particles enter the cell and the nucleocapsid is released into the cytoplasm. Primary transcription results in positive sense mRNA transcripts from viral genes. **3.** Translation of the mRNA takes place using host cell machinery, **4.** In the case of EBOV a soluble glycoprotein sGP is secreted. **5.** The viral RNA (vRNA) is replicated with the aid of the viral RNP complex proteins NP, VP30, VP35, and L into a positive sense, complementary RNA (cRNA), anti-gemone. **6.** The cRNA is used as a template to generate progeny vRNA which are encapsided by RNP complex proteins. **7.** In the final step of replication the progeny nucleocapsids are united at the plasma membrane with VP24, VP40, and $GP_{1,2}$ and mature viruses bud from the cell surface.

1.4. Clinical Representation and Pathogenesis

1.4.1 Clinical Representation of EBOV Hemorrhagic Fever

Ebola hemorrhagic fever (EHF) begins with the abrupt onset of fever and malaise, followed by a decrease in blood pressure, which leads to profound shock symptoms and severe coagulation defects. For some patients their Humoral and cell-mediated immune responses develop in time to restrict viral replication and bring about survival, otherwise death occurs in 7-14 days after the onset of symptoms (138). Thus far no antiviral drugs are available for filoviral infections; however nicotin-monosulfate has showed some capability to inhibit viral replication with pre or posttreatment of the drug in nanomolar concentrations (unpublished data). Another treatment which has shown some promise in the treatment of filoviral infections is the administration of a recombinant nematode anticoagulant protein c2 (rNAPc2). This is a potent inhibitor of tissue factor-initiated blood coagulation and has shown some ability to decrease viral pathogenesis during an infection. Post-exposure protection with rNAPc2 against Ebola virus in primates provides a new foundation for therapeutic regimens that target the disease process rather than viral replication. Since human clinical studies within the filoviral field has only yielded fragmented and often-contradictory information. I will focus mainly on laboratory studies of the uniformly lethal infection caused by ZEBOV in cynomolgus and rehesus macaques. Fatal illness in human cases include fever, high circulating viral load, a marked rise in blood neutrophil count and a fall in lymphocytes and platelets, hypotension and shock, coagulopathy and hemorrhage, and biochemical alterations suggestive of massive lymophocyte apoptosis (4, 26, 138, 160). All of these symptoms have also been demonstrated in a filoviral infection within macaques. The coagulopathy

in macaques conforms to the definition of disseminated intravascular coagulation (DIC) (58).

1.4.2 Overview of Pathogenesis:

Filoviruses are able to infect a broad range of primate cells, partly due to their heavily glycosylated surface glycoprotein which can bind to a variety of target molecules including cell surface lectins (154). With replication of the virus resulting in necrosis of infected cells. Studies have demonstrated that two cell types macrophages and dendritic cells are the major targets of filoviruses. Interestingly, both of these cell types elicit specific immune responses which when impaired allow for the systemic spread of the virus. Both cell types are only partially impaired to allow the initiation of the inflammatory and coagulation responses, which will bring more macrophages and dendritic cells to the infected area (Figure. 5). This consequence results in the dissemination of virus to reside in macrophages and dendritic cells in tissue throughout the body, causing massive release of proinflammatory mediators and vasoactive substances (75). These host responses produce a syndrome of refractory hypertension and DIC resembling septic shock (22). The extensive tissue damage caused by the replication of ZEBOV in macrophages and dendritic cells as well as in parenchymal cells of the liver and other organs also plays a role in fatal disease (Figure. 5). Natural killer cells and T lymphocytes remain uninfected, but undergo apoptosis, further impairing the immune function allowing for viral replication (130). Macrophages play a central role in inducing the hypotension and shock of EHF, the binding of double-stranded RNA as well as other viral products, result in the release of proinflammatory cytokines, such as TNF-α and IL-1β, chemokines, such as MIP-1α, and nitric oxide (NO) (70, 75, 153). These

mediators elicit help and attract additional monocytes and macrophages to the site of infection causing vasodilatation, increased endothelial permeability and expression of endothelial cell-surface adhesion molecules. This occurrence of vascular changes throughout the body as a result of the systemic spread of ZEBOV leads to catastrophic circulatory collapse (22, 61, 102). Macrophages which have been infected by EBOV also play an important role in initiating DIC by synthesizing cell-surface tissue factor which can stimulate the extrinsic coagulation pathway leading to deposition of fibrin on the surface of infected cells (62). Thrombocytopenia does not become evident until day 3-4, as platelets attach to activated endothelium or become part of nascent thrombi.

The ability of ZEBOV to disseminate rapidly from its site of entry suggests that infected cells are unable to produce sufficient amounts of interferon (INF)-α/β or respond adequately to exogenous types I of II IFN. It has been suggested that VP35 of ZEBOV blocks IFN production by virus-infected cells by preventing the recognition of dsRNA that normally leads to phosphorylation of IRF-3 (73). Interestingly, VP24 may contribute to this process by blocking responses to exogenous IFN (11) These inhibitions would profoundly impair the anti-viral response, since types I and II IFN are needed to activate NK cells, assist with the adaptive immunity through stimulation of major histocompatibility complex and activate macrophages and dendritic cells for effective anti-microbial function.

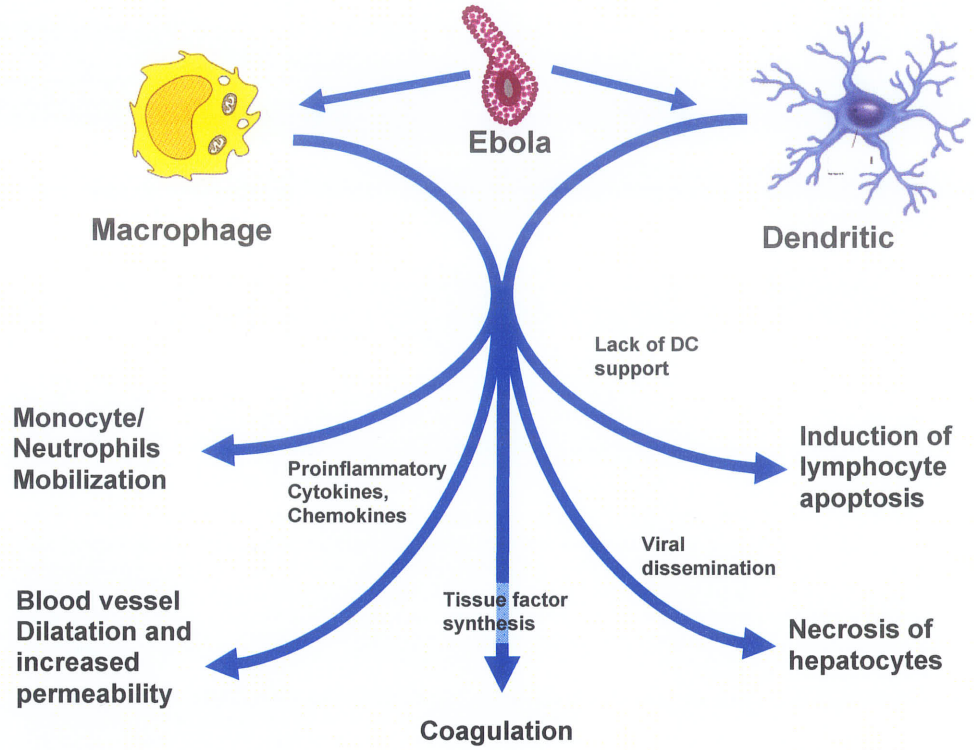


Figure. 5 ZEBOV infection in macrophages and dendritic cells. Macrophages and dendritic cells infected with ZEBOV induce the role of many clinical factors which represent an ebola infection. Secreted cytokines, chemokines and other mediators alter blood vessel function and elicit an influx of inflammatory cells, including additional monocytes and macrophages. Synthesis of cell-surface tissue factor stimulates systemic coagulopathy. Release of virus form infected cells spreads through out the body infecting similar cell types. This causes infected parenchymal cells in many organs to develop multifocal tissue necrosis. This type of infection also decreases the host's ability to develop an effective adaptive immune response by depleting lymphocytes action. [Modified from (21)]

1.4.3 Transmission of EBOV virus:

Human outbreaks of filoviral hemorrhagic fever are propagated via person to person contact, usually in a nosicomial fashion (141). Health care providers are at greatest risk with 20% of the total infections during the 1995 ZEBOV Kikwit being primary care givers (76), (89). Transmission though the aerosol route still remains some what controversial. Virus has been experimentally transmitted in non-human primates by the aerosol route with virions being isolated in the lung alveoli of infected monkeys, however the actual role of this transmission during a human outbreak has never been proven (60).

1.5. INFECTIOUS ANIMAL MODLES FOR FILOVIRUS

1.5.1 Introduction and evidence of filoviral host: Since the first ebola out-break in 1976 the wild reservoir has remained a mystery. Recently a publication in nature has shown evidence that fruit bats may be the elusive host. Three fruit bat species *Hypsignathus monstrosus, Epomops franqueti*, and *Myonycteris torquata* are suspect to carrying ebola virus like RNA. Further evidence supporting fruit bats as the possible host is illustrated by migration patterns which follow most regions of sporadic filoviral infections within central Africa. One troubling result which should be address is the in ability of researchers to isolate live infectious ebola virus from these bat species. However, the virus may need an outside stimulus to cause viral replication and virion production to take place (98).

1.5.2 In Mice: Adult immunocompetent mice are resistant to filovirus infection. Serial passageing of an Ebola Zaire '76 virus variant initially inoculated in suckling mice

and then transferred to progressively older mice resulted in a mouse adapted virus variant. This adapted virus which was developed in BALB/C mice is lethal for adult SCID and adult BALB/C mice, as well as C57BL/6 inbred, and ICR (CD-1) mice. However, the induced disease differs from the human, primate and guinea pig infection models (20).

Adult immunocompetent mice infected with adapted virus develop ruffled fur, progressive lethargy and weight loss. Death occurs around day six to eight post-infection. Hemorrhagic manifestations like bleeding from the orbits, bladder, gastrointestinal tract, and abdominal cavity before death can also be observed in some infected animals. Mice previously inoculated with one plague forming unit (PFU) of non-adapted Zaire Ebola virus are protected against challenge with adapted Zaire Ebola virus, whereas it has been demonstrated that heat-inactivated or irradiated virus does not induce immunity (19).

There is currently a reverse genetic system which was adapted from the original ZEBOV reverse genetic system (114) for the mouse adapted virus (36).

1.5.3 Guinea Pig infection: When experimentally infected with ZEBOV Dunkin-Hartley guinea pigs develop only a mild febrile disease, and most survive the infection. Serial passaging of the virus in guinea pigs leads to virus adaptation. The incubation time decreases, and the disease becomes more severe from passage to passage. By the fourth passage every newly infected guinea pig dies (17).

The developing leukocytosis is due to increasing neutrophilia. Concomitant absolute lymphopenia and thrombocytopenia develops as seen in other animal models and during a EHF infections. Hemoglobin, hematocrit, and erythrocyte counts remain normal. Serum alkaline phosphatase activity levels rise sharply in the terminal stage of

disease. Blood urea nitrogen and serum creatinine levels also rise with progression of the animal's disease to death (28). Since the guinea pig adapted model demonstrated similarities to a human EHF infections the development of more advanced methodologies and experimental procedures are needed to full utilize this small animal model.

1.5.4 In Primates: After the initial outbreak in 1967 of Marburg, researchers rapidly learned that Marburg virus caused severe lethal infection in nonhuman primates that resembled a human infection (146, 150). It has been demonstrated that experimental infection with ZEBOV caused illness in vervet monkeys (Cercopithecus aethiops), rhesus macaques (Macaca mulatta), the common squirrel monkeys (Saimiri sciureus), and Cynomolgus macaques (Macaca fascimlans) and proved to be 100% fatal between days seven and nine post-infection. All monkeys developed a febrile illness independent of dose or route of inoculation (8, 17, 50, 51).

1.5.5 Viral course in humans and non-human primates: Since primates and humans are only dead end host for EBOV infection, there has been no opportunity for the evolution of effective defenses mechanism against the filoviruses; this is demonstrated by the poor immune response or in some cases a damaging immune response. Since much experimentation has been carried out using macaques, they provide us with an excellent 'worse case scenario' as only a very small dose of ZEBOV causes uniform lethality (83, 87). In contrast to some human infections which begin to show clinical improvements during the second week of illness and ultimately recover from their infection. Survival seems to depend on the development of antigen-specific immune response (4, 138).

The different disease course may reflect differences seem in the host response to a ZEBOV infection. For example, ZEBOV infection, of macaques results in a continuing

increase in circulating proinflammatory cytokines over the course of illness, in the absence of anti-inflammatory mediators, such as IL-10, while blood samples from human cases have shown the presence of both proinflammatory cytokines such as TNF-α and IL-6, and anti-inflammatory mediators, such as IL-10 and IL-1β receptor antagonist (4, 5). As seen in cases of bacterial septic shock, fatal infection of humans appears to be associated with an elevation of anti-proinflammatory cytokines. This suggests that the balance and timing of early responses may play a critical role in determining disease outcome (22).

1.6. REVERSE GENETICS

Reverse genetic systems can be broadly grouped into two categories: minigenome and infectious clone systems. Both utilize cloned cDNA to either mediate expression of reporter genes (minigenome systems) or produce infectious virus (infectious clone systems). Together they provide excellent tools for studying replication and transcription as well as infectivity and pathogenicity. The first reverse genetic systems were established for positive-sense, single stranded RNA viruses (127, 156) where transfection of the full-length genomic RNA transcripts into eukaryotic cells resulted in viral protein expression, viral replication, particle formation and release. The development of reverse genetic systems for these viruses was favored by the fact that the genomic RNA of positive-sense RNA viruses can directly serve as the template for the expression of viral proteins through the cellular machinery. In contrast, negative-sense RNA viruses first need to transcribe their genomes into positive-sense RNA prior to translation of viral proteins; this step that is dependent on the presence of a functional viral replicase

complex, since cells lack the necessary enzymes to mediate (-)RNA → (+)RNA synthesis.

In the past decade, several reverse genetic systems have been developed for negative-sense RNA viruses (117) with the establishment of a minigenome system generally preceding the development of the infectious clone system, although this is not always the case. Reverse genetics systems have been developed for representatives of the negative-strand RNA virus families *Orthomyxoviridae*, *Bunyaviridae*, *Arenaviridae*, *Rhabdoviridae*, *Paramyxoviridae*, *Filoviridae* and *Bornaviridae* (29, 30, 34, 54, 65, 81, 95, 110, 111, 122, 131, 132) either in the form of minigenome systems and/or infectious clone systems. In each case the technology used reflects both the particular requirements of the virus as well as the availability of established methodologies.

The first negative-sense RNA virus minigenome system was developed by Palese and colleagues in 1989 (101), in which they modified the influenza A virus by the addition of a reporter chloramphenicol acetyltransferase (CAT) gene, cloned between the 5' and 3' non-coding viral RNA segment sequences. The reporter gene construct was flanked by a promoter region for the T7 RNA polymerase and a restriction enzyme recognition site, which allowed for the formation of authentic viral 3' ends. Following runoff *in vitro* transcription of the viral-like RNA and the addition of purified polymerase and nucleoprotein, a reconstituted ribonucleoprotein (RNP) complex was produced. Subsequent transfection of the RNP complexes and infection with helper influenza virus was undertaken and a virus containing the virus-like RNA encoding CAT, in addition to the other influenza vRNAs, was generated. Although these experiments demonstrated the

ability to generate modified infectious viruses, the particular system used requires a selection system to distinguish modified viruses from helper viruses (101).

The generation of recombinant rabies virus by Conzelmann and colleagues in 1994 (143) demonstrated for the first time that generating a non-segmented negative-sense single-stranded (NNS) RNA virus entirely from a cDNA was possible. The cDNA encoding the full-length positive-sense genome of rabies virus along with the nucleoprotein, phosphoprotein and RNA-dependent RNA polymerase, all under the control of a T7 RNA polymerase promoter, were transfected into eukaryotic cells. Infecting the transfected cells with recombinant vaccinia virus provided the T7 polymerase. This study helped initiate the development of other genetic systems for members of *Mononegavirales*, including other members of *Rhabdoviridae* (29, 30, 95, 122, 131, 132) as well as *Paramyxoviridae* (6, 24, 27, 57, 74, 121, 129), which were based on T7 polymerase supplied by infection with 'modified vaccinia virus Ankara' (MVA-T7). In addition, more effective methods of providing the T7 RNA polymerase were also investigated and included the use of plasmids expressing the protein (97, 114), expression of the polymerase from a stably transfected cell line (24, 129), and a heat shock method, which was shown to increase rescue efficiencies (121). The utility of these alternative methods of supplying T7 RNA polymerase were illustrated by the development of infectious clone systems for filoviruses. The first system developed for *Zaire ebolavirus* by Volchkov et al. in 2001 (166) was based on transfection of T7-driven plasmids encoding the genome as well as the nucleoprotein (NP), virion protein (VP) 35, VP30 and RNA-dependent RNA polymerase (L) into BSR T7/5 cells, which stably expressed the T7 polymerase. Shortly thereafter, Neumann and colleagues (114)

demonstrated that T7 could be supplied via an additional plasmid, which encodes the T7 polymerase under the control of a eukaryotic promoter. Recently an infectious clone system for Borna disease virus has also been generated in which it was illustrated that cDNA constructs carrying the published genome sequence are functional but that the rescued viruses are strongly attenuated, and that regulatory sequences of the viral genome determine virulence (142). This illustrates a possibility of developing an attenuated ZEBOV virus using the reverse genetic system, although the usage of such a virus would most certainly be questioned!

Infectious clone systems for segmented RNA viruses, although more complex, saw a breakthrough in 1996 when Bridgen and Elliott showed that the segmented Bunyamwera virus could be rescued using three anti-genome encoding plasmids in addition to viral protein expression plasmids (23). This was the first rescue of a segmented negative-sense virus solely from cDNA. The establishment of minigenome and infectious clone system for influenza virus based on this technology followed these results. However, this system needed to be modified to deliver the vRNAs to the nucleus of transfected cells, where influenza virus replication naturally occurs. This was overcome by the establishment of an RNA polymerase I based system (55, 116, 118). In the case of the infectious clone system the optimized system required eight RNA polymerase I driven plasmids encoding the eight vRNA segments, in addition to four RNA polymerase II driven plasmids encoding the polymerase components (PA, PB1 and PB2) and the nucleoprotein (55, 116). Recently Hoffmann and Webster have modified the RNA polymerase system allowing both negative vRNA and positive-sense mRNA transcripts to be synthesized from the same template, and thereby decreasing the plasmids

required for influenza rescue to eight (78). Reverse genetics systems (minigenomes) have been published for other bunyaviruses such as Toscana and Rift Valley fever virus (genus *Phlebovirus*) (3, 100) and, more recently, Hantaan virus (52) and Crimean-Congo hemorrhagic fever (CCHF) virus (53), members of the genera *Hantavirus* and *Nairovirus*, respectively. While no infectious clone system has been developed to date for members of the family *Arenaviridae*, the recent development of a minigenome system for lymphocytic choriomengitis virus (LCMV) (96) and Tacaribe virus (99) is promising for the development of such a system in the near future.

1.7. MINIGENOME SYSTEMS FOR FILOVIRUSES

In the case of filoviruses, minigenome systems for MARV (strain Musoke) and ZEBOV (strain Mayinga) were developed based on the T7 RNA polymerase to synthesize negative-sense vRNA transcripts from cDNA (110, 111). Initially, the cDNA constructs contained the leader 3', the non-coding region of the NP gene, the 5' non-coding region of the L gene and the 5' trailer sequences of the genome flanking the single reporter gene CAT (Figure. 6). These minigenome systems were driven either by a helper virus infection or transfection of plasmid DNA encoding the RNP complex proteins to provide the necessary machinery for transcription and replication. In the past 5 years, both systems have allowed the study of different aspects of filovirus transcription and replication.

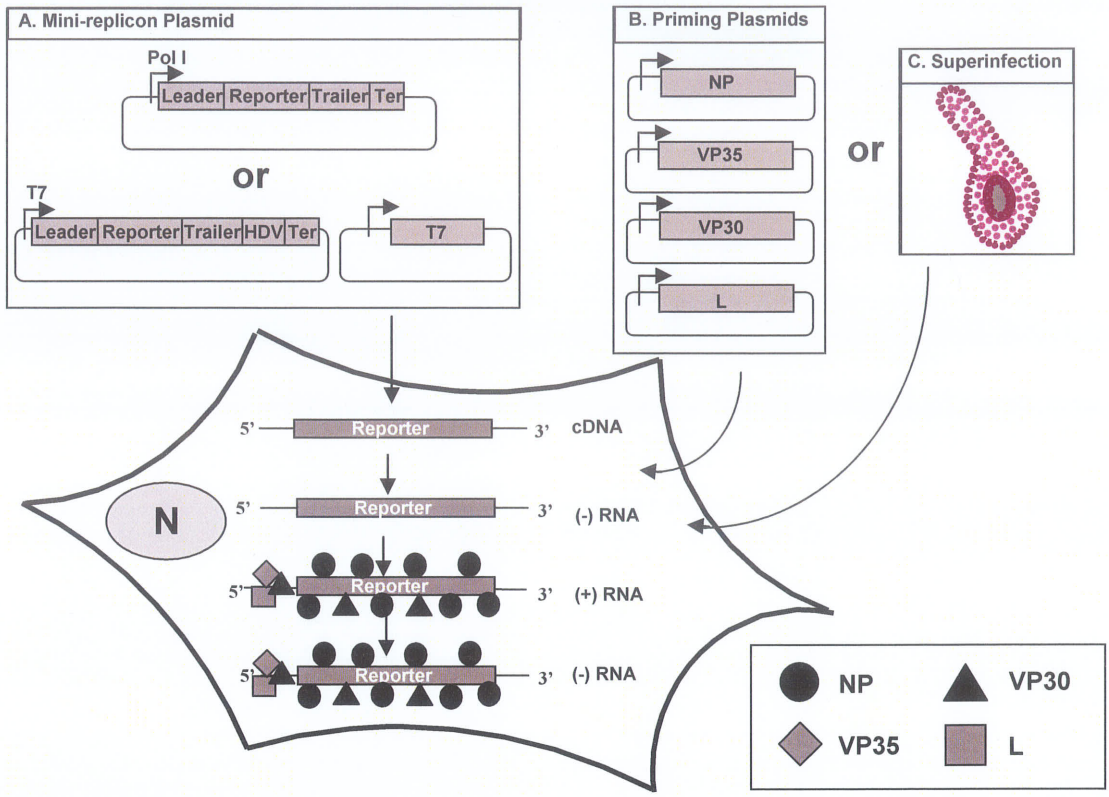


Figure. 6 Transcription and replication steps in a minigenome system. To examine replication and transcription in a minigenome system, cells have to be transfected with a minigenome plasmid (A) that contains a reporter gene flanked by the genomic leader and trailer regions under control of either a T7 or a Pol I promoter. The T7-driven minigenome contains an additional hepatitis delta virus ribozyme sequence, which results in transcript cleavage to generate and authentic genome end. In the case the T7-driven constructs, T7 polymerase has to be present in the cells either by transfecting a plasmid coding for it, infection with a recombinantly T7 expressing vaccinia virus, or by using cell lines that express this protein constitutively. Once generated further transcription and replication of viral RNA-like species can be driven either by helper plasmid encoded ribonucleoprotein (RNP) complex components [nucleoprotein (NP), virion protein (VP) 35, VP30, RNA-dependent RNA polymerase (L)] (B) or by infection with live virus (helper virus) (C). The transfected cDNA is transcribed into a virus-like vRNA in negative orientation by either the T7 or the Pol I polymerase. Subsequently, this vRNA is replicated by the proteins of the viral RNP complex into cRNAs in positive and vRNAs in negative orientation. Transcription of reporter mRNA transcripts from the vRNA-like minigenomes by the RNP complex proteins and subsequent translation leads to expression of the reporter gene. Key: HDV = hepatitis delta virus ribozyme; L = RNA-dependent RNA polymerase; N = nucleus; NP = nucleoprotein; Pol I = RNA polymerase I; T7 = bacteriophage T7 RNA polymerase; Ter = terminator; 24, 30, 35, 40 = virion proteins (number indicates the molecular weight in kDa).

Using the minigenome systems it was established that only three of the four-nucleocapsid proteins, NP, VP35 and L, were necessary to support replication and transcription of the monocistronic MARV minigenomes (110). This is in agreement with data obtained from various paramyxovirus and rhabdovirus systems where others had determined that the nucleoprotein (N), phosphoprotein (P) and the RNA-dependent RNA polymerase (L) are the minimum proteins required for replication (31, 35, 68, 88, 128, 151, 178). Minigenome systems developed for members of the family *Bunyaviridae* (tri-segmented negative-sense RNA viruses) could be transcribed and replicated using the NP and L protein (34, 52-54, 100).

The minigenome system for ZEBOV, which was subsequently generated using the same strategy as for the MARV system, required all four of the nucleocapsid proteins NP, VP35, VP30 and L for efficient replication and transcription of the monocistronic minigenomes (111). It was further shown utilizing the plasmid based minigenome system that VP30 of ZEBOV could efficiently enhance transcription of ZEBOV minigenome (106). The presence of VP30 might resolve or cover RNA secondary structures either by RNA binding or by directing an additional co-factor to the folding RNA. However, thus far, RNA binding activity for VP30 has not been described, and since the ZEBOV genome is bound by NP, secondary structure formation may not occur. The only naked RNA species present are the positive-sense mRNA transcripts which have been shown to form secondary structures. Thus, it could be possible for VP30 to have an effect on these secondary structures at the mRNA level (171). Our knowledge of VP30 was extended when VP30 was found to contain two N-terminal serine clusters, which positively regulated the binding of VP30 to NP, and in doing so negatively regulated the

transcription activation function of VP30. It was also shown that VP30 is a target for cellular protein phosphatases PP1 and PP2A. In a reconstituted minigenome system, ZEBOV specific transcripts were blocked by okadaic acid, which is known to inhibit PPI and PP2A. Treatment of ZEBOV infected cells with okadaic acid also inhibited ZEBOV growth, which could be compensated for by the expression of a non-phosphorylated VP30 *in trans* (106). All these results taken together illustrate that VP30 phosphorylation is a regulatory factor in the replication cycle of ZEBOV.

1.8. Minigenome Systems for other Mononegavirales

A literature search revealed that several attempts have been made to investigate whether non-segmented negative-sense, single strand RNA virus replication complexes were able to recognize heterologous RNA templates *in vivo*. It was reported that human parainfluenza virus (hPIV) type 1 and type 3 could accept a Sendai virus minigenome as a template for replication, whereas measles virus could not. The same rescue results were also seen when others utilized a plasmid-based artificial replication system (124). However, the rescue of the hPIV type 3 minigenome could not be supported by respiratory syncytial virus (RSV) or, unexpectedly, by bovine PIV type 3 (32). For Toscana and Rift Valley fever virus (genus *Phlebovirus*, family *Bunyaviridae*) it was demonstrated that the transcription complexes were active on heterologous template (3), and for VSV it was shown that replication of defective interfering particle RNAs from serotypes New Jersey and Indiana was possible but only when the replication complex was supplied by VSV Indiana (108). These data illustrate that the specificity of the replicase complex to the target sequences is not absolute and depends on the virus

system. In general, MARV/ZEBOV minigenome systems supported transcription/replication of homologous but not heterologous RNA templates, regardless if helper virus or transfected nucleocapsid complex protein expression plasmids were used for transcription/replication of the minigenome system. As an exception, MARV VP30 could replace ZEBOV VP30 in the ZEBOV minigenome system, although this switch did result in lower activity than the native VP30 ZEBOV protein (111). However, a chimeric minigenome system containing the ZEBOV leader and the MARV trailer was shown to be encapsidated, replicated, transcribed, and packaged by both viruses (111). Recently, our group has developed a minigenome system for REBOV (64) employing an alternative to the classical T7-driven approach used with MARV and ZEBOV. This system was based on an initial transcription step mediated by RNA polymerase (Pol) I, an endogenous host polymerase. Thus, minigenome transcription by Pol I eliminates the need to introduce a source of the polymerase into mammalian cells (see Table 1 for a comparison of the respective polymerase properties). While introduction of T7 into mammalian cells can be achieved in a number of ways, the need to do so presents a potential limitation if the entire population of cells is not targeted. Minigenome transcription by Pol I also overcomes a number of limitations of the T7-driven system relating to the production of authentic, non-modified transcripts which have correct sequences at their termini (Table 1).

Table 1:

	T7 RNA Polymerase	RNA Polymerase I
Origin	Bacteriophage	Eukaryotic
Methods of Introduction	MVA-T7 infection, Transient or stable transfection	N/A
Localization	Cytoplamic	Nuclear
mRNA modification	5'capping and 3' poly A*	No
Initiation/Termination	Addition 5' and 3' nucleotides	No additional nucleotides

Table 1: Comparison of polymerase properties. Characteristics of the bacteriophage T7 RNA polymerase and the eukaryotic RNA polymerase I with respect to origin, localization and transcriptional properties are outlined. Key: MVA-T7 = modified vaccina virus Ankara recombinantly expressing the bacteriophage T7 polymerase.

Despite the exclusively cytoplasmic replication of filoviruses, the nucleolar localization of Pol I within host cells does not appear to present a barrier to successful rescue of high levels of reporter activity from this system as has been shown previously with several bunyavirus systems (52-54). To the contrary, our data indicate that a Pol I-driven REBOV minigenome generates both a higher level of reporter activity and a higher signal-to-noise ratio than did a comparable T7-driven REBOV construct and, thus, this system seems to help overcome some of the technical limitations of T7-mediated transcription (64). The successful establishment of a REBOV minigenome system is of particular interest as a tool to study transcription and replication of this virus in comparison to other more pathogenic filoviruses.

1.9. Infectious Clone Systems for Filoviruses

As noted earlier, minigenome systems are the building blocks for an infectious clone. These systems utilize the same principles as a minigenome system, but rather than a

reporter gene being expressed, the full viral genome is replicated and transcribed producing live infectious, fully functional viruses (Figure. 7). Two such reverse genetic systems have been developed for ZEBOV. The first system developed by Volchkov and colleagues, utilized a cell line which stably expressed the T7 polymerase (BSR T7/5) (166). The T7 polymerase drives the transcription of the cDNA copy of ZEBOV producing a negative-sense RNA molecule. This RNA species can be used for the replication of an antigenomic template (positive-sense RNA) producing both RNA species used in viral transcription and replication. Neumann and colleagues (114) developed a ZEBOV reverse genetic system, which successfully utilized a plasmid driven T7-RNA polymerase rather than the BSR T7/5 cell line (Figure. 7). These two methods are both sufficient to provide the T7 RNA polymerase. For the development of previous systems, the T7 RNA polymerase has been commonly provided by infection with a recombinant vaccinia virus (113, 117). However, this system has the disadvantage of requiring separation of the recombinant viruses of interest from progeny of the recombinant vaccinia virus. Recently we have optimized the infectious clone system developed by Neumann and colleagues (114) to a rescuability of nearly 100 % (158). This system can now be more reliably used for the generation and analysis of mutants, particularly if rescues are unsuccessful due to incompatibility of the mutations with virus replication.

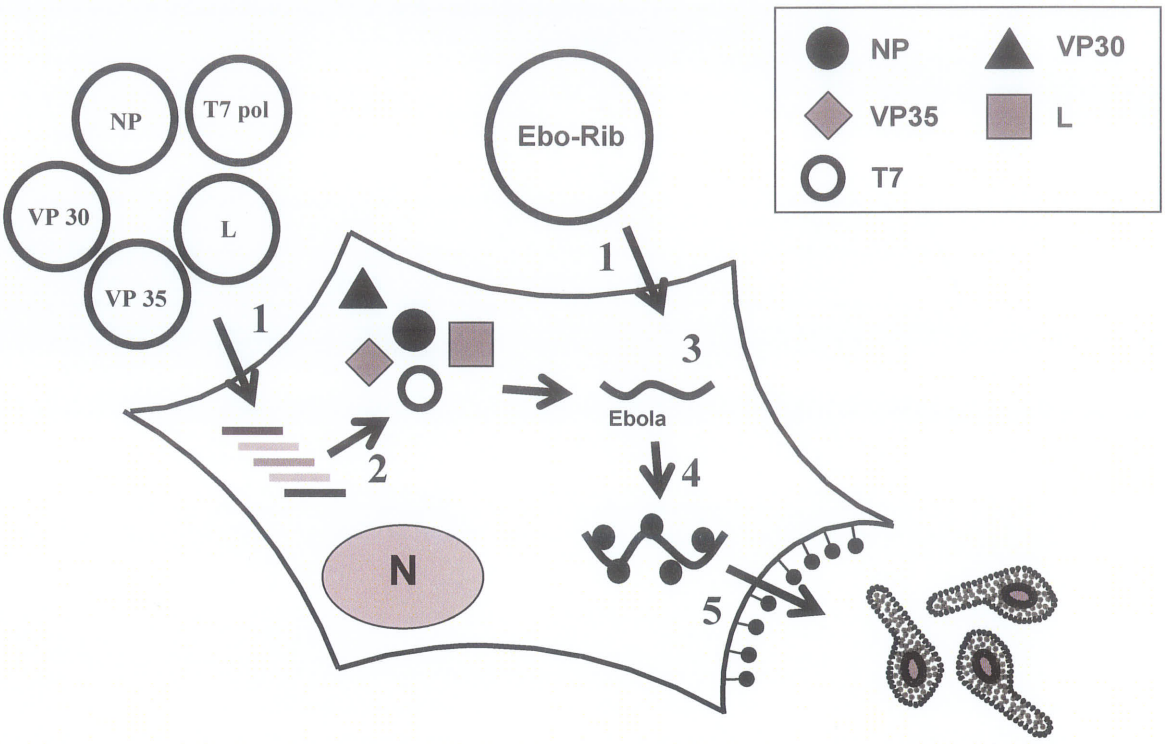


Figure. 7 *Zaire ebolavirus* infectious clone system. The scheme illustrates the components of the system and the steps involved in the rescue of infectious virus. (1) Co-transfection of the plasmid carrying the full-length ZEBOV genome and the expression plasmids for the bacteriophage T7 RNA polymerase (T7 Pol) and the four ZEBOV proteins associated in the ribonucleoprotein complex (L, NP, VP30, VP35); (2) – expression of the viral support proteins and the bacteriophage T7 RNA polymerase under the control of the chicken β-actin promoter; (3) – transcription of the ZEBOV genome under the control of the bacteriophage T7 RNA polymerase promoter; (4) – formation of the ribonucleoprotein complex, transcription and replication; (5) – virus maturation at the plasma membrane and subsequent budding of infectious virus particles. *Key*: L = RNA-dependent RNA polymerase; N = nucleus; NP = nucleoprotein, VP = virion protein 30 and 35 kDa. [altered from (158)]

The ZEBOV infectious clone systems have been used in the past to address questions regarding the pathogenic potential of the transmembrane glycoprotein (GP), which is encoded by gene 4 of the linear arranged genome (Figure. 3).

1.10. Past and Future Challenges using a reverse genetic system

Major problems with the infectious clone systems relate to the handling of the larger plasmids such as the genomic plasmid and the plasmid encoding the L protein. One common occurrence is spontaneous mutations, which have occurred in the generation process of the two existing systems. The first system (166) carried a single mutation in the genomic clone at nucleotide position 18227 (within the L-gene) which can be attributed to a polymerase error during RT-PCR in the original development. This mutation was silent and did not have a recognizable effect in viral rescue or viral transcription or replication. The cDNA clone that was developed, by Neumann and colleagues (114) showed three nucleotide changes. The first mutation was an A insertion between nucleotide positions 9744 and 9745. Another A insertion was found between nucleotides 18495 and 18496 and an A-to-T replacement was detected at position 18226. Interestingly, all of the mutations found in the cDNA full-length clone had been reported to be present in the functional ZEBOV minigenome (111) or have been found with other ZEBOV strains (see data bank sequences) and are, therefore, considered naturally occurring variants. Neither the insertions nor the replacement mutations had an effect on the virus once rescued, indicating that some minor mutations seem to be tolerated within the cDNA full length constructs without effecting rescuability. Volchkov and colleagues (166) took advantage of this by intentionally inserting a silent mutation at nucleotide position 6180 to create a unique SalI restriction enzyme site, which subsequently was used to identify rescued mutant viruses from wild-type ZEBOV.

These mutational problems not only occur when developing a cDNA viral genome copy but have been a recurring event when cassette mutagenesis and full length cDNA re-cloning are carried out. To alleviate this problem a new cDNA full-length reverse genetic clone has been established for ZEBOV, which encodes a pBR322 origin (36). This change decreased the plasmid copy number and has thus decreased the probability of a spontaneous mutation occurring when mutagenizing cDNA fragments. Another method, which has been used in mutatgensis strategies when working with these large plasmids, is developing smaller cassettes of the cDNA plasmid. These cassettes can then be easily mutagenized and re-cloned into the full-length cDNA viral plasmid allowing for easier cloning and development of mutant viruses.

In conclusion, reverse genetics for filoviruses are likely to become extremely valuable research tools in the future. The existing minigenome systems for ZEBOV (111), REBOV (66), and MARV (110) are first choices for deciphering the mechanisms of viral replication and transcription; work that has already made great progress since the development of the systems (see above). In addition, they will be helpful for screening antiviral drugs targeting the replicase complex of filoviruses, a priority for the response capacity against A List bioterrorism agents such as EBOV and MARV (15, 18). The infectious clone systems, existing for ZEBOV (wild-type) and since 2006 for Marburg virus (114, 166) will become key elements for pathogenesis studies and might be helpful for vaccine development. Pathogenesis studies are dependent on animal models. Since the rodent models for filoviruses are dependent on adapted virus strains, it will be important to develop infectious clone systems for the mouse-adapted ZEBOV (20), and the guinea

pig-adapted ZEBOV (28), which are both currently under development (36). Recently, a Marburg virus reverse genetic system was developed,(41) this was exciting news as this tool not only will advance our knowledge about the Marburg virus itself; it will allow for the comparison of all genera within the family *Filoviridae*. Comparisons of these viruses may help to answer many long standing questions from viral evolution to host cell adaptation.

2.0 Materials and Methods

2.1 Cells

2.1.1 Eukaryotic Cell lines:

2.1.1.a Vero E6 (African green monkey kidney) epithelial cells (ATCC CRL-1586) were grown in Dulbecco's Modified Eagle's Medium (DMEM, Sigma) supplemented with 10% (w/v) fetal bovine serum (FBS) and 1% (w/v) penicillin/streptomycin (Gibco/BRL, Rockville, MD, USA). Cells were maintained at 37°C in 5% CO_2.

2.1.1.b 293T (Human embryonic kidney) cells were maintained in DMEM which was supplemented with 10% FBS and 1% penicillin/streptomycin. When working with 293T cells plates were coated with Poly-D-lysine (1mg/ml, Sigma) for 30min at 37°C to increase cell attachment by steric interaction. Poly-D-lysine was removed and culture flasks or culture plates were washed twice with sterile water before cells were seeded. Cells were maintained at 37°C in 5% CO_2.

2.1.1.c U-937 (Human monocyte) cells (ATCC CRL-1593.2) were supplemented with RPM1 1640, 10% FBS, 1% penicillin/streptomycin and maintained at 37°C in 5% CO_2.

2.1.2. Prokaryotic cell lines - Escherichia coli (E.coli):

All *E.coli* strains were made competent by adding an overnight culture to 200ml Lauria broth (LB) Lenox (0.5% (w/v) NaCl). Cells were incubated at 37°C with shaking for approximately 4 hours until an optical density of 0.5 to 0.7 at 600nm has been reached.

Cells were then stored on ice for 10 min and centrifuged at 4000 rpm, 10 min, at 4°C. Supernatants were discarded and the pellet was resuspended in ice-cold 0.1 M $CaCl_2$, 16% (w/v) glycerol, 5% (w/v) dimethyl sulfoxide (DMSO) and 50 mM $MgCl_2$. Resuspending cells in 2 ml per 50 ml original culture, cells were aliquoted as 200 µl samples and stored at -80°C for later use.

Table 2: *E.coli* strains used in these projects.

Cell line	Genotype	Application	Source/Reference
DH5α	F⁻supE44 Δ*lac*U169 (*Φ*80 *lac*ΔM15) *hsd*R17 *rec*A1 endA1 *gyr*A96 *thi*-1 *rel*A1	Maintain clones Common cloning Glycerol stocks	Invitrogen (71)
XL-1 Blue	*rec*A1 *end*A1 gyrA96 *thi*-1 *hsd*R17 *sup*E44 relA1 lac[F' *pro*AB lacI^{q}ZΔM15 Tn10 (tet^r)]	Site directed Mutagenesis	Stratagene (25)
BL21-Gold	F⁻ ompT hsdS($r_B^-m_B^-$) dcm$^+$ Tet^r gal endA Hte	Site directed Mutagenesis Large plasmid mutagenesis	Stratagene (172)
XL10-Gold	*rec*A1 *end*A1 *gyr*A96 thi-1 hsdR17 *sup*E44 relA1 lac[F' *pro*AB lacI^{q}ZΔM15 Tn10 (tet^r)]	Site directed Mutagenesis Large plasmid mutagenesis	Stratagene (86)
TOP-10	F⁻[LacI, Tn10(Tet^R)] mcrA Δ(mrr-hsdRMS-mcrBC) *Φ*80*lac*ZΔM15 D*lac*X74 *rec*A1 *ara*D139 D(*ara-leu*)7697 *gal*U *gal*K *rps*L (Str^R) *end*A1 *nup*G	TOPO cloning Common cloning Glycerol stocks	Invitrogen (145)

2.2 Viruses

2.2.1 Virus Strains:

The following viruses were used for this study: *Zaire ebolavirus*, strain Mayinga; *Reston ebolavirus*, strain Pennsylvania; *Lake Victoria marburgvirus*, strain Musoke (accession numbers AF272001, AF522874, Z12132, respectively). Viruses were propagated in biocontainment level 4 (BSL-4) in Vero E6 cells supplemented with DMEM (10% FBS 1% (w/v) penicillin/Streptomycin). Infection was performed with a multiplicity of infection (MOI) of 0.01 (unless otherwise specified). Infections were carried out for 30 min at 37°C, cells were then washed with DMEM (no additives) to remove any unbound virus. Viruses were harvested when the cytopathic effect (CPE) was approximately 70% or higher, purified by low speed centrifugation to remove cellular debris followed by ultra centrifugation through a 20% (w/v) sucrose cushion.

2.2.2 Virus Inactivation Protocols and RNA isolation protocols:

Biocontainment and inactivated viruses are the only two ways with which to work with these highly pathogenic viruses. Thus our inactivation protocols are extremely stringent.

2.2.2.a Trizol LS (Gibco/BRL, Rockville, MD, USA):

Virus was inactivated following treatment with Trizol LS reagent (Gibco/BRL, Rockville, MD, USA), samples were diluted in a 3:1 ratio or 1 ml Trizol LS per $1x10^6$ cells. Once removed from biocontainment samples were frozen overnight before processing, this step has allowed for increased RNA concentrations compared to processing the sample without freezing. 0.2 ml of chloroform per 0.75 ml of Trizol LS used in sample was added to disrupt protein structures. Samples were centrifuged (12000

xg, 5 min, 4°C) and the aqueous phase was collected. Aqueous RNA was precipitated and resuspended using normal isolation procedures (136). Samples were inactivated before removal from biocontainment.

2.2.2.b RLT buffer Quanidinium isothiocyanate (QIAGEN, Mississauga, Ontario, Canada):

RNeasy kits are designed to isolate total RNA from but not inclusive to, quanidinium isothiocyanate inactivated viruses. The RNA can be isolated from small quantities of starting material such as tissue culture or serum from infected animals. Following the manufacturer's procedures, RNA from infected samples was isolated and RNA concentration was determined using a NanoDrop® ND-1000 UV-Vis spectrophotometer (NanoDrop Technologies).

2.2.2.c SDS Loading buffer inactivation:

For western blot analysis, samples were heat treated at 100°C in 5x SDS gel-loading buffer (100mM Tris-Cl (pH 6.8), 200 mM dithiothreitol, 10% SDS, 0.2% bromophenol blue, 20% glycerol) for 15 min prior to a tube change and removal from level 4 for molecular analysis.

2.2.2.d Formalin/paraformaldehyde fixation:

In order to visualize infected samples using fluorescence microscopy, infected cells were fixed and inactivated with 4% paraformaldehyde or 10% formalin in PBS for 2 days with one fixative exchange. Samples were then bagged and flooded in 10% formalin, heat sealed in level 4 and removed for analysis.

2.3 Molecular Techniques:

2.3.1 Plasmid Cloning Vectors:

Table 3: Vector Descriptions

Vector	Base Pairs	Promoter	Resistance	Reference
pBR322	4,361	Prokaryotic	Ampicillin, Tetracyclin	(14)
pUC19	2,686	Prokaryotic With lacZα	Ampicillin	(177)
pBK-CMV	4,518	Eukaryotic With lacZα	Neomycin Kanamycin	(144)
pBluescript II $SK^{+/-}$	2,958	Prokaryotic With lacZα	Ampicillin	(144)
pTM1	5,357	Prokaryotic	Ampicillin Kanamycin	(107)
pCAGGS	4,746	Eukaryotic	Ampicillin	(93, 119)
pSP72	2,462	Prokaryotic	Ampicillin	(94)
pSP64	3,030	Prokaryotic	Ampicillin	(104)
pCR4Blunt-TOPO	3,957	Prokaryotic With lacZα	Ampicillin Kanamycin	(145)
pCR2.1-TOPO	3,931	Prokaryotic With lacZα	Ampicillin Kanamycin	(145)

All vectors which were stored in glycerol were transformed in DH5α bacterial cell line for storage at -80°C.

2.3.2 Cloning

2.3.2b ZEBOV clone development:

ZEBOV Filovirus plasmid constructs expressing viral support proteins (NP, VP35, VP30, and L) were cloned or sub-cloned into the eukaryotic expression vectors or prokaryotic expression vectors (Table 3). The generation of the full length ZEBOV, strain Mayinga, reverse genetic clone used for the infectious clone system was developed by Neumann et. al. in 2002 (114).

2.3.3 Ligation reaction

Ligation were generally carried out using insert to vector rations of 3:1, 5:1, and 10:1 in a molar ratio with a T7 DNA ligase concentration of 5U/μl. Ligation were incubated between 12° C-16° C overnight.

2.3.4. Polymerase Chain Reaction (PCR)

PCR reactions were performed using Pfu Turbo polymerase (Stratgene). This polymerase provides robust amplification of long, complex genomic targets. This enhanced version of polymerase is a mixture of Pfu DNA polymerase and a thermostable ArchaeMAxx polymerase enhancing factor which enhances PCR product yields and increases target length capability without altering DNA replication fidelity.

Table 4: Optimized parameters for PCR reaction conditions:

Parameter	**Target DNA size <10 kb**	**Target DNA size >10kb**
Extension time	1 min per kb	2 min per kb
Pfu polymerase	2.5U	5.0U
Input template	100ng	200ng
Forward primer Reverse primer	150ng 150ng	150ng 150ng
dNTP	250μM	250μM
Sterile distilled water	Up to 50μl	Up to 50μl

Table 5: PCR cycling parameters:

Number of cycles	**Temperature**	**Duration**	**Function**
1X	95°C	2 minutes	Denaturation
25-50X	95°C 45-70°C* 72°C	1 minute 1 minute 1 – 12 minutes	Denaturation Annealing Elongation
1X	72°C	10 minutes	Elongation
Pause	4°C	1 minute – overnight	Stabilization

* Temperature is dependant on melting temperature of primers used

2.3.5 Reverse Transcription PCR

RT-PCR was carried out using a Qiagen OneStep RT-PCR kit. This kit utilizes two reverse transcriptases Omniscript and Sensiscript which together provide a highly efficient reverse transcription of RNA quantities in the 1 pg – 2 μg range. Once the RT reaction is complete and RNA template has been copied into cDNA, HotStartTaq DNA polymerase becomes active. This enzyme must be heated to 95°C for 15 minutes before becoming active, which in turn inactivates the reverse transcriptases. Reaction conditions are listed below in Table 6

Table 6: Standard protocol for RT-PCR reaction

Components	**Amount**
RNase-free water	Up to 50μl
5X QIAGEN Onestep RT-PCR Buffer	10.0μl
dNTP Mix (10mM each dNTP0	2.0μl
Forward primer	200ng
Reverse primer	200ng
Qiagen OneStep RT-PCR Enzyme mix	2.0μl
Template DNA	10-100ng

All tubes were kept on ice while all components were added to ensure full fidelity of the enzymes. A typical thermocycling protocol is listed below in Table 7

Table 7. Standard cycling protocol for RT-PCR reaction

Number of Cycles	**Conditions**
1X	50°C for 30 min
1X	95°C for 15 min
30-40X	94°C for 30 sec 50-68°C for 30 sec (range is for melting temperature of primers) 72°C for 1 min
1X	72°C for 7 min
HOLD	4°C

2.3.6 PCR Screening:

Since most of the cloning within these projects revolves around the generation of cDNA mutants, we used PCR screening to analyze large numbers of transformants. Briefly, A PCR master mix was prepared as described in Table 8

Table 8: Standard reaction components

Amount	Reaction components
3 µl	10X reaction buffer (500mM KCL; 100mM Tris-HCL pH 8.3; 0.01% (w/v) Triton X-100; 15mM $MgCl_2$; ddH_2O
2.5 µl	dNTP mix (10mM stock solution)
0.3 µl	Forward primer 200ng/µl
0.3 µl	Reverse primer 200ng/µl
0.2 µl	Taq DNA polymerase 1U/µl
23.7 µl	Sterile distilled water

This master mix was used to lyse bacterial cells which had been picked from the transformation plates and placed into a sterile eppendoff tube. Once colonies were picked and sub-colonized on a LB+ ampicillin (100 mg/ml) plate which has been grided to ensure proper order of colonies picked and PCR samples run using the standard PCR protocol. Any positive PCR samples were further analysed using restriction digest and concentrated using MINI prep technologies (QIAGEN).

2.3.7 Site Directed Mutagenesis:

The QuikChange$_{TM}$ XL site directed mutagenesis system was used to make point mutations within the cDNA fragment of interest. Following the manufacturer's protocols we generated all mutations using this PCR based system.

2.3.8 Taq Polymerase generation:

Since most of the work carried out within this project revolved around the generation of PCR mutants we generated our own Taq polymerase to carryout PCR screening. Briefly, two pre-cultures of strain DH1/pTaq was grown in LB medium with 100 mg/ml of ampicillin. Cultures were grown at 37°C centrifuged to collect the bacterial pellet and washed once with Buffer A (50mM Tris-HCL pH 7.9, 50mM glucose, 1mM EDTA) and frozen at -80°C. Cells were then thawed and resuspended in 50ml of buffer A with 200mg of lysozyme. Equal amounts of lysis buffer (10 mM Tris-HCL pH 7.9, 50 mM KCL, 1 mM EDTA, 1 mM PMSF, 0.5% (w/v) Tween 20, 0.5% (w/v) Nonidet P40) was added to mixture and incubated at 75°C for 1 hour. Solution was clarified by centrifugation (15000 rpm, 10min) removing cellular debris. Solid streptomycin was added to the crude extract to a final concentration of 2.5% (w/v). Supernatants were removed after centrifugation. Solid ammonium sulfate was added to a final concentration of 15% (w/v) (15 g/ 100 ml), supernatants were removed after centrifugation. And an additional 22 g / 100 ml of solid ammonium sulfate was added for a final concentration of 50% (w/v). Centrifugation of the sample separated the precipitated protein from the supernatant. The precipitated protein was resuspended in buffer A and the sample was dialzed with two changes of storage buffer (50 mM Tris-HCL pH 7.9, 50 mM KCL, 0.1 mM EDTA, 1 mM DTT, 0.5 mM PMSF, 50% glycerol) at 4°C. After dialysis the solution was diluted to 10 ml final volume and stored at -80°C.

2.3.9 Immunofluorescence assay (IFA):

Vero E6 cells were grown on cover slips and infected with rescued ZEBOV at an MOI of 0.1. Following an incubation of 4 days, the infected cells were fixed and inactivated with 2% (w/v) paraformaldehyde in PBS for 2 days with one fixative exchange. For immunoflorescence analyses, the following protocol was applied. Briefly, cells were permeabilized with 0.1% (w/v) Triton X-100 in PBS for 15 minutes, washed three times with PBS, incubated with a polyclonal rabbit serum directed against ZEBOV VP40 (1:200 dilution in PBS) for 1 hour at 37°C, washed three times with PBS, incubated with a FITC-labeled anti-rabbit conjugate (1:500 dilution in PBS) (Sigma, St. Louis, USA) for 1 hour at 37°C, and washed three times with PBS. The cover slips were mounted and analyzed using an Axioplan 2 Fluorescence microscope (Zeiss, Germany).

2.3.10 Dot-matrix comparison:

Dot-matrix comparisons were carried out using DNA Star (Megaline Lasergene program). Comparisons between ZEBOV and REBOV or MARV were based on the percent matches, in which similarities are represented by the color of the line. Blue indicates the weakest matches and red the strongest.

2.4 Fluorescence detection systems:

2.4.1 Immunoplaque assay:

The detection and titration of infectious ZEBOV-GFP was performed by infecting Vero E6 cells which were grown on cover slips using a 10-fold dilution series. Following virus

adsorption for 30 min at 37°C, the cells were washed three times with PBS. The infected cells were then overlaid with DMEM containing 1.5% (w/v) carboxymethyl-cellulose (CMC) and 2% (w/v) fetal calf serum with 1% (w/v) penicillin/streptomycin (GIBCO). Following an incubation of 5 to 7 days, the infected cells were fixed and inactivated with 4% (w/v) paraformaldehyde in PBS for 2 days with one fixative exchange. The cover slips were mounted and analyzed using an Axioplan 2 Fluorescence microscope (Zeiss, Germany).

2.4.2 FACS analysis: Blood samples were collected from infected STAT-1 mice at defined days 1, 3 & 5 and placed in EDTA tubes during blood extraction. Blood samples were processed by normal FACS lysis protocol. Samples were resuspended and inactivated by adding 4% (w/v) paraformaldehyde and stored at 4°C for 24 hours. Samples were pelleted and fresh 4% (w/v) paraformaldehyde was added to resuspended fixed cells. Tissues from infected mice were homogenized and supernatants from homogenized tissue were added to FACS lysis solution, which removes all red blood cells. Inactivated samples were then assayed using the FACS caliber program.

2.5 Reverse Genetics

2.5.1 ZEBOV reverse genetic system:

The generation of infectious ZEBOV was performed in BSL-4 containment. Using a 50/50 split of $5x10^5$ Vero E6/293T cells, we transfected this 80% confluent layer of cells with 1 μg of the ZEBOV genomic clone cDNA and 1 μg of the T7 RNA polymerase expression plasmid. Simultaneously, the support proteins (NP- 1 μg, VP35- 0.5 μg,

VP30- 0.3 μg, and L- 1 μg) were transfected. The plasmid ratios were kept the same for all rescue attempts using support proteins. Three days post-transfection, the supernatants were collected and used to infect fresh Vero E6 cells. Subsequently, the cells were monitored for CPE over a period of 14 days. After positive rescue virus stocks were prepared on Vero E6 cells (T-75) for seven days, aliquoted and stored in liquid nitrogen within level 4.

3.1 Heterologous Protein Switching

3.1.1 Introduction

The filovirus genome encodes seven genes, which are transcribed into seven (MARV) or eight (EBOV) monocistronic polyadenylated mRNA transcripts encoding seven structural proteins and, in the case of EBOV, a single additional non-structural protein. Four of these proteins, the nucleoprotein (NP), the virion structural proteins (VP) 30 and 35 and the RNA-dependent RNA polymerase (L), constitute the ribonucleoprotein (RNP) complex in association with the genomic RNA (Figure. 1). The other three structural proteins are the surface glycoprotein (GP), the matrix protein (VP40), and VP24. The risk posed by these agents has prompted research and development of tools to identify steps in virus replication and viral pathogenicity. Minigenome-based reverse genetic systems for filoviruses were recently developed (110, 111) (Figure 6). On the basis of these studies, the infectious clone system for ZEBOV has been established, which now allows for the study of virus biology and pathogenesis in the context of infectious virus mutants in tissue culture and animal models (114, 166)

Hypothesis and Objectives of this study:

Transcription and replication of negative-stranded RNA viruses is thought to be a highly specific driven only by the homologous RNA polymerase and its cofactors. The current lack of an infectious clone system for any other EBOV species raises the question to what extent the existing ZEBOV system would be useful in rescuing heterologous EBOV genomes and, thus, rescue wild-type or mutant REBOV, SEBOV, or ICEBOV. For this, the ZEBOV infectious clone system was used to switch the proteins which drive

viral replication. This would allow us to gain insight into specificity of transcription/replication and further might allow us to rescue different EBOV species in the future.

3.1.2 RESULTS and DISCUSSION

3.1.2a Molecular Clones:

REBOV strain Pennsylvania (AF522874) was grown in 10xT-150 tissue culture flasks on Vero E6 cells for approximately 10 days or until viral CPE was 70%. Viral genomic RNA was extracted using the Qiagen RNA extraction protocol (see Materials and Methods). For REBOV, following extraction of viral genomic RNA the open reading frames of NP, VP35, VP30, and L were transcribed with Superscript II reverse transcriptase (Invitrogen) and amplified with Pwo DNA polymerase using specific primers for the amplification of viral proteins. These fragments were then cloned into a eukaryotic expression vector pCAGGS, DNA was isolated from positive transformants and standardized to 1μg/μl for use in a reverse genetic rescue.

ZEBOV strain Mayinga (AF272001) was used as a template for the amplification of viral open reading frames (NP, VP35, VP30, and L). The full length clone was also derived from this Ebola template which was previously described (114).

The open reading frames of MARV Musoke (AFZ12132) were sub-cloned from previously existing plasmid constructs (45), into the eukaryotic expression vector pCAGGS.

Plasmids containing the open reading frames of VSV P, N, L were kindly provided by J.K. Rose, Yale University (95). These plasmids are under the control of the bacteriophage T7 RNA polymerase promoter.

3.1.2b Optimizing the Reverse Genetic System

In order to better interpret negative rescues, particularly when heterologous support proteins are used, the rescue efficiency of the ZEBOV infectious clone system was first optimized. Plasmid concentrations were standardized to 1μg/μl when used in a reverse genetic rescue attempts. Plasmids were transfected into a 1:1 mixture of 293T and Vero E6 cells. 293T cells were used as they are highly transfectable and Vero E6 cells have been shown to promote virus replication efficiently (114). Transformations were carried out using 2μl *trans*IT®-LT1 (Mirus) transfection reagent per 1 μg of the standardized DNA. In order to optimize the reverse genetic system a wide range of plasmid concentrations were tested (0.2 μg – 2 μg) to determine what ratios of RNP components would yield a 100 % rescue efficiency. Using the RNP complex proteins at the following concentrations, pCEZ-NP 1 μg/μl, pCEZ-30 0.3 μg/μl, pCEZ-35 0.5 μg/μl, pCEZ-L 1 μg/μl, pCT7 Pol 1 μg/μl, and pTM Ebo-Rib 1 μg/μl gave 100 % efficiency in more then fifty rescue attempts using the ZEBOV reverse genetic system. If one of the RNP complex proteins are removed, virus rescue cannot be attained using the reverse genetic system (Figure. 8).

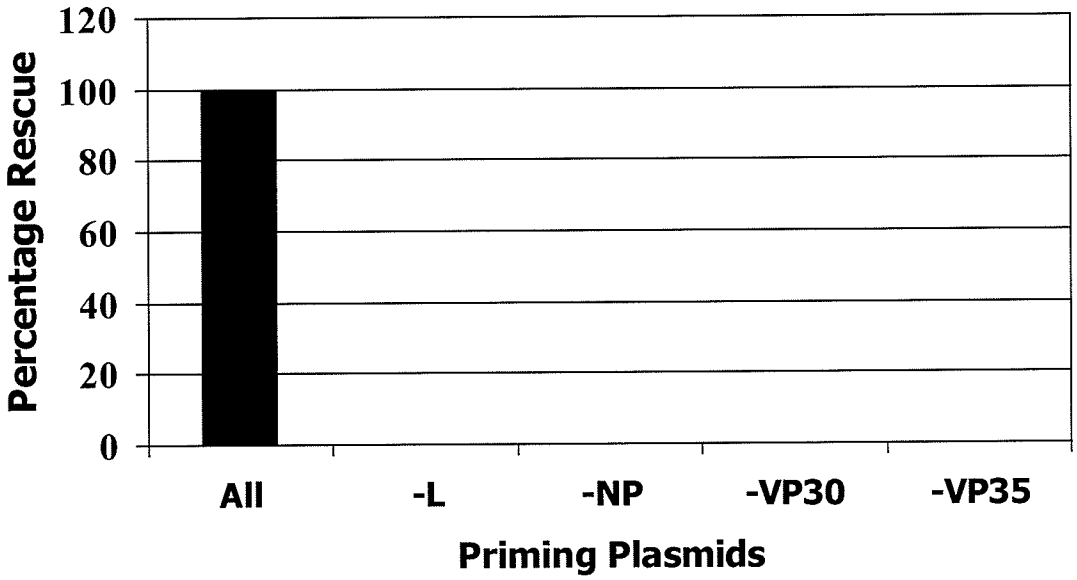


Figure. 8 Rescue efficiencies using ZEBOV RNP component proteins for optimization of the ZEBOV reverse genetic system. The presence of all four support proteins is indicated by 'All' and the lack of support proteins by '-'. When one RNP protein is removed from the rescue attempt no detectable virus was present 14 days post transformation passage. Positive rescue of the ZEBOV reverse genetic system was attend using all RNP complex proteins 100% out of 50 rescues. In contrast, 50 attempts per priming plasmid removal experiments never resulted in a positive rescue.

Samples were brought into level 4 for evaluation of virus rescue immediately after plasmid transformation. In general, the supernatants of transfected cells were passaged onto Vero E6 cells on day 3 post-transfection due to the acidification of the media from the death of the 293T cells. Also, at this point virus proteins produced solely from the infectious clone could already be detected by indirect immunofluorescence assay (IFA) (Figure. 9A) in the cells. With the ZEBOV infectious clone system, positive rescue as determined by the presence of cytopathic effect was achieved between day 4 and 6 after passaging (Figure. 9B). The identity of the rescued virus was verified by IFA using rabbit serum against the ZEBOV-VP40 (Figure. 9A).

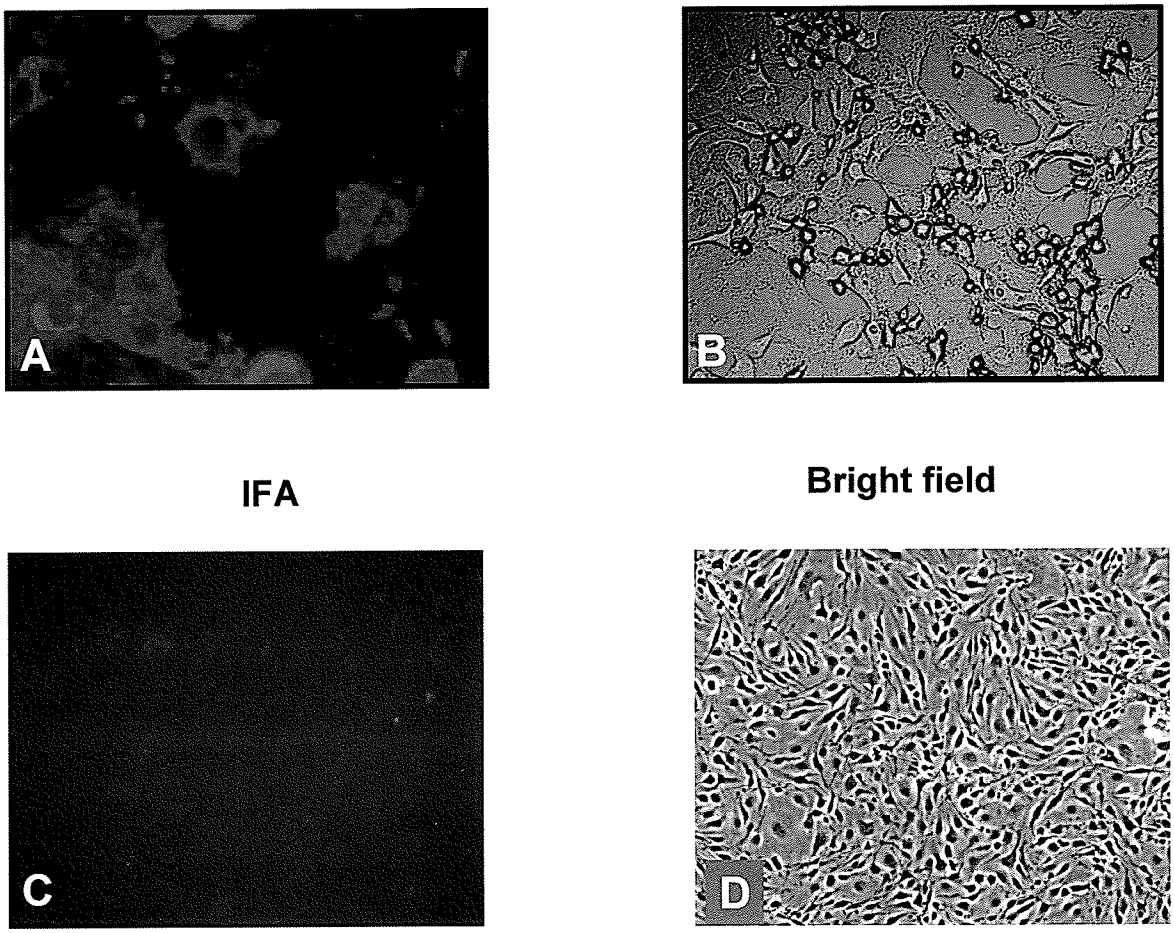


Figure. 9 Positive and negative rescue using the ZEBOV reverse genetic system A) Indirect immunofluorescence assay of a culture, 5 days post-passaging onto Vero E6 cells. The identity of the rescued virus was verified using an anti-ZEBOV-VP40 (1:200 dilution in PBS). A FITC anti-rabbit secondary antibody (dilution 1:500 in PBS) was used for the fluorescence assay. Background for the secondary labelled antibody was tested and proved to be minimal (data not shown). B) Brightfield view of cytopathic effect in a Vero E6 cell culture 6 days post-passaging of the transfected supernatant onto Vero E6 cells. C) Indirect immunofluorescence assay of a negative rescue attempt. Rescue attempts were deemed negative after a 14 day incubation period with no resulting cytopathic effect or positive immunofluorescence. D) Bright field view of negative control in a Vero E6 cell culture 6 days post passage. Photos in A - B and C – D respectively were taken from the same plate but in different locations.

A rescue attempt was deemed negative after an incubation period of 14 days without showing CPE or positive IFA results (Figure. 9 C, D).

To assess the rescuability of the ZEBOV reverse genetic system using REBOV-derived expression plasmids, the same amounts and ratios among support proteins which was established and optimized for the rescue of the ZEBOV reverse genetic system were applied (158). The following rescue attempts using heterologous RNP support proteins were performed: Within each table the far left column represents the RNP complex proteins, and within the rows preceding the protein name is the species name of the virus type used in the rescue.

TABLE 9												
NP	**REBOV**	**ZEBOV**	**ZEBOV**	**ZEBOV**	**REBOV**	**ZEBOV**	**REBOV**	**REBOV**	**ZEBOV**	**ZEBOV**	**REBOV**	**REBOV**
VP30	**ZEBOV**	**REBOV**	**ZEBOV**	**ZEBOV**	**ZEBOV**	**REBOV**	**REBOV**	**ZEBOV**	**ZEBOV**	**REBOV**	**REBOV**	**------**
VP35	**ZEBOV**	**ZEBOV**	**REBOV**	**ZEBOV**	**ZEBOV**	**REBOV**	**ZEBOV**	**REBOV**	**REBOV**	**ZEBOV**	**REBOV**	**REBOV**
L	**ZEBOV**	**ZEBOV**	**ZEBOV**	**REBOV**	**REBOV**	**ZEBOV**	**ZEBOV**	**ZEBOV**	**REBOV**	**REBOV**	**REBOV**	**REBOV**
Rescue CPE	**2 of 6**	**6 of 6**	**6 of 6**	**2 of 6**	**2 of 6**	**6 of 6**	**2 of 6**	**3 of 6**	**3 of 6**	**3 of 6**	**6 of 6**	**0 of 6**
IFA Results	Positive	Positive	Positive	Positive	Positive	Positive	Positive	Positive	Positive	Positive	Positive	Negative

TABLE 10												
NP	**MARV**	**ZEBOV**	**ZEBOV**	**ZEBOV**	**MARV**	**ZEBOV**	**MARV**	**MARV**	**ZEBOV**	**ZEBOV**	**MARV**	**MARV**
VP30	**ZEBOV**	**MARV**	**ZEBOV**	**ZEBOV**	**ZEBOV**	**MARV**	**MARV**	**ZEBOV**	**ZEBOV**	**MARV**	**MARV**	**------**
VP35	**ZEBOV**	**ZEBOV**	**MARV**	**ZEBOV**	**ZEBOV**	**MARV**	**ZEBOV**	**MARV**	**MARV**	**ZEBOV**	**MARV**	**MARV**
L	**ZEBOV**	**ZEBOV**	**ZEBOV**	**MARV**	**MARV**	**ZEBOV**	**ZEBOV**	**ZEBOV**	**MARV**	**MARV**	**MARV**	**MARV**
Rescue CPE	**0 of 6**	**0 of 6**	**0 of 6**	**0 of 6**	**0 of 6**	**0 of 6**	**0 of 6**	**3 of 6**	**3 of 6**	**4 of 6**	**4 of 6**	**0 of 6**
IFA Results	**Negative**	**Negative**	**Negative**	**Negative**	**Negative**	**Negative**	**Negative**	**Positive**	**Positive**	**Positive**	**Positive**	**Negative**

Table 9 and 10: ZEBOV rescue using REBOV and MARV RNP complex plasmids.

A) REBOV plasmids are in orange, ZEBOV plasmids are in black. Positive rescues out of six possible rescues are illustrated in the rescue CPE row. Positive or negative IFA results are illustrated in the bottom row. REBOV VP30 and 35 resulted in 100% or (6/6) rescue efficiency where as NP and L single switches resulted in only 33% or (2/6) rescue efficiency. ZEBOV rescue using double REBOV RNP complex plasmids. The combination of REBOV VP30 and VP35 resulted in 100% rescue, NP and VP35 only showed 50% or (3/6) rescue efficiency. NP in combination with L and VP30 resulted in 33% or (2/6) rescue efficiency. The combination of REBOV L and VP35 or VP30 resulted in 33% or (2/6) rescue efficiency, When all REBOV RNP complex plasmids were used 100% rescue efficiency was achieved six out of six times. Removal of one REBOV RNP complex protein resulted in 0% (0/6) rescue. All rescues were analysised after one passage in Vero E6 cells.

B) ZEBOV rescue using single MARV RNP complex plasmids. MARV plasmids are in orange, ZEBOV plasmids are in black. Positive rescues out of six possible rescues are illustrated in the rescue CPE row. Positive and negative IFA results are shown in the bottom row. No single MARV protein when substituted for a ZEBOV protein, could rescue the ZEBOV reverse genetic system. ZEBOV rescue using double MARV RNP complex plasmids. The combinations of MARV NP and VP30 or VP35 resulted in 0% or (0/6) rescue. This result was also seen with the combination of VP30 and VP35. The combination of NP and VP35 did result in a rescue efficiency of 50% (3/6). The combination of MARV L and VP35 or VP30 resulted in 50% or (3/6) and 66% or (4/6) rescue efficiency respectively. When all MARV RNP complex plasmids were used 66%% or (4/6) rescue efficiency was achieved. When one MARV RNP complex protein was removed (VP30) 0% or (0/6) rescue was seen.

Infectious clone systems have been developed for several members of *Mononegavirales* over the past ten years, beginning with the infectious clone for rabies virus (143) and followed by several other systems (6, 27, 95). The infectious clone system developed by our group (114) is based on the simultaneous transfection of a plasmid containing the ZEBOV genome under the control of the bacteriophage T7 RNA polymerase promoter and five plasmids expressing the four RNP complex-associated viral proteins (NP, VP35, VP30, L) and the bacteriophage T7 RNA polymerase under the control of the chicken β-actin promoter (119). After successful transcription of negative sense RNA by the T7 RNA polymerase, the viral nucleoprotein encapsidates the template and associates with the other protein components to form a functional RNP complex, which is subsequently used to generate infectious ZEBOV (Figure. 7).

In total, 100 % rescue efficiency for ZEBOV wild-type (50/50) was obtained. As expected from previous work (111), the lack of a single protein component (NP, VP35, VP30 or L) of the ZEBOV RNP complex (Figure. 8) completely abolished the rescue of the ZEBOV genome confirming the necessity of all four proteins for transcription and/or replication (Figure. 8). In addition, the transfection of the ZEBOV genome without any support proteins but with the plasmid encoding for the T7 RNA polymerase did not result in the rescue of infectious ZEBOV.

Apart from the ZEBOV and MARV sequences, only the REBOV full-length genome sequence is available (67, 80). REBOV belongs to the same genus as ZEBOV but represents a different species with an amino acid similarity to ZEBOV ranging from 71 to 81 % for the RNP complex associated proteins. Therefore, it seemed logical to first study rescue efficiency of ZEBOV using heterologous support proteins derived from this

virus species. The corresponding open reading frames were cloned under the control of the chicken β-actin promoter and the expression of NP, VP35, VP30, and L was verified using the REBOV minigenome system which has recently been established within our group (65).

Rescue of ZEBOV was achieved in 6/6 attempts with only REBOV-derived support proteins indicating that, in general, the functional domains of these proteins are highly conserved and can substitute for their counterparts. In comparison to the ZEBOV system, the rescue efficiency was slightly reduced as can be seen by indirect immunofluorescence assay results 3 days post transfection (Figure. 10). Single substitutions of ZEBOV support proteins by a heterologous protein derived from REBOV resulted in a rescue efficiency of 100 % (6/6) for VP30 and VP35 but dropped to 33 % (2/6) for NP and L (Figure. 11). This trend continued if a combination of 2 support proteins from each species was used. In the case of a combination of VP30 and VP35, ZEBOV was rescued in 100 % of the experiments (6/6), whereas all other combinations resulted in lower rescue efficiencies ranging from 33 to 50 % (Figure. 11). Collectively, these data indicate that the function of the replicase complex may be more dependent on specific protein-protein interactions than on protein-RNA interactions (6/6 positive rescues if all REBOV support proteins were used).

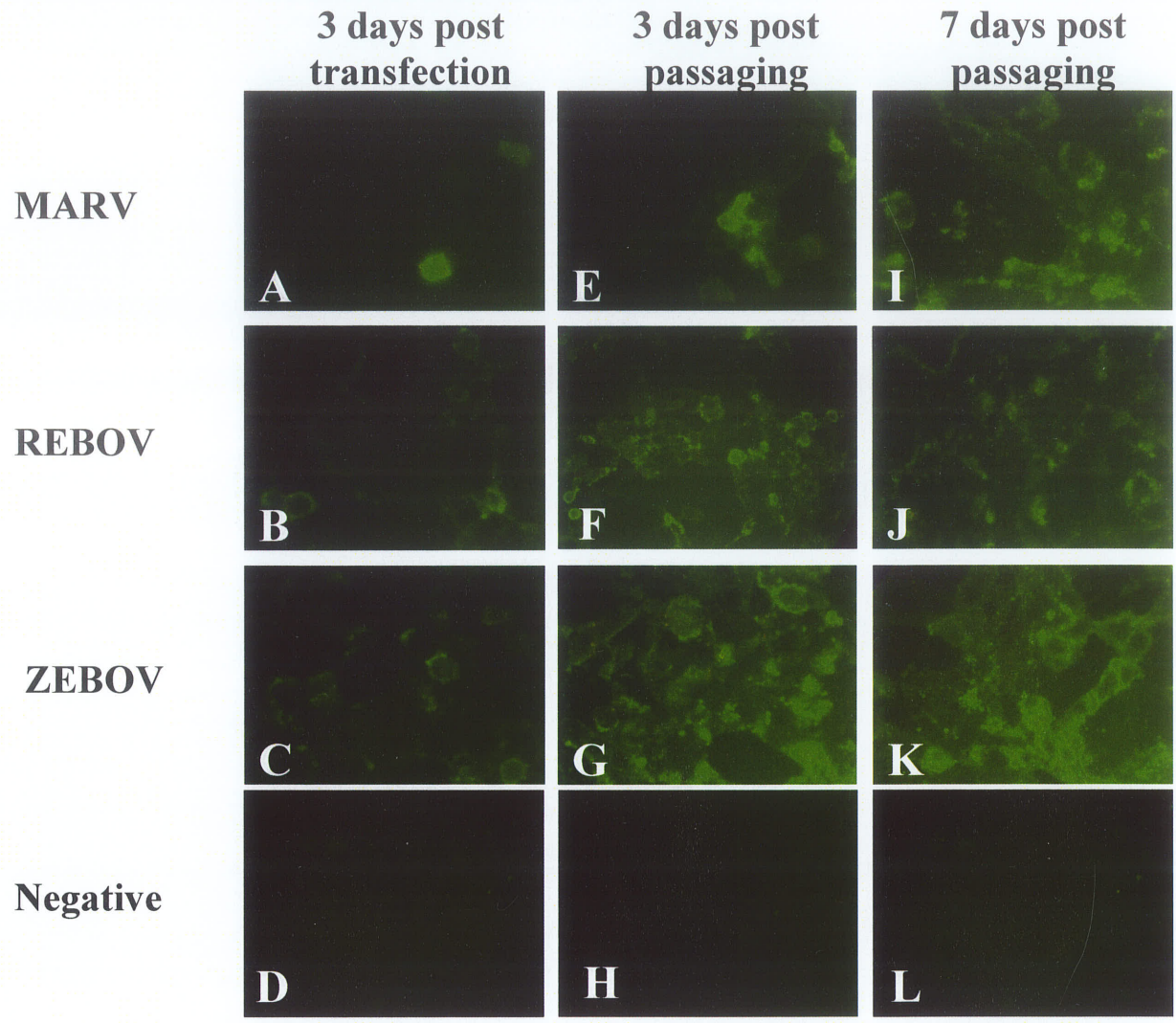


Figure. 10. Indirect immunofluorescence assay demonstrating the rescue of the ZEBOV reverse genetic system using homologous and heterologous support proteins. The rescue presented here were performed with four support proteins derived from the same virus (MARV, *lake Victoria marburgvirus* strain Musoke; REBOV Reston *ebolavirus* strain Pennsylvania; ZEBOV, Zaire *ebolavirus* strain Mayinga). The IFA was performed on formalin-fixed cells (2%) at the indicated time points. For immunodetection a rabbit antiserum directed against ZEBOV VP40 (dilution 1:200 in PBS) with a FITC anti-rabbit secondary antibody (dilution 1:500 in PBS) used for fluorescence. (A-C) demonstrated the rescue efficacy 3 days after transfection of the original 293T/Vero E6 cell culture. (E-G) and (I-K) show virus growth on 3 and 7 days after passaging the transfection supernatants onto Vero E6 cells, respectively. D, H, and L are the corresponding negative controls.

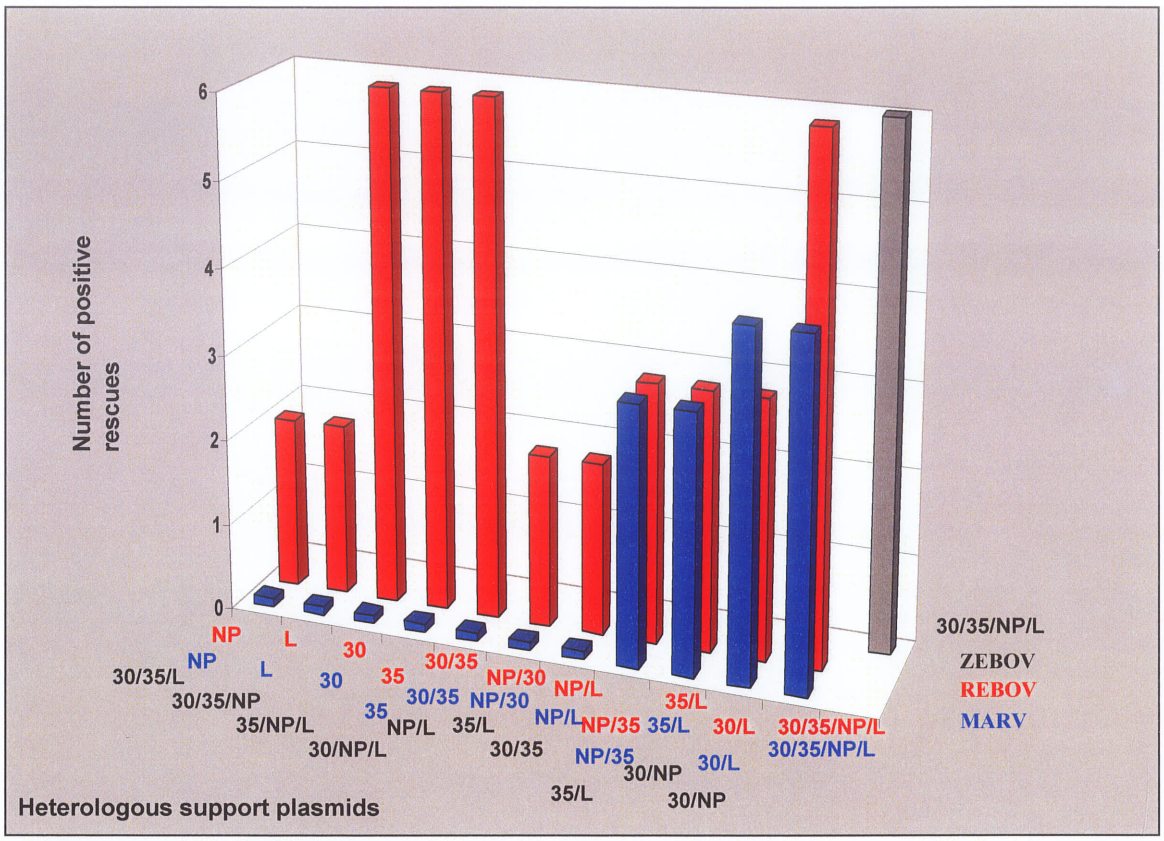


Figure. 11. Rescue efficiency of ZEBOV using different combinations of homologous and heterologous support proteins. Plasmids supplied from a heterologous virus species are indicated along the x-axis in red (REBOV) and blue (MARV) with the remaining plasmids (black) form ZEBOV. Each transfection introduced NP, VP35, VP30, and L in various combinations. Plasmid supplied from a heterologous virus species are indicated along the x-axis with remaining plasmids derived from ZEBOV. All experiments were performed six times using independent transfections. Data for substitutions involving MARV-derived proteins are shown as blue bars while those for REBOV are shown as red bars. Control rescues involving only ZEBOV proteins are shown as a gray bar. Key: MARV, *lake Victoria marburgvirus* strain Musoke; REBOV Reston *ebolavirus* strain Pennsylvania; ZEBOV, Zaire *ebolavirus* strain Mayinga; L, RNA-dependent RNA polymerase; NP, nucleoprotein; VP virion protein 30 and 35 KDa.

Subsequently, we investigated the rescue efficiency of ZEBOV using heterologous support proteins from MARV, a more distantly related filovirus with amino acid similarity to ZEBOV ranging from 46 to 58 % for the RNP complex associated

proteins. The corresponding open reading frames were cloned under the control of the chicken β-actin promoter and the expression was verified by immunoblot and/or immunofluorescence analysis. The rescue of ZEBOV using only MARV-derived support proteins was successful in 66 % of experiments (4/6) (Figure. 11). The lower rescue efficiency is also demonstrated by IFA results 3 days post transfection (Figure. 10). No rescue was achieved when single support proteins were exchanged (Figure. 11). The combination of two support proteins from each virus resulted in rescue efficiencies ranging from 0 to 66 % with the highest values obtained for the combination of VP30/L (66 %), followed by VP35/L and VP35/NP (both 50 %). No rescue was obtained using the combinations VP30/VP35, NP/L, or VP30/NP (Figure. 11). Replacement of VP30 in the ZEBOV reverse genetics system by its MARV counterpart completely abolished rescue (Figure. 11). This is in contrast to the ZEBOV minigenome system where MARV VP30 was able to maintain a low level of transcriptional activity (111). Previously, it was reported that VP30 was not needed for transcription and replication in the MARV minigenome system but required for transcription in the ZEBOV minigenome system (110, 111). However, no ZEBOV rescue was observed (0/6) when only NP, VP35 and L derived from MARV were used for rescue attempts (Table 10). Taken together, the data for MARV-derived support proteins again seem to support the notion that protein-protein interactions are more critical than protein-RNA interactions as indicated by the higher rescue efficiency when more than one heterologous protein was added to the system.

Amino acid identities for the RNP complex associated proteins were compared between ZEBOV and REBOV as well as MARV using the National Center for Biotechnology Information BLAST search programs. In general, ZEBOV and REBOV

shared a higher degree of amino acid identities (62 %, 64 %, 65 % and 71 % for NP, VP35, VP30 and L, respectively) than ZEBOV and MARV (35 %, 36 %, 34 % and 43 % for NP, VP35, VP30 and L, respectively). Despite clear distinctions among the three viruses in the nucleotide and amino acid sequences, all proteins seemed to show similarities as demonstrated by dot-matrix comparison. These conserved regions likely contain the functional domains of these proteins (Figure. 12). Interestingly, VP35 and VP30 of REBOV and ZEBOV, but not NP and L, were very similar in dot-matrix comparison explaining the unaltered rescue efficacy when VP35 and VP30 were switched (100 %) versus a drastic reduction when NP and L were exchanged (33 %) (Figure. 11). Dot-matrix comparisons between ZEBOV and MARV proteins resulted in reduced similarities for NP, VP35 and L (Figure. 10). This may explain the overall lower rescue efficiency when heterologous support proteins derived from MARV were used to rescue ZEBOV (Figure. 11). Surprisingly, MARV VP30 showed a remarkable similarity to ZEBOV VP30 despite its inability to rescue ZEBOV (Figures. 11, 12). In general, rescue was more likely to be successful if NP and VP35 and/or L and VP35 were derived from the same virus species. This supports the model previously hypothesized for MARV transcription and replication (13) which was based on independent interactions of NP and L with VP35, with VP35 being the link between NP and L which otherwise did not interact. On the basis of the dot-matrix comparison (Figure. 12) one might assume that critical domains for protein-protein interactions might be located in the amino-terminal portion of VP35 and the carboxyl-termini of NP and L.

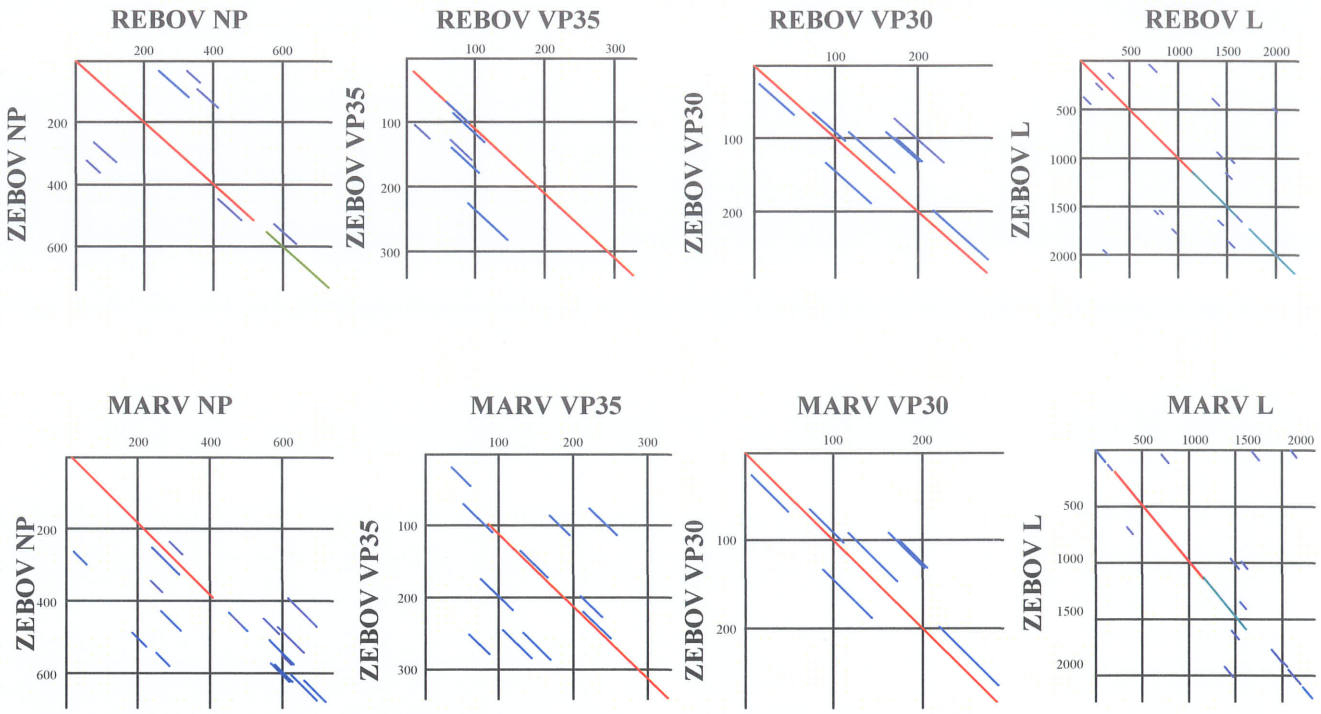


Figure. 12. Dot-matrix comparisons of ZEBOV vs. REBOV or MARV support proteins. In the dot-matrix comparison, each match meeting the specific similarity within the specified group of residues is displayed as a line on the plot. The comparison was carried out using the Lipmann and Wilbur methods with the following parameters: % match = 20; minimum window = 1; and window = 30 using the PAM 250 matrix. The color of the line reflects the degree of similarity, taking into account both the percent match and the match length. Blue indicates the weakest matches and red the strongest.

The three exceptions from the model [VP30 & VP35 of REBOV and NP & L of ZEBOV (6/6 rescues); REBOV VP35 and NP, VP30 & L of ZEBOV (6/6 rescues; REBOV VP30 and NP, VP35 & L of ZEBOV (6/6 rescues)] (Figure. 12) could be explained by the high similarity of VP35 and VP30 among the two EBOV species as discussed above (Figure.

11). In contrast to previously published work (111), REBOV and MARV RNP associated proteins were clearly able to recognize heterologous ZEBOV RNA reemphasizing that protein-protein interactions are more critical than protein-RNA interactions.

The rescue of ZEBOV using heterologous support proteins from other filoviruses was surprising. Therefore, we investigated if support proteins from even more distantly related viruses could substitute for ZEBOV support proteins. For this purpose vesicular stomatitis virus (VSV), (genus *Vesiculovirus*, family *Rhabdoviridae*, order *Mononegavirales*) was used. Exchanging the ZEBOV support proteins in various combinations, as described above for REBOV and MARV, did not result in any positive rescue of the ZEBOV genome. This was also the case if all support proteins were derived from VSV (Figure. 13).

Our data clearly demonstrate that rescue with heterologous support proteins derived from a virus of a different species or different genus within the same family is achievable. Earlier it was reported that rescue of Sendai virus cDNA templates could be achieved with cDNA clones expressing parainfluenza virus type 1 and 3 N, P and L proteins (125). However, rescue of a recombinant virus from plasmid DNA using heterologous support proteins derived from a virus of a different genus has never been reported to our knowledge. The results further demonstrate that transcription and/or replication are not strictly species-specific but require a certain degree of specificity. By evaluating all the different combinations that were tested in this study, the interaction between homologous NP/VP35 and L/VP35 (proteins from the same virus species, not necessarily ZEBOV) seems to be the most critical step. The importance of these interactions was further demonstrated by co-expression studies; using RNP complex

protein, derived from a eukaryotic expression vector. Interaction between NP/VP35, L/VP35 as well as L/VP30 were demonstrated in both heterologous and homologous RNP complex switching which, were demonstrated by co-localization (63).

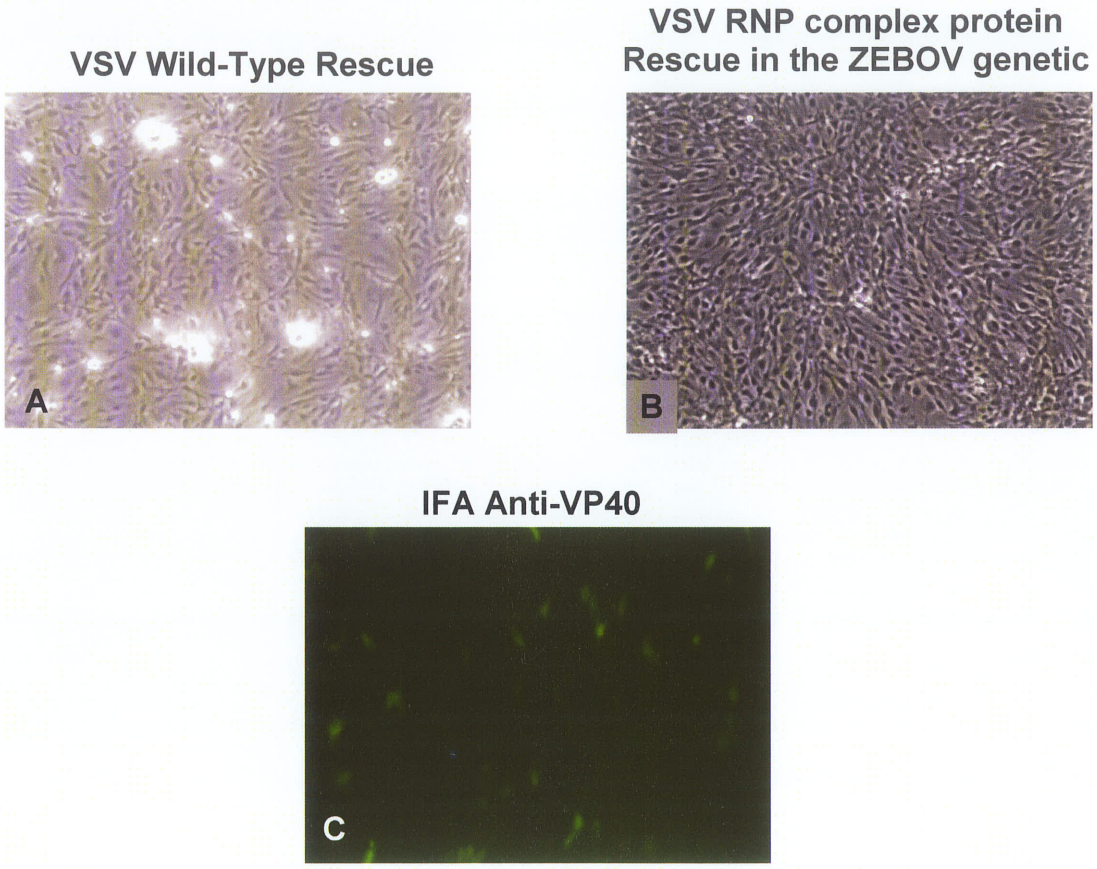


Figure. 13. ZEBOV rescue using all VSV RNP complex plasmids. N,P, and L were transfected in 1 μg, 1 μg, and 2 μg plasmid concentration respectively. ZEBOV rescues using VSV plasmid could not be produced. A) VSV rescue using the VSV reverse genetic system, CPE was seen 24-48 hours after transfection. B) Determination of rescue ability of VSV RNP complex protein to drive the ZEBOV revere genetic system, VSV RNP complex plasmids were transfected in 1 mg (N), 1 μg (P), and 2 μg (L). Cytopathic effect was not seen in all rescue attempts, demonstrating that VSV RNP complex proteins cannot cause ebola replication in the ZEBOV reverse genetic system. C) Indirect immunofluorescence assay was carried out to ensure no live ZEBOV is present. Samples were treated with an anti-VP40 antibody (1:200 dilution in PBS) and FITC (1:500 dilution in PBS)staining for visualization. Some auto-fluorescence is seen in floating dead cells.

The presence of VP30 was critical for the rescue of ZEBOV, which confirms previous observations that found VP30 to be important for the replication of ZEBOV but not for the replication of MARV minigenome systems (111). In addition, the use of only MARV-derived NP, VP35 and L, which are sufficient to drive transcription in the MARV minigenome system, are not sufficient to rescue ZEBOV using the infectious clone system, which confirms recently published data obtained from the MARV infectious clone system(41). In contrast to previous data derived from the filoviral minigenome systems (111), heterologous RNP complex associated proteins (support proteins) were able to complement heterologous filovirus RNA and rescue ZEBOV. The differences between the minigenome and the infections clone system using heterologous support proteins [e.g., lack of CAT activity as shown in (111) and positive rescue as demonstrated here] may be explained by the sensitivities of the read-out systems. CAT activity (minigenome read-out) must be expressed in a specific quantity to result in a positive signal whereas a single event could drive the rescue of virus in an infectious clone system because it is based on multiple amplification cycles thereafter. Thus, the infectious clone system appears to be more sensitive in this respect. However, the minigenome genetic systems are more easily manipulated (BSL2 vs. BSL4) and, thus, both systems are extremely helpful tools for many applications.

3.1.3 Future Direction

Heterologous protein switching has demonstrated that reverse genetic systems are less stringent then previously thought. One future goal from this work is to generate chimeric

proteins, which would allow us to determine regions critical for both transcription and replication of ZEBOV. Also since the development of the Marburg reverse genetic system the possibility to determine the interactions between MARV viral transcription and replication with ZEBOV RNP complex proteins is now possible. These experiments will give us more insight into the transcription and replication patterns of filoviruses. Another direction this research could take is to determine the functional regions within the RNP complex proteins. Since all filoviruses share sequence similarities nucleotide alignment may provide possible unique binding motifs within the RNP complex protein.

The possibility for future studies using genome clones from REBOV, SEBOV, and ICEBOV could potentially be rescued using the well established ZEBOV system. This would avoid the issues associated with a functional L protein which is difficult to express and to characterize.

3.2 Cell Free Cloning

3.2.1 Introduction

In the past decade, several reverse genetic systems have been developed for negative-sense RNA viruses (101) with the establishment of a mini-replicon system generally preceding the development of the infectious clone system. In each case the technology used reflects both the particular requirements of the virus as well as the availability of established methodologies. Both utilize cloned cDNA to either mediate expression of reporter genes (mini-replicon systems) or produce infectious virus particles (infectious clone systems). Together they provide excellent tools for studying replication and transcription as well as infectivity and pathogenicity. Since the establishment of the first ZEBOV reverse genetic systems (114, 166) an inherent problem with the infectious clone system has been identified with regard/respect to the handling of the larger plasmids such as the genomic plasmid and the plasmid encoding the RNA dependant RNA polymerase protein. One common occurrence is spontaneous mutations, which have occurred in the process of generating the two existing ZEBOV systems as well as other genetic systems. The first system (166) carried a single mutation in the genomic clone at nucleotide position 18227 (within the L-gene) which can be attributed to a polymerase error during RT-PCR in the original development. This mutation was silent and did not have a recognizable effect in viral rescue or viral transcription or replication. The cDNA clone that was developed by Neumann and colleagues (114) showed three nucleotide changes. The first mutation was an A insertion between nucleotide positions 9744 and 9745. Another A insertion was found between nucleotides 18495 and 18496 and an A-to-T

replacement was detected at position 18226. Interestingly, all of the mutations found in the cDNA full-length clone had been reported to be present in the functional ZEBOV minigenome (111) or have been found with other ZEBOV strains (see data bank sequences) and are, therefore, considered naturally occurring variants. These natural variants represent regions known as mutational hotspot within the ZEBOV genome. Neither the insertions nor the replacement mutations had an effect on the virus once rescued, indicating that some minor mutations seem to be tolerated within the full length cDNA construct without effecting rescuability. Volchkov and colleagues (166) took advantage of this by intentionally inserting a silent mutation at nucleotide position 6180 to create a unique SalI restriction enzyme site which subsequently was used to identify rescued mutant viruses from wild-type ZEBOV. Although some minor mutations are tolerated in this system, generally an authentic sequence is desired. These mutational problems not only occur when developing a cDNA viral genome copy but have been a recurring event when cassette mutagenesis and full length cDNA re-cloning are carried out. To help alleviate this problem a new cDNA full-length reverse genetic clone has been established for ZEBOV, which encodes a pBR322 origin of replication. This change decreased the plasmid copy number during replication due to expression controls which are found within the pBR322 origin of replication. This decrease in expression will remove some selective pressure on the ZEBOV cDNA construct decreased the probability of a spontaneous mutation occurring when mutagenizing cDNA fragments. Another method which has been used in mutatgenesis strategies when working with these large plasmids is the development of smaller cassettes of the cDNA plasmid. These cassettes can be easily mutagenized and re-cloned into the full-length cDNA viral

plasmid allowing for easier cloning and development of mutant viruses. Regardless of the methodologies used to establish the genetic system, mutagenizing the cDNA follows the same general protocols. Using the second method, cassette building, we developed a new method of mutant clone generation, called 'cell free cloning'. This new cloning strategy removed the propagation steps in normal cloning practices thereby decreasing the chance of spontaneous mutations and allowing for quicker clone generation.

Main hypothesis and objectives of study:

Our main hypothesis was to generate a novel 'cell free cloning system' that would alleviate the occurrence of unwanted spontaneous mutation when manipulating large plasmids using the reverse genetic system for ZEBOV. Our objective was to develop a system which limited the amount of bacterial propagation in generating a mutant cDNA construct. Resulting in fewer spontaneous mutations which are usually generated by the semi-error prone bacterial polymerase.

3.2.1a Current Techniques for Viral Mutagenesis:

A common current method used to generate a mutant cDNA reverse genetic clone is shown schematically in Figure. 14. Briefly, the region of interest is digested with compatible restriction enzymes and sub-cloned into a replicating plasmid. This is to facilitate propagation in bacteria increasing the amount of plasmid construct.

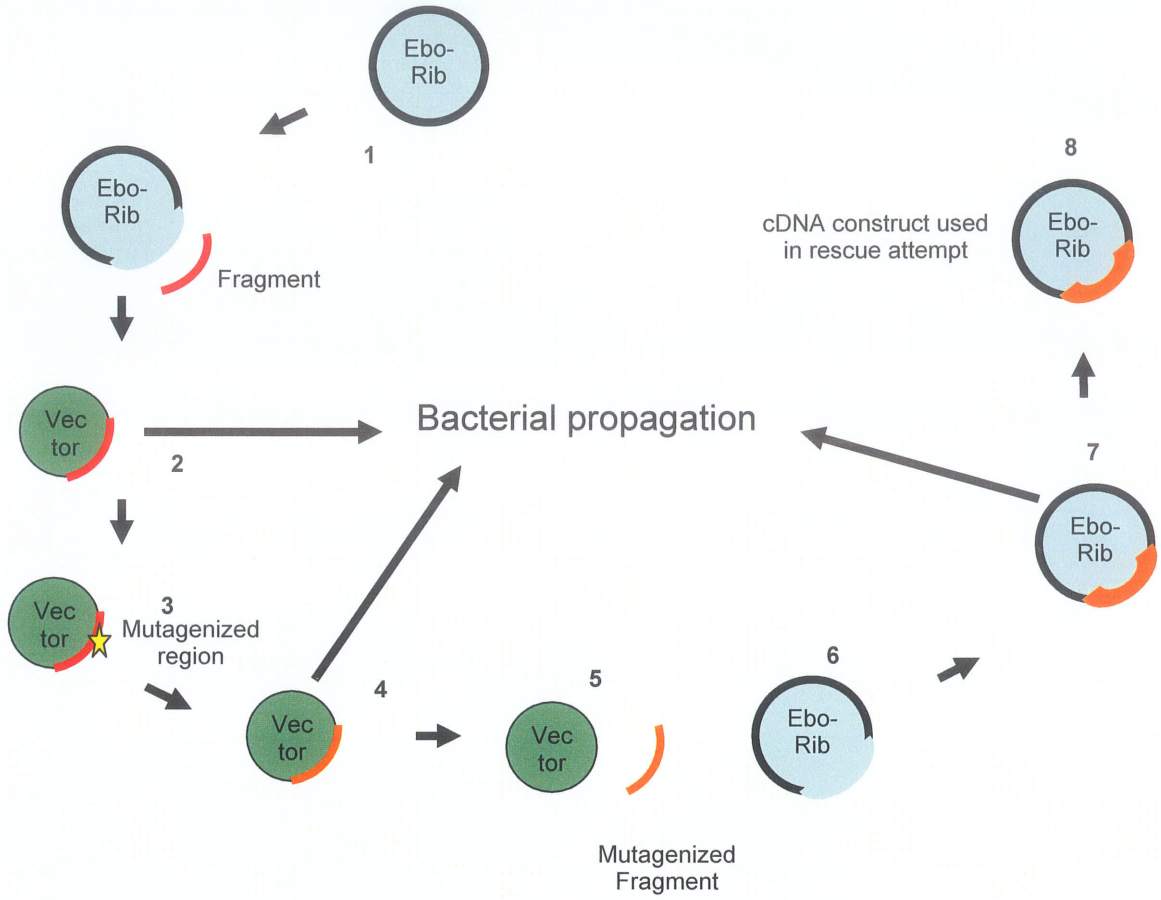


Figure. 14. An illustration of a normal cloning technique used in generating cDNA mutants using large plasmids. 1) Cut the cDNA construct with compatible restriction enzymes. 2) Agarose gel purify the desired fragment, and clone into a prokaryotic expression vector [a bacterial propagation step is needed here to increase plasmid DNA concentration. 3) Develop mutant using site directed mutagenesis (PCR based) 4) propagate in bacteria. 5) Restriction digest the mutagenized fragment and agarose gel purify. 6) Ligation of original cDNA construct which was digested and the mutagenized fragment. 7) This generates a cDNA construct which carries a specific mutation, 8) which can be used for virus rescue and mutant characterization.

The plasmid is re-isolated using transformation and plasmid DNA isolation kits (Qiagen), and a mutation is generated using PCR based site directed mutagenesis. Breaking down the large cDNA ZEBOV full-length clone into smaller fragments (sub-cloning) is required for generating correct mutations due to polymerase restrictions. Unique

restrictions sites and the fidelity of polymerases become the limiting factors when developing mutations in any plasmid. An example of this is demonstrated by the ZEBOV cDNA plasmid, which has only seventeen unique restriction sites in 23313 bp, making this system extremely difficult to work with. Additionally most polymerases can only replicate small fragments (1000-4000 bp) of DNA efficiently. The longer the fragment size, the greater the chance of non-specific mutations resulting within the cDNA copy. In a reverse genetic system this usually results in the generation of a cDNA clone that cannot successfully produce virus. After mutagenesis, increasing the DNA concentration via propagation within a bacterial cell is required. Following isolation of plasmid DNA the insert fragment is removed from the vector using the same restriction enzymes used to clone the insert. Once the insert is isolated using DEAE cellulose it can be re-introduced into the original Ebo-Rib clone. The construct is propagated again in bacteria to increase DNA concentrations for rescue attempts.

3.2.2 Results and Discussion

3.2.2a Steps in cell free cloning:

Cell free cloning, (Figure. 15) which we have termed this new cloning method, utilizes normal procedures in cloning with various modifications of currently used protocols and buffering systems. This new protocol was developed to allow mutagenesis in large cDNA plasmids, without the need for propagation in bacteria and thus abolished the need for sub-cloning into a replicating plasmid. This decreases the chance of spontaneous mutations when generating cDNA constructs.

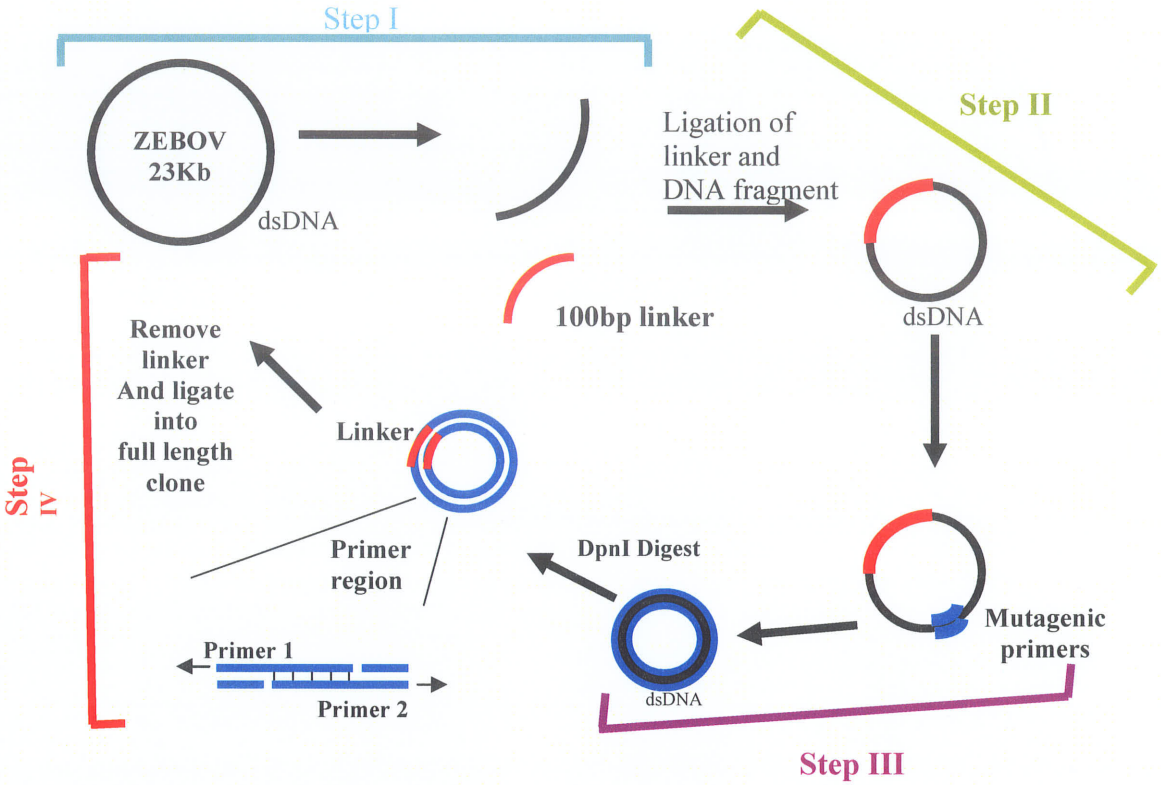


Figure. 15. Schematic diagram of 'Cell free cloning'. Step 1) Using unique restriction sites within the ZEBOV cDNA construct, digestion of a fragment with isolation and purification of both band fragments. The fragment and the remaining cDNA construct **Step 2)** Ligation of a DNA fragment to a linker piece of DNA which carries restriction sites at both the 3' and 5' ends. **Step 3)** Using PCR based mutagenesis, mutations were made with a region of interest within the fragment. Removal of parental DNA was carried out by the addition of DpnI. **Step 4)** Removal of the linker from the mutated fragment using the same restriction enzymes used in the excision. Once isolated the fragment can be reinserted back into the ZEBOV full-length clone. Patent number # IP 80.02.508

This system utilizes conventional cloning techniques with the help of a DNA linker (a small piece of DNA with multiple restriction enzyme recognition sites) and different buffer concentrations/components to digest, ligate, mutate, re-digest, and re-ligate the

DNA fragment in a matter of a few steps. The first step in cell-free cloning involves the restriction digest of the cDNA plasmid using two enzymes that will allow for the removal of the desired region. The second step involves the ligation of the cDNA fragment to a linker piece of DNA generating a circular fragment. Step three involves site-directed mutagenesis of the fragment of interest using normal PCR based mutagenesis. Once the mutations are generated, the parental cDNA is removed using the DpnI restriction enzyme that only digests methylated DNA or DNA which has been propagated in bacteria. Step four involves removing the linker from the cDNA fragment using the same restriction enzymes that were used to ligate them together. Finally, purification of the fragment for final ligation back into the full length cDNA replicating plasmid is carried out. As illustrated above, both systems work for generating mutant cDNA viral genome copies, but current methods are more time consuming and require sub-cloning steps which increase the possibility of spontaneous mutations. The greatest advantage of cell-free cloning is that it can be used to generate mutants very quickly by eliminating the sub-cloning steps and combining the ligation and mutagenesis steps. Utilizing this new system we have generated numerous mutant constructs which are devoid of polymerase error mutations.

In developing this cloning system we first determined where some key problems existed in normal cloning methods which were used in the development of mutant constructs (Figure. 14). One main problem was identified with the propagation of large plasmids in *E.coli* cells and the time involved in generating mutations in the reverse genetic system clones. The schematic in Figure. 15 represents the experimental design for cell free cloning. This method was developed to allow mutagenesis in large cDNA

plasmids, without the need for propagation in bacteria and abolishing the need for cassette building. This system utilizes conventional cloning techniques with the help of a linker along with different buffer concentrations/components to digest, ligate, mutate, re-digest, and re-ligate the DNA fragment in a few steps.

Step I: The first step in linker cloning is to restriction digest the full length cDNA construct using unique restriction enzymes. In these experiments the generation of a Ebo-Rib clone which lacks the KpnI site at 14677 was used. This clone was developed using the sub-cloning vector pBluescriptSK+'EcoRV'5220' (Figure. 16), which was developed by restriction digest of Ebo-Rib with EcoRV (Figure. 17), gel extracted, and cloned into pBluescriptSK+. Using this clone, site directed mutagenesis was carried out to generate a KpnI'14677' deletion mutant using the following primers.

KpnID'14677'f: CATTGGCACTAGCAGTACCGCAGGTGCTTGGAGGG

KpnID'14677'r: CCCTCCAAGCACCTGCGGTACTGCTAGTGCCAATG

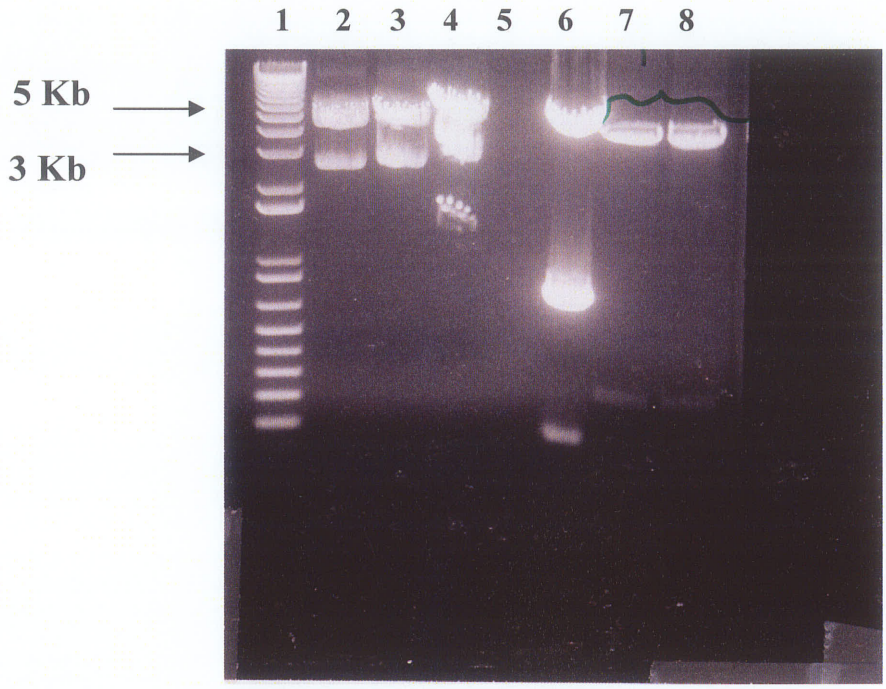


Figure. 16. Generation of the sub-cloning vector pBlue'EcoRV'5220'. This clone was generated by digestion of the full length cDNA construct Ebo-Rib with EcoRV. Isolating the band at 5220 and sub-cloning it into pBlueScript SK+. **Lane 1)** 1Kb+ ladder (Invitrogen), **Lanes 2-4)** contain three clones of pBlue'EcoRV'5220, digest with EcoRV lane 1 was used for further experiments. **Lane 5)** negative control, **Lane 6)** PCR amplification and cloning of GFP open reading frame into TOPO 2.1 sample was cut with SacI to verify restriction enzyme cutting. **Lanes 7, 8)** digestion of pCAGGS with SacI for GFP insertion.

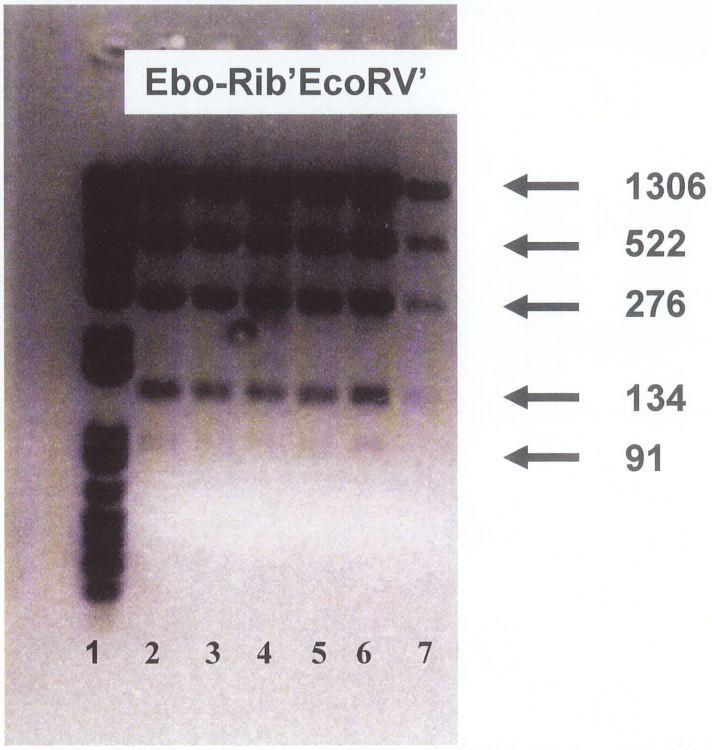


Figure 17. Restriction digest of Ebo-Rib with EcoRV. Lane 1) Marker 1Kb ladder (Invitrogen), Lanes 1-7) Ebo-Rib cDNA construct was restriction digested with EcoRV, Once digested, samples were electrophoresed using a 1 % agarose gel. Resulting bands were located at 13069, 5221, 2762, 1346, and 915 bp

After mutagenesis, the EcoRV fragment was re-cloned into Ebo-Rib. To verify the full length clone carried the newly introduced mutation, a digestion of the Ebo-RibΔ(KpnI'14677bp') clones with KpnI and BssSI (Figure. 18) was carried out. This band (3.5 kb long) was agarose gel extracted using DEAE cellulose (136) for the ligation reaction in step 2. This fragment corresponds to a region within Ebo-Rib that carries most of the glycoprotein and VP30 open reading frames.

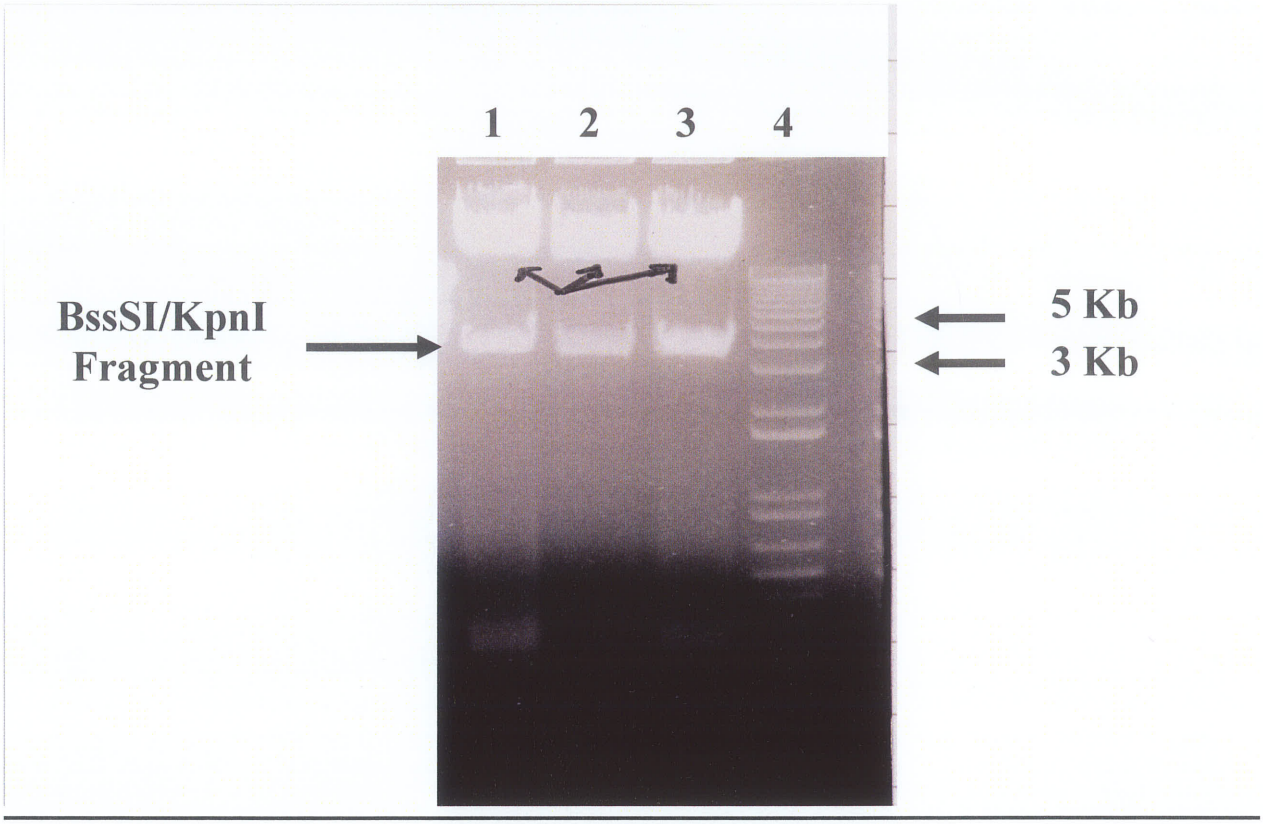


Figure. 18. Agarose electrophoresis gel of Ebo-Rib(KpnI'14677bp' deletion). Lanes 1-3. Restriction digest of Ebo-Rib with BssSI and KpnI. This digestion excised a fragment which corresponds to the 3.3 Kb band. Both bands were removed from the agarose gel and purified using Qiagen DNA gel extraction kit. Lane 4. 1 Kb+ ladder (Invitrogen). Arrows indicate the 5Kb and 3Kb bands from the 1Kb ladder.

Step 2: The KpnI/BssSI fragment was ligated using 5U of Roche ligase to a linker DNA fragment with corresponding restriction enzyme sites. The linker fragment used in this experimental assay was a 106 bp fragment, which contained the restriction enzymes BssSI and KpnI on 5' and 3' ends respectfully (Figure. 19). This linker allows for the circularization of the KpnI/BssSI fragment in preparation for the mutagenesis of the plasmid.

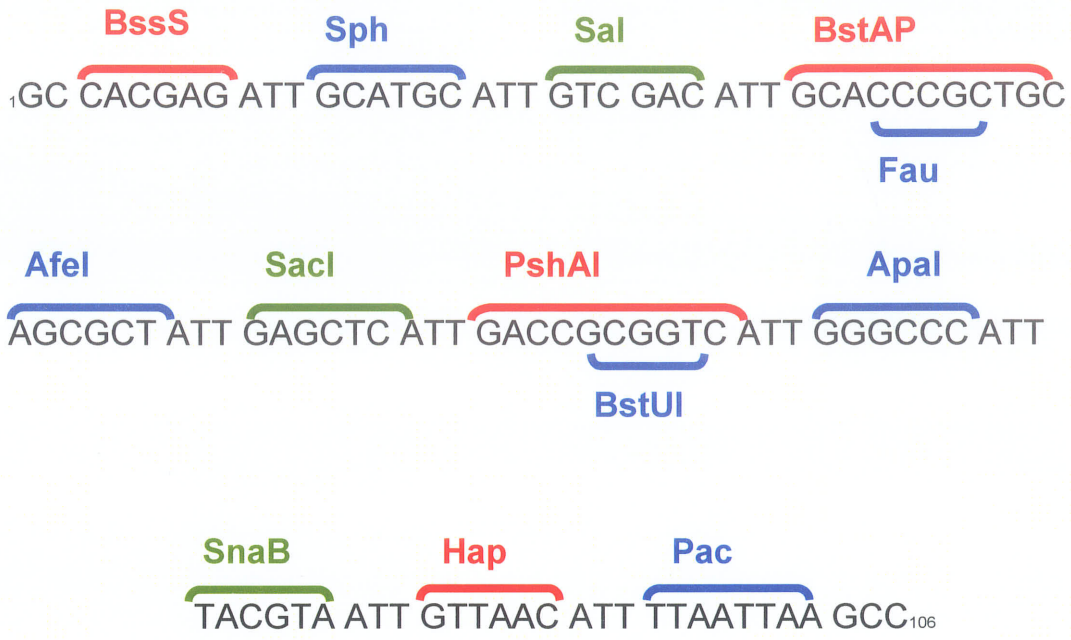


Figure. 19. Linker design for cell free cloning. Illustrated in color are restriction recognition sites for digestion of this linker. This linker was cloned into the multiple cloning site of TOPO-2.1. This was used as a vector to allow for the propagation of the linker DNA as well as allowing for the proper removal and thus restriction site presentation of the linker to the fragment.

Step 3: The ligated linker and KpnI/BssSI fragment was used to generate mutations within the glycoprotein region of the plasmid. Using the following primer sets that were generated four glycoprotein mutant constructs using site directed mutagenesis.

sGP'6915'f	ACTAAAAAAACCTGACTAGAAAAATTCG
sGP'6915'r	CGAATTTTTCTAGTCAGGTTTTTTTAGT
Δ-peptide'7017'f	CTTCCGACCCAGTAGCCAACACAACAAC
Δ-peptide'7017'r	GTTGTTGTGTTGGCTACTGGGTCGGAAG
Editing'6908'f	TGGGAAACTAAGAAGAACCTCACTAGAA

Editing'6908'r TTCTAGTGAGGTTCTTCTTAGTTTCCCA

Cleavage'7520'f GGCGGGGCAGGTACTCGCAGCAGAAGCA

Cleavage'7520'r TGCTTCTGCTGCGAGTACCTGCCCCGCC

Mutagenic primers were extended using PFU turbo (Stratagene). Once the PCR reaction was completed the parental DNA was removed by adding 1µl of DpnI (10 U/µl), and incubated at 37°C for 1 hour. After 1 hour the temperature was raised to 65°C to inactivate any enzyme. Since buffer 2 is not removed and ligase inactivation showed varying results, the following components were added to ensure the highest fidelity of mutagenesis. Using 1 µg of DNA the following reaction components were added. PCR primers were added at a concentration of 250 ng with 2.5 unit (or 1 µl) of PFU turbo polymerase enzyme and 2 ul dNTPs (100 µg/µl) per 50 µl volume. The following ions were added to stabilize the DNA and ensure divalent cation activity was at its maximum for polymerase activity. (10X reaction buffer, 20 mM $MgCl_2$, 100 mM KCL, 100 mM $(NH_4)_2SO_4$)

Step 4: In the last step the linker was removed from the DNA fragment using the KpnI/BssSI restriction enzymes and the fragment was purified using DEAE cellulose (136). The mutagenized fragment concentrations were determined (Nanodrop) and ligations were carried out with the Ebo-RibΔ(KpnI'14677bp') cut with KpnI/BssSI. Ligations were set up in vector to insert ratios of 1:5 and 1:10. Ligation reactions were transformed into XL1 Blue cells and grown on LB-AMP plates at 30°C. DNA was isolated (Mini-prep kit Qiagen) from transformed colonies and digested with EcoRV to ensure proper full length orientation was seen (Figure. 20).

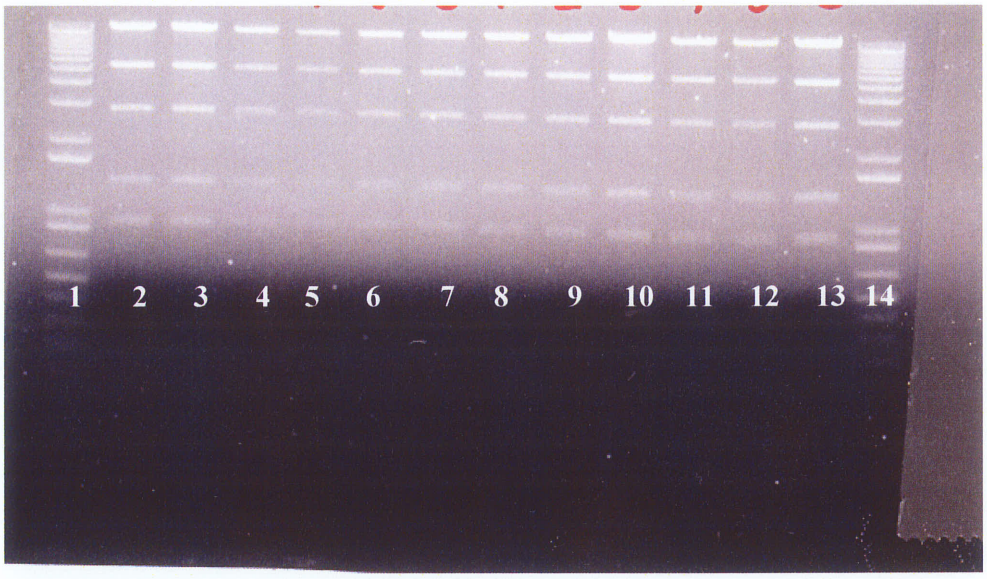


Figure. 20. Glycoprotein fragment religation into ZEBOV cDNA construct using cell free cloning method. Lanes 1 and 14) 1Kb ladder (Invitrogen). All positive religations will result in a five band pattern when a complete ZEBOV cDNA construct is digested with EcoRV. Lanes 2-4) sGP mutant within the Ebo-Rib backbone digested with EcoRV. Lanes 5-7) Delta-peptide mutant within the Ebo-Rib backbone digested with EcoRV. Lane 8-10) Editing site mutation within the glycoprotein region within the Ebo-Rib backbone digested with EcoRV. Lanes 11-13) Glycoprotein proteolytic cleavage site mutation within the Ebo-Rib backbone digested with EcoRV.

Positive clones were prepared and sent for in-house sequencing (Figure. 21). Once sequences were confirmed mutant ZEBOV cDNA constructs were transformed into eukaryotic cells Vero E6/293T along with Ebo-Rib (full length) and rescue was examined for 14 days.

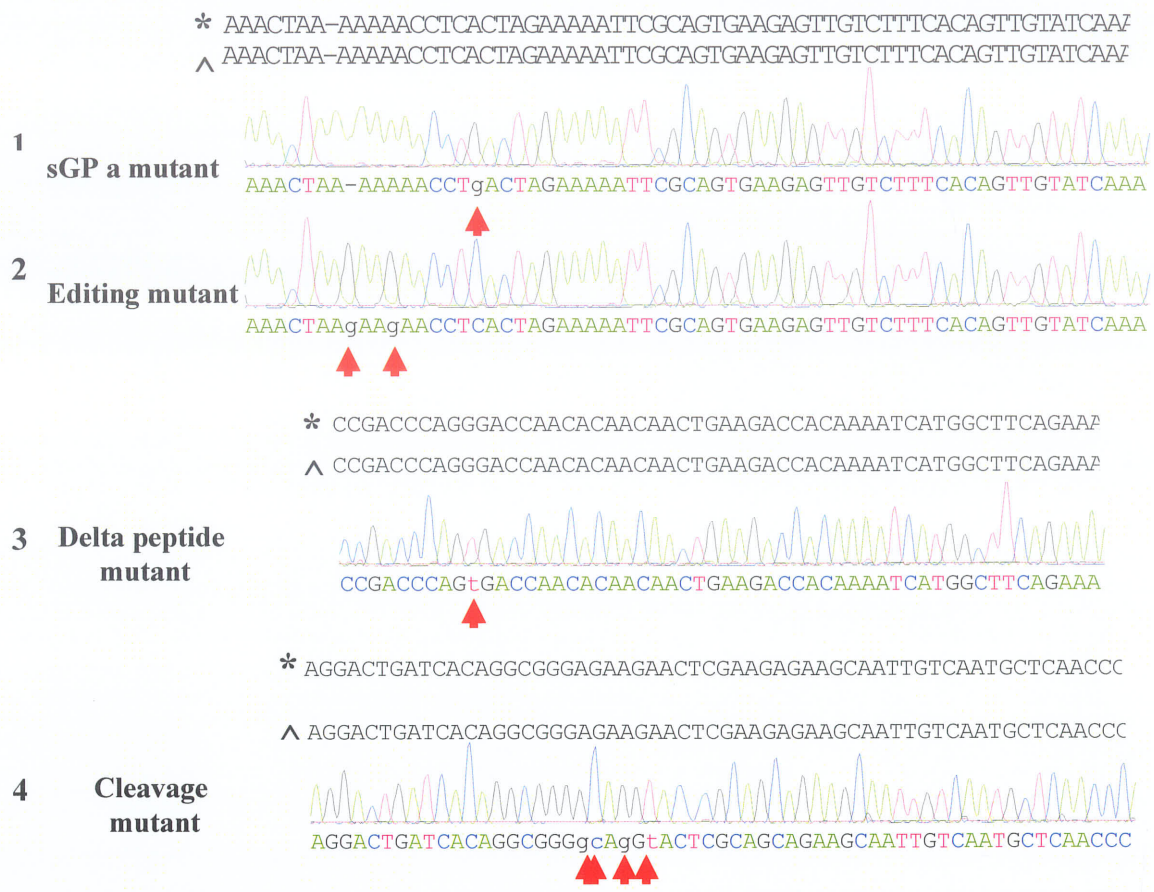


Figure. 21. DNA sequence of glycoprotein mutants generated using cell free cloning. Sequence of ZEBOV 76 strain Mayinga, and recombinant ZEBOV virus (Ebo-Rib) were used as reference sequences. Mutations which were generated within the glycoprotein are illustrated by red arrows. **1)** A stop codon generated within the sGP region. This will abolish the production of sGP during glycoprotein transcription. **2)** Editing mutation generated by the insertion of two guanine residues in place of two adenine residues, disrupting normal glycoprotein processing. **3)** A stop codon was generated within the delta peptide region of the glycoprotein, truncating the peptide possibly disrupting it function **4)** A Cleavage mutation was generated by the insertion of multiple nucleotides disrupting the binding motif for the proteolytic enzyme furin. **Key:** '*' represents the ZEBOV 76 viral sequence as well as '^' which represents the ZEBOV reverse genetic system recombinant virus.

Once positive rescue was assessed by cytopathic effect and indirect immunofluorescence assay (anti-VP40) virus stocks were grown and virus titers were determined using an immunoplaque focus forming unit assay (Figure. 22).

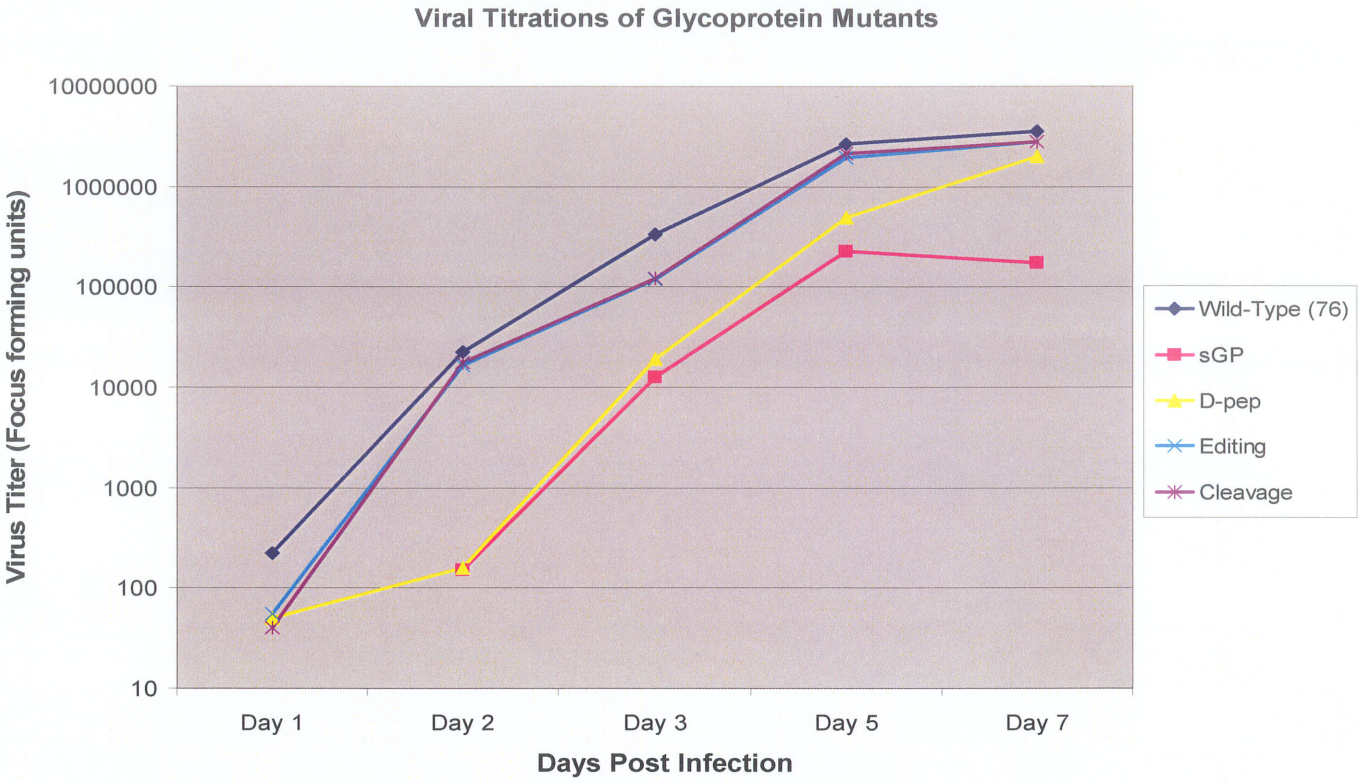


Figure. 22. Viral titration of glycoprotein mutants. Titrations were performed in biocontainment level 4 in Vero E6 cells with samples being assayed for viral titres on days 1-3, 5, and 7. Titrations were performed using a 10-fold dilution series of viral stocks, 1 ml of each dilution series were used to infect Vero E6 cells in a 96 well plate. Infected cells were overlaid with DMEM containing 1.5% carboxymethyl-cellulose (CMC) and 2% (v/v) fetal calf serum. Indirect immunofluorescence assay was performed on formalin-fixed cells (2%) 5 to 7 days post infection. For immunodetection a rabbit antiserum directed against ZEBOV VP40 (dilution 1:200 in PBS) with a FITC anti-rabbit secondary antibody (dilution 1:500 in PBS) was used for fluorescence. Infected cells were analyzed using an Axioplan 2 fluorescence microscope, titers were determined by counting positive fluorescence plagues (Focus forming unit). Four glycoprotein mutants were assayed (sGP, delta peptide, editing site mutant, and cleavage site mutant).

Recent developments in mutagenesis technologies now allow for the mutagenesis of specific regions within viral cDNA copies. The manipulation of viral receptors and viral pathogenic factors are currently being studied, in systems for members of the negative-stranded RNA virus families *Orthomyxoviridae, Bunyaviridae, Arenaviridae,*

Rhabdoviridae, Paramyxoviridae, Filoviridae, and *Bornaviridae* (114, 157). Inherent problems exist within all negative stranded RNA groups when generating mutations within the large cDNA constructs and many methods have been adopted to increase the probability of successful clone development. Within this project we have demonstrated a new method of mutant clone development named 'cell free cloning'. As the name indicates, this method of cloning lacks the requirement of multiple clonal selection steps which decreases the chance of spontaneous mutations and plasmid rejection. This new method of cloning, when used to make reverse genetic mutants provides tremendous advantages in the time it takes to generate clones and the reduction of spontaneous mutation rates.

Defining the pathogenic effects of ZEBOV glycoprotein products has been a focal point in Ebola research for the past decade. With this in mind, and the need to test the new cloning method of 'cell free cloning', we developed four cDNA viral clones encoding mutations within the glycoprotein gene. Utilizing the ZEBOV reverse genetic system and cell free cloning, sGP and Δ-peptide mutants were generated by inserting stop codons in the open reading frame of the gene products. This would abolish the production of sGP and Δ-peptide. Editing and cleavage site mutants were also generate which, would change the (Figure. 21). Interestingly, all mutant viruses grew to a fairly high titer ~10^6 - 10^7 (FFU), with the exception of sGP which could only be grown to a titer of 10^5 (FFU) (Figure. 22). In generating these glycoprotein mutant clones we have developed tools which could assay the physiological difference between wild-type and mutant viruses both in biology and pathogenesis.

3.2.3 Future studies

Since the generation of the glycoprotein mutants was achieved to assess the ability of cell free cloning, there are many aspects to study using these mutant viruses, including:

1. Assessment of mutant viruses for changes in levels of production of mutated protein.
2. Assessment of mutants *in vivo* (immunodeficient mice)

As my studies progressed, it soon became apparent that developing glycoprotein mutants in the ZEBOV' 76 (strain Mayinga) reverse genetic system limited us to experiments using tissue culture methodologies, immunodeficient mice, or primates (*Cynomolgus macaque*). We realized this inherent problem could be resolved by developing a reverse genetic system for a small animal model. Which would allow for pathogenic studies to be carried out using a small animal model which closely resembles a EHF infection in humans. We utilized the system described above to generate a reverse genetic clone for the ZEBOV' 76 strain Mayinga guinea pig adapted virus described in the following chapter.

3.3 Guinea Pig Adapted ZEBOV reverse genetic clone

3.3.1 Introduction

Historically guinea pigs have an interesting past with filoviruses. The first Marburg patient who was infected in 1967 was suspected to have either rickettsia or one of the leptospira species. Common practices for diagnosis of these diseases in the 1960's was to injected Dunkin-Hartly guinea pigs with patient blood and monitor for clinical signs of either disease. Interestingly, the guinea pigs developed a fever, lost weight and soon recovered (150). These results as well as electron microscopy results demonstrated that this was a new virus. With these results, passaging of the virus within guinea pigs was carried out. It was determined that three passages of the filovirus within guinea pigs resulted in a virus that caused 100% lethality (150).

A few decades later, increases in the number of animal models and newly developed techniques have advanced the study of filoviral infection in guinea pigs. Currently, little is known about the pathogenesis of filovirus infections. Previous studies of natural human outbreaks or of experimentally infected primates have described the clinical signs, histopathologic lesions, pathophysiologic changes, coagulopathies, and viral antigen distribution (9, 16, 51). These investigations have shed little light on the pathogenesis of filovirus infection before death. It has become apparent during the last decade of study in the filoviral field that a reliable small animal model is needed to determine the physiological effects of filoviruses infections.

The adaptation of Ebola subtype Zaire (Mayinga) to produce a lethal infection in guinea pigs was carried out by serial passage of the virus within guinea pig spleens. To achieve uniform lethality, guinea pigs were inoculated subcutaneously with a passage one

virus from a human serum specimen (057931). Seven days after inoculation guinea pigs were euthanized and their spleens were homogenized and used as an inoculum for infection of more guinea pigs. This process was repeated until uniform lethality was achieved (28).

Hypothesis and objectives of study:

Guinea pig adaptation is based on specific mutations in the genes encoding interferon antagonistic gene products such as VP35 and VP24. Our hypothesis was to determine if these mutations are required for interferon inhibition allowing for the infection of a new host. Our main objectives were to sequence the guinea pig adapted virus and develop a reverse genetic system based on that sequence. Once complete we could test the specific mutations needed for adaptation to the new host. Since filoviruses normally do not cause disease in guinea pigs, one or all of the mutations found in the guinea pig adapted virus could be needed to adapt to its new host. With this in mind we will develop single, double, and triple mutations of the guinea pig adapted changes, using the ZEBOV guinea pig reverse genetic system.

3.3.2 Results and Discussion

3.3.2a Sequencing of the guinea pig adapted virus

Vero E6 cells were infected with the guinea pig adapted virus until, it has caused apoptosis in 60% of the cells infected, (2+ cytopathic effect), which usually occurs between days 7-10 p.i.. Viral RNA was isolated using Trizol LS extraction, and RNA concentration was determined using NanoDrop ND-1000 UV-Vis spectrophotometer (NanoDrop Technologies). To determine the entire sequence of the guinea pig adapted

genome we subjected the RNA to complete nucleotide sequence analysis. Using the primary sequence of the full-length genomic RNA for ZEBOV (GenBank Accession No.AF086833) (165), we generated primers which were designed to generate 1 Kb fragments covering the complete genome of the guinea pig adapted virus (See Appendix).

The complete nucleotide sequence was determined in triplicate by sequencing the RT-PCR fragments and compiling the sequence data in SeqMan (DNA Star) (See appendix for complete sequence). Sequencing demonstrated ten single mutations covering five genes, of which seven caused amino acid changes (Table 11).

Table 11:

Position in genome	Nucleotide change	Amino acid change
NP (2411)	T → C	Phe → Leu
VP40 (4437)	T → C	Ile → Thr
VP40 (5053)	C → T	Thr → Ile
GP (6424)	T → C	Ile → Thr
GP (7668)	T → C	Ile → Thr
VP24 (10420)	C → T	Leu → Phe
VP24 (10741)	C → A	His → Asn
L (13053)	G → A	Arg → Arg
L (13968)	A → G	Ser → Ser
L (14040)	A → C	Thr → Thr

Table 11. Single nucleotide changes in the full sequence of the guinea pig adapted virus. Position within genome (see appendix) is in the first column, nucleotide and amino acid changes follow. Ten mutations were found within five of the genes encoding structural proteins within the ZEBOV genome. Three of these mutations were contained within the L region and were silent mutants with no amino acid change. Since silent mutations unusually have no effect on protein production or function we decided to omit these three mutations from the signal nucleotide change study.

3.3.2b Generation of reverse genetic guinea pig adapted cDNA construct:

To generate the full length guinea pig adapted reverse genetic clone, we first generated primers which encompassed regions containing mutations (Table 11), flanked by unique restriction sites (Figure. 23). Primers used to generate fragments are as follows:

SphI '2534'f GCTCAACCAGCCCTCGCATGCTGACACCAATTAACG

SphI '6159'r CAGATGTAAGCATGCAGGCAATTTGAGG

SalI '6566'f GGTTAGTGATGTCGACAAACTAG

SacI'11327'r CGGGTTCTTGGAGCTCCACCAGAAAACCC

SacII'12245' CCTGTTGAGCCGCGGTGCCAACAGTTC

HpaI'14626' GGTTTGCCGAGTGTTAACTGTCCAAGG

Using RNA from the guinea pig adapted virus we amplified three fragments designated 1C, 4C, and 5C using RT-PCR (QIAGEN) which were sub-cloned into the existing ZEBOV reverse genetic cDNA construct cut with the corresponding restriction enzymes (Figure. 23). Once all mutant fragments were generated and cloned into the Ebo-Rib backbone, we confirmed the sequence of each fragment and attempted to rescue the viruses using the ZEBOV reverse genetic system.

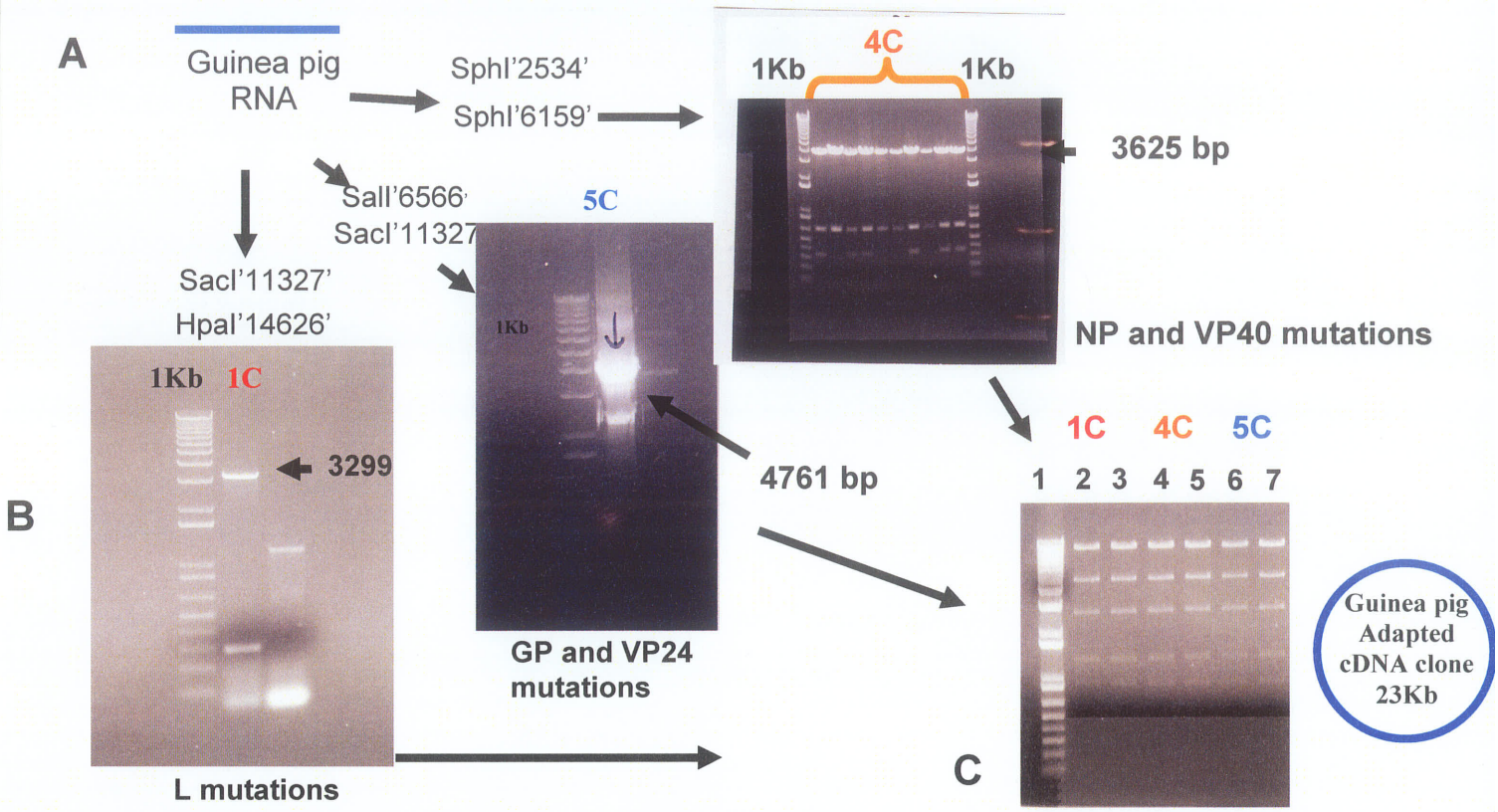


Figure. 23. The generation of the Guinea pig adapted cDNA construct. A) Using ZEBOV guinea pig adapted RNA derived form infected Vero E6 cells, three fragments were amplified using RT-PCR. **B)** Fragments were amplified using the listed primers, the corresponding bands were isolated using Qiagen gel extraction kit. **C)** These fragments were sub-cloned into the ZEBOV cDNA (Ebo-Rib) backbone using the prescribed restriction sites assembling the guinea pig adapted cDNA construct which, was confirmed by EcoRV restriction digest. **Key:** 1C = mutations found in the L region, 4C = mutations found in the GP and VP24 regions, 5C = mutations found in the NP and VP40 regions.

Transfections were carried out using 2µl *trans*IT®-LT1 (Mirus) transfection reagent per 1µg of the standardized DNA, using the optimized plasmid concentration (Figure. 24). Transfections were carried out using a 1:1 ratio of Vero E6 and 293T cells and incubated

at 37°C in the BSL4. Supernatants were collected three days post transfection and passaged onto 80 % confluent Vero E6 cells.

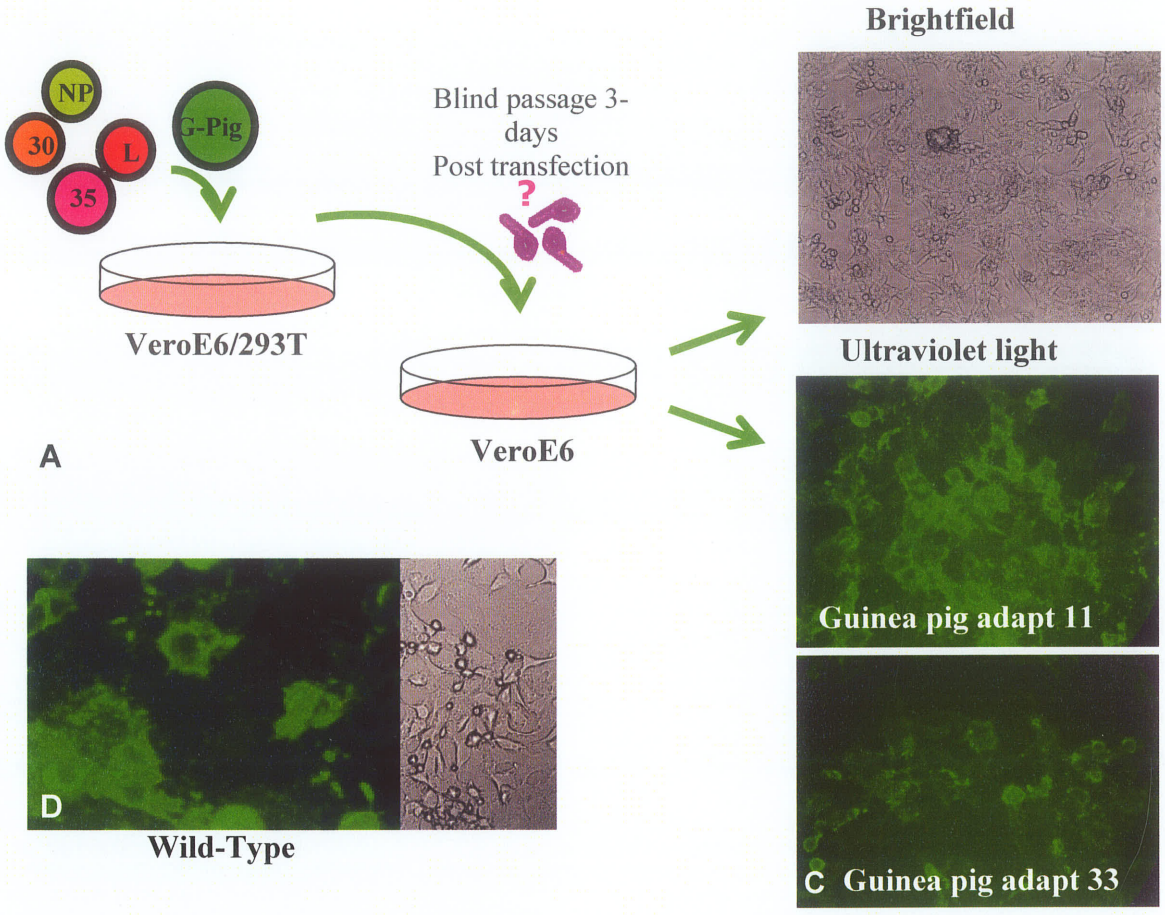


Figure. 24. Illustration of reverse genetic rescue of ZEBOV guinea pig adapted cDNA constructs. A) Reverse genetic rescue of ZEBOV guinea pig adapted viruses. The RNP complex proteins (NP, VP35, VP30, and L) were transfected into VeorE6/293T cells using the optimized plasmid ratios (1ug, 0.5ug, 0.3ug, and 1ug respectively) of DNA. **B)** Brightfield view of reverse genetic rescue 10-days post transfection. Cytopathic effect is seen demonstrating viral infection **C)** Indirect immunofluorescence assay of rescued ZEBOV guinea pig adapted viruses. Infectious particles were detected using FITC anti-rabbit (1:500 dilution in PBS) staining against Anti-VP40 (1:200 diluted in PBS). **D)** ZEBOV '76' reverse genetic system positive control, infectious particles were detected using indirect immunofluorescence assay against anti-VP40. Brightfield and Indirect immunofluorescence assay are presented.

Cells were monitored for 14 days for CPE. The guinea pig adapted rescue showed two clones which caused CPE on day 10 post passage (Figure. 24). Both viruses, G-pig adapt 11, and G-pig adapt 33, could grow to a viral titer between $10^6 - 10^7$ (FFU) (Figure. 25). Once rescued and titered, both newly developed viruses were assayed for their ability to infect and cause disease in guinea pigs.

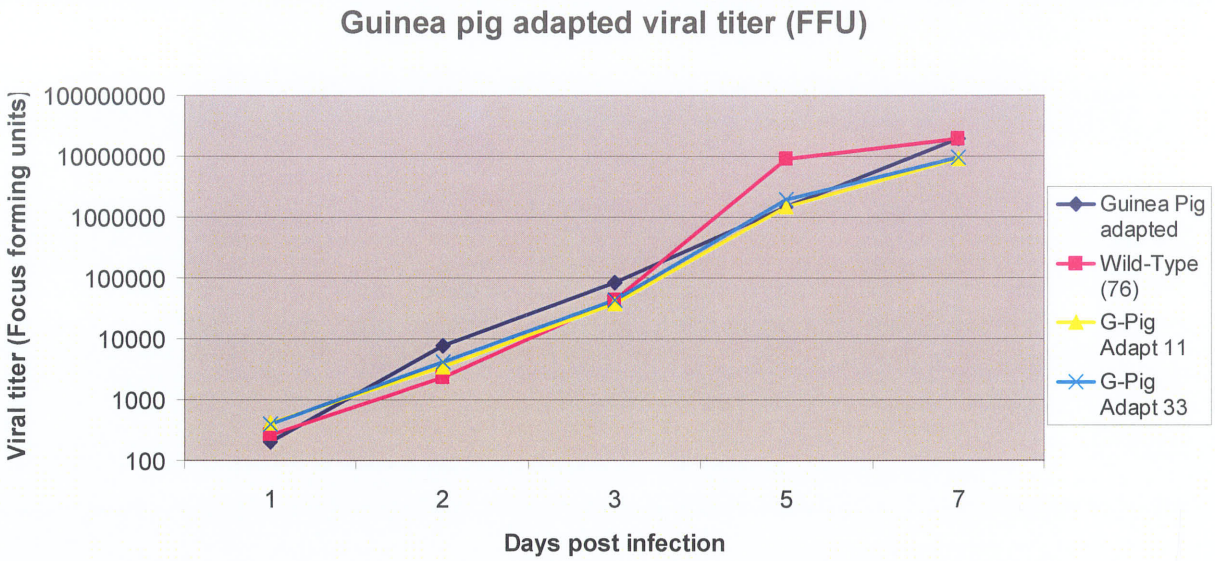


Figure. 25. In vitro titration of guinea pig adapted viruses. Kinetic growth studies guinea pig adapt 11, guinea pig adapt 33, guinea pig adapted, and ZEBOV 76 (wild-type). Viruses were titered in 96 well plates using Vero E6. Samples were assayed for viral titres on days 1- 3, 5, and 7. Viruses were propagated in biocontainment level 4, using a ten-fold dilution series of viral stocks, 1 ml of each dilution series were used to infect Vero E6 cells. Infected cells were overlaid with DMEM containing 1.5% carboxymethyl-cellulose (CMC) and 2% (v/v) fetal calf serum. Following an incubation of seven days, the infected cells were fixed and inactivated with 4% paraformaldehyde buffered in PBS for two days with one fixative exchange. Indirect immunofluorescence assay of infected cells using anti-VP40 (1:200 dilution in PBS) was carried out and fluorescence was produced with a FITC anti-rabbit antibody (1:500 dilution in PBS). Infected cells were analyzed using an Axioplan 2 fluorescence microscope, titers were determined by counting positive fluorescence plagues (Focus forming unit).

Three groups of six Dunkin Hartley Guinea pigs were infected with 1000 FFU of either G-Pig Adapt 11, G-Pig Adapt 33, or guinea pig adapted EBOV. Disease progression was monitored by weight and death rates were determined (Figure. 26).

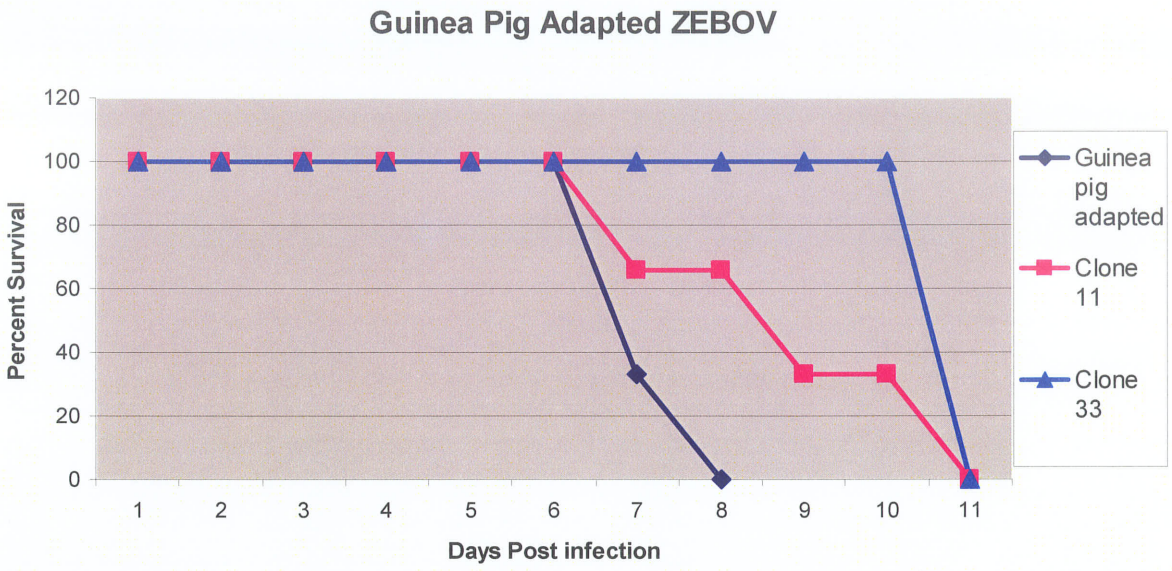


Figure. 26. Kaplan-Meier survival curve of virulence for the recombinant guinea pig adapted viruses. *In vivo* infection of six Dunkin-Hartley guinea pigs with 1000 FFU of either Guinea pig adapt 11, Guinea pig adapt 33, and guinea pig adapted virus. Guinea pigs were monitored for weight loss and lethality caused by virus infection. Illustrated above is the percentage survival rate for all three viruses. Blood and organs (spleen, liver, and lung) were removed to determine serology for infection and virus isolation for sequence comparison between input and output viruses. Positive RT-PCR from virus infection was positive in sera and organs for all three viruses.

The guinea pig adapted virus was seen to be uniformly lethal by day 8 post infection, where as both recombinant guinea pig adapted viruses demonstrated lethality by day 11 post infection. All animals followed the same general disease pattern, a decrease in water and food consumption resulting in weight loss and death, in the most severe cases a prolapsed rectum would occur (animals were euthanized) demonstrating DIC and internal hemorrhage. The difference in time between the adapted virus and the recombinant virus

producing death in this model may be explained by clonal selection which would have occurred in the adapted virus. Since a rescued virus will only have one virus type this could play a role in producing a slight lag in viral production.

3.3.2c Determination of key mutations for guinea pig conversion.

To determine the key changes in the nucleotide sequence of the guinea pig adapted virus, which may be responsible for adaptation, we performed single mutational analysis using seven of the ten mutations. Using Ebo-Rib sub-clones (Figure. 23) which correspond to sections 1C, 4C, and 5C, we generated single mutations by cell free cloning as described in Chapter Two. Briefly, using a linker piece of DNA which carried three unique restriction sites (SphI... SacII ...SalI) on the 5'and three (HpaI...SacI...SphI) on the 3' end separated by 90 nucleotides. The following primers (forward primers only) were used to generate fragments of DNA which were ligated to the above linker. Single mutations were generated within the fragments using 'cell free cloning' and PCR based site directed mutagenesis.

NP(2411) GTCAGAACACT**C**TCTTGAGGAGATGTATCGCC

VP40(4437) GAGAGTGTTTTTTCA**C**TAACCTTCATCTTG

VP40(5053) GCTGCAACATGGA**T**CGATGACACTCCAACAGG

GP(6424) CCAGACGGGA**C**TCGGGGCTTCCCCCGGTG

GP(7668) GCAGCCGAGGGAATTTACA**C**AGAGGGGCTAATGC

VP24(10420) GGTTGTCTTAAGCGAC**T**TCTGTAACTTCTTAG

VP24(10741) CCAACACTAAC**A**ATTTCAACARGCGAACAC

All single mutations were re-cloned into the Ebo-Rib backbone using unique restriction sits corresponding to sections 1C, 4C, and 5C (Figure. 27).

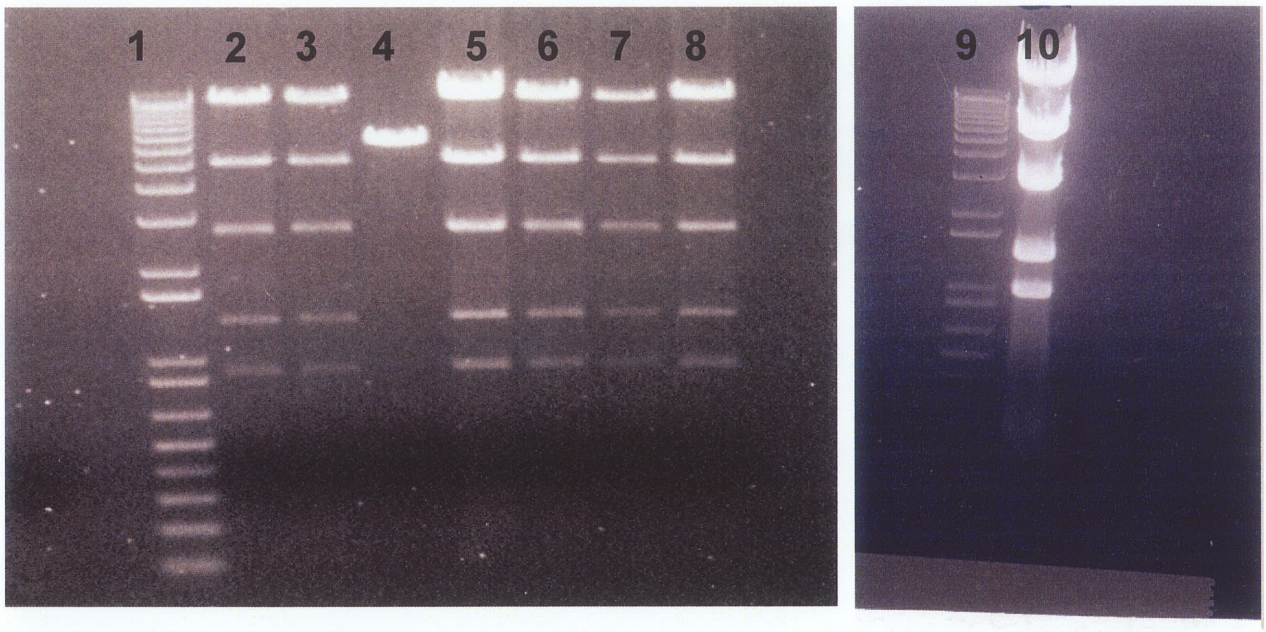


Figure. 27. A 1% agarose gel representing the re-insertion of all single mutations for the development of a recombinant guinea pig adapted virus. cDNA constructs were digested with EcoRV and visualized on an agarose gel. Lane 1,9) DNA Marker 1Kb ladder (Invitrogen) Lanes 2-8 demonstrates reinsertion of single mutagenized fragments for, (2) NP, (3) VP40(4437), (4) unsuccessful reinsertion, (5) GP(6425), (6) GP(7669), (7) VP24(10421), (8) VP24(10767) (10) VP40(5053) back into the Ebo-Rib backbone for analysis.

Rescue attempts were made as described in Figure. 24 using the ZEBOV reverse genetics system. Once rescued, samples were assayed for the production of ZEBOV VP40 (Figure. 28), and titered using a Focus Forming Unit assay described in Chapter 2 (Figure. 29).

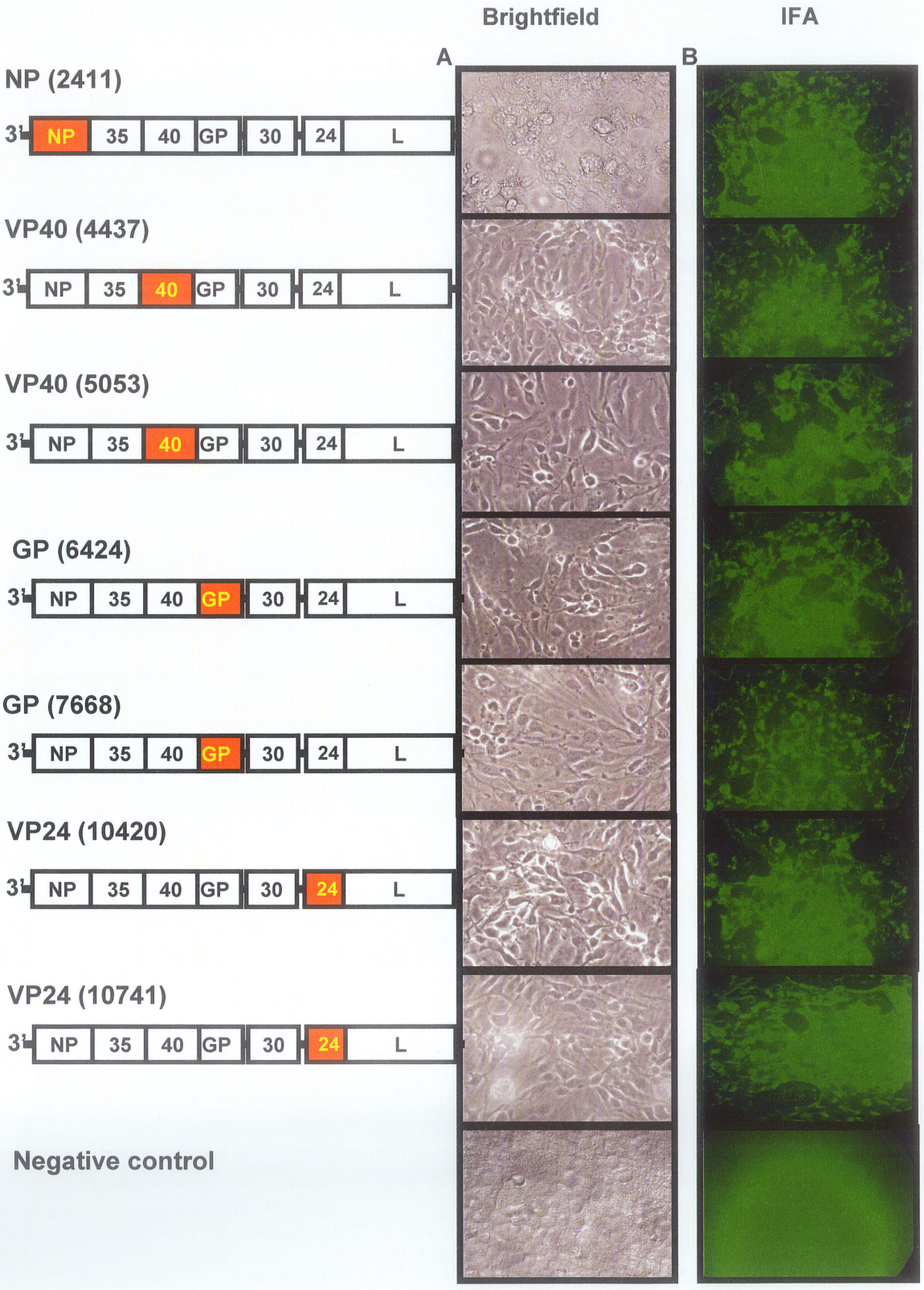

Brightfield
IFA
A
B
NP (2411)
3'
NP
35
40
GP
30
24
L
VP40 (4437)
VP40 (5053)
GP (6424)
GP (7668)
VP24 (10420)
VP24 (10741)
Negative control

Figure. 28. Development of single guinea pig adapted nucleotide changes within the reverse genetic system. Positive rescue after the development of (NP, VP40(4437), VP40(5053), GP(6424), GP(7668), VP24(10420), and VP24(10741) single guinea pig adapted mutations was determined by indirect immunofluorescence assay. All single mutations were rescued using the ZEBOV reverse genetic system demonstrated by the positive immunofluorescence results which were detected using FITC staining (1:500 dilution of antibody in PBS) against anti-VP40 (1:200 dilution in PBS) in column B. Column A represents a bright field view verifying cell confluence and cytopathic effect. Column A and B represent the same infected plate with pictures taken at different locations.

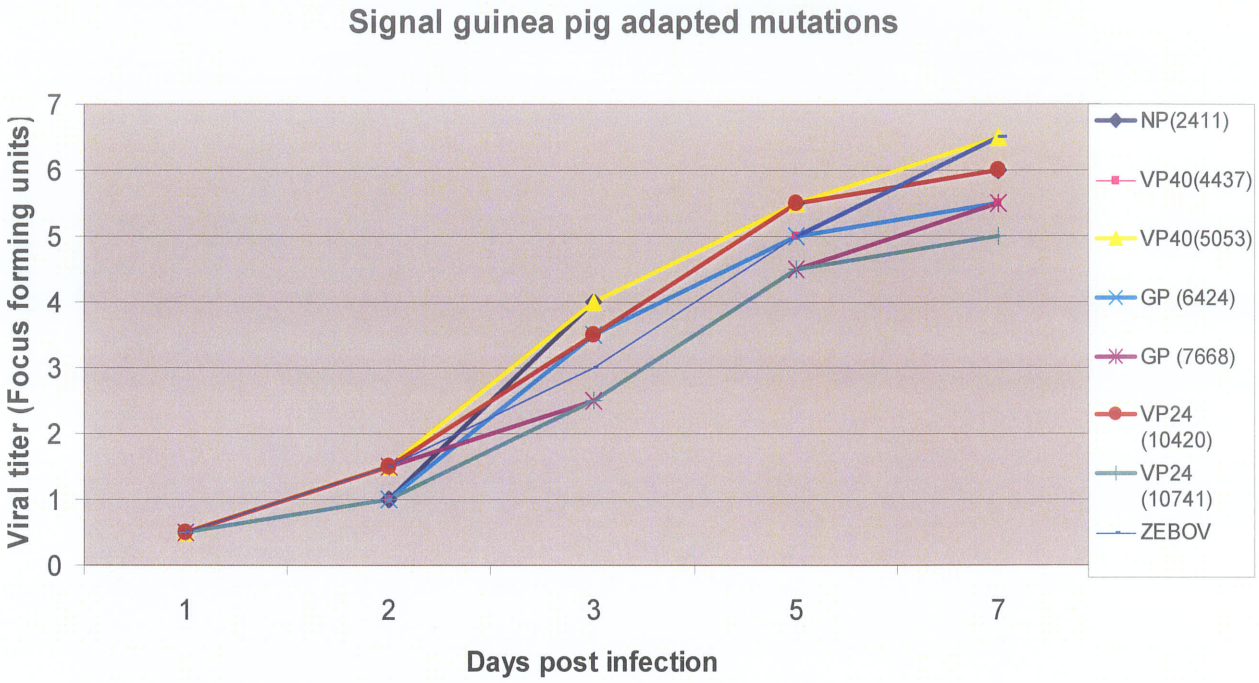


Figure. 29. Kinetic growth studies of all single guinea pig adapted mutations within the ZEBOV reverse genetic backbone. Viruses were propagated in Vero E6 cells with samples being assayed for viral titres on days 1-3, 5, and 7. Titrations were performed using ten-fold dilution series of viral stocks, 1 ml of each dilution series were used to infect Vero E6 cells. Infected cells were overlaid with DMEM containing 1.5 % (v/v) carboxymethyl-cellulose (CMC) and 2 % (v/v) fetal calf serum. Following an incubation of seven days, the infected cells were fixed and inactivated with 4% paraformaldehyde buffered in PBS for two days with one fixative exchange. Indirect immunofluorescence assay of infected cells using anti-VP40 (1:200 dilution in PBS) and FITC anti-rabbit staining (1:500 dilution in PBS) was used for fluorescence detection. Infected cells were analyzed using an Axioplan 2 fluorescence microscope, titers were determined by counting positive fluorescence plagues (Focus forming unit).

Viruses were re-sequenced to ensure mutation were still present in the infectious virus.

All single mutations were tested for infectivity and lethality in guinea pigs. To assess this six guinea pigs were injected with 1000 FFU of one of the seven single mutations. Animals were weighed for the first 14 days and monitored for a total of 21 days after

infection (Figure. 30). All animals infected with single guinea pig mutations survived with no weight loss for the full 21 days.

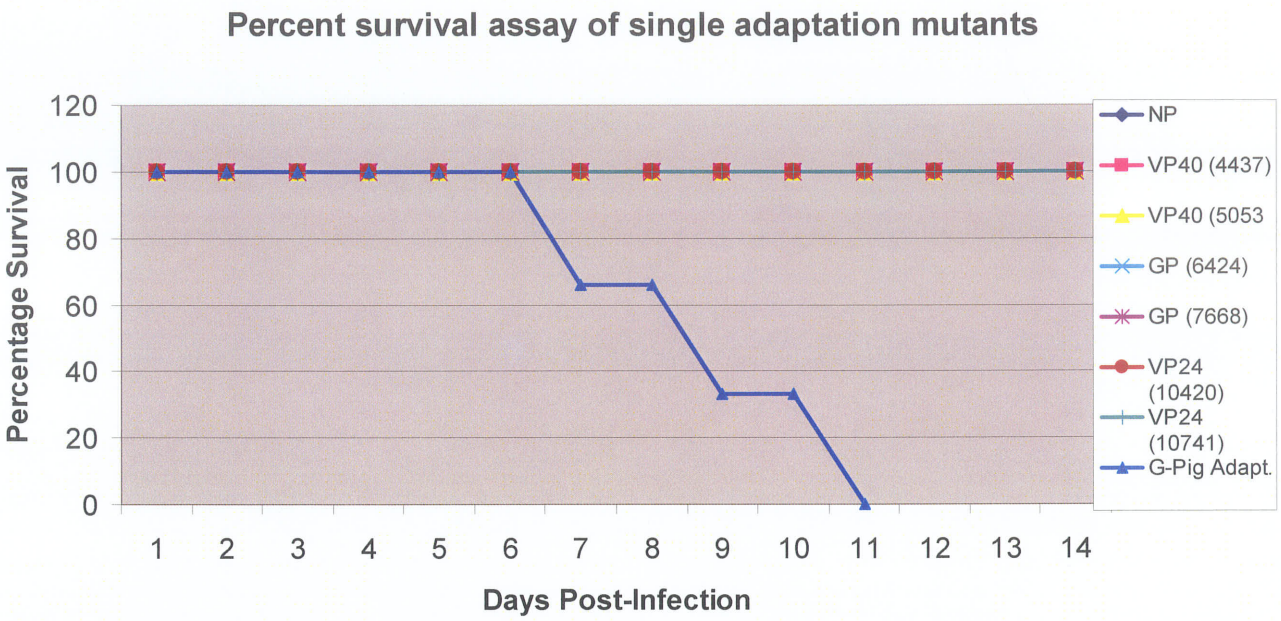


Figure. 30. Kaplan-Meier survival curve for single guinea pig adapted mutants. *In vivo* infection of six Dunkin-Hartley guinea pigs with 1000 FFU, of all seven single adapting nucleotide changes and the guinea pig adapted virus. Guinea pigs were monitored for weight loss and lethality caused by virus infection. Illustrated above is the percentage survival rate for all seven viruses and guinea pig adapted. Blood and organs (spleen, liver, and lung) were removed to determine serology from the infection. RT-PCR from virus infections was negative in sera for all seven single adapting viruses.

Double mutations were generated as described above, using the single mutation sub-clones. We generated four double mutants and one triple mutant (Table 12) to help define the nucleotides which allow for the adaptation of ZEBOV to guinea pigs.

Table 12:

Mutant	Virus rescue	Virus titer
NP, VP24(10420)	Day 9	10^6
NP, VP24(10741)	Day 10	$10^{5.3}$
GP(6424),GP(7668)	Day 9	$10^{5.3}$
VP24(10420), VP24(10741)	Day 8	10^6
NP,VP24(10420), VP24(10741)	Day 10	$10^{5.5}$

Table 12. To define the nucleotides responsible for the conversion from ZEBOV-wild-type to guinea pig adapted ZEBOV. We generated five mutant variations containing two or three guinea pig nucleotide conversions. Virus constructs were generated using RT-PCR and cell free cloning, transfection were carried out using optimized rescue transfections (see material and methods). Viral titers were carried out in BSL-4; briefly, samples were diluted in ten-fold increments, and were used to infect a 96 well plate of 80% confluent Vero E6 cells. All wells were cover with DMEM containing 1.5 % carboxymethyl-cellulose (CMC) and 2 % (v/v) fetal calf serum. Following an incubation of seven days, the infected cells were fixed and inactivated with 4% paraformaldehyde buffered in PBS for two days with one fixative exchange. Indirect immunofluorescence assay was used to detect infectious particles using FITC anti-rabbit (1:500 dilution in PBS) staining against Anti-VP40 (1:200 diluted in PBS). Infected cells were analyzed using an Axioplan 2 fluorescence microscope, titers were determined by counting positive fluorescence plagues (Focus forming unit).

All of the double mutant viruses could be rescued (Figures. 31) and could grow to a high titer (Figure. 32) in Vero E6 cells.

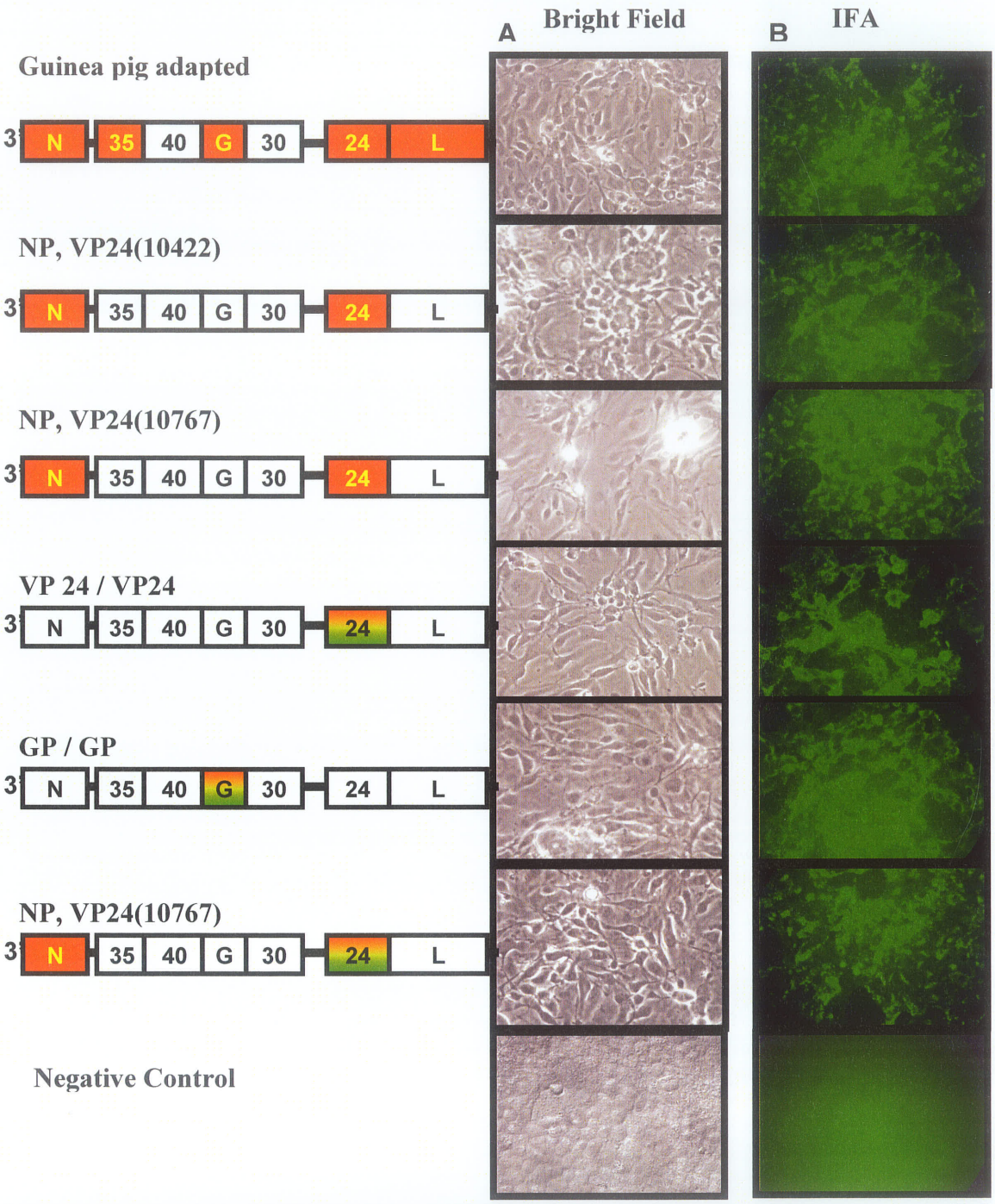

A
Bright Field
B
IFA
Guinea pig adapted
3'
N
35
40
G
30
24
L
NP, VP24(10422)
NP, VP24(10767)
VP 24 / VP24
GP / GP
NP, VP24(10767)
Negative Control

Figure. 31. Development of multiple guinea pig adapted nucleotide changes within the ZEBOV reverse genetic backbone. Positive rescue identification for all mutations was carried out by indirect immunofluorescence assay using anti-VP40 (dilution 1:200 in PBS). Fluorescence was generated using FITC labelled antibody (1:500 dilution in PBS). All mutations were rescued using the ZEBOV reverse genetic system demonstrated by the positive immunofluorescence results in column B. Column A represents a bright field view verifying cell confluence and cytopathic effect. Column A and B represent the same infected plate with pictures taken at different locations within the plate.

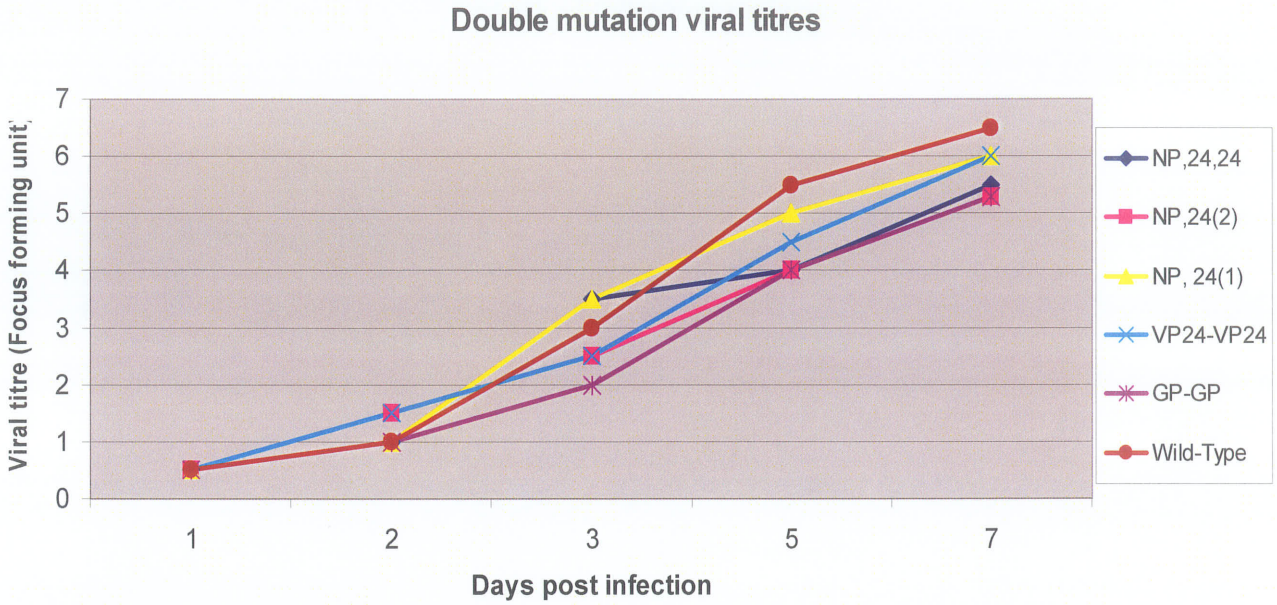


Figure. 32. Kinetic growth studies of multiple guinea pig adapted mutations within the ZEBOV reverse genetic backbone. Viruses were propagated in Vero E6 cells with samples being assayed for viral titres on days 1-3, 5, and 7. Infections was performed using a ten-fold dilution series of viral stocks, 1 ml of each dilution series was used to infect 80% confluent Vero E6 cells. Infected cells were overlaid with DMEM containing 1.5 % (v/v) carboxymethyl-cellulose (CMC) and 2 % (v/v) fetal calf serum. Following an incubation of seven days, the infected cells were fixed and inactivated with 4% paraformaldehyde buffered in PBS for two days with one fixative exchange. Indirect immunofluorescence assay of rescued ZEBOV guinea pig adapted viruses. Infectious particles were detected using FITC anti-rabbit (1:500 dilution in PBS) staining against Anti-VP40 (1:200 diluted in PBS). Infected cells were analyzed using an Axioplan 2 fluorescence microscope, titers were determined by counting positive fluorescence plagues (Focus forming unit).

Using these viruses we injected six groups of six Dunkin-Hartley guinea pigs with 1000 FFU of the viruses listed in Table 12. Guinea pigs were weighed for the first 14 days and monitored for 21 days after infection (Figure. 33). Generally, guinea pigs infected with the guinea pig adapted virus succumbed to the infection by day seven, while animals infected with mutant viruses succumb to infection between seven and nine days post-infection.

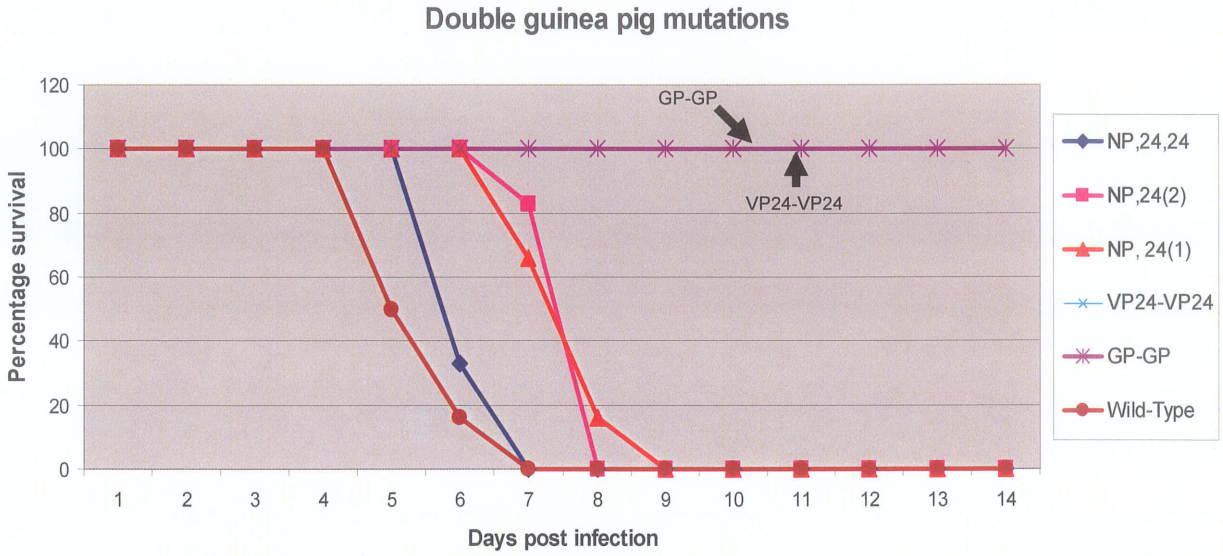


Figure. 33. Kaplan-Meier survival curve for guinea pig adapted double and triple mutations within the Ebo-Rib background using Dunkin-Hartly guinea pigs. *In vivo* infection of six Dunkin-Hartley guinea pigs with 1000 FFU of either Guinea pig adapted mutation listed in the key. Guinea pigs were monitored for weight loss and lethality caused by virus infection. Illustrated above is the percentage survival rate for all six viruses. Blood and organs (spleen, liver, and lung) were removed to determine serology for infection and virus isolation for sequence comparison between input and output viruses. RT-PCR from animals who had visible signs of disease was positive. Virus isolation was also possible from the sera and organs of animals showing signs of disease. As illustrated above double mutations containing NP and VP24 in combination produced a lethal infection within seven to nine days. Wild-type guinea pig virus resulted in a lethal infection within 4 to 7 days. There was 100% survival for double mutations (VP24) and (GP).

Summarizing our experiments we used the existing ZEBOV reverse genetic system to generate a reverse genetic copy which carries all nucleotide mutations seen in the guinea pig adapted virus (Table 11). Using RT-PCR and unique restriction sites we generated three fragments which cover all regions of mutations. Two positive clones were used to

test their ability to infect and grow in Vero E6 cells (Figure. 25). Both G-pig adapt 11 and G-pig adapt 33 could infect and replicate in Vero E6 cell producing a viral titer of $10^{6.5}$ - $10^{6.3}$(FFU), respectively. Since these guinea pig adapted constructs produced CPE in tissue culture and were positive for the presence of ZEBOV VP40 (Figure. 24), we attempted to determine the lethality of these rescued viruses *in vivo*. Three groups of six Dunkin-Hartley guinea pigs were infected with 1000 FFU of G-pig adapt 11, G-pig adapt 33, and guinea pig adapted virus to determine the percentage death rate (Figure. 26). Both guinea pig adapted recombinant viruses produced 100 % lethality when injected i.p. into guinea pigs demonstrating the successful development of a guinea pig adapted cDNA clone. Interestingly, these two clones had varying rates of death, G-pig adapt 33 produced 100% lethality by day 11, whereas G-pig adapt 11 had a progression of death starting at day 6 and progressing until the last death on day 11 (Figure. 26). The time of death varies with the guinea pig adapted virus, which produced 100% lethality by day 8. The phenomena of a delay in lethality between the adapted guinea pig virus and the recombinant guinea pig viruses can be explained by clonal selection. Within our infectious dilutions for both recombinant viruses only one genotype will exists, thus causing a lag in replication. Whereas the guinea pig adapted virus will have many quasispecies which will allow for a more rapid replication pattern.

Once the guinea pig adapted virus was constructed we wanted to determine, which and how many of the nucleotide changes were required for the adaptation from a non-infectious guinea pig virus to an infectious virus. To answer this question we developed seven single guinea pig adaptive mutations within the ZEBOV reverse genetic backbone. Each construct carried one nucleotide change. After rescuing these mutant

viruses (Figure. 28) we determined that all rescued viruses could grow to high viral titers indicating that the single mutations did not interfere with the normal replication processes of the virus (Figure. 29). Using these viruses we determined that no single mutation could cause lethality in guinea pigs (Figure. 30). From information gathered from the ZEBOV mouse adapted virus it was determined that the nucleoprotein and VP24 were two of the key proteins involved in the adaptation from primates to mice (36). Using this information, we concentrated our first efforts on developing double mutants which involved VP24 and NP (Table 12). Since NP contained only one mutation and VP24 contained two, we developed a double mutant of VP24(10420) and VP24(10741). This viral construct was rescued and could produce a high viral titer *in vitro* (Figure. 32), but was non-lethal to guinea pigs (Figure. 33). We generated three combinations (NP, VP24 (10420); NP, VP24 (10741); NP, VP24 (10420), VP24 (10741)). To further test the roles of these two proteins in the adaptation of the guinea pig adapted virus. When two mutations were present with combinations of NP and either VP24, the virus achieved 100 % lethality in infected guinea pigs (Figure. 33). Interestingly, viruses with a single NP and a single VP24 mutation seemed to display an attenuated phenotype (Figure. 33). When NP and both VP24 mutations were added the resulting virus was lethal in guinea pigs, which was similar to wild-type. This illustrates that mutations in NP and VP24 are also important in the adaptation to guinea pigs.

Since the glycoprotein has been linked to pathogenicity in filoviruses, we also generated a double mutant containing both glycoprotein mutations. As seen with the double VP24 mutations, the double glycoprotein mutation was non-lethal in guinea pigs. Taken together these results demonstrate the importance of NP and VP24 in the

adaptation from species to species. Interestingly, it has been suggested that VP24 may act as an inhibitor of the signally pathway for the JAK-STAT pathway. The mechanism of inhibition was due to blocks in the phosphorylation of both Janus kinases, Jak1 and Tyk2, during IFN-α signaling and at least a failure of Jak1 phosphorylation following IFN-γ stimulation (77). This result of IFN inhibition may be in direct relation with the adaptation results seen within this thesis. Mutations seen within VP24 of both mouse and guinea pig adapted models demonstrate the importance of this inhibition to allow viral infection and propagation within a normally non-infectable host. The responsibility, of NP is clearly a factor but is yet unknown, as both guinea pig adapted VP24 mutations when expressed in the reverse genetic system did not produce a lethal genotype. However, recent and previous research has demonstrated that NP plays a central role in virus replication. As well NP together with the minor matrix protein VP24 and polymerase cofactor VP35, are necessary and sufficient for the formation of nucleocapsid-like structures that are morphologically indistinguishable from those seen in EBOV-infected cells (79). It was also shown that NP is O glycosylated, and sialylated, and that these modifications are important for interaction between NP and VP35 (79) as well as VP24 (personal communication Kawaoka). The roll of NP interaction with other RNP complex proteins and the importance in viral adaptation is certainly linked to VP24 and VP35. The results presented here are only the beginning in our understanding of the complex interactions between NP, VP24, and the other RNP complex proteins.

3.3.3 Future goals using the guinea pig adapted reverse genetic system

The development of this reverse genetic system will allow researchers to investigate pathogenic factors in a small animal model which closely represents human

infection. The next step for this project would be investigate the interactions of all mutations. Constructs which are currently being constructed include:

(NP, VP40(4437) VP40(5053);

VP40 (4437), VP40 (5053), VP24 (10420), VP24 (10741);

NP, GP (6424), GP (7668);

VP40 (4437), VP40 (5053), GP (6424), GP (7668).

These mutant constructs will allow for the characterization of the remaining mutations within the guinea pig adapted virus.

Another interesting project which has arisen form this research includes the determination and characterization of the possible INF inhibiting domains which, seem to be present in VP24 and will very likely be the key region for host adaptation. The interaction between NP and VP24 is also part of the key elements for viral adaptation, defining these regions of interaction and mutagenizing them to determine the physiological effect *in vivo* may help us in our understanding of filoviral adaptation.

3.4 Construction of a recombinant Ebola virus expressing GFP

3.4.1 Introduction

In order to study the processes of viral entry and infection we developed a recombinant ZEBOV virus expressing GFP. The ability to detect viral particles *in vitro* or *in vivo* has become a commonly sought after tool when studying viral entry and localization. Currently, indirect immunofluorescence assays (IFA) are the favored tools for experimental assays of filovirus infection (38, 149). Since filoviruses do not hemagglutinate human and guinea pig thrombocytes (148), immunofluorescence assays are frequently used to detect filoviral antigen in tissue cultures and impression smears from organs of nonhuman primates (133). Monoclonal antibodies to filoviral antigen of specific strains have been developed for IFA and ELISA, further enhancing these techniques. The technology of immunofluorescence has advanced since the initial development of fluorescence techniques. Break through work being carried out using the canine distemper virus (CDV) has demonstrated that these tools could be used for any application which would allow for GFP visualization of quantification. Descried in a recent publication a CDV which expresses GFP was used to detect replication in lymphoid organs which allowed for the spread to the epithelial{Messling, 2004 #1053}. Advances have also been seen in the filoviral field one of these advancements was illustrated in a recent publication, were Towner et al. describes the generation of a recombinant ZEBOV which was engineered to express a foreign protein, eGFP. This type of virus was developed to provide a rapid and sensitive means to monitor virus replication in infected cells. This genetically engineered virus represents the first insertion of a foreign gene into ZEBOV. This virus was used to demonstrate that EBOV-

eGFP virus can be detected in known infected cells and serves as an ideal model to screen antiviral compounds in less time than any previously published assay (159).

Hypothesis and current objectives:

With current successes in the development of tools to study aspects of ZEBOV infection and cellular topism, we propose to further these successes by developing a ZEBOV cDNA construct which carries a unique open reading frame. Using this open reading frame we propose to insert a GFP gene which, when expressed will allow us to study viral tropism, viral entry, and, viral pathogenesis.

3.4.2 Results and Discussion

Reverse genetics has provided a new way to develop diagnostic and quantitative tools for the study of filoviruses. The ability to mutate specific regions or even add entire unique open reading frames is now possible. A key component to studying viruses is the development of tools which characterize virus infection.

Generation of ZEBOV-GFP full length clone:

To determine the effectiveness of an EBOV virus which expresses a green fluorescence protein as a readout method; we generated a ZEBOV reverse genetic clone which carries an additional ORF encoding a GFP. The original full-length clone was used as a template (114), to insert a unique open reading frame and the GFP gene. We used extension PCR mutagenesis to introduce a BssHII site in the untranslated region between NP and VP35 (Figure. 34) with the primers:

Ebo3005(f) 5'CGTTTTATAATTAAGAAAAAAGCGCGC 3'

Ebo3048(r) 5' GGTTTTAATCTTCATCGCGCGC 3'

where the BssHII site is shown in red.

The sub-cloning vector pUC19ZE'Sph3.6' was generated by the digestion of Ebo-Rib with SphI, and this fragment was sub-cloned into the vector pUC19 which was digested with SphI. The GFP insert was generated using a sub-cloning vector TOPO2.1'BssHII'; which contained the NP transcriptional start and stop sequences and separated by a SacI site flanked by BssHII sites (Figure. 34A). The GFP ORF was amplified using the following primers:

GFP'SacI'(f) GCG TCC GAG CTC ATG GTG AGC AAG GGC GAG GAG C

GFP'SacI'(r) GCG TCC GAG CTC TTA GTG ATG GTG ATG GTG ATG

This fragment was ligated into the SacI site of TOPO2.1'BssHII'. The GFP ORF was cloned into the pUC19'SphI' construct using the SacI restriction site. Once constructed the pUC19'GFP' construct was digested with BssHII and cloned into the ZEBOV full length clone (Figure. 34B).

Using the ZEBOV reverse genetic system transfections were performed using 2 µl *trans*IT®-LT1 (Mirus) transfection reagent per 1 µg of the standardized DNA, as described in section 3.1.2 (Figure. 35). Transfections were performed with a 1:1 ratio of Vero E6 and 293T cells and incubated at 37°C in BSL4. Supernatants were collected three days post transfection and passaged onto 80 % confluent Vero E6 cells. Cells transfected with ZEBOV 'wild-type' and ZEBOV'GFP' showed cytopathic effect before day 14. ZEBOV-Wild-Type was harvested on day 5 and ZEBOV-GFP was harvested on

day 10. Using an aliquot from the rescued ZEBOV Wild-Type and ZEBOV-GFP, we infected Vero E6 cells to determine if ZEBOV-GFP could be detected under UV light.

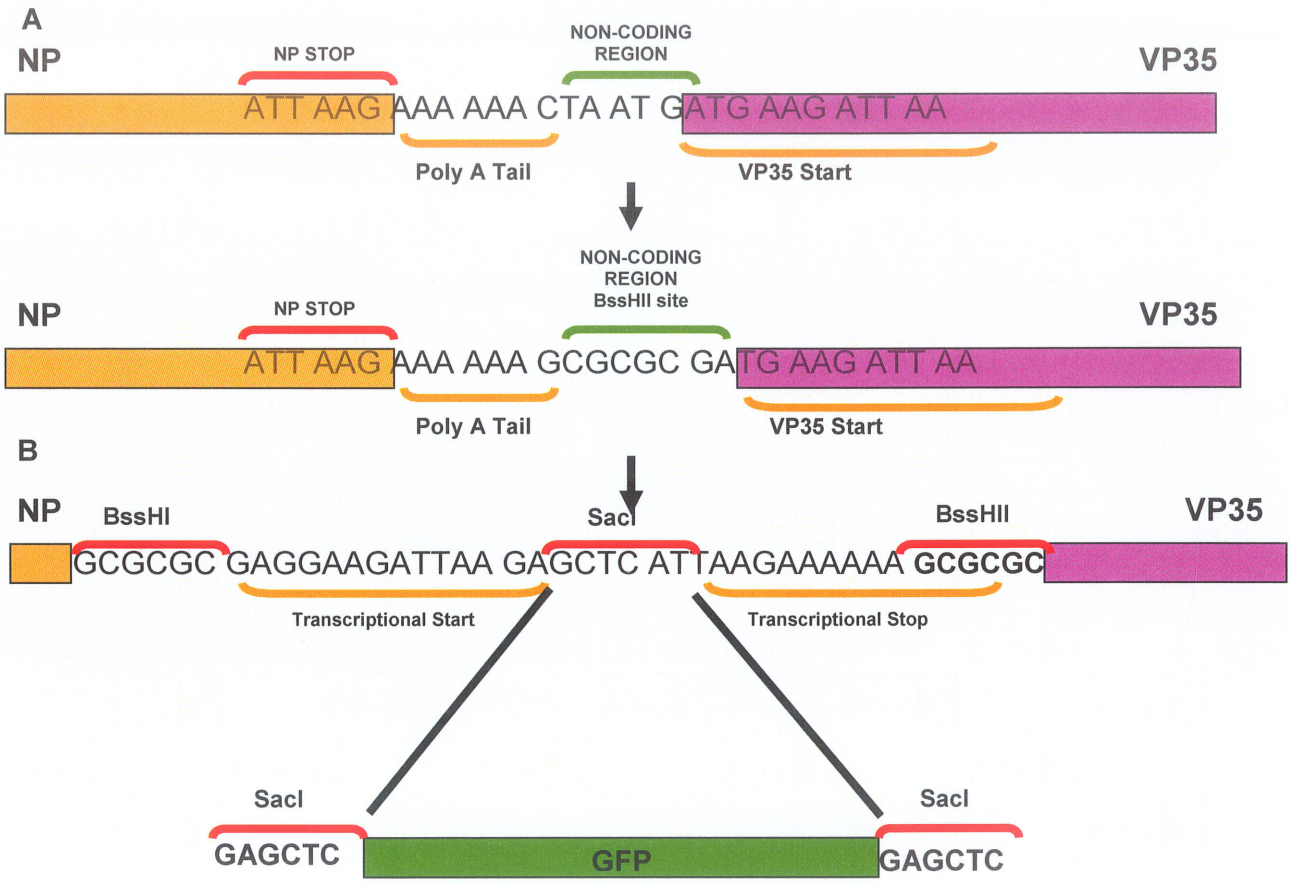


Figure. 34. Schematic of BssHII development and GFP insertion for the generation of a ZEBOV-GFP cDNA construct. A) This clone was developed by extension PCR using the primers which carried the restriction site BssHII in place of the non-coding region found between NP and VP35. The restriction sit was used as the point of insertion for a new transcriptional start and a transcriptional stop separated by the open reading frame of the eGFP. The final construct was digested to ensure proper GFP orientation as well as being digested with BssHII to ensure single GFP insertion.

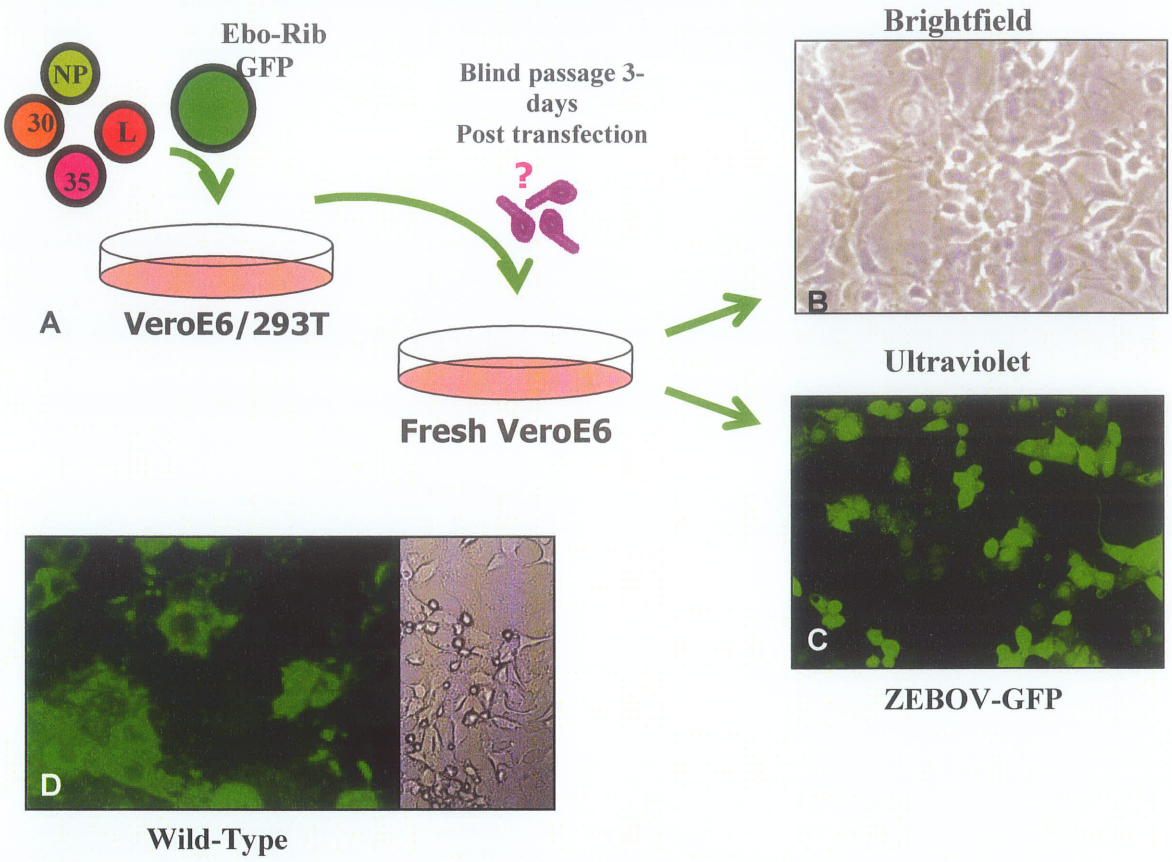


Figure. 35. Illustration of reverse genetic rescue of ZEBOV-GFP cDNA constructs. A) Reverse genetic rescue of ZEBOV-GFP, the RNP complex proteins (NP, VP35, VP30, and L) were transfected into Vero E6/293T cells using the optimized plasmid ratios (1ug, 0.5ug, 0.3ug, and 1ug respectively) of DNA. B) Brightfield view of reverse genetic rescue 10-days post transfection demonstrating slight cytopathic effect. C) Florescence microscopy of cells transfected with ZEBOV-GFP. This demonstrates the expression of GFP from infected cells D) ZEBOV '76' reverse genetic system positive control, infected cells were assayed using indirect immunofluorescence against anti-VP40 (1:200 dilution in PBS), and 1:500 dilution of FITC labelled secondary antibody in PBS. Brightfield and FITC stained(VP40) are illustrated together.

Following an incubation of 4 days, the infected cells were fixed and inactivated with 4 % paraformaldehyde in PBS for 2 days with one fixative exchange before being removed from level 4. For indirect immunofluorescence analyses of ZEBOV wild-type, see materials and methods section 2.3.9 for full description of protocol.

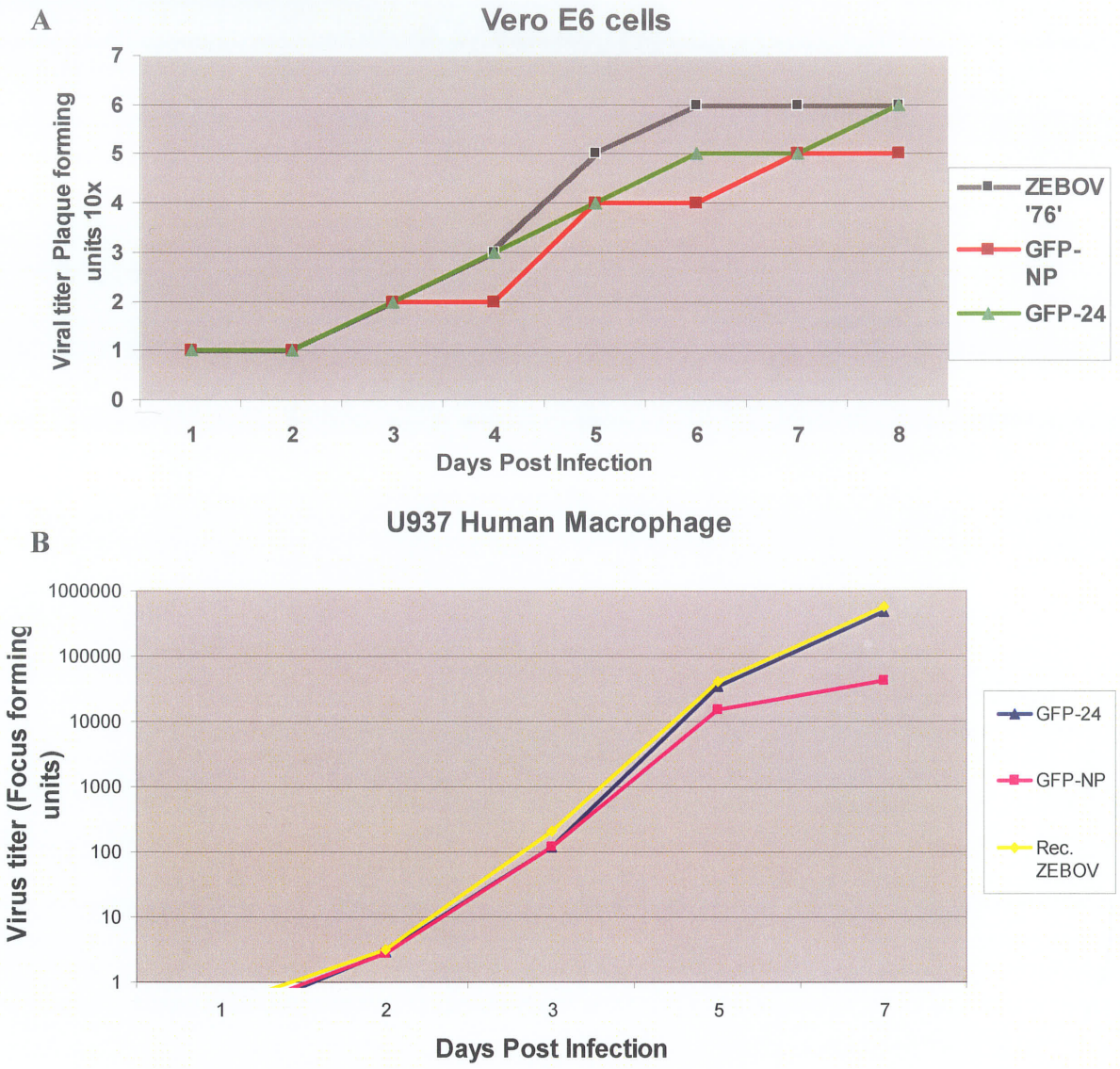


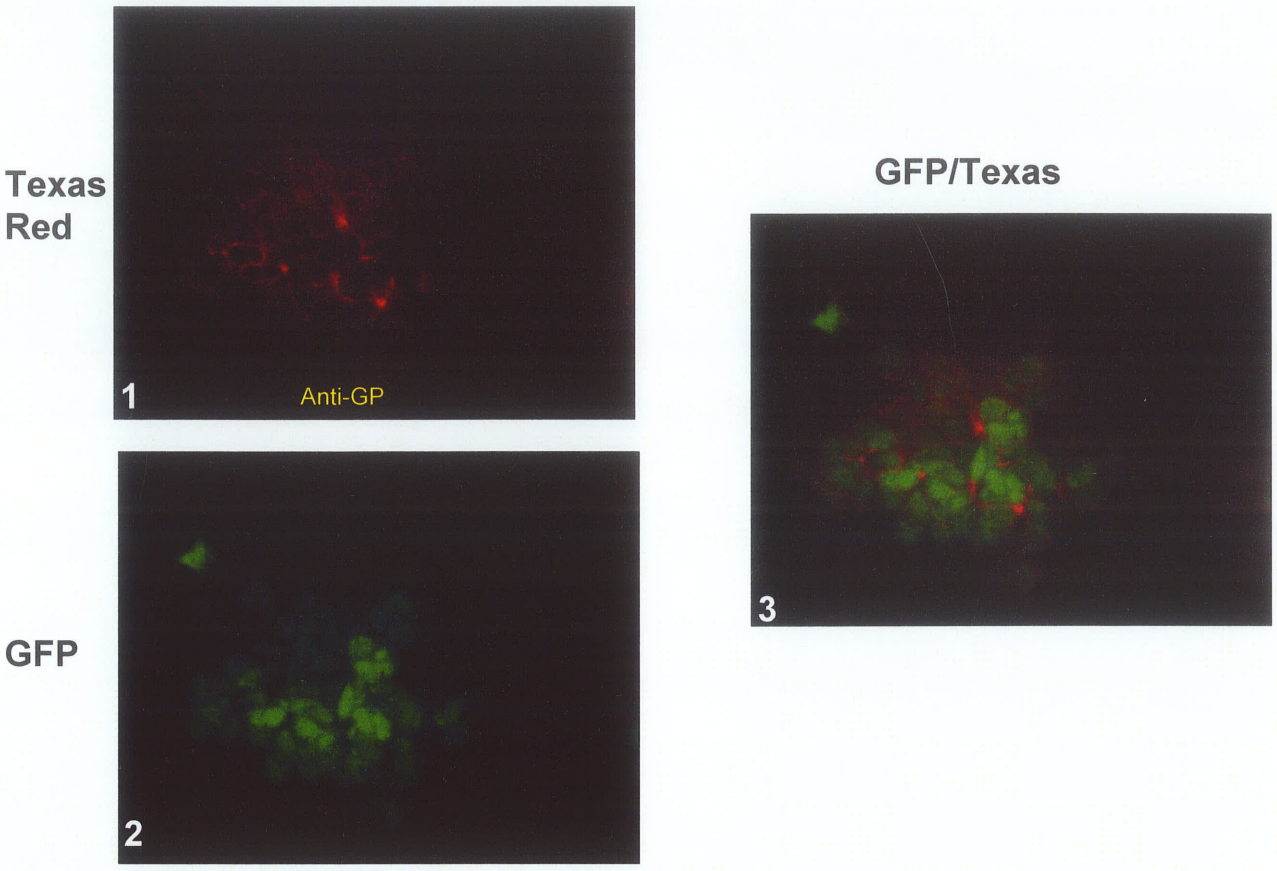
Figure. 36. Viral growth kinetics of ZEBOV-GFP viruses in Vero E6 African green monkey (A) and U937 human macrophage (B) cell lines. Vero E6 and U937 cells were infected at an MOI of 0.1 with samples being assayed for viral titres on days 1 through 7 for Vero E6 cells and days 1-3, 5, and 7 for U937 cells. Viruses were propagated in biocontainment level 4. Titrations were performed using a 10-fold dilution series of viral stocks, 1 ml of each dilution series were used to infect Vero E6 cells in a 96 well plate. Infected cells were overlaid with DMEM containing 1.5% carboxymethyl-cellulose (CMC) and 2% (v/v) fetal calf serum. Following an incubation of eight days, the infected cells were fixed and inactivated with 4% paraformaldehyde. Indirect immunofluorescence assay of ZEBOV '76' infected cells using anti-VP40 (1:200 dilution in PBS) was carried out. ZEBOV'76' and ZEBOV-GFP recombinant viruses were analyzed using an Axioplan 2 fluorescence microscope, titers were determined by counting positive fluorescence plagues (Focus forming unit). All virus types (ZEBOV-GFP's and recombinant ZEBOV) grew within one log of each other in both cell types.

Analysis was carried using an Axioplan 2 Fluorescence microscope (Zeiss, Germany). For ZEBOV-GFP viruses cells were washed with PBS and analyzed under UV (Figure. 35C).

Characterization of ZEBOV-GFP: To our advantage our collaborators (Y. Kawaoka University of Wisconsin) also developed a ZEBOV-GFP reverse genetic construct. This virus construct contains the eGFP open reading frame between within the intergenic region between VP30 and VP24. Titrations of both viruses were carried out using an immunoplaque assay. Vero E6 cells and a human macrophage cell lines (U937) were grown in 96 well plates and titrated (Figure. 36). Both cell lines could be infected with ZEBOV and produced high viral titers which ranged from $10^{4.5}$ to 10^{7} FFU. This demonstrates that both ZEBOV-GFP viruses can infect and propagate as well as wild-type with no adverse effects because of their cDNA origin.

Stablility of GFP open reading g frame: In order to ensure the GFP construct was stable in the full-length construct we carried out stability assays to determine the percentage of infected cell which expressed GFP. We infected 80% confluent Vero E6 cells with an MOI of 0.1, samples were visualized using indirect immunofluorescence analyses for ZEBOV-GFP stability, (Figure. 37). Stability was assayed since the introduction of a foreign gene product into a replicating virus has been previously shown to remove the foreign product by polymerase exclusion (152).

A


Texas Red
1
Anti-GP
GFP/Texas
3
GFP
2

B

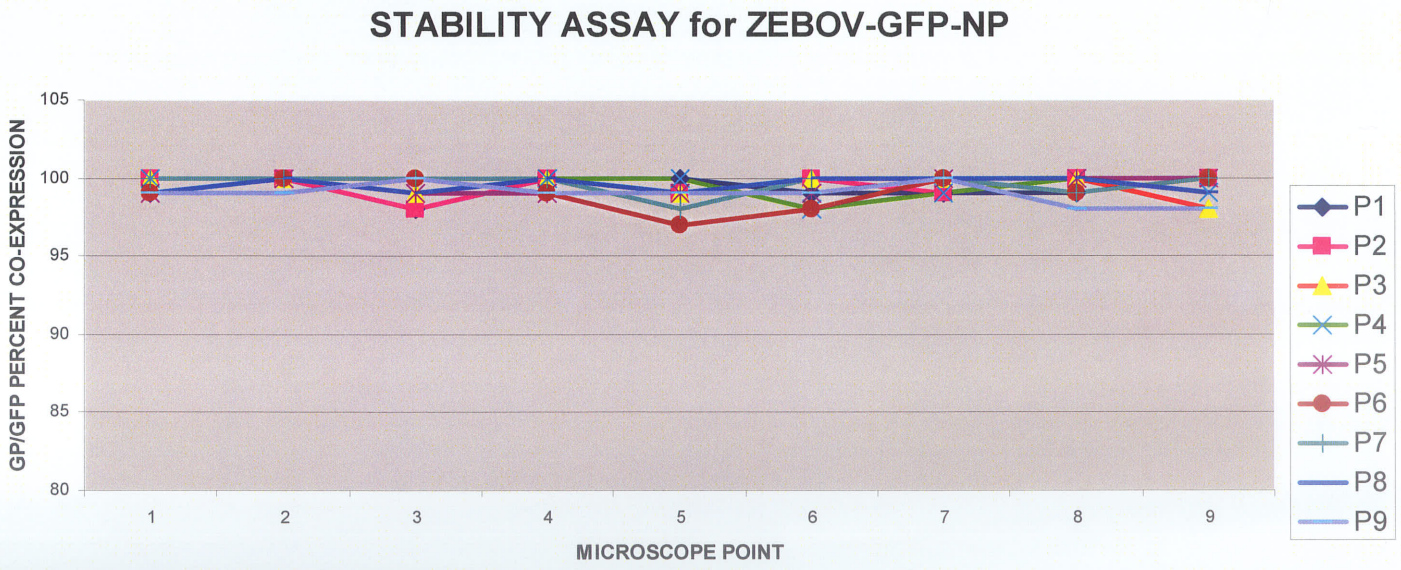

STABILITY ASSAY for ZEBOV-GFP-NP
GP/GFP PERCENT CO-EXPRESSION
105
100
95
90
85
80
1
2
3
4
5
6
7
8
9
MICROSCOPE POINT
P1
P2
P3
P4
P5
P6
P7
P8
P9

Figure. 37. ZEBOV-GFP (NP insertional version) stability assay in Vero E6 cells. A) Panel 1, represents an indirect immunofluorescence assay of ZEBOV-(NP) GFP infected cells detected with anti-ZEBOV glycoprotein (1:1000 dilution in PBS) and visualized using Texas Red. Panel 2, Infected cells seen through UV light. Panel 3, represents a merge of panel A and B illustrating the stability of this intracellular expressed GFP protein during an infection.
B) Vero E6 cells were infected with ZEBOV-GFP-NP, cells were grown for 5 days and assayed for GFP expression and ZEBOV glycoprotein stained with texas red (1:500 dilution in PBS) using indirect immunofluorescence assay. GP/GFP percentage co-expression was determined in a microscopic view point by counting cells which co-expressed GFP and texas red. 'Key' represents passage number of Vero E6 cell infected with ZEBOV-GFP-NP.

Demonstrated in figure. 37 stability of the ZEBOV-GFP viruses was stable up to and past nine cellular passages showing great stability of this foreign insertion. Interestingly, this result and the fact that EBOV have a long filamentous structure, we can hypothesis that EBOV should be able to house much larger gene insertions simple to its structure.

Animal experiments using ZEBOV-GFP: STAT-1 knockout mice were housed in level 4 for fourteen days prior to injection for acclimation. Three groups of nine mice were each injected IP with 1000 FFU of ZEBOV-GFP-NP, ZEBOV-GFP-24, and wild-type.

Blood, liver, and spleens were collected from both GFP viruses and the control mice positive and negative on days 1, 3, and 5. These samples were divided into two groups, the first group was homogenized and fixed for FACS analysis (Figure. 38).

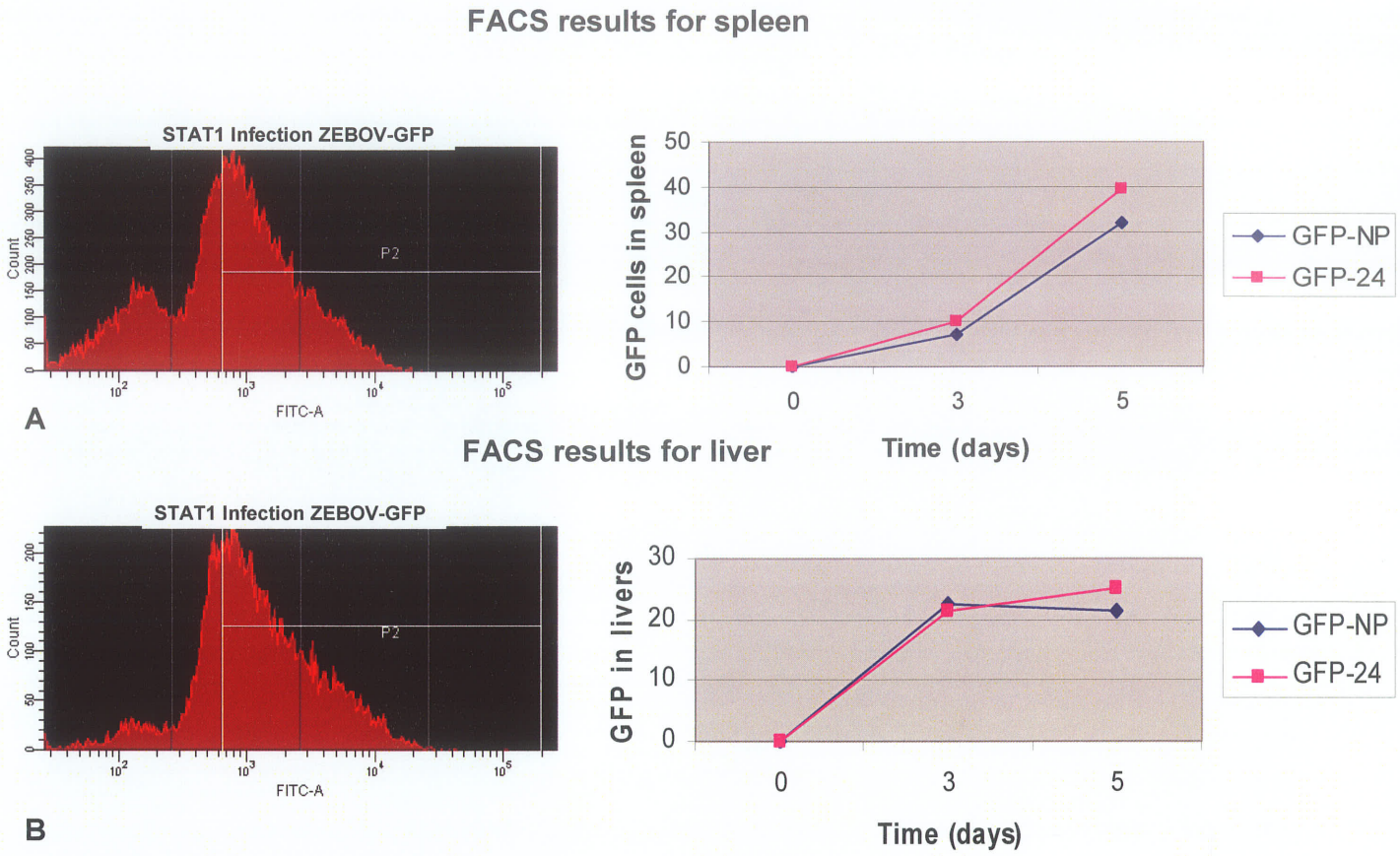


Figure. 38. FACS analysis of organs from infected STAT-1 mice. Animals were infected with 1000 PFU and sampled on days 1, 3, and 5. Organs including the liver and spleen were assayed for GFP expression. Organs were removed from infected animals on prescribed days, homogenized and fixed using 4% paraformaldehyde. **Panel A**, illustrates the FACS result from spleen infected with ZEBOV-GFP(NP). Graph shows both GFP viruses could be detected in the spleens of infected animals. **Panel B,** illustrates the FACS results for infected livers with ZEBOV-GFP(NP).

The second sample was used for organ smears for GFP detection under UV light in BSL-4. Briefly, organs were pressed between two glass slides, breaking organs into single cells. Samples were mounted and visualized under UV light (Figure. 39).

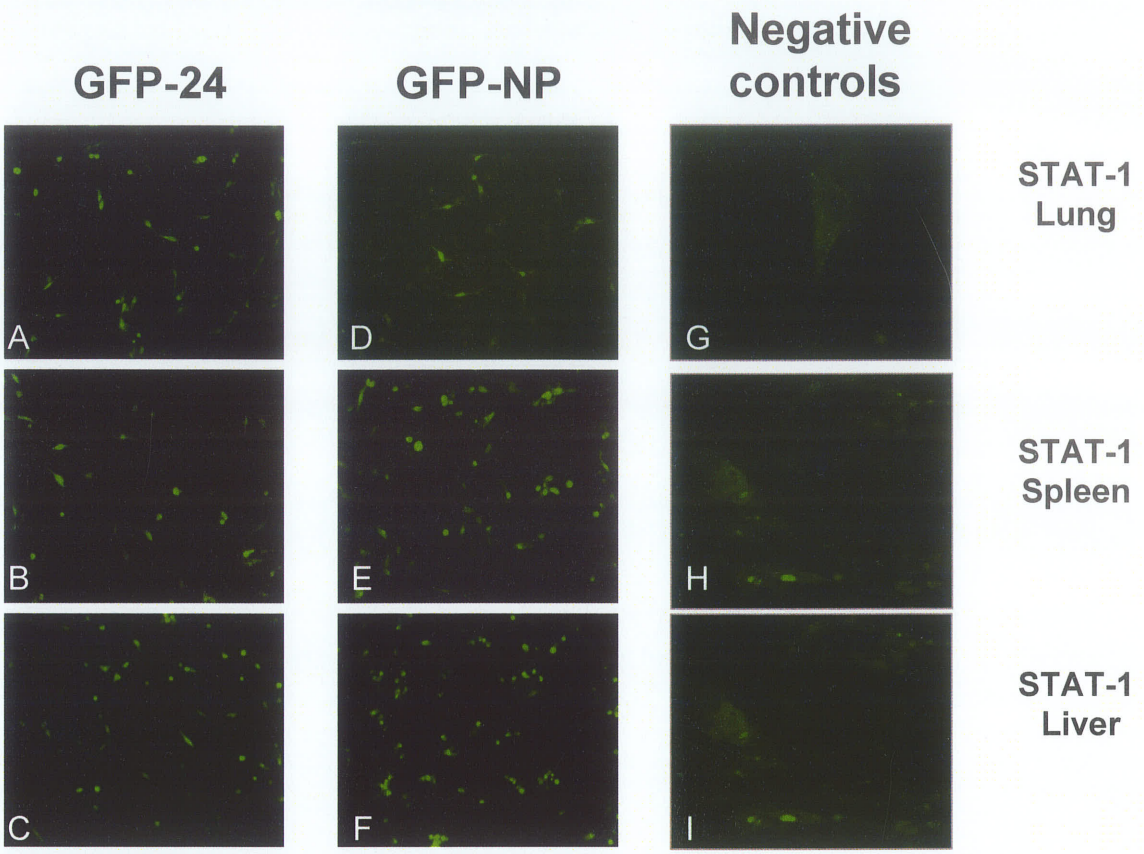
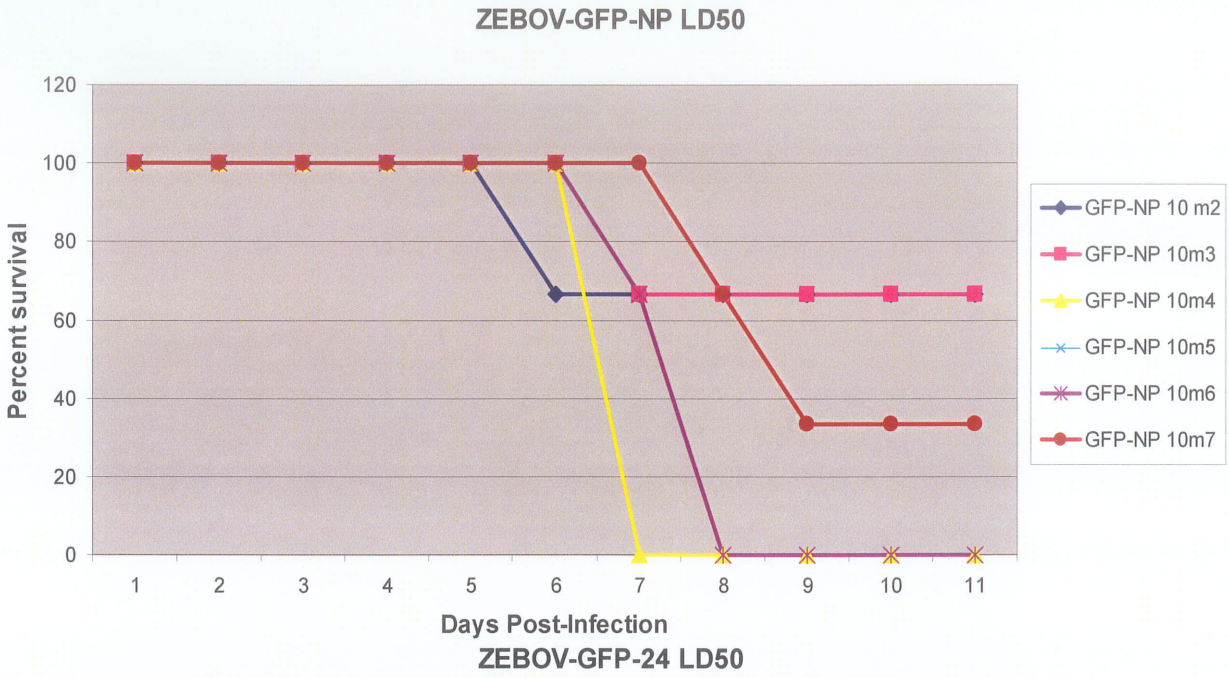


Figure. 39. Glass slide smears of ZEBOV-GFP-NP and ZEBOV-GFP-24 infected organs. Blood samples and organs were collected at defined days (1, 3 & 5). Blood samples were aliquoted in EDTA tubes during blood extraction periods, and processed by normal FACS lysis protocol. Samples were resuspended by adding 4% (w/v) paraformaldehyde and stored at 40°C for 24 hours. Samples were pelleted and fresh 4% (w/v) paraformaldehyde was added to resuspended fixed cells. Tissues from infected mice were collected on days 1, 3, and 5 and homogenized. Supernatants from homogenized tissue were added to FACS lyse solution to lysis red blood cells.
STAT-1 mice who were infected with either ZEBOV-GFP-NP or ZEBOV-GFP-24. Organ sections were smeared on glass slides and photographed within BSL 4 using a (Zeiss) fluorescent microscopy. Organs samples above were collected on day 5. Both ZEBOV-GFP viruses showed organ infection and GFP expression, Panels A, B and C for ZEBOV-GFP-24 and D, E and F for ZEBOV-GFP-NP. Negative control for all organs are found in panels G, H, and I.

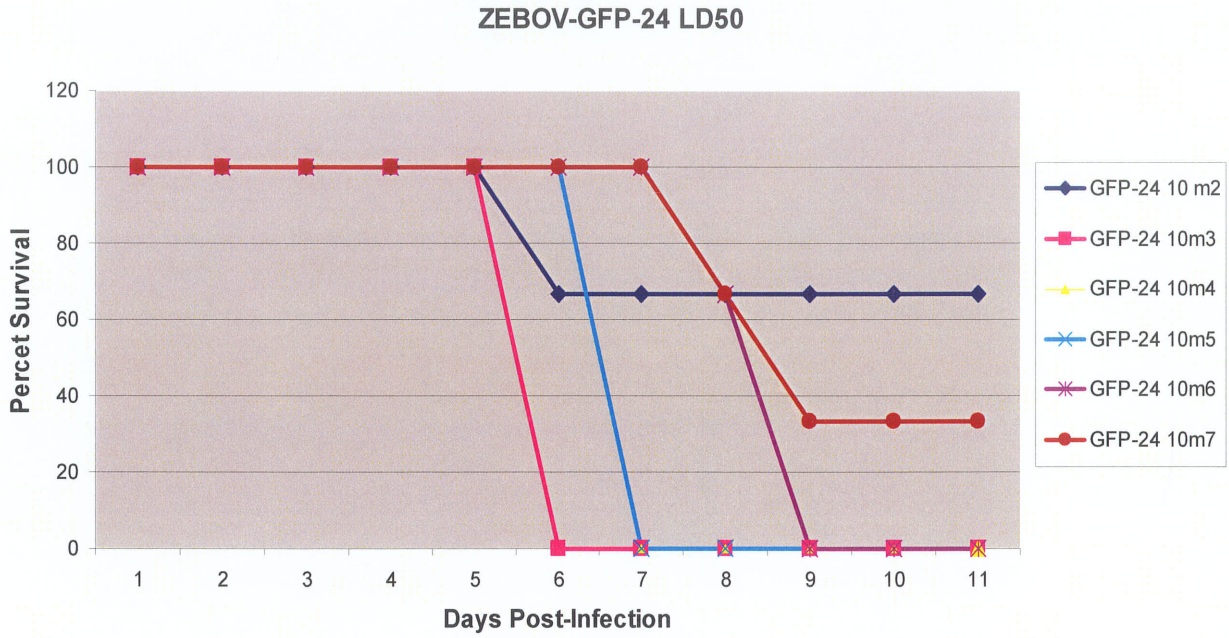
FACS analysis of infected organs and blood demonstrated a very useful application for these ZEBOV-GFP viruses. Since no other staining procedure is needed to visualize GFP expression autofluorescence is eliminated from secondary antibody staining, leaving only cellular autofluorescence to contend with. Also since these ZEBOV-GFP viruses show the same cellular topism as wild-type ZEBOV Figure. 36 we could follow the natural viral path *in-vivo* using online imaging technologies. Since the ability to stain many different cell types using commercial antibodies is available, the possibility to isolate and study specific cell types which preferentially harbor EBOV infection is now possible using these GFP viruses. Organ smears have demonstrated that detection of infectious ZEBOV-GFP is possible within immunodeficient mice. The use of immunodeficient was needed as mice with an intact IFN response pathway will ultimately remove infectious virus. This is another point which further illustrates the importance of IFN anatagonism in viral adaptation and infection of mice and guinea pigs. Virus RNA and virus isolation was possible from blood of all infected mice as well as in all organs (spleen, liver, and lung) sampled.

Also demonstrated here is the Lethal Dose (LD) 50 of both GFP-expressing viruses. Viruses were diluted from $10^{-2} - 10^{-7}$ (FFU) and inoculated IP into three groups of three STAT-1 knockout mice. Both, ZEBOV-GFP viruses, ZEBOV-GFP-NP, ZEBOV-GFP-24, as well as ZEBOV-Wild-Type were assayed (Figure. 40)..

A


ZEBOV-GFP-NP LD50
Percent survival
Days Post-Infection
GFP-NP 10 m2
GFP-NP 10m3
GFP-NP 10m4
GFP-NP 10m5
GFP-NP 10m6
GFP-NP 10m7

B


ZEBOV-GFP-24 LD50
Percet Survival
Days Post-Infection
GFP-24 10 m2
GFP-24 10m3
GFP-24 10m4
GFP-24 10m5
GFP-24 10m6
GFP-24 10m7

C

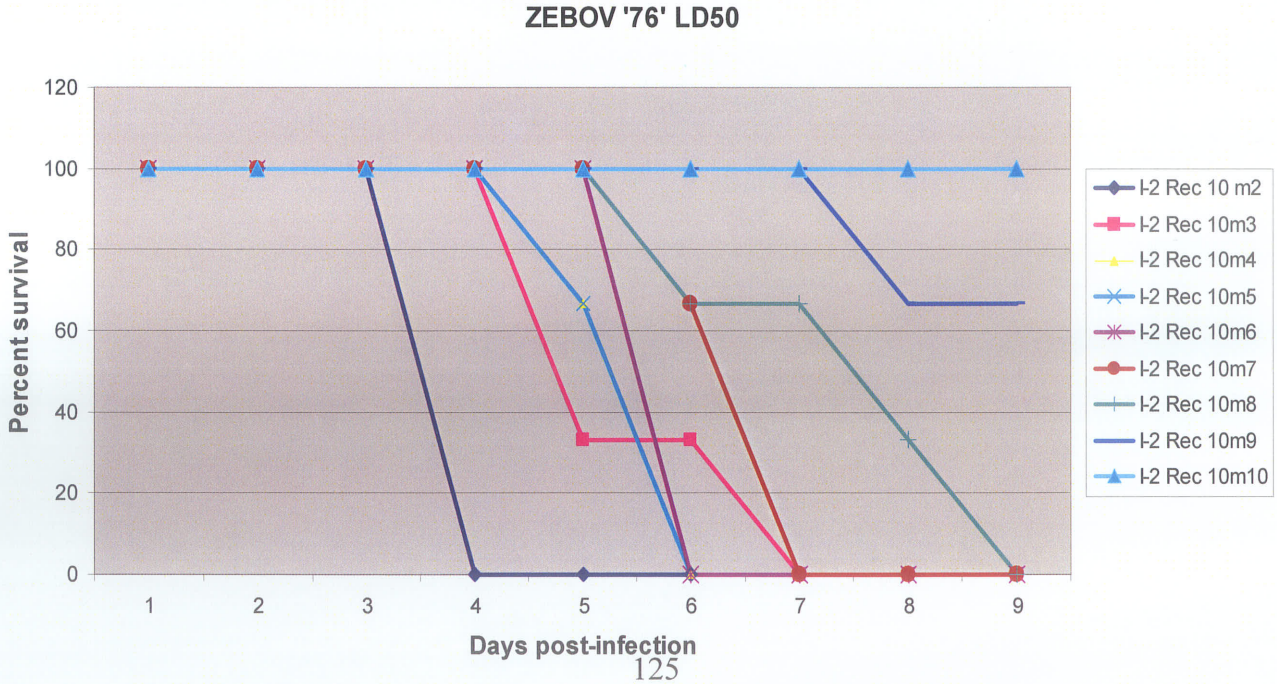

ZEBOV '76' LD50
Percent survival
Days post-infection
I-2 Rec 10 m2
I-2 Rec 10m3
I-2 Rec 10m4
I-2 Rec 10m5
I-2 Rec 10m6
I-2 Rec 10m7
I-2 Rec 10m8
I-2 Rec 10m9
I-2 Rec 10m10

Figure. 40. Kaplan-Meier survival curve demonstrating the Lethal Dose (LD) 50 of ZEBOV-GFP-24 and ZEBOV-GFP-NP viruses and ZEBOV wild-type 76'. A) Groups of three STAT-1 mice were infected IP with six serial dilutions (10^2-10^7) of ZEBOV-GFP-NP virus. The LD_{50} for ZEBOV-GFP-NP was determined to be $10^{6.3}$ (FFU). **B)** Groups of three STAT-1 mice were infected IP six serial dilutions (10^2-10^7) of ZEBOV-GFP-24 virus. The LD_{50} for ZEBOV-GFP-24 was determined to be 10^6 (FFU). Also as seen with ZEBOV-GFP-NP higher viral load allows for some survival of animals. **C)** Groups of three STAT-1 mice were infected IP with six serial dilutions (10^2-10^{10}) of ZEBOV '76' virus (aka I-2). The LD_{50} for ZEBOV '76' was determined to be $10^{8.9}$ (FFU). Also as seen with the other LD_{50} experiments a higher viral load allowed for the survival of some animals. All animals were monitored for weight loss and signs of disease, death was from acute shock.

The results seem in the all the LD50 experiments using STAT-1 demonstrate an interesting phenomenon. It is seen that a high viral load has limited ability to generate disease and thus death in this mouse model. This can be partial explained by the activation of the mouse innate immunity which includes natural killer cells. These cells when activated by a large intake of foreign material (antigen) are unregulated and activated to travel to the infection site. With this mobilization of these innate response elements, mice seem to be able to advert the infection and clear it from their system. This process seems to differ on an individual bases, most likely being dependant on how fit the mouse is immunologically. With lethality being present uniformly in lower dilution one can ascertain that with a lower initial concentration of antigen being infected into the mice does not over stimulate the innate natural killer cell response.

In order to build advanced tools to study the viral life cycle and the localization of infected cells *in vivo*, we developed a ZEBOV reverse genetic clone which expresses a green fluorescent protein (ZEBOV-GFP). This construct has given us the ability to detect virus infection *in vitro* and *in vivo* using fluorescence microscopy as seen in the CDV system{Messling, 2004 #1053}. One of the greatest advantages in using a virus which expresses its own fluorescent marker is the ease of detection and the speed with which virus infection can be confirmed, making this an extremely useful tool for studying anti-virals and *in vivo* infections.

Our collaborators have also generated a ZEBOV-GFP virus. Using these two ZEBOV-GFP (ZEBOV-NP and ZEBOV-24) we determined the effect of 5' or 3' proximal gene insertion in the viral genome. Both viruses were characterized for GFP stability (Figure. 37), to demonstrate that all infected cells express GFP. Stability was attained in both viruses 98% of the time (Figure. 37). To determine the replicative fitness of both GFP viruses growth kinetics experiments were performed and compared to ZEBOV-wild-type. Interestingly, ZEBOV-GFP-NP was slightly attenuated when compared to ZEBOV-wild-type and ZEBOV-GFP-24 which both grew to a titer greater then 10^6 (FFU) (Figure. 36). Viruses were also assayed for their ability to infect and grow in macrophages (Figure. 36). As illustrated in both kinetic experiments there was a slight attenuation observed for ZEBOV-GFP-NP when compared to ZEBOV-GFP-24 and ZEBOV-wild-type, but both viruses could infect human macrophages and grow to a high titers (Figure. 36). The attenuation seen with ZEBOV-NP may be the result of the insertion of GFP within the non-coding region between NP and VP35. This attenuated phenotype is seen in vitro (Figure. 36). However, when the lethal dose was determined in

STAT-1 mice, there was a characteristic high titer survival among infected mice (Figure. 40). This phenomenon has been observed in previous experiments when dealing with high concentrations of virus were inoculated into mice (unpublished data). This result is thought to be indirect relation to the innate immune response and the activaton of natural killer cells.

In our initial design of these GFP viruses we were building tools which could be used to infect and detect infectious virus *in vivo* and *in vitro* allowing for the use as a diagnostic tool. We developed a ZEBOV-GFP-NP construct which could be detected in both tissue culture and within a STAT-1 knockout mouse model. STAT1 knockout mice were used as they are deficient in the INF α/β response, which allows for the infection with wild-type filoviruses. As seen above both ZEBOV-GFP viruses where infectious and detectable in tissue culture and the STAT-1 knockout model illustrating that this system can be used to quantitatively measure infectious virus within infected samples.

3.4.3 Future Work

The next step in generating detection tools for the study of filoviruses would be to generate fusion protein constructs. Fusion of Green Fluorescence Protein, Blue Fluorescence Protein, or Red Fluorescence Protein to viral proteins 24, 30, and 35 as well as NP could allow for tracking of these viral proteins *in vivo*. A fusion protein would be

retained with the virus, in contrast to the current GFP virus which exogenously produces GFP within the infected cell.

The generation of a green fluorescent fusion protein has been attempted with VP24. Unfortunately, all virus rescue attempts failed, illustrating that VP24 role in the virus life cycle was being interrupted by the addition of GFP.

4.0 Closing Statements

The main focus of this thesis was the utilization of the ZEBOV reverse genetic system to analysis the pathogenic factors relating to EBOV infection. Our first goal was to optimize the reverse genetic system and determine what type of efficacy it had? In doing so we determined that heterologous proteins from closely related family members were able to drive the replication of the reverse genetic system and thus produce an infections ZEBOV virus. These were interesting results as replication and transcription was always thought to be species specific! These results also presented a valuable tool, since there is cross reaction between the RNP complex protein within the genera *filovirus.* We can use the use the ZEBOV system to assay proteins from REBOV, ICEBOV, and SEBOV which will not only help in the development of new reverse genetic and mini-genome systems but, will answer key questions about filovirus structural protein interactions.

While working with the ZEBOV reverse genetic system it soon became apparent that generation of mutant cDNA constructs was limited because of spontaneous mutations which, were generated during bacterial propagation. This problem drove us to develop a cloning system which limited the use of bacteria to propagate the cDNA construct. Cell free cloning which we have named this method utilizes common mutagenic techniques to mutagenize a cDNA construct in 4 step within one eppendoff tube. This method when

used with the ZEBOV system has allow us to generate mutant in $1/10^{th}$ of the time normally needed and with very little to no spontaneous mutations. Currently this method is patented and under review for commercial production.

A demonstration of its use was presented in this thesis for the development of four glycoprotein mutations. Since the glycoprotein of filoviruses, and its gene products have been proposed to be in some part responsible for the overwhelming pathogenic effect seen during infection. To help answer these questions more fully we generated four glycoprotein mutations which would remove key structures of the glycoprotein which maybe responsible for pathogenicity. Further studies are currently being carried out to determine the effect of these mutations *in vivo*.

With the advent of cell free cloning we were able to develop two independent systems to study filovirus infection and cellular topism. The ZEBOV-GFP was generated to assist in our understanding of cellular infection and also tropism. Using this virus we demonstrated that FACS analysis and fluorescence microscopy can be used to follow viral infection patterns. As well this virus has provided us with a tool which could be used for fast readout of infection within experimental methods.

In the development and utilization of the reverse genetics system we realized there was a limitation in the animal work which we could carry out. Realizing that guinea pigs presented an infection which is similar to human infection we choose to develop a guinea pig adapted reverse genetics system to study the effects of viral tropism and pathogenicity. During the development of this system it became apparent that NP and VP24 are key components for the generation of a lethal genotype. This result is linked to IFN inhibition which may be in direct relation to the adaptation results seen in the

nucleotide sequence. Mutations seen within VP24 of both mouse and guinea pig adapted models demonstrate the importance of this inhibition to allow viral infection and propagation within a normally non-infectable host. The responsibility, of NP is clearly a factor but is yet unknown, as both guinea pig adapted VP24 mutations when expressed in the reverse genetic system did not produce a lethal genotype.

This thesis has advanced the filoviral field by demonstrating the usefulness of the reverse genetic system, by answering key questions about cellular tropism and viral adaptation to a new host.

5.0 Bibliography:

1. 1978. Ebola haemorrhagic fever in Sudan, 1976. Report of a WHO/International Study Team. Bull World Health Organ **56:**247-70.
2. 1978. Ebola haemorrhagic fever in Zaire, 1976. Bull World Health Organ **56:**271-93.
3. **Accardi, L., C. Prehaud, P. Di Bonito, S. Mochi, M. Bouloy, and C. Giorgi.** 2001. Activity of Toscana and Rift Valley fever virus transcription complexes on heterologous templates. J Gen Virol:781-785.
4. **Baize, S., E. M. Leroy, M. C. Georges-Courbot, M. Capron, J. Lansoud-Soukate, P. Debre, S. P. Fisher-Hoch, J. B. McCormick, and A. J. Georges.** 1999. Defective humoral responses and extensive intravascular apoptosis are associated with fatal outcome in Ebola virus-infected patients. Nat Med **5:**423-6.
5. **Baize, S., E. M. Leroy, A. J. Georges, M. C. Georges-Courbot, M. Capron, I. Bedjabaga, J. Lansoud-Soukate, and E. Mavoungou.** 2002. Inflammatory responses in Ebola virus-infected patients. Clin Exp Immunol **128:**163-8.
6. **Baron, M., and T. Barrett.** 1997. Rescue of rinderpest virus from cloned cDNA. J Virol:1265-1271.
7. **Barr, J., P. Chamers, C. R. Pringle, and A. Easton.** 1991. Sequence of the major nucleocapsid gene of pneumonia virus of mice: sequence comparison suggests structural homolgy between nucleocapsid proteins of pneumoviruses, paramyxoviruses, rhabdoviruses, and filoviruses. J. Gen. Virol.:677-685.
8. **Baskerville, A., E. T. Bowen, G. S. Platt, L. B. McArdell, and D. Simpson.** 1977. The pathology of experimental Ebola virus infections in monkeys. The Journal of Pathology:131-138.
9. **Baskerville, A., E. T. Bowen, G. S. Platt, L. B. McArdell, and D. I. Simpson.** 1978. The pathology of experimental Ebola virus infection in monkeys. J Pathol **125:**131-8.
10. **Basler, C. F., A. Mikulasova, L. Martinez-Sobrido, J. Paragas, E. Muhlberger, M. Bray, H. D. Klenk, P. Palese, and A. Garcia-Sastre.** 2003. The Ebola virus VP35 protein inhibits activation of interferon regulatory factor 3. J Virol **77:**7945-56.
11. **Basler, C. F., and P. Palese.** 2004. Modulation of innate immunity by filoviruses. Horizon Bioscience, Norfolk, UK.
12. **Basler, C. F., X. Wang, E. Muhlberger, V. Volchkov, J. Paragas, H. D. Klenk, A. Garcia-Sastre, and P. Palese.** 2000. The Ebola virus VP35 protein functions as a type I IFN antagonist. Proc Natl Acad Sci U S A **97:**12289-94.
13. **Becker, S., C. Rinne, U. HofsaB, H. D. Klenk, and E. Muhlberger.** 1998. Interactions of Marburg virus nucleocapsid proteins. Virology:406-417.
14. **Bolivar, F., R. L. Rodriguez, P. J. Greene, M. C. Betlach, H. L. Heyneker, and H. W. Boyer.** 1977. Construction and characterization of new cloning vehicles. II. A multipurpose cloning system. Gene **2:**95-113.
15. **Borio, L., T. Inglesby, C. Peters, A. Schmaljohn, H. J., P. Jahrling, T. Ksiazek, K. Johnson, A. Meyerhoff, T. Toole, M. Ascher, J. Bartlett, J. Breman, E. J. Eitzen, M. Hamburg, J. Hauer, D. K. Henderson, R. Johnson, G. Kwik, M. Layton, S. Lillibridge, G. Nabel, M. Osterholm, T. Perl, P.**

Russell, and K. Tonat. 2002. Working Group on Civilian Biodefense; Hemorrhagic fever viruses as biological weapons: medical and public health management. JAMA:2391-2405.

16. **Bowen, E. T., G. Lloyd, W. J. Harris, A. Baskerville, and E. E. Vella.** 1977. Viral haemorrhagic fever in southern Sudan and northern Zaire. Preliminary studies on the aetiological agent. Lancet:571-573.
17. **Bowen, E. T., G. S. Platt, D. I. Simpson, L. B. McArdell, and R. T. Raymond.** 1978. Ebola haemorrhagic fever: experimental infection of monkeys. Trans R Soc Trop Med Hyg **72:**188-91.
18. **Bray, M.** 2003. Defense against filoviruses used as biological weapons. Antiviral Res **57:**53-60.
19. **Bray, M.** 2001. The role of the Type I interferon response in the resistance of mice to filovirus infection. J Gen Virol **82:**1365-73.
20. **Bray, M., K. Davis, T. Geisbert, C. Schmaljohn, and J. Huggins.** 1998. A mouse model for evaluation of prophylaxis and therapy of Ebola hemorrhagic fever. J Infect Dis **178:**651-61.
21. **Bray, M., and T. W. Geisbert.** 2005. Ebola virus: the role of macrophages and dendritic cells in the pathogenesis of Ebola hemorrhagic fever. Int J Biochem Cell Biol **37:**1560-6.
22. **Bray, M., and S. Mahanty.** 2003. Ebola hemorrhagic fever and septic shock. J Infect Dis **188:**1613-7.
23. **Bridgen, A., and R. M. Elliott.** 1996. Rescue of a segmented negative-stranded RNA virus entirely from cloned complementary DNAs. Proc Natl Acad Sci U S A:15400-15404.
24. **Buchholz, U. J., S. Finke, and K. Conzelmann.** 1999. Generation of bovine respiratory syncytial virus (BRSV) from cDNA: BRSV NS2 is not essential for virus replication in tissue culture, and the human RSV leader region acts as a functional BRSV genome promoter. J Virol:251-259.
25. **Bullock, W. O., J. M. Fernandez, and J. M. Short.** 1987. Biotechniques:376-379.
26. **Bwaka, M. A., M. J. Bonnet, P. Calain, R. Colebunders, A. De Roo, Y. Guimard, K. R. Katwiki, K. Kibadi, M. A. Kipasa, K. J. Kuvula, B. B. Mapanda, M. Massamba, K. D. Mupapa, J. J. Muyembe-Tamfum, E. Ndaberey, C. J. Peters, P. E. Rollin, E. Van den Enden, and E. Van den Enden.** 1999. Ebola hemorrhagic fever in Kikwit, Democratic Republic of the Congo: clinical observations in 103 patients. J Infect Dis **179 Suppl 1:**S1-7.
27. **Collins, P., M. Hill, E. Camargo, H. Grosfeld, R. Chanock, and B. Murphy.** 1995. Production of infectious human repitatory syncytial virus from cloned cDNA confirms an essential role for the transcription elongation factor from the 5' proximal open reading frame of the M2 mRNA in gene expression and provides a capability for vaccine development. Proc Natl Acad Sci U S A:11563-11567.
28. **Connolly, B. M., K. E. Steele, K. J. Davis, T. W. Geisbert, W. M. Kell, N. K. Jaax, and P. B. Jahrling.** 1999. Pathogenesis of experimental Ebola virus infection in guinea pigs. J Infect Dis **179 Suppl 1:**S203-17.
29. **Conzelmann, K.** 1996. Genetic manipulation of non-segmented negative-stranded RNA viruses. J Gen Virol:381-389.

30. **Conzelmann, K.** 1998. Non-segmented negative-strand RNA viruses: genetics and manipulation of viral genomes. Annual Review of Genetics:123-162.
31. **Conzelmann, K., and M. J. Schnell.** 1994. Rescue of synthetic genome RNA analogs of rabies virus by plasmid-encoded proteins. J Virol:713-719.
32. **Dimock, K., and P. Collins.** 1993. Rescue of synthetic analogs of genomic RNA and replicative-intermediate RNA of human parainfluenza virus type 3. J Virol:2772-2778.
33. **Dolnik, O., V. Volchkova, W. Garten, C. Carbonnelle, S. Becker, J. Kahnt, U. Stroher, H. D. Klenk, and V. Volchkov.** 2004. Ectodomain shedding of the glycoprotein GP of Ebola virus. Embo J **23:**2175-84.
34. **Dunn, E. F., D. C. Pritlove, H. Jin, and R. M. Elliott.** 1995. Transcription of a recombinant bunyavirus RNA template by transiently expresses bunyavirus proteins. Virology:133-143.
35. **Durbin, A. P., J. W. Siew, B. R. Murphy, and P. L. Collins.** 1997. Minimum protein requirements for transcripton and RNA replication of a minigenome of human parainfluenza virus type 3 and evaluation of the rule of six. Virology:74-83.
36. **Ebihara, H., A. Takada, D. Kobasa, H. Feldmann, S. Theriault, M. Bray, and Y. Kawaoka.** 2004. Presented at the Annual Meeting of the American Society for Virology, Montreal, Quebec, Canada.
37. **Elliott, L., M. Kiley, and J. McCormick.** 1985. Descriptive analysis of Ebola virus proteins. Virology:169-176.
38. **Elliott, L. H., S. P. Bauer, G. I. Perez-Oronoz, and E. Lloyd.** 1993. Improved specificity of testing methods for filovirus antibodies. Journal of Virological Methods **43:**85-100.
39. **Elliott, L. H., M. P. Kiley, and J. B. McCormick.** 1985. Descriptive analysis of Ebola virus proteins. Virology **147:**169-76.
40. **Elliott, L. H., A. Sanchez, B. P. Holloway, M. P. Kiley, and J. B. McCormick.** 1993. Ebola protein analyses for the determination of genetic organization. Arch Virol **133:**423-36.
41. **Enterlein, S., V. Volchkov, M. Weik, L. Kolesnikova, V. Volchkova, H.-D. Klenk, and E. Muhlberger.** 2006. Rescue of recombinant Marburg virus from cDNA is dependent on nucleocapsid protein VP30. J Virol:1038-1043.
42. **Fauquet CM, M. M., Maniloff J, Desselerger U, Ball LA.** 2004. In: Virus Taxononmy, Eighth report of the International Committee on Taxonomy of Viruses. Elsevier.
43. **Feldmann, H., and M. Kiley.** 1999. Classification, structure, and replication of filoviruses. Curr Top Microbiol Immunol:1-21.
44. **Feldmann, H., and H. D. Klenk.** 1996. Marburg and Ebola viruses. Adv Virus Res **47:**1-52.
45. **Feldmann, H., E. Muhlberger, A. Randolf, C. Will, M. P. Kiley, A. Sanchez, and H.-D. Klenk.** 1992. Marburg virus, a filovirus: messenger RNAs, gene order, and regulatory elements of the replication cycle. Virus Res:1-19.
46. **Feldmann, H., V. Volchkov, V. Volchkova, U. Stroeher, and H. D. Klenk.** 2001. Biosynthesis and role of filoviral glycoproteins. J Gen Virol:2839-2848.

47. **Feldmann, H., V. E. Volchkov, V. A. Volchkova, and H. D. Klenk.** 1999. The glycoproteins of Marburg and Ebola virus and their potential roles in pathogenesis. Arch Virol Suppl **15:**159-69.
48. **Feldmann, H., V. Wahl-Jensen, S. M. Jones, and U. Stroher.** 2004. Ebola virus ecology: a continuing mystery. Trends Microbiol **12:**433-7.
49. **Feldmann, H., C. Will, M. Schikore, W. Slenczka, and H. D. Klenk.** 1991. Glycosylation and oligomerization of the spike protein of Marburg virus. Virology:353-356.
50. **Fisher-Hoch, S. P., G. I. Perez-Oronoz, E. L. Jackson, L. M. Hermann, and B. G. Brown.** 1992. Filovirus clearance in non-human primates. Lancet **340:**451-3.
51. **Fisher-Hoch, S. P., G. S. Platt, G. H. Neild, T. Southee, A. Baskerville, R. T. Raymond, G. Lloyd, and D. I. Simpson.** 1985. Pathophysiology of shock and hemorrhage in a fulminating viral infection (Ebola). J Infect Dis **152:**887-94.
52. **Flick, K., J. Hooper, A. Schmaljohn, R. Pettersson, H. Feldmann, and R. Flick.** 2003. Rescue of Hantaan virus minigenomes. Virology:219-224.
53. **Flick, R., K. Flick, H. Feldmann, and F. Elgh.** 2003. Reverse genetics for Crimean-Congo hemorrhagic fever virus. J Virol:5997-6006.
54. **Flick, R., and R. Pettersson.** 2001. Reverse genetics system for Uukuniemi virus (Bunyaviridae): RNA polymerase I-catalyzed expression of chimeric viral RNAs. J Virol:1643-1655.
55. **Fodor, E., L. Devenish, O. G. Engelhardt, P. Palese, G. G. Browniee, and A. Garcia-Sastre.** 1999. Rescue of influenza A virus from recombinant DNA. J Virol:9679-9682.
56. **Gallaher, W. R.** 1996. Similar structural models of the transmembrane proteins of Ebola and avian sarcoma viruses. Cell **85:**477-8.
57. **Garcin, D., T. Pelet, P. Calain, L. Roux, J. Curran, and D. Kolakofsky.** 1995. A highly recombinogenic system for the recovery of infectious Sendai paramyxovirus from cDNA: generation of a novel copy-back nondefective interfering virus. Embo J:6087-6094.
58. **Geisbert, T. W., L. E. Hensley, J. B. Geisbert, and P. B. Jahrling.** 2002. Evidence against an important role for infectivity-enhancing antibodies in Ebola virus infections. Virology **293:**15-9.
59. **Geisbert, T. W., and P. B. Jahrling.** 1995. Differentiation of filoviruses by electron microscopy. Virus Res **39:**129-50.
60. **Geisbert, T. W., P. Pushko, K. Anderson, J. Smith, K. J. Davis, and P. B. Jahrling.** 2002. Evaluation in nonhuman primates of vaccines against Ebola virus. Emerg Infect Dis **8:**503-7.
61. **Geisbert, T. W., H. A. Young, P. B. Jahrling, K. J. Davis, E. Kagan, and L. E. Hensley.** 2003. Mechanisms underlying coagulation abnormalities in ebola hemorrhagic fever: overexpression of tissue factor in primate monocytes/macrophages is a key event. J Infect Dis **188:**1618-29.
62. **Geisbert, T. W., H. A. Young, P. B. Jahrling, K. J. Davis, T. Larsen, E. Kagan, and L. E. Hensley.** 2003. Pathogenesis of Ebola hemorrhagic fever in primate models: evidence that hemorrhage is not a direct effect of virus-induced cytolysis of endothelial cells. Am J Pathol **163:**2371-82.

63. **Groseth, A.** 2006. University of Manitoba, Winnipeg.
64. **Groseth, A., H. Feldmann, S. Theriault, G. Mehmetoglu, and R. Flick.** 2005. RNA polymerase I-driven minigenome system for Ebola viruses. J Virol **79:**4425-33.
65. **Groseth, A., H. Feldmann, S. Theriault, G. Mehmetoglu, and R. Flick.** 2005. An RNA Polymerase I-driven minigenome system for *Reston ebolavirus*. J Virol:4425-4433.
66. **Groseth, A., H. Feldmann, S. Theriault, G. Mehmetoglu, and R. Flick.** 2004. An RNA Polymerase I-drivern minigenome system for *Reston ebolavirus*. Submitted.
67. **Groseth, A., U. Stroher, S. Theriault, and H. Feldmann.** 2002. Molecular characterization of an isolate from the 1989/90 epizootic of Ebola virus Reston among macaques imported into the United States. Virus Res **87:**155-63.
68. **Grosfeld, H., M. Hill, and P. Collins.** 1995. RNA replication by respiratory syncytial virus (RSV) is directed by N, P, and L proteins; transcription also occurs under these conditions but requires RSV superinfection for efficient synthesis of full-lenght mRNA. J Virol:5677-5686.
69. **Guenno., B. L., P. Formentry., M. Wyers., P. Gounon., F. Walker., and C. Boesch.** 1995. Isolation and partial characterisation of a new strain of Ebola virus. The Lancet **345:**1271-1274.
70. **Gupta, M., S. Mahanty, R. Ahmed, and P. E. Rollin.** 2001. Monocyte-derived human macrophages and peripheral blood mononuclear cells infected with ebola virus secrete MIP-1alpha and TNF-alpha and inhibit poly-IC-induced IFN-alpha in vitro. Virology **284:**20-5.
71. **Hanahan, D.** 1983. Studies on transformation of Escherichia coli with plasmids. J. Mol. Biol:557-580.
72. **Hartlieb, B., J. Modrof, E. Muhlberger, H. D. Klenk, and S. Becker.** 2003. Oligomerization of Ebola virus VP30 is essential for viral transcription and can be inhibited by a synthetic peptide. J Biol Chem **278:**41830-6.
73. **Hartman, A. L., J. S. Towner, and S. T. Nichol.** 2004. A C-terminal basic amino acid motif of Zaire ebolavirus VP35 is essential for type I interferon antagonism and displays high identity with the RNA-binding domain of another interferon antagonist, the NS1 protein of influenza A virus. Virology **328:**177-84.
74. **He, B., R. Paterson, C. Ward, and R. Lamb.** 1997. Recovery of infectious SV5 from cloned DNA and expression of a foreign gene. Virology:249-260.
75. **Hensley, L. E., H. A. Young, P. B. Jahrling, and T. W. Geisbert.** 2002. Proinflammatory response during Ebola virus infection of primate models: possible involvement of the tumor necrosis factor receptor superfamily. Immunol Lett **80:**169-79.
76. **Heymann, D. L., D. Barakamfitiye, M. Szczeniowski, J. J. Muyembe-Tamfum, O. Bele, and G. Rodier.** 1999. Ebola hemorrhagic fever: lessons from Kikwit, Democratic Republic of the Congo. J Infect Dis **179 Suppl 1:**S283-6.
77. **Hoenen, T., A. Groseth, H. Ebihara, B. Hartlieb, H. Feldmann, U. Stroher, and S. Becker.** 2006. Infectious of naive target cells with virus like particles Implications for functions of ebola virus VP24. Under revision (Journal of Virology).

78. **Hoffmann, E. G., and R. G. Webster.** 2000. Unidirectional RNA polymerase I-polymerase II transcription systems for the generation of influenza A virus from eight plasmids. J Gen Virol:2843-2847.
79. **Huang, Y., L. Xu, Y. Sun, and G. J. Nabel.** 2002. The assembly of Ebola virus nucleocapsid requires virion-associated proteins 35 and 24 and posttranslational modification of nucleoprotein. Mol Cell **10:**307-16.
80. **Ikegami, T., A. B. Calaor, M. E. Miranda, M. Niikura, M. Saijo, I. Kurane, Y. Yoshikawa, and S. Morikawa.** 2001. Genome structure of Ebola virus subtype Reston: differences among Ebola subtypes. Brief report. Arch Virol **146:**2021-7.
81. **Ito, H., S. Watanabe, A. Sanchez, M. A. Whitt, and Y. Kawaoka.** 1999. Mutational analysis of the putative fusion domain of Ebola virus glycoprotein. J Virol **73:**8907-12.
82. **Ito, H., S. Watanabe, A. Takada, and Y. Kawaoka.** 2001. Ebola virus glycoprotein: proteolytic processing, acylation, cell tropism, and detection of neutralizing antibodies. J Virol **75:**1576-80.
83. **Jaax, N. K., K. J. Davis, T. J. Geisbert, P. Vogel, G. P. Jaax, M. Topper, and P. B. Jahrling.** 1996. Lethal experimental infection of rhesus monkeys with Ebola-Zaire (Mayinga) virus by the oral and conjunctival route of exposure. Arch Pathol Lab Med **120:**140-55.
84. **Jahrling, P., T. Geisbert, and D. Dalgard.** 1990. Isolation of ebola virus from monkeys imported to U.S.A. Lancet **335:**502-505.
85. **Jasenosky, L. D., and Y. Kawaoka.** 2004. Filovirus budding. Virus Res **106:**181-8.
86. **Jerpseth, B., M. Callahan, and A. Greener.** 1997. Strategies:3-4.
87. **Johnson, E., N. Jaax, J. White, and P. Jahrling.** 1995. Lethal experimental infections of rhesus monkeys by aerosolized Ebola virus. Int J Exp Pathol **76:**227-36.
88. **Kato, A., Y. Sakai, T. Shioda, T. Kondo, M. Nakanishi, and Y. Nagai.** 1996. Initiation of Sendi virus multiplication from transfected cDNA or RNA with negative or positive sense. Genes Cells:569-579.
89. **Kerstiens, B., and F. Matthys.** 1999. Interventions to control virus transmission during an outbreak of Ebola hemorrhagic fever: experience from Kikwit, Democratic Republic of the Congo, 1995. J Infect Dis **179 Suppl 1:**S263-7.
90. **Kiley, M., N. Cox, L. Elliott, A. Sanchez, R. DeFries, M. Buchmeier, D. Richman, and J. McCormick.** 1988. Physicochemical properties of marburg virus: evidence for three distinct virus strains and their relationship to ebola virus. J. Gen. Virol.:1957-1967.
91. **Kiley, M. P., E. T. Bowen, G. A. Eddy, M. Isaacson, K. M. Johnson, J. B. McCormick, F. A. Murphy, S. R. Pattyn, D. Peters, O. W. Prozesky, R. L. Regnery, D. I. Simpson, W. Slenczka, P. Sureau, G. van der Groen, P. A. Webb, and H. Wulff.** 1982. Filoviridae: a taxonomic home for Marburg and Ebola viruses? Intervirology **18:**24-32.
92. **Klenk, H. D., and W. Garten.** 1994. Host cell proteases controlling virus pathogenicity. Trends in Microbiology:39-43.

93. **Kobasa, D., M. E. Rodgers, K. Wells, and Y. Kawaoka.** 1997. Neuraminidase hemadsorption activity, conserved in avian influenza A viruses, does not influence viral replication in ducks. J Virol:6706-6713.
94. **Krieg, P. A., and D. A. Melton.** 1987. In vitro RNA synthesis with SP6 RNA polymerase. Meth. Enzymol.:397-415.
95. **Lawson, N. D., E. A. Stillmann, M. A. Whitt, and J. K. Rose.** 1995. Recombinant vesicular stomatitis viruses from cDNA. Proc Natl Acad Sci U S A:4477-4481.
96. **Lee, K., I. Novella, M. Teng, M. Oldstone, and J. De la Torre.** 2000. NP and L proteins of lymphocytic chorimeningitis virus (LCMV) are sufficient for efficient transcription and replication of LCMV genomic RNA analogs. J Virol:3470-3477.
97. **Lee, K. H., M. Perez, D. Pinschewer, and J. De la Torre.** 2002. Identification of the lymphocytic choriomeningitis virus (LCMV) proteins required to rescue LCMV RNA analogs into LCMV-like particles. J Virol:6393-6397.
98. **Leroy, E., B. Kumulungui, X. Pourrut, P. Rouquet, A. Hassanin, P. Yaba, A. Delicat, J. Paweska, J. P. Gonzalez, and R. Swanapoel.** 2005. Fruit bats as reservoirs of ebola virus. Nature **438:**Online version.
99. **Lopez, N., R. Jacamo, and M. T. Franze-Fernandez.** 2001. Transcription and RNA replication of tacaribe virus genome and antigenomic analogs require N and L proteins: Z protein is an inhibitor of these processes. J Virol:12241-12251.
100. **Lopez, N., R. Muller, C. Prehaud, and M. Bouloy.** 1995. The L protein of Rift Vally fever virus can rescue viral ribonucleoproteins and transcribe synthesis genome-like RNA molecule. J Virol:3972-3979.
101. **Luytjes, W., M. Krystal, M. Enami, J. Pavin, and P. Palese.** 1989. Amplication, expression, and packaging of foreign gene by influenza virus. Cell **59**.
102. **Mahanty, S., and M. Bray.** 2004. Pathogenesis of filoviral haemorrhagic fevers. Lancet Infect Dis **4:**487-98.
103. **Martini, G. A., H. G. Knauff, H. A. Schmidt, G. Mayer, and G. Baltzer.** 1968. A hitherto unknown infectious disease contracted from monkeys. "Marburg-virus" disease. Ger Med Mon.:457-470.
104. **Melton, D., P. A. Krieg, M. R. Rebagliati, T. Maniatis, K. Zinn, and M. R. Green.** 1984. Efficient in vitro synthesis of biologically active RNA and RNA hybridization probes form plasmids containing a bacteriophage SP6 promoter. Nucl. Acids Res.:7035.
105. **Modrof, J., C. Moritz, L. Kolesnikova, T. Konakova, B. Hartlieb, A. Randolf, E. Muhlberger, and S. Becker.** 2001. Phosphorylation of Marburg virus VP30 at serines 40 and 42 is critical for its interaction with NP inclusions. Virology:171-182.
106. **Modrof, J., E. Muhlberger, H. D. Klenk, and S. Becker.** 2002. Phosphorylation of VP30 impairs ebola virus transcription. J Biol Chem **277:**33099-104.
107. **Moss, B., O. Elroy-Stein, T. Mizukami, W. A. Alexander, and T. R. Fuerst.** 1990. New mammalian expression vectors. Nature **348:**91.
108. **Moyer, S. A.** 1989. Replication of the genome RNAs of defective interfering particles of vesicular stomatitis and sendai viruses using heterologous viral proteins. Virology:341-345.

109. **Muhlberger, E.** 2004. Genome organization replication and transcription of filoviruses. Horizon Bioscience, Wymondham, norfolk.
110. **Muhlberger, E., B. Lotfering, H. D. Klenk, and S. Becker.** 1998. Three of the four nucleocapsid proteins of Marburg virus, NP, VP35, and L are sufficient to mediate replication and transcription of Marburg virus-specific monocistronic minigenomes. J Virol:8756-8764.
111. **Muhlberger, E., M. Weik, V. E. Volchkov, H. D. Klenk, and S. Becker.** 1999. Comparison of the transcription and replication strategies of marburg virus and Ebola virus by using artificial replication systems. J Virol **73:**2333-42.
112. **Murphy FA, v. d. G. G., Whitfield SG, Lange JV.** 1978. Ebola virus hemorrhagic fever, 1st edn. ed. Elsevier, North-Holland.
113. **Nagai, Y., and A. Kato.** 1999. Paramyxovirus reverse genetics is coming of age. Microbiol Immunol:613-624.
114. **Neumann, G., H. Feldmann, S. Watanabe, I. Lukashevich, and Y. Kawaoka.** 2002. Reverse genetics demonstrates that proteolytic processing of the Ebola virus glycoprotein is not essential for replication in cell culture. J Virol **76:**406-10.
115. **Neumann, G., T. Noda, A. Takada, L. D. Jasenosky, and Y. Kawaoka.** 2004. Roles of filoviral matrix-and glycoproteins in the viral life cycle. Horizon Bioscience, Wymondham, Norfolk.
116. **Neumann, G., T. Watanabe, H. Ito, S. Watanabe, H. Goto, P. Gao, M. Hughes, D. R. Perez, R. Donis, E. Hoffmann, G. Hobom, and Y. Kawaoka.** 1999. Generation of influenza A virus entirely from cloned cDNAs. Proc Natl Acad Sci U S A:9345-9350.
117. **Neumann, G., M. A. Whitt, and Y. Kawaoka.** 2002. A decade after the generation of a negative-sense RNA virus from cloned cDNA - what have we learned? J Gen Virol **83:**2635-62.
118. **Neumann, G., A. Zobel, and G. Hobom.** 1994. RNA polymerase I-mediated expression of influenza viral RNA molecules. Virology:477-479.
119. **Niwa, H., K. Yamamura, and J. Miyazaki.** 1991. Efficient selection for high-expression transfectants with a novel eukaryotic vector. Gene **108.**
120. **Noda, T., H. Sagara, E. Suzuki, A. Takada, H. Kida, and Y. Kawaoka.** 2002. Ebola virus VP40 drives the formation of virus-like filamentous particles along with GP. J Virol **76:**4855-65.
121. **Parks, C. L., R. A. Lerch, P. Walpita, M. S. Sidhu, and S. A. Udem.** 1999. Enhanced maesles virus cDNA rescue and gene expression after heat shock. J Virol:3560-3566.
122. **Pattnaik, A. K., and G. W. Wertz.** 1990. Replication and amplification of defective interfering particles RNA's of vesicular stomatitis virus in cells expressing viral proteins from vectors containing cloned cDNAs. J Virol:2948-2957.
123. **Pattyn, S.** 1978. Ebola virus hemorrhagic fever, 1 ed. Elsevier/North-Holland, Amsterdam.
124. **Pelet, T., C. Delenda, O. Gubbay, D. Garcin, and D. Kolakofsky.** 1995. Partial characterization of a Sendi virus replication promoter and the rule of six. Virology:405-414.

125. **Pelet, T., J. B. Marq, Y. Sakai, S. Wakao, H. Gotoh, and J. Curran.** 1996. Rescue of Sendi virus cDNA templates with cDNA clones expressing parainfluenza virus type 3, N, P, and L proteins. J Gen Virol **77**.
126. **Peters CJ, M. G., Slenczka W.** 1971. Morphology, development, and classification of marburg virus, 1st edn. ed. Springer, Heidelberg Berline New York.
127. **Racaniello, V., and D. Baltimore.** 1981. Cloned polivirus complementry DNA if infectious in mammalian cells. Science:916-919.
128. **Radecke, F.**
129. **Radecke, F., P. Spielhofer, H. Schneider, K. Kaelin, M. Huber, C. Dotsch, G. Christiansen, and M. A. Billeter.** 1995. Rescue of measles viruses from cloned cDNA. Embo J:5773-5784.
130. **Reed, D. S., L. E. Hensley, J. B. Geisbert, P. B. Jahrling, and T. W. Geisbert.** 2004. Depletion of peripheral blood T lymphocytes and NK cells during the course of ebola hemorrhagic Fever in cynomolgus macaques. Viral Immunol **17:**390-400.
131. **Roberts, A., L. Buonocore, R. Price, J. Forman, and J. K. Rose.** 1999. Attenuated vesicular stomatitis viruses as vaccine vectors. J Virol:3723-3732.
132. **Roberts, A., E. Kretzschmar, A. S. Perkins, J. Forman, R. Price, L. Buonocore, Y. Kawaoka, and J. K. Rose.** 1998. Vaccination with a recombinate vesicular stomatitis virus expressing an influenza virus hemaggulutinin provides complete protection from influenza virus challenge. J Virol:4704-4711.
133. **Rollin, P. E., T. G. Ksiazek, P. B. Jahrling, M. Haines, and C. J. Peters.** 1990. Detection of Ebola-like viruses by immunofluorescence. Lancet **336:**1591.
134. **Rotz, L. D., A. S. Khan, S. R. Lillibridge, S. M. Ostroff, and J. M. Hughes.** 2002. Public health assessment of potential biological terrorism agents. Emerg Infect Dis:225-230.
135. **Ruiz-Arguello, M. B., F. M. Goni, F. B. Pereira, and J. L. Nieva.** 1998. Phosphatidylinositol-dependent membrane fusion induced by a putative fusogenic sequence of Ebola virus. J Virol **72:**1775-81.
136. **Sambrook, J., E. F. Fritsh, and T. Maniatis.** 1989. "Molecular cloning, A laboratory manual."
137. **Sanchez, A., and M. P. Kiley.** 1987. Identification and analysis of Ebola virus messenger RNA. Virology **157:**414-20.
138. **Sanchez, A., M. Lukwiya, D. Bausch, S. Mahanty, A. J. Sanchez, K. D. Wagoner, and P. E. Rollin.** 2004. Analysis of human peripheral blood samples from fatal and nonfatal cases of Ebola (Sudan) hemorrhagic fever: cellular responses, virus load, and nitric oxide levels. J Virol **78:**10370-7.
139. **Sanchez, A., S. G. Trappier, B. W. Mahy, C. J. Peters, and S. T. Nichol.** 1996. The virion glycoproteins of Ebola viruses are encoded in two reading frames and are expressed through transcriptional editing. Proc Natl Acad Sci U S A **93:**3602-7.
140. **Sanchez, A., Z. Y. Yang, L. Xu, G. J. Nabel, T. Crews, and C. J. Peters.** 1998. Biochemical analysis of the secreted and virion glycoproteins of Ebola virus. J Virol **72:**6442-7.

141. **Sauchez, A., A. S. Khan, S. Zaki, G. Nabel, T. Ksiazek, and P. C.** 2001. 'Filoviridae' Marburg and Ebola Viruses, 4th edn ed, vol. 1. Lippincott Williams and Wilkins, Philadelphia.
142. **Schneider, U., A. Ohnemus, M. Schwemmle, and P. Staeheli.** 2004. Presented at the Annuall Meeting of the American Society for Virology, Montreal, Quebec, Canada.
143. **Schnell, M. J., T. Mebatsion, and K. Conzelmann.** 1994. Infectious rabies viruses from cloned cDNA. EMBO Journal:4195-4203.
144. **Short, J. M., J. M. Fernandez, J. A. Sorge, and W. Huse.** 1988. Lambda ZAP: a bacteriophage lambda expression vector with in vivo excision properties. Gene:7583-7599.
145. **Shuman, S.** 1991. Recombination mediated by vaccinia virus DNA topoisomerase I in Escherichia coli is sequence specific. Proc Natl Acad Sci U S A:10104-10108.
146. **Simpson, D., I. Zoltnik, and D. A. Rutter.** 1968. Vervet monkey disease. Experiment infection of guinea pigs and monkeys with the causative agent. British Journal of experimental pathology:458-464.
147. **Sleat, D., and A. Banerjee.** 1993. Transcriptional activity and mutational analysis of recombinant vesicular stomatitis virus RNA polymerase. J. Virol.:1334-1339.
148. **Slenczka, W.** 1999. The Marburg Virus Outbreak of 1967 and subsequent Episodes, vol. 235. Springer-Verlag, Berlin, Germany.
149. **Slenczka, W., H.-L. Shu, G. Piepenberg, and R. Siegert.** 1968. Detection of the antigen of the "Marburg Virus" in the organs of infected Guinea-Pigs by immunoflorescence. German Medical Monthly (Stuttgart):524-529.
150. **Smith, C. E., D. Simpson, E. T. Bowen, and I. Zoltnik.** 1967. Fatel Human disease from Vervet monkeys. Lancet:1119-1121.
151. **Stillmann, E. A., J. K. Rose, and M. A. Whitt.** 1995. Replication and amplification of novel vesicular stomatitis virus minigenomes encoding viral structural proteins. J Virol:2946-2953.
152. **Strachan, T., and A. Read.** 2000. Human Molecular Genetics, 2 ed. BIOS Scientific Publishers Ltd., New York.
153. **Stroher, U., E. West, H. Bugany, H. D. Klenk, H. J. Schnittler, and H. Feldmann.** 2001. Infection and activation of monocytes by Marburg and Ebola viruses. J Virol **75:**11025-33.
154. **Takada, A., K. Fujioka, M. Tsuiji, A. Morikawa, N. Higashi, H. Ebihara, D. Kobasa, H. Feldmann, T. Irimura, and Y. Kawaoka.** 2004. Human macrophage C-type lectin specific for galactose and N-acetylgalactosamine promotes filovirus entry. J Virol **78:**2943-7.
155. **Takada, A., C. Robison, H. Goto, A. Sanchez, K. G. Murti, M. A. Whitt, and Y. Kawaoka.** 1997. A system for functional analysis of Ebola virus glycoprotein. Proc Natl Acad Sci U S A **94:**14764-9.
156. **Taniguchi, T., M. Palmieri, and C. Weissmann.** 1978. QB DNA containing hybrid plasmids giving rise to QB phage formation in the bacterial host. Nature:223-228.

157. **Theriault, S., A. Groseth, H. Artsob, and H. Feldmann.** 2005. The role of reverse genetics systems in determining filovirus pathogenicity. Arch Virol **[Suppl]:**157-177.
158. **Theriault, S., A. Groseth, G. Neumann, Y. Kawaoka, and H. Feldmann.** 2004. Rescue of Ebola virus from cDNA using heterologous support proteins. Virus Res **106:**43-50.
159. **Towner, J. S., J. Paragas, J. E. Dover, M. Gupta, C. S. Goldsmith, J. W. Huggins, and S. T. Nichol.** 2005. Generation of eGFP expressing recombinant Zaire ebolavirus for analysis of early pathogenesis events and high-throughput antiviral drug screening. Virology **332:**20-7.
160. **Towner, J. S., P. E. Rollin, D. G. Bausch, A. Sanchez, S. M. Crary, M. Vincent, W. F. Lee, C. F. Spiropoulou, T. G. Ksiazek, M. Lukwiya, F. Kaducu, R. Downing, and S. T. Nichol.** 2004. Rapid diagnosis of Ebola hemorrhagic fever by reverse transcription-PCR in an outbreak setting and assessment of patient viral load as a predictor of outcome. J Virol **78:**4330-41.
161. **Volchkov, V., V. Volchkova, O. Dolnik, H. Feldmann, and H. D. Klenk.** 2004. Structural and functional polymorphism of the glycoprotein of filoviruses. Horizon Bioscience, Wymondham Norfolk.
162. **Volchkov, V. E., S. Becker, V. A. Volchkova, V. A. Ternovoj, A. N. Kotov, S. V. Netesov, and H. D. Klenk.** 1995. GP mRNA of Ebola virus is edited by the Ebola virus polymerase and by T7 and vaccinia virus polymerases. Virology **214:**421-30.
163. **Volchkov, V. E., A. A. Chepurnov, V. A. Volchkova, V. A. Ternovoj, and H. D. Klenk.** 2000. Molecular characterization of guinea pig-adapted variants of Ebola virus. Virology **277:**147-55.
164. **Volchkov, V. E., H. Feldmann, V. A. Volchkova, and H. D. Klenk.** 1998. Processing of the Ebola virus glycoprotein by the proprotein convertase furin. Proc Natl Acad Sci U S A **95:**5762-7.
165. **Volchkov, V. E., V. A. Volchkova, A. A. Chepurnov, V. M. Blinov, O. Dolnik, S. V. Netesov, and H. Feldmann.** 1999. Characterization of the L gene and 5' trailer region of Ebola virus. J Gen Virol **80 (Pt 2):**355-62.
166. **Volchkov, V. E., V. A. Volchkova, E. Muhlberger, L. V. Kolesnikova, M. Weik, O. Dolnik, and H. D. Klenk.** 2001. Recovery of infectious Ebola virus from complementary DNA: RNA editing of the GP gene and viral cytotoxicity. Science **291:**1965-9.
167. **Volchkov, V. E., V. A. Volchkova, W. Slenczka, H. D. Klenk, and H. Feldmann.** 1998. Release of viral glycoproteins during Ebola virus infection. Virology **245:**110-9.
168. **Volchkova, V. A., H. Feldmann, H. D. Klenk, and V. E. Volchkov.** 1998. The nonstructural small glycoprotein sGP of Ebola virus is secreted as an antiparallel-orientated homodimer. Virology **250:**408-14.
169. **Volchkova, V. A., H. D. Klenk, and V. E. Volchkov.** 1999. Delta-peptide is the carboxy-terminal cleavage fragment of the nonstructural small glycoprotein sGP of Ebola virus. Virology **265:**164-71.

170. **Watanabe, S., A. Takada, T. Watanabe, H. Ito, H. Kida, and Y. Kawaoka.** 2000. Functional importance of the coiled-coil of the Ebola virus glycoprotein. J Virol **74:**10194-201.
171. **Weik, M., J. Modrof, H. D. Klenk, S. Becker, and E. Muhlberger.** 2002. Ebola virus VP30-mediated transcription is regulated by RNA secondary structure formation. J Virol **76:**8532-9.
172. **Weiner, M. P., C. Anderson, B. Jerpseth, S. Wells, and B. Johnson-Browne.** 1994. Strategies:41-43.
173. **Weissenhorn, W.** 2004. Structure of viral proteins. Horizon Bioscience, Wymondham Norfolk.
174. **Wool-Lewis, R. J., and P. Bates.** 1999. Endoproteolytic processing of the ebola virus envelope glycoprotein: cleavage is not required for function. J Virol **73:**1419-26.
175. **Yang, Z., R. Delgado, L. Xu, R. F. Todd, E. G. Nabel, A. Sanchez, and G. J. Nabel.** 1998. Distinct cellular interactions of secreted and transmembrane Ebola virus glycoproteins. Science **279:**1034-7.
176. **Yang, Z. Y., H. J. Duckers, N. J. Sullivan, A. Sanchez, E. G. Nabel, and G. J. Nabel.** 2000. Identification of the Ebola virus glycoprotein as the main viral determinant of vascular cell cytotoxicity and injury. Nat Med **6:**886-9.
177. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **1:**103-119.
178. **Yu, Q., R. W. Hardy, and G. W. Wertz.** 1995. Functional cDNA clones of the human respiratory syncytial (RS) virus N, P, and L proteins support replication of RS virus genomic RNA analogs and define minimal *trans-acting* requirements for RNA replication. J Virol:2412-2419.

6.0 Appendix:

Primer list for full clone sequencing project

Primer name	Location base pair	Sequence	Region within genome
S1f	1~	CTA TAG CGG ACA CAC	Leader region
S2r	~897	CGG CAT TAG CTT CAG	NP
S3f	483~	CTC AGA AAA TCT GGA TGC	NP
S4r	~483	CCA TCC AGA TTT TCT GAG	Leader region
S5f	897~	CTG AAG CTA ATG CCG	NP
S6r	~1895	ATC TGG TGC ATT CAT GCC	NP
S7f	1531~	CCA ACA ATA TGC CGA GT	NP
S8r	~1531	CTC TGC ATA TTG TTG G	NP
S9f	1895~	GGC ATG AAT GCA CCA GAT	NP
S10r	~2840	GGC TAG TAA TAA TAA GC	Intergenic region
S11f	2380~	GGA CAG TGA CAA CAC CC	Intergenic region
S12r	~2380	GGG TGT TGT CAC TGT CC	NP
S13f	2840~	GCT TAT TAT TAC TAG CC	VP35
S14r	~3834	GGT GGA AAG CAG TTC C	VP35
S15f	3331~	CGC GCA ACA GTC AAA CCC	VP35
S16r	~3331	GGG TTT GAC TGT TGC	VP35

S17f	3834~	GGA ACT GCT TTC CAC C	VP35
S18r	~4830	GGT GAT AGT GTA TGA AGC	VP40
S19f	4331~	GGA GCT ATA TCT CTG ACA G	VP40
S20r	~4331	CTG TCA GAG ATA TAG CTC C	VP40
S21f	4830~	GCT TCA TAC ACT ATC ACC	VP40
S22r	~5831	CCT AGA AAT GGG TCC CG	VP40
S23f	5333~	CCC TGT TCT TTT GCC	VP40
S24r	~5333	GGC AAA AGA ACA GGG	VP40
S25f	5831~	CGG GAC CCA TTT CTA GG	Intergenic region
S26r	~6829	GGT ATT GCT CCT TTT CC	GP
S27f	6326~	GGT CAA TTA TGA AGC	GP
S28r	~6326	GCT TCA TAA TTG ACC	GP
S29f	6829~	GGA AAA GGA GCA ATA CC	GP
S30r	~7840	GGT CCC AGA ATG TGG	GP
S31f	7331~	GCA GAG AAC ACC AAC ACG	GP
S32r	~7331	CGT GTT GGT GTT CTC TGC	GP
S33f	7840~	CCA CAT TCT GGG ACC	GP
S34r	~8823	CCA CAA GTC TTA CGG	VP30
S35f	8336~	CCT TGA TTC TAC AAT C	VP30
S36r	~8336	GAT TGT AGA ATC AAG G	VP30
S37f	8823~	CCG TAA GAC TTG TGG	VP30
S38r	~9842	CCT CTT GAA ACA AG	Intergenic region
S39f	9321~	GGA AGC TTC AAC CAA CC	Intergenic region

S40r	~9321	GGT TGG TTG AAG CTT CC	VP30
S41f	9842~	CTT GTT TCA AGA GG	Intergenic region
S42r	~10836	CGA CAT GTA GAG CAT CC	VP24
S43f	10321~	CCA AGC AAG ACC TGA G	VP24
S44r	~10321	CTC AGG TCT TGC TTG G	VP24
S45f	10836~	GGA TGC TCT ACA TGT CG	VP24
S46r	~11826	GCC TGA CAG TGC C	Intergenic region
S47f	11338~	GCT TTA TTA TAT G	Intergenic region
S48r	~11338	CAT ATA ATA AAG C	Intergenic region
S49f	11826~	GGC ACT GTC AGG C	Polymerase
S50r	~12850	CTG TAC CAA GAT CC	Polymerase
S51f	12319~	CGA TTC AAC ACA AC	Polymerase
S52r	~12319	GTT GTG TTG AAT CG	Polymerase
S53f	12850~	GGA TCT TGG TAC AG	Polymerase
S54r	~13839	GCT CGT CTG CGT C	Polymerase
S55f	13336~	GCT AAA GCA TTT CC	Polymerase
S56r	~13336	GGA AAT GCT TTA GC	Polymerase
S57f	13839~	GAC GCA GAC GAG C	Polymerase
S58r	~14822	CCA TAG GCT CCA CC	Polymerase
S59f	14318~	CCT TCT TGA ATC CTG	Polymerase
S60r	~14318	CAG GAT TCA AGA AGG	Polymerase
S61f	14822~	GGT GGA GCC TAT GG	Polymerase
S62r	~15828	CCA GAT AAG TGA GGC	Polymerase

S63f	15332~	GTG TTC AAG AAA TAC	Polymerase
S64r	~15332	GTA TTT CTT GAA CAC	Polymerase
S65f	15828~	GCC TCA CTT ATC TGG	Polymerase
S66r	~16826	CCT TCC CTC TTG G	Polymerase
S67f	16234~	GAG GGT CAA AAC CC	Polymerase
S68r	~16324	GGG TTT TGA CCC TC	Polymerase
S69f	16826~	CCA AGA GGG AAG G	Polymerase
S70r	~17818	GAG ACT AGT GGA CC	Polymerase
S71f	17323~	GCA AGG CTA CCT AAG C	Polymerase
S72r	~17323	GCT TAG GTA GCC TTG C	Polymerase
S73f	17818~	GGT CCA CTA GTC TC	Polymerase
S74r	~18841	GCG TGG TCA ATG TC	Trailer
S75f	18343~	CGA AAG GAG TCC C	Trailer
S76r	~18343	GGG ACT CCT TTC G	Trailer

Table 15. List of sequencing primers used to sequence the entire guinea pig adapted virus genome

12 Jan 2006 Alignment Results

Alignment: Global DNA alignment against reference molecule
Parameters: Scoring matrix: Linear (Mismatch 2, OpenGap 4, ExtGap 1)

Reference molecule: I-2 seq, Region 1-23313
Number of sequences to align: 2
Settings: Similarity significance value cutoff: >= 60%

Summary of Percent Matches:

Reference:	I-2 seq	1 - 23313	(23313 bps)	--
Sequence 2:	G-pig Win	1 - 18957	(18957 bps)	81%

```
I-2 seq          1 aagcttcttgaagacaaacaacgtctgtagcgacccttgcaggcagcggaacccccac
G-pig Win        1 ------------------------------------------------------------

I-2 seq         61 ctggcgacaggtgcctctgcggccaaaagccacgtgtataagatacacctgcaaaggcgg
G-pig Win        1 ------------------------------------------------------------

I-2 seq        121 cacaaccccagtgccacgttgtgagttggatagttgtggaaagagtcaaatggctctcct
G-pig Win        1 ------------------------------------------------------------

I-2 seq        181 caagcgtattcaacaaggggctgaaggatgcccagaaggtaccccattgtatgggatctg
G-pig Win        1 ------------------------------------------------------------

I-2 seq        241 atctggggcctcggtgcacatgctttacatgtgtttagtcgaggttaaaaaacgtctagg
G-pig Win        1 ------------------------------------------------------------

I-2 seq        301 ccccccgaaccacggggacgtggttttcctttgaaaaacacgataataccatgggcggcc
G-pig Win        1 ------------------------------------------------------------

I-2 seq        361 gcgaattctaatacgactcactatagcggacacacaaaaagaaagaagaatttttaggat
G-pig Win        1 --------------------------cggacacacaaaaagaaagaagaatttttaggat

I-2 seq        421 cttttgtgtgcgaataactatgaggaagattaataatttcctctcattgaaatttatat
G-pig Win       35 cttttgtgtgcgaataactatgaggaagattaataatttcctctcattgaaatttatat

I-2 seq        481 cggaatttaaattgaaattgttactgtaatcacacctggtttgtttcagagccacatcac
G-pig Win       95 cggaatttaaattgaaattgttactgtaatcacacctggtttgtttcagagccacatcac

I-2 seq        541 aaagatagagaacaacctaggtctccgaagggagcaagggcatcagtgtgctcagttgaa
G-pig Win      155 aaagatagagaacaacctaggtctccgaagggagcaagggcatcagtgtgctcagttgaa

I-2 seq        601 aatcccttgtcaacacctaggtcttatcacatcacaagttccacctcagactctgcaggg
G-pig Win      215 aatcccttgtcaacacctaggtcttatcacatcacaagttccacctcagactctgcaggg

I-2 seq        661 tgatccaacaaccttaatagaaacattattgttaaaggacagcattagttcacagtcaaa
G-pig Win      275 tgatccaacaaccttaatagaaacattattgttaaaggacagcattagttcacagtcaaa

I-2 seq        721 caagcaagattgagaattaaccttggttttgaacttgaacacttaggggattgaagattc
G-pig Win      335 caagcaagattgagaattaaccttggttttgaacttgaacacttaggggattgaagattc

I-2 seq        781 aacaaccctaaagcttggggtaaaacattggaaatagttaaaagacaaattgctcggaat
G-pig Win      395 aacaaccctaaagcttggggtaaaacattggaaatagttaaaagacaaattgctcggaat

I-2 seq        841 cacaaaattccgagtatggattctcgtcctcagaaaatctggatggcgccgagtctcact
G-pig Win      455 cacaaaattccgagtatggattctcgtcctcagaaaatctggatggcgccgagtctcact

I-2 seq        901 gaatctgacatggattaccacaagatcttgacagcaggtctgtccgttcaacaggggatt
G-pig Win      515 gaatctgacatggattaccacaagatcttgacagcaggtctgtccgttcaacaggggatt

I-2 seq        961 gttcggcaaagagtcatcccagtgtatcaagtaaacaatcttgaagaaatttgccaactt
G-pig Win      575 gttcggcaaagagtcatcccagtgtatcaagtaaacaatcttgaagaaatttgccaactt
```

```
I-2 seq        1021 atcatacaggcctttgaagcaggtgttgattttcaagagagtgcggacagtttccttctc
G-pig Win       635 atcatacaggcctttgaagcaggtgttgattttcaagagagtgcggacagtttccttctc

I-2 seq        1081 atgctttgtcttcatcatgcgtaccagggagattacaaacttttcttggaaagtggcgca
G-pig Win       695 atgctttgtcttcatcatgcgtaccagggagattacaaacttttcttggaaagtggcgca

I-2 seq        1141 gtcaagtatttggaagggcacgggttccgttttgaagtcaagaagcgtgatggagtgaag
G-pig Win       755 gtcaagtatttggaagggcacgggttccgttttgaagtcaagaagcgtgatggagtgaag

I-2 seq        1201 cgccttgaggaattgctgccagcagtatctagtggaaaaaacattaagagaacacttgct
G-pig Win       815 cgccttgaggaattgctgccagcagtatctagtggaaaaaacattaagagaacacttgct

I-2 seq        1261 gccatgccggaagaggagacaactgaagctaatgccggtcagtttctctcctttgcaagt
G-pig Win       875 gccatgccggaagaggagacaactgaagctaatgccggtcagtttctctcctttgcaagt

I-2 seq        1321 ctattccttccgaaattggtagtaggagaaaaggcttgccttgagaaggttcaaaggcaa
G-pig Win       935 ctattccttccgaaattggtagtaggagaaaaggcttgccttgagaaggttcaaaggcaa

I-2 seq        1381 attcaagtacatgcagagcaaggactgatacaatatccaacagcttggcaatcagtagga
G-pig Win       995 attcaagtacatgcagagcaaggactgatacaatatccaacagcttggcaatcagtagga

I-2 seq        1441 cacatgatggtgattttccgtttgatgcgaacaaatttttctgatcaaatttctcctaata
G-pig Win      1055 cacatgatggtgattttccgtttgatgcgaacaaatttttctgatcaaatttctcctaata

I-2 seq        1501 caccaagggatgcacatggttgccgggcatgatgccaacgatgctgtgatttcaaattca
G-pig Win      1115 caccaagggatgcacatggttgccgggcatgatgccaacgatgctgtgatttcaaattca

I-2 seq        1561 gtggctcaagctcgtttttcaggcttattgattgtcaaaacagtacttgatcatatcctа
G-pig Win      1175 gtggctcaagctcgtttttcaggcttattgattgtcaaaacagtacttgatcatatcctа

I-2 seq        1621 caaaagacagaacgaggagttcgtctccatcctcttgcaaggaccgccaaggtaaaaaat
G-pig Win      1235 caaaagacagaacgaggagttcgtctccatcctcttgcaaggaccgccaaggtaaaaaat

I-2 seq        1681 gaggtgaactccttaaggctgcactcagctcccctggccaagcatggagagtatgctcct
G-pig Win      1295 gaggtgaactccttaaggctgcactcagctcccctggccaagcatggagagtatgctcct

I-2 seq        1741 ttcgcccgacttttgaacctttctggagtaaataatcttgagcatggtcttttccctcaa
G-pig Win      1355 ttcgcccgacttttgaacctttctggagtaaataatcttgagcatggtcttttccctcaa

I-2 seq        1801 ctatcggcaattgcactcggagtcgccacagcacacgggagtaccctcgcaggagtaaat
G-pig Win      1415 ctatcggcaattgcactcggagtcgccacagcacacgggagtaccctcgcaggagtaaat

I-2 seq        1861 gttggagaacagtatcaacaactcagagaggctgccactgaggctgagaagcaactccaa
G-pig Win      1475 gttggagaacagtatcaacaactcagagaggctgccactgaggctgagaagcaactccaa

I-2 seq        1921 caatatgcagagtctcgcgaacttgaccatcttggacttgatgatcaggaaaagaaaatt
G-pig Win      1535 caatatgcagagtctcgcgaacttgaccatcttggacttgatgatcaggaaaagaaaatt

I-2 seq        1981 cttatgaacttccatcagaaaaagaacgaaatcagcttccagcaaacaaacgctatggta
G-pig Win      1595 cttatgaacttccatcagaaaaagaacgaaatcagcttccagcaaacaaacgctatggta

I-2 seq        2041 actctaagaaaagagcgcctggccaagctgacagaagctatcactgctgcgtcactgccc
G-pig Win      1655 actctaagaaaagagcgcctggccaagctgacagaagctatcactgctgcgtcactgccc

I-2 seq        2101 aaaacaagtggacattacgatgatgatgacgacattccctttccaggacccatcaatgat
G-pig Win      1715 aaaacaagtggacattacgatgatgatgacgacattccctttccaggacccatcaatgat

I-2 seq        2161 gacgacaatcctggccatcaagatgatgatccgactgactcacaggatacgaccattccc
G-pig Win      1775 gacgacaatcctggccatcaagatgatgatccgactgactcacaggatacgaccattccc

I-2 seq        2221 gatgtggtggttgatcccgatgatggaagctacggcgaataccagagttactcggaaaac
G-pig Win      1835 gatgtggtggttgatcccgatgatggaagctacggcgaataccagagttactcggaaaac
```

```
I-2 seq        2281 ggcatgaatgcaccagatgacttggtcctattcgatctagacgaggacgacgaggacact
G-pig Win      1895 ggcatgaatgcaccagatgacttggtcctattcgatctagacgaggacgacgaggacact

I-2 seq        2341 aagccagtgcctaatagatcgaccaagggtggacaacagaagaacagtcaaaagggccag
G-pig Win      1955 aagccagtgcctaatagatcgaccaagggtggacaacagaagaacagtcaaaagggccag

I-2 seq        2401 catatagagggcagacagacacaatccaggccaattcaaaatgtcccaggccctcacaga
G-pig Win      2015 catatagagggcagacagacacaatccaggccaattcaaaatgtcccaggccctcacaga

I-2 seq        2461 acaatccaccacgccagtgcgccactcacggacaatgacagaagaaatgaaccctccggc
G-pig Win      2075 acaatccaccacgccagtgcgccactcacggacaatgacagaagaaatgaaccctccggc

I-2 seq        2521 tcaaccagccctcgcatgctgacaccaattaacgaagaggcagacccactggacgatgcc
G-pig Win      2135 tcaaccagccctcgcatgctgacaccaattaacgaagaggcagacccactggacgatgcc

I-2 seq        2581 gacgacgagacgtctagccttccgcccttggagtcagatgatgaagagcaggacagggac
G-pig Win      2195 gacgacgagacgtctagccttccgcccttggagtcagatgatgaagagcaggacagggac

I-2 seq        2641 ggaacttccaaccgcacacccactgtcgccccaccggctcccgtatacagagatcactct
G-pig Win      2255 ggaacttccaaccgcacacccactgtcgccccaccggctcccgtatacagagatcactct

I-2 seq        2701 gaaaagaaagaactcccgcaagacgagcaacaagatcaggaccacactcaagaggccagg
G-pig Win      2315 gaaaagaaagaactcccgcaagacgagcaacaagatcaggaccacactcaagaggccagg

I-2 seq        2761 aaccaggacagtgacaacacccagtcagaacactcttttgaggagatgtatcgccacatt
G-pig Win      2375 aaccaggacagtgacaacacccagtcagaacactctcttgaggagatgtatcgccacatt

I-2 seq        2821 ctaagatcacaggggccatttgatgctgttttgtattatcatatgatgaaggatgagcct
G-pig Win      2435 ctaagatcacaggggccatttgatgctgttttgtattatcatatgatgaaggatgagcct

I-2 seq        2881 gtagttttcagtaccagtgatggcaaagagtacacgtatccagactcccttgaagaggaa
G-pig Win      2495 gtagttttcagtaccagtgatggcaaagagtacacgtatccagactcccttgaagaggaa

I-2 seq        2941 tatccaccatggctcactgaaaaagaggctatgaatgaagagaatagatttgttacattg
G-pig Win      2555 tatccaccatggctcactgaaaaagaggctatgaatgaagagaatagatttgttacattg

I-2 seq        3001 gatggtcaacaattttattggccggtgatgaatcacaagaataaattcatggcaatcctg
G-pig Win      2615 gatggtcaacaattttattggccggtgatgaatcacaagaataaattcatggcaatcctg

I-2 seq        3061 caacatcatcagtgaatgagcatggaacaatgggatgattcaaccgacaaatagctaaca
G-pig Win      2675 caacatcatcagtgaatgagcatggaacaatgggatgattcaaccgacaaatagctaaca

I-2 seq        3121 ttaagtagtcaaggaacgaaaacaggaagaatttttgatgtctaaggtgtgaattattat
G-pig Win      2735 ttaagtagtcaaggaacgaaaacaggaagaatttttgatgtctaaggtgtgaattattat

I-2 seq        3181 cacaataaaagtgattcttatttttgaatttaaagctagcttattattactagccgtttt
G-pig Win      2795 cacaataaaagtgattcttatttttgaatttaaagctagcttattattactagccgtttt

I-2 seq        3241 tcaaagttcaatttgagtcttaatgcaaataggcgttaagccacagttatagccataatt
G-pig Win      2855 tcaaagttcaatttgagtcttaatgcaaataggcgttaagccacagttatagccataatt

I-2 seq        3301 gtaactcaatattctaactagcgatttatctaaattaaattacattatgcttttataact
G-pig Win      2915 gtaactcaatattctaactagcgatttatctaaattaaattacattatgcttttataact

I-2 seq        3361 tacctactagcctgcccaacatttacacgatcgttttataattaagaaaaaactaatgat
G-pig Win      2975 tacctactagcctgcccaacatttacacgatcgttttataattaagaaaaaactaatgat

I-2 seq        3421 gaagattaaaaccttcatcatccttacgtcaattgaattctctagcactcgaagcttatt
G-pig Win      3035 gaagattaaaaccttcatcatccttacgtcaattgaattctctagcactcgaagcttatt

I-2 seq        3481 gtcttcaatgtaaaagaaaagctggtctaacaagatgacaactagaacaaagggcagggg
G-pig Win      3095 gtcttcaatgtaaaagaaaagctggtctaacaagatgacaactagaacaaagggcagggg
```

```
I-2 seq        3541 ccatactgcggccacgactcaaaacgacagaatgccaggccctgagctttcgggctggat
G-pig Win      3155 ccatactgcggccacgactcaaaacgacagaatgccaggccctgagctttcgggctggat

I-2 seq        3601 ctctgagcagctaatgaccggaagaattcctgtaagcgacatcttctgtgatattgagaa
G-pig Win      3215 ctctgagcagctaatgaccggaagaattcctgtaagcgacatcttctgtgatattgagaa

I-2 seq        3661 caatccaggattatgctacgcatcccaaatgcaacaaacgaagccaaacccgaagacgcg
G-pig Win      3275 caatccaggattatgctacgcatcccaaatgcaacaaacgaagccaaacccgaagacgcg

I-2 seq        3721 caacagtcaaacccaaacggacccaatttgcaatcatagttttgaggaggtagtacaaac
G-pig Win      3335 caacagtcaaacccaaacggacccaatttgcaatcatagttttgaggaggtagtacaaac

I-2 seq        3781 attggcttcattggctactgttgtgcaacaacaaaccatcgcatcagaatcattagaaca
G-pig Win      3395 attggcttcattggctactgttgtgcaacaacaaaccatcgcatcagaatcattagaaca

I-2 seq        3841 acgcattacgagtcttgagaatggtctaaagccagtttatgatatggcaaaaacaatctc
G-pig Win      3455 acgcattacgagtcttgagaatggtctaaagccagtttatgatatggcaaaaacaatctc

I-2 seq        3901 ctcattgaacagggtttgtgctgagatggttgcaaaatatgatcttctggtgatgacaac
G-pig Win      3515 ctcattgaacagggtttgtgctgagatggttgcaaaatatgatcttctggtgatgacaac

I-2 seq        3961 cggtcgggcaacagcaaccgctgcggcaactgaggcttattgggccgaacatggtcaacc
G-pig Win      3575 cggtcgggcaacagcaaccgctgcggcaactgaggcttattgggccgaacatggtcaacc

I-2 seq        4021 accacctggaccatcactttatgaagaaagtgcgattcggggtaagattgaatctagaga
G-pig Win      3635 accacctggaccatcactttatgaagaaagtgcgattcggggtaagattgaatctagaga

I-2 seq        4081 tgagaccgtccctcaaagtgttagggaggcattcaacaatctaaacagtaccacttcact
G-pig Win      3695 tgagaccgtccctcaaagtgttagggaggcattcaacaatctaaacagtaccacttcact

I-2 seq        4141 aactgaggaaaattttgggaaacctgacatttcggcaaaggatttgagaaacattatgta
G-pig Win      3755 aactgaggaaaattttgggaaacctgacatttcggcaaaggatttgagaaacattatgta

I-2 seq        4201 tgatcacttgcctggttttggaactgctttccaccaattagtacaagtgatttgtaaatt
G-pig Win      3815 tgatcacttgcctggttttggaactgctttccaccaattagtacaagtgatttgtaaatt

I-2 seq        4261 gggaaaagatagcaactcattggacatcattcatgctgagttccaggccagcctggctga
G-pig Win      3875 gggaaaagatagcaactcattggacatcattcatgctgagttccaggccagcctggctga

I-2 seq        4321 aggagactctcctcaatgtgccctaattcaaattacaaaaagagttccaatcttccaaga
G-pig Win      3935 aggagactctcctcaatgtgccctaattcaaattacaaaaagagttccaatcttccaaga

I-2 seq        4381 tgctgctccacctgtcatccacatccgctctcgaggtgacattccccgagcttgccagaa
G-pig Win      3995 tgctgctccacctgtcatccacatccgctctcgaggtgacattccccgagcttgccagaa

I-2 seq        4441 aagcttgcgtccagtcccaccatcgcccaagattgatcgaggttgggtatgtgttttca
G-pig Win      4055 aagcttgcgtccagtcccaccatcgcccaagattgatcgaggttgggtatgtgttttca

I-2 seq        4501 gcttcaagatggtaaaacacttggactcaaaatttgagccaatctcccttccctccgaaa
G-pig Win      4115 gcttcaagatggtaaaacacttggactcaaaatttgagccaatctcccttccctccgaaa

I-2 seq        4561 gaggcgaataatagcagaggcttcaactgctgaactatagggtacgttacattaatgata
G-pig Win      4175 gaggcgaataatagcagaggcttcaactgctgaactatagggtacgttacattaatgata

I-2 seq        4621 cacttgtgagtatcagccctggataatataagtcaattaaacgaccaagataaaattgtt
G-pig Win      4235 cacttgtgagtatcagccctggataatataagtcaattaaacgaccaagataaaattgtt

I-2 seq        4681 catatctcgctagcagcttaaaatataaatgtaataggagctatatctctgacagtatta
G-pig Win      4295 catatctcgctagcagcttaaaatataaatgtaataggagctatatctctgacagtatta

I-2 seq        4741 taatcaattgttattaagtaacccaaaccaaaagtgatgaagattaagaaaaacctacct
G-pig Win      4355 taatcaattgttattaagtaacccaaaccaaaagtgatgaagattaagaaaaacctacct
```

```
I-2 seq        4801 cggctgagagagtgtttttttcattaaccttcatcttgtaaacgttgagcaaaattgttaa
G-pig Win      4415 cggctgagagagtgtttttttcactaaccttcatcttgtaaacgttgagcaaaattgttaa

I-2 seq        4861 aaatatgaggcgggttatattgcctactgctcctcctgaatatatggaggccatataccc
G-pig Win      4475 aaatatgaggcgggttatattgcctactgctcctcctgaatatatggaggccatataccc

I-2 seq        4921 tgtcaggtcaaattcaacaattgctagaggtggcaacagcaatacaggcttcctgacacc
G-pig Win      4535 tgtcaggtcaaattcaacaattgctagaggtggcaacagcaatacaggcttcctgacacc

I-2 seq        4981 ggagtcagtcaatggggacactccatcgaatccactcaggccaattgccgatgacaccat
G-pig Win      4595 ggagtcagtcaatggggacactccatcgaatccactcaggccaattgccgatgacaccat

I-2 seq        5041 cgaccatgccagccacacaccaggcagtgtgtcatcagcattcatccttgaagctatggt
G-pig Win      4655 cgaccatgccagccacacaccaggcagtgtgtcatcagcattcatccttgaagctatggt

I-2 seq        5101 gaatgtcatatcgggccccaaagtgctaatgaagcaaattccaatttggcttcctctagg
G-pig Win      4715 gaatgtcatatcgggccccaaagtgctaatgaagcaaattccaatttggcttcctctagg

I-2 seq        5161 tgtcgctgatcaaaagacctacagctttgactcaactacggccgccatcatgcttgcttc
G-pig Win      4775 tgtcgctgatcaaaagacctacagctttgactcaactacggccgccatcatgcttgcttc

I-2 seq        5221 atacactatcacccatttcggcaaggcaaccaatccacttgtcagagtcaatcggctggg
G-pig Win      4835 atacactatcacccatttcggcaaggcaaccaatccacttgtcagagtcaatcggctggg

I-2 seq        5281 tcctggaatcccggatcatcccctcaggctcctgcgaattggaaaccaggctttcctcca
G-pig Win      4895 tcctggaatcccggatcatcccctcaggctcctgcgaattggaaaccaggctttcctcca

I-2 seq        5341 ggagttcgttcttccgccagtccaactaccccagtatttcacctttgatttgacagcact
G-pig Win      4955 ggagttcgttcttccgccagtccaactaccccagtatttcacctttgatttgacagcact

I-2 seq        5401 caaactgatcacccaaccactgcctgctgcaacatggaccgatgacactccaacaggatc
G-pig Win      5015 caaactgatcacccaaccactgcctgctgcaacatggatcgatgacactccaacaggatc

I-2 seq        5461 aaatggagcgttgcgtccaggaatttcatttcatccaaaacttcgccccattcttttacc
G-pig Win      5075 aaatggagcgttgcgtccaggaatttcatttcatccaaaacttcgccccattcttttacc

I-2 seq        5521 caacaaaagtgggaagaaggggaacagtgccgatctaacatctccggagaaaatccaagc
G-pig Win      5135 caacaaaagtgggaagaaggggaacagtgccgatctaacatctccggagaaaatccaagc

I-2 seq        5581 aataatgacttcactccaggactttaagatcgttccaattgatccaaccaaaaatatcat
G-pig Win      5195 aataatgacttcactccaggactttaagatcgttccaattgatccaaccaaaaatatcat

I-2 seq        5641 gggaatcgaagtgccagaaactctggtccacaagctgaccggtaagaaggtgacttctaa
G-pig Win      5255 gggaatcgaagtgccagaaactctggtccacaagctgaccggtaagaaggtgacttctaa

I-2 seq        5701 aaatggacaaccaatcatccctgttcttttgccaaagtacattgggttggacccggtggc
G-pig Win      5315 aaatggacaaccaatcatccctgttcttttgccaaagtacattgggttggacccggtggc

I-2 seq        5761 tccaggagacctcaccatggtaatcacacaggattgtgacacgtgtcattctcctgcaag
G-pig Win      5375 tccaggagacctcaccatggtaatcacacaggattgtgacacgtgtcattctcctgcaag

I-2 seq        5821 tcttccagctgtgattgagaagtaattgcaataattgactcagatccagttttatagaat
G-pig Win      5435 tcttccagctgtgattgagaagtaattgcaataattgactcagatccagttttatagaat

I-2 seq        5881 cttctcagggatagtgataacatctatttagtaatccgtccattagaggagacactttta
G-pig Win      5495 cttctcagggatagtgataacatctatttagtaatccgtccattagaggagacactttta

I-2 seq        5941 attgatcaatatactaaaggtgctttacaccattgtcttttttctctcctaaatgtagaa
G-pig Win      5555 attgatcaatatactaaaggtgctttacaccattgtcttttttctctcctaaatgtagaa

I-2 seq        6001 cttaacaaaagactcataatatacttgtttttaaaggattgattgatgaaagatcataac
G-pig Win      5615 cttaacaaaagactcataatatacttgtttttaaaggattgattgatgaaagatcataac
```

```
I-2 seq        6061 taataacattacaaataatcctactataatcaatacggtgattcaaatgttaatctttct
G-pig Win      5675 taataacattacaaataatcctactataatcaatacggtgattcaaatgttaatctttct

I-2 seq        6121 cattgcacatacttttgcccttatcctcaaattgcctgcatgcttacatctgaggatag
G-pig Win      5735 cattgcacatacttttgcccttatcctcaaattgcctgcatgcttacatctgaggatag

I-2 seq        6181 ccagtgtgacttggattggaaatgtggagaaaaaatcgggacccatttctaggttgttca
G-pig Win      5795 ccagtgtgacttggattggaaatgtggagaaaaaatcgggacccatttctaggttgttca

I-2 seq        6241 caatccaagtacagacattgcccttctaattaagaaaaaatcggcgatgaagattaagcc
G-pig Win      5855 caatccaagtacagacattgcccttctaattaagaaaaaatcggcgatgaagattaagcc

I-2 seq        6301 gacagtgagcgtaatcttcatctctcttagattatttgttttccagagtaggggtcgtca
G-pig Win      5915 gacagtgagcgtaatcttcatctctcttagattatttgttttccagagtaggggtcgtca

I-2 seq        6361 ggtccttttcaatcgtgtaaccaaaataaactccactagaaggatattgtggggcaacaa
G-pig Win      5975 ggtccttttcaatcgtgtaaccaaaataaactccactagaaggatattgtggggcaacaa

I-2 seq        6421 cacaatgggcgttacaggaatattgcagttacctcgtgatcgattcaagaggacatcatt
G-pig Win      6035 cacaatgggcgttacaggaatattgcagttacctcgtgatcgattcaagaggacatcatt

I-2 seq        6481 ctttctttgggtaattatccttttccaaagaacattttccatcccacttggagtcatcca
G-pig Win      6095 ctttctttgggtaattatccttttccaaagaacattttccatcccacttggagtcatcca

I-2 seq        6541 caatagcacattacaggttagtgatgtcgacaaactagtttgtcgtgacaaactgtcatc
G-pig Win      6155 caatagcacattacaggttagtgatgtcgacaaactagtttgtcgtgacaaactgtcatc

I-2 seq        6601 cacaaatcaattgagatcagttggactgaatctcgaagggaatggagtggcaactgacgt
G-pig Win      6215 cacaaatcaattgagatcagttggactgaatctcgaagggaatggagtggcaactgacgt

I-2 seq        6661 gccatctgcaactaaaagatggggcttcaggtccggtgtcccaccaaaggtggtcaatta
G-pig Win      6275 gccatctgcaactaaaagatggggcttcaggtccggtgtcccaccaaaggtggtcaatta

I-2 seq        6721 tgaagctggtgaatgggctgaaaactgctacaatcttgaaatcaaaaaacctgacgggag
G-pig Win      6335 tgaagctggtgaatgggctgaaaactgctacaatcttgaaatcaaaaaacctgacgggag

I-2 seq        6781 tgagtgtctaccagcagcgccagacgggattcggggcttcccccggtgccggtatgtgca
G-pig Win      6395 tgagtgtctaccagcagcgccagacgggactcggggcttcccccggtgccggtatgtgca

I-2 seq        6841 caaagtatcaggaacgggaccgtgtgccggagactttgccttccataaagagggtgcttt
G-pig Win      6455 caaagtatcaggaacgggaccgtgtgccggagactttgccttccataaagagggtgcttt

I-2 seq        6901 cttcctgtatgatcgacttgcttccacagttatctaccgaggaacgactttcgctgaagg
G-pig Win      6515 cttcctgtatgatcgacttgcttccacagttatctaccgaggaacgactttcgctgaagg

I-2 seq        6961 tgtcgttgcatttctgatactgccccaagctaagaaggacttcttcagctcacacccctt
G-pig Win      6575 tgtcgttgcatttctgatactgccccaagctaagaaggacttcttcagctcacacccctt

I-2 seq        7021 gagagagccggtcaatgcaacggaggacccgtctagtggctactattctaccacaattag
G-pig Win      6635 gagagagccggtcaatgcaacggaggacccgtctagtggctactattctaccacaattag

I-2 seq        7081 atatcaggctaccggttttggaaccaatgagacagagtacttgttcgaggttgacaattt
G-pig Win      6695 atatcaggctaccggttttggaaccaatgagacagagtacttgttcgaggttgacaattt

I-2 seq        7141 gacctacgtccaacttgaatcaagattcacaccacagtttctgctccagctgaatgagac
G-pig Win      6755 gacctacgtccaacttgaatcaagattcacaccacagtttctgctccagctgaatgagac

I-2 seq        7201 aatatatacaagtgggaaaaggagcaataccacgggaaaactaatttggaaggtcaaccc
G-pig Win      6815 aatatatacaagtgggaaaaggagcaataccacgggaaaactaatttggaaggtcaaccc

I-2 seq        7261 cgaaattgatacaacaatcggggagtgggccttctgggaaactaaaaaaacctcactaga
G-pig Win      6875 cgaaattgatacaacaatcggggagtgggccttctgggaaactaaaaaaacctcactaga
```

```
I-2 seq        7321 aaaattcgcagtgaagagttgtctttcacagttgtatcaaacggagccaaaaacatcagt
G-pig Win      6935 aaaattcgcagtgaagagttgtctttcacagttgtatcaaacggagccaaaaacatcagt

I-2 seq        7381 ggtcagagtccggcgcgaacttcttccgacccagggaccaacacaacaactgaagaccac
G-pig Win      6995 ggtcagagtccggcgcgaacttcttccgacccagggaccaacacaacaactgaagaccac

I-2 seq        7441 aaaatcatggcttcagaaaattcctctgcaatggttcaagtgcacagtcaaggaagggaa
G-pig Win      7055 aaaatcatggcttcagaaaattcctctgcaatggttcaagtgcacagtcaaggaagggaa

I-2 seq        7501 gctgcagtgtcgcatctaacaacccttgccacaatctccacgagtccccaatccctcaca
G-pig Win      7115 gctgcagtgtcgcatctaacaacccttgccacaatctccacgagtccccaatccctcaca

I-2 seq        7561 accaaaccaggtccggacaacagcacccataatacacccgtgtataaacttgacatctct
G-pig Win      7175 accaaaccaggtccggacaacagcacccataatacacccgtgtataaacttgacatctct

I-2 seq        7621 gaggcaactcaagttgaacaacatcaccgcagaacagacaacgacagcacagcctccgac
G-pig Win      7235 gaggcaactcaagttgaacaacatcaccgcagaacagacaacgacagcacagcctccgac

I-2 seq        7681 actccctctgccacgaccgcagccggacccccaaaagcagagaacaccaacacgagcaag
G-pig Win      7295 actccctctgccacgaccgcagccggacccccaaaagcagagaacaccaacacgagcaag

I-2 seq        7741 agcactgacttcctggaccccgccaccacaacaagtccccaaaaccacagcgagaccgct
G-pig Win      7355 agcactgacttcctggaccccgccaccacaacaagtccccaaaaccacagcgagaccgct

I-2 seq        7801 ggcaacaacaacactcatcaccaagataccggagaagagagtgccagcagcgggaagcta
G-pig Win      7415 ggcaacaacaacactcatcaccaagataccggagaagagagtgccagcagcgggaagcta

I-2 seq        7861 ggcttaattaccaatactattgctggagtcgcaggactgatcacaggcgggagaagaact
G-pig Win      7475 ggcttaattaccaatactattgctggagtcgcaggactgatcacaggcgggagaagaact

I-2 seq        7921 cgaagagaagcaattgtcaatgctcaacccaaatgcaaccctaatttacattactggact
G-pig Win      7535 cgaagagaagcaattgtcaatgctcaacccaaatgcaaccctaatttacattactggact

I-2 seq        7981 actcaggatgaaggtgctgcaatcggactggcctggataccatatttcgggccagcagcc
G-pig Win      7595 actcaggatgaaggtgctgcaatcggactggcctggataccatatttcgggccagcagcc

I-2 seq        8041 gagggaatttacatagaggggctaatgcacaatcaagatggtttaatctgtgggttgaga
G-pig Win      7655 gagggaatttacacagaggggctaatgcacaatcaagatggtttaatctgtgggttgaga

I-2 seq        8101 cagctggccaacgagacgactcaagctcttcaactgttcctgagagccacaactgagcta
G-pig Win      7715 cagctggccaacgagacgactcaagctcttcaactgttcctgagagccacaactgagcta

I-2 seq        8161 cgcacctttttcaatcctcaaccgtaaggcaattgatttcttgctgcagcgatggggcggc
G-pig Win      7775 cgcacctttttcaatcctcaaccgtaaggcaattgatttcttgctgcagcgatggggcggc

I-2 seq        8221 acatgccacattctgggaccggactgctgtatcgaaccacatgattggaccaagaacata
G-pig Win      7835 acatgccacattctgggaccggactgctgtatcgaaccacatgattggaccaagaacata

I-2 seq        8281 acagacaaaattgatcagattattcatgattttgttgataaaacccttccggaccagggg
G-pig Win      7895 acagacaaaattgatcagattattcatgattttgttgataaaacccttccggaccagggg

I-2 seq        8341 gacaatgacaattggtggacaggatggagacaatggataccggcaggtattggagttaca
G-pig Win      7955 gacaatgacaattggtggacaggatggagacaatggataccggcaggtattggagttaca

I-2 seq        8401 ggcgttataattgcagttatcgctttattctgtatatgcaaatttgtcttttagtttttc
G-pig Win      8015 ggcgttataattgcagttatcgctttattctgtatatgcaaatttgtcttttagtttttc

I-2 seq        8461 ttcagattgcttcatggaaaagctcagcctcaaatcaatgaaaccaggatttaattatat
G-pig Win      8075 ttcagattgcttcatggaaaagctcagcctcaaatcaatgaaaccaggatttaattatat

I-2 seq        8521 ggattacttgaatctaagattacttgacaaatgataatataatacactggagctttaaac
G-pig Win      8135 ggattacttgaatctaagattacttgacaaatgataatataatacactggagctttaaac
```

```
I-2 seq        8581 atagccaatgtgattctaactcctttaaactcacagttaatcataaacaaggtttgacat
G-pig Win      8195 atagccaatgtgattctaactcctttaaactcacagttaatcataaacaaggtttgacat

I-2 seq        8641 caatctagttatctctttgagaatgataaacttgatgaagattaagaaaaaggtaatctt
G-pig Win      8255 caatctagttatctctttgagaatgataaacttgatgaagattaagaaaaaggtaatctt

I-2 seq        8701 tcgattatctttaatcttcatccttgattctacaatcatgacagttgtctttagtgacaa
G-pig Win      8315 tcgattatctttaatcttcatccttgattctacaatcatgacagttgtctttagtgacaa

I-2 seq        8761 gggaaagaagccttttttattaagttgtaataatcagatctgcgaaccggtagagtttagt
G-pig Win      8375 gggaaagaagccttttttattaagttgtaataatcagatctgcgaaccggtagagtttagt

I-2 seq        8821 tgcaacctaacacacataaagcattggtcaaaaagtcaatagaaatttaaacagtgagtg
G-pig Win      8435 tgcaacctaacacacataaagcattggtcaaaaagtcaatagaaatttaaacagtgagtg

I-2 seq        8881 gagacaacttttaaatggaagcttcatatgagagaggacgcccacgagctgccagacagc
G-pig Win      8495 gagacaacttttaaatggaagcttcatatgagagaggacgcccacgagctgccagacagc

I-2 seq        8941 attcaagggatggacacgaccaccatgttcgagcacgatcatcatccagagagaattatc
G-pig Win      8555 attcaagggatggacacgaccaccatgttcgagcacgatcatcatccagagagaattatc

I-2 seq        9001 gaggtgagtaccgtcaatcaaggagcgcctcacaagtgcgcgttcctactgtatttcata
G-pig Win      8615 gaggtgagtaccgtcaatcaaggagcgcctcacaagtgcgcgttcctactgtatttcata

I-2 seq        9061 agaagagagttgaaccattaacagttcctccagcacctaaagacatatgtccgaccttga
G-pig Win      8675 agaagagagttgaaccattaacagttcctccagcacctaaagacatatgtccgaccttga

I-2 seq        9121 aaaaaggatttttgtgtgacagtagtttttgcaaaaaagatcaccagttggagagtttaa
G-pig Win      8735 aaaaaggatttttgtgtgacagtagtttttgcaaaaaagatcaccagttggagagtttaa

I-2 seq        9181 ctgatagggaattactcctactaatcgcccgtaagacttgtggatcagtagaacaacaat
G-pig Win      8795 ctgatagggaattactcctactaatcgcccgtaagacttgtggatcagtagaacaacaat

I-2 seq        9241 taaatataactgcacccaaggactcgcgcttagcaaatccaacggctgatgatttccagc
G-pig Win      8855 taaatataactgcacccaaggactcgcgcttagcaaatccaacggctgatgatttccagc

I-2 seq        9301 aagaggaaggtccaaaaattaccttgttgacactgatcaagacggcagaacactgggcga
G-pig Win      8915 aagaggaaggtccaaaaattaccttgttgacactgatcaagacggcagaacactgggcga

I-2 seq        9361 gacaagacatcagaaccatagaggattcaaaattaagagcattgttgactctatgtgctg
G-pig Win      8975 gacaagacatcagaaccatagaggattcaaaattaagagcattgttgactctatgtgctg

I-2 seq        9421 tgatgacgaggaaattctcaaaatcccagctgagtcttttatgtgagacacacctaaggc
G-pig Win      9035 tgatgacgaggaaattctcaaaatcccagctgagtcttttatgtgagacacacctaaggc

I-2 seq        9481 gcgaggggcttgggcaagatcaggcagaacccgttctcgaagtatatcaacgattacaca
G-pig Win      9095 gcgaggggcttgggcaagatcaggcagaacccgttctcgaagtatatcaacgattacaca

I-2 seq        9541 gtgataaaggaggcagttttgaagctgcactatggcaacaatgggaccgacaatccctaa
G-pig Win      9155 gtgataaaggaggcagttttgaagctgcactatggcaacaatgggaccgacaatccctaa

I-2 seq        9601 ttatgtttatcactgcattcttgaatattgctctccagttaccgtgtgaaagttctgctg
G-pig Win      9215 ttatgtttatcactgcattcttgaatattgctctccagttaccgtgtgaaagttctgctg

I-2 seq        9661 tcgttgtttcagggttaagaacattggttcctcaatcagataatgaggaagcttcaacca
G-pig Win      9275 tcgttgtttcagggttaagaacattggttcctcaatcagataatgaggaagcttcaacca

I-2 seq        9721 acccggggacatgctcatggtctgatgagggtacccccttaataaggctgactaaaacact
G-pig Win      9335 acccggggacatgctcatggtctgatgagggtacccccttaataaggctgactaaaacact

I-2 seq        9781 atataaccttctacttgatcacaatactccgtatacctatcatcatatatttaatcaaga
G-pig Win      9395 atataaccttctacttgatcacaatactccgtatacctatcatcatatatttaatcaaga
```

```
I-2 seq        9841 cgatatcctttaaaacttattcagtactataatcactctcgtttcaaattaataagatgt
G-pig Win      9455 cgatatcctttaaaacttattcagtactataatcactctcgtttcaaattaataagatgt

I-2 seq        9901 gcatgattgccctaatatatgaagaggtatgatacaaccctaacagtgatcaaagaaaat
G-pig Win      9515 gcatgattgccctaatatatgaagaggtatgatacaaccctaacagtgatcaaagaaaat

I-2 seq        9961 cataatctcgtatcgctcgtaatataacctgccaagcatacctcttgcacaaagtgattc
G-pig Win      9575 cataatctcgtatcgctcgtaatataacctgccaagcatacctcttgcacaaagtgattc

I-2 seq       10021 ttgtacacaaataatgttttactctacaggaggtagcaacgatccatcccatcaaaaaat
G-pig Win      9635 ttgtacacaaataatgttttactctacaggaggtagcaacgatccatcccatcaaaaaat

I-2 seq       10081 aagtatttcatgacttactaatgatctcttaaaatattaagaaaaactgcggaacataaa
G-pig Win      9695 aagtatttcatgacttactaatgatctcttaaaatattaagaaaaactgcggaacataaa

I-2 seq       10141 ttctttatgcttcaagctgtggaggaggtgtttggtattggctattgttatattacaatc
G-pig Win      9755 ttctttatgcttcaagctgtggaggaggtgtttggtattggctattgttatattacaatc

I-2 seq       10201 aataacaagcttgtaaaaatattgttcttgtttcaagaggtagattgtgaccggaaatgc
G-pig Win      9815 aataacaagcttgtaaaaatattgttcttgtttcaagaggtagattgtgaccggaaatgc

I-2 seq       10261 taaactaatgatgaagattaatgcggaggtctgataagaataaaccttattattcagatt
G-pig Win      9875 taaactaatgatgaagattaatgcggaggtctgataagaataaaccttattattcagatt

I-2 seq       10321 aggccccaagaggcattcttcatctccttttagcaaagtactatttcagggtagtccaat
G-pig Win      9935 aggccccaagaggcattcttcatctccttttagcaaagtactatttcagggtagtccaat

I-2 seq       10381 tagtggcacgtcttttagctgtatatcagtcgcccctgagatacgccacaaaagtgtctc
G-pig Win      9995 tagtggcacgtcttttagctgtatatcagtcgcccctgagatacgccacaaaagtgtctc

I-2 seq       10441 taagctaaattggtctgtacacatcccatacattgtattaggggcaataatatctaattg
G-pig Win     10055 taagctaaattggtctgtacacatcccatacattgtattaggggcaataatatctaattg

I-2 seq       10501 aacttagccgtttaaaatttagtgcataaatctgggctaacaccaccaggtcaactccat
G-pig Win     10115 aacttagccgtttaaaatttagtgcataaatctgggctaacaccaccaggtcaactccat

I-2 seq       10561 tggctgaaaagaagcttacctacaacgaacatcactttgagcgccctcacaattaaaaaa
G-pig Win     10175 tggctgaaaagaagcttacctacaacgaacatcactttgagcgccctcacaattaaaaaa

I-2 seq       10621 taggaacgtcgttccaacaatcgagcgcaaggtttcaaggttgaactgagagtgtctaga
G-pig Win     10235 taggaacgtcgttccaacaatcgagcgcaaggtttcaaggttgaactgagagtgtctaga

I-2 seq       10681 caacaaaatattgatactccagacaccaagcaagacctgagaaaaaaccatggctaaagc
G-pig Win     10295 caacaaaatattgatactccagacaccaagcaagacctgagaaaaaaccatggctaaagc

I-2 seq       10741 tacgggacgatacaatctaatatcgcccaaaaaggacctggagaaaggggttgtcttaag
G-pig Win     10355 tacgggacgatacaatctaatatcgcccaaaaaggacctggagaaaggggttgtcttaag

I-2 seq       10801 cgacctctgtaacttcttagttagccaaactattcaggggtggaaggtttattgggctgg
G-pig Win     10415 cgacttctgtaacttcttagttagccaaactattcaggggtggaaggtttattgggctgg

I-2 seq       10861 tattgagtttgatgtgactcacaaaggaatggccctattgcatagactgaaaactaatga
G-pig Win     10475 tattgagtttgatgtgactcacaaaggaatggccctattgcatagactgaaaactaatga

I-2 seq       10921 ctttgcccctgcatggtcaatgacaaggaatctctttcctcatttatttcaaaatccgaa
G-pig Win     10535 ctttgcccctgcatggtcaatgacaaggaatctctttcctcatttatttcaaaatccgaa

I-2 seq       10981 ttccacaattgaatcaccgctgtgggcattgagagtcatccttgcagcagggatacagga
G-pig Win     10595 ttccacaattgaatcaccgctgtgggcattgagagtcatccttgcagcagggatacagga

I-2 seq       11041 ccagctgattgaccagtctttgattgaacccttagcaggagcccttggtctgatctctga
G-pig Win     10655 ccagctgattgaccagtctttgattgaacccttagcaggagcccttggtctgatctctga
```

```
I-2 seq       11101 ttggctgctaacaaccaacactaaccatttcaacatgcgaacacaacgtgtcaaggaaca
G-pig Win     10715 ttggctgctaacaaccaacactaaccatttcaacatgcgaacacaacgtgtcaaggaaca

I-2 seq       11161 attgagcctaaaaatgctgtcgttgattcgatccaatattctcaagtttattaacaaatt
G-pig Win     10775 attgagcctaaaaatgctgtcgttgattcgatccaatattctcaagtttattaacaaatt

I-2 seq       11221 ggatgctctacatgtcgtgaactacaacggattgttgagcagtattgaaattggaactca
G-pig Win     10835 ggatgctctacatgtcgtgaactacaacggattgttgagcagtattgaaattggaactca

I-2 seq       11281 aaatcatacaatcatcataactcgaactaacatgggttttctggtggagctccaagaacc
G-pig Win     10895 aaatcatacaatcatcataactcgaactaacatgggttttctggtggagctccaagaacc

I-2 seq       11341 cgacaaatcggcaatgaaccgcatgaagcctgggccggcgaaattttccctccttcatga
G-pig Win     10955 cgacaaatcggcaatgaaccgcatgaagcctgggccggcgaaattttccctccttcatga

I-2 seq       11401 gtccacactgaaagcatttacacaaggatcctcgacacgaatgcaaagtttgattcttga
G-pig Win     11015 gtccacactgaaagcatttacacaaggatcctcgacacgaatgcaaagtttgattcttga

I-2 seq       11461 atttaatagctctcttgctatctaactaaggtagaatacttcatattgagctaactcata
G-pig Win     11075 atttaatagctctcttgctatctaactaaggtagaatacttcatattgagctaactcata

I-2 seq       11521 tatgctgactcaatagttatcttgacatctctgctttcataatcagatatataagcataa
G-pig Win     11135 tatgctgactcaatagttatcttgacatctctgctttcataatcagatatataagcataa

I-2 seq       11581 taaataaatactcatatttcttgataatttgtttaaccacagataaatcctcactgtaag
G-pig Win     11195 taaataaatactcatatttcttgataatttgtttaaccacagataaatcctcactgtaag

I-2 seq       11641 ccagcttccaagttgacacccttacaaaaaccaggactcagaatccctcaaacaagagat
G-pig Win     11255 ccagcttccaagttgacacccttacaaaaaccaggactcagaatccctcaaacaagagat

I-2 seq       11701 tccaagacaacatcatagaattgctttattatatgaataagcattttatcaccagaaatc
G-pig Win     11315 tccaagacaacatcatagaattgctttattatatgaataagcattttatcaccagaaatc

I-2 seq       11761 ctatatactaaatggttaattgtaactgaacccgcaggtcacatgtgttaggtttcacag
G-pig Win     11375 ctatatactaaatggttaattgtaactgaacccgcaggtcacatgtgttaggtttcacag

I-2 seq       11821 attctatatattactaactctatactcgtaattaacattagataagtagattaagaaaaa
G-pig Win     11435 attctatatattactaactctatactcgtaattaacattagataagtagattaagaaaaa

I-2 seq       11881 agcctgaggaagattaagaaaaactgcttattgggtctttccgtgttttagatgaagcag
G-pig Win     11495 agcctgaggaagattaagaaaaactgcttattgggtctttccgtgttttagatgaagcag

I-2 seq       11941 ttgaaattcttcctcttgatattaaatggctacacaacatacccaatacccagacgctag
G-pig Win     11555 ttgaaattcttcctcttgatattaaatggctacacaacatacccaatacccagacgctag

I-2 seq       12001 gttatcatcaccaattgtattggaccaatgtgacctagtcactagagcttgcgggttata
G-pig Win     11615 gttatcatcaccaattgtattggaccaatgtgacctagtcactagagcttgcgggttata

I-2 seq       12061 ttcatcatactcccttaatccgcaactacgcaactgtaaactcccgaaacatatctaccg
G-pig Win     11675 ttcatcatactcccttaatccgcaactacgcaactgtaaactcccgaaacatatctaccg

I-2 seq       12121 tttgaaatacgatgtaactgttaccaagttcttgagtgatgtaccagtggcgacattgcc
G-pig Win     11735 tttgaaatacgatgtaactgttaccaagttcttgagtgatgtaccagtggcgacattgcc

I-2 seq       12181 catagatttcatagtcccagttcttctcaaggcactgtcaggcaatggattctgtcctgt
G-pig Win     11795 catagatttcatagtcccagttcttctcaaggcactgtcaggcaatggattctgtcctgt

I-2 seq       12241 tgagccgcggtgccaacagttcttagatgaaatcattaagtacacaatgcaagatgctct
G-pig Win     11855 tgagccgcggtgccaacagttcttagatgaaatcattaagtacacaatgcaagatgctct

I-2 seq       12301 cttcttgaaatattatctcaaaaatgtgggtgctcaagaagactgtgttgatgaacactt
G-pig Win     11915 cttcttgaaatattatctcaaaaatgtgggtgctcaagaagactgtgttgatgaacactt
```

```
I-2 seq       12361 tcaagagaaaatcttatcttcaattcagggcaatgaatttttacatcaaatgtttttctg
G-pig Win     11975 tcaagagaaaatcttatcttcaattcagggcaatgaatttttacatcaaatgtttttctg

I-2 seq       12421 gtatgatctggctattttaactcgaaggggtagattaaatcgaggaaactctagatcaac
G-pig Win     12035 gtatgatctggctattttaactcgaaggggtagattaaatcgaggaaactctagatcaac

I-2 seq       12481 atggtttgttcatgatgatttaatagacatcttaggctatggggactatgtttttggaa
G-pig Win     12095 atggtttgttcatgatgatttaatagacatcttaggctatggggactatgtttttggaa

I-2 seq       12541 gatcccaatttcaatgttaccactgaacacacaaggaatccccccatgctgctatggactg
G-pig Win     12155 gatcccaatttcaatgttaccactgaacacacaaggaatccccccatgctgctatggactg

I-2 seq       12601 gtatcaggcatcagtattcaaagaagcggttcaagggcatacacacattgtttctgtttc
G-pig Win     12215 gtatcaggcatcagtattcaaagaagcggttcaagggcatacacacattgtttctgtttc

I-2 seq       12661 tactgccgacgtcttgataatgtgcaaagatttaattacatgtcgattcaacacaactct
G-pig Win     12275 tactgccgacgtcttgataatgtgcaaagatttaattacatgtcgattcaacacaactct

I-2 seq       12721 aatctcaaaaatagcagagattgaggatccagtttgttctgattatcccaattttaagat
G-pig Win     12335 aatctcaaaaatagcagagattgaggatccagtttgttctgattatcccaattttaagat

I-2 seq       12781 tgtgtctatgctttaccagagcggagattacttactctccatattagggtctgatgggta
G-pig Win     12395 tgtgtctatgctttaccagagcggagattacttactctccatattagggtctgatgggta

I-2 seq       12841 taaaattattaagttcctcgaaccattgtgcttggccaaaattcaattatgctcaaagta
G-pig Win     12455 taaaattattaagttcctcgaaccattgtgcttggccaaaattcaattatgctcaaagta

I-2 seq       12901 cactgagaggaagggccgattcttaacacaaatgcatttagctgtaaatcacaccctaga
G-pig Win     12515 cactgagaggaagggccgattcttaacacaaatgcatttagctgtaaatcacaccctaga

I-2 seq       12961 agaaattacagaaatgcgtgcactaaagccttcacaggctcaaaagatccgtgaattcca
G-pig Win     12575 agaaattacagaaatgcgtgcactaaagccttcacaggctcaaaagatccgtgaattcca

I-2 seq       13021 tagaacattgataaggctggagatgacgccacaacaactttgtgagctattttccattca
G-pig Win     12635 tagaacattgataaggctggagatgacgccacaacaactttgtgagctattttccattca

I-2 seq       13081 aaaacactgggggcatcctgtgctacatagtgaaacagcaatccaaaaagttaaaaaaca
G-pig Win     12695 aaaacactgggggcatcctgtgctacatagtgaaacagcaatccaaaaagttaaaaaaca

I-2 seq       13141 tgctacggtgctaaaagcattacgccctatagtgattttcgagacatactgtgtttttaa
G-pig Win     12755 tgctacggtgctaaaagcattacgccctatagtgattttcgagacatactgtgtttttaa

I-2 seq       13201 atatagtattgccaaacattattttgatagtcaaggatcttggtacagtgttacttcaga
G-pig Win     12815 atatagtattgccaaacattattttgatagtcaaggatcttggtacagtgttacttcaga

I-2 seq       13261 taggaatctaacaccgggtcttaattcttatatcaaaagaaatcaattccctccgttgcc
G-pig Win     12875 taggaatctaacaccgggtcttaattcttatatcaaaagaaatcaattccctccgttgcc

I-2 seq       13321 aatgattaaagaactactatgggaattttaccaccttgaccaccctccacttttctcaac
G-pig Win     12935 aatgattaaagaactactatgggaattttaccaccttgaccaccctccacttttctcaac

I-2 seq       13381 caaaattattagtgacttaagtatttttataaaagacagagctaccgcagtagaaaggac
G-pig Win     12995 caaaattattagtgacttaagtatttttataaaagacagagctaccgcagtagaaagaac

I-2 seq       13441 atgctgggatgcagtattcgagcctaatgttctaggatataatccacctcacaaatttag
G-pig Win     13055 atgctgggatgcagtattcgagcctaatattctaggatataatccacctcacaaatttag

I-2 seq       13501 tactaaacgtgtaccggaacaatttttagagcaagaaaacttttctattgagaatgttct
G-pig Win     13115 tactaaacgtgtaccggaacaatttttagagcaagaaaacttttctattgagaatgttct

I-2 seq       13561 ttcctacgcacaaaaactcgagtatctactaccacaatatcggaacttttctttctcatt
G-pig Win     13175 ttcctacgcacaaaaactcgagtatctactaccacaatatcggaacttttctttctcatt
```

```
I-2 seq       13621 gaaagagaaagagttgaatgtaggtagaaccttcggaaaattgccttatccgactcgcaa
G-pig Win     13235 gaaagagaaagagttgaatgtaggtagaaccttcggaaaattgccttatccgactcgcaa

I-2 seq       13681 tgttcaaacactttgtgaagctctgttagctgatggtcttgctaaagcatttcctagcaa
G-pig Win     13295 tgttcaaacactttgtgaagctctgttagctgatggtcttgctaaagcatttcctagcaa

I-2 seq       13741 tatgatggtagttacggaacgtgagcaaaaagaaagcttattgcatcaagcatcatggca
G-pig Win     13355 tatgatggtagttacggaacgtgagcaaaaagaaagcttattgcatcaagcatcatggca

I-2 seq       13801 ccacacaagtgatgattttggtgaacatgccacagttagagggagtagctttgtaactga
G-pig Win     13415 ccacacaagtgatgattttggtgaacatgccacagttagagggagtagctttgtaactga

I-2 seq       13861 tttagagaaatacaatcttgcatttagatatgagtttacagcacctttttatagaatattg
G-pig Win     13475 tttagagaaatacaatcttgcatttagatatgagtttacagcacctttttatagaatattg

I-2 seq       13921 caaccgttgctatggtgttaagaatgtttttaattggatgcattatacaatcccacagtg
G-pig Win     13535 caaccgttgctatggtgttaagaatgtttttaattggatgcattatacaatcccacagtg

I-2 seq       13981 ttatatgcatgtcagtgattattataatccaccacataacctcacactggagaatcgaga
G-pig Win     13595 ttatatgcatgtcagtgattattataatccaccacataacctcacactggagaatcgaga

I-2 seq       14041 caacccccccgaagggcctagttcatacaggggtcatatgggagggattgaaggactgca
G-pig Win     13655 caacccccccgaagggcctagttcatacaggggtcatatgggagggattgaaggactgca

I-2 seq       14101 acaaaaactctggacaagtatttcatgtgctcaaatttctttagttgaaattaagactgg
G-pig Win     13715 acaaaaactctggacaagtatttcatgtgctcaaatttctttagttgaaattaagactgg

I-2 seq       14161 ttttaagttacgctcagctgtgatgggtgacaatcagtgcattactgttttatcagtctt
G-pig Win     13775 ttttaagttacgctcagctgtgatgggtgacaatcagtgcattactgttttatcagtctt

I-2 seq       14221 ccccttagagactgacgcagacgagcaggaacagagcgccgaagacaatgcagcgagggt
G-pig Win     13835 ccccttagagactgacgcagacgagcaggaacagagcgccgaagacaatgcagcgagggt

I-2 seq       14281 ggccgccagcctagcaaaagttacaagtgcctgtggaatctttttaaaacctgatgaaac
G-pig Win     13895 ggccgccagcctagcaaaagttacaagtgcctgtggaatctttttaaaacctgatgaaac

I-2 seq       14341 atttgtacattcaggttttatctattttggaaaaaaacaatatttgaatggggtccaatt
G-pig Win     13955 atttgtacattcgggttttatctattttggaaaaaaacaatatttgaatggggtccaatt

I-2 seq       14401 gcctcagtcccttaaaacggctacaagaatggcaccattgtctgatgcaatttttgatga
G-pig Win     14015 gcctcagtcccttaaaacggctaccagaatggcaccattgtctgatgcaatttttgatga

I-2 seq       14461 tcttcaagggaccctggctagtataggcactgcttttgagcgatccatctctgagacacg
G-pig Win     14075 tcttcaagggaccctggctagtataggcactgcttttgagcgatccatctctgagacacg

I-2 seq       14521 acatatctttccttgcaggataaccgcagctttccatacgtttttttcggtgagaatctt
G-pig Win     14135 acatatctttccttgcaggataaccgcagctttccatacgtttttttcggtgagaatctt

I-2 seq       14581 gcaatatcatcatctcgggttcaataaaggttttgaccttggacagttaacactcggcaa
G-pig Win     14195 gcaatatcatcatctcgggttcaataaaggttttgaccttggacagttaacactcggcaa

I-2 seq       14641 acctctggatttcggaacaatatcattggcactagcggtaccgcaggtgcttggagggtt
G-pig Win     14255 acctctggatttcggaacaatatcattggcactagcggtaccgcaggtgcttggagggtt

I-2 seq       14701 atccttcttgaatcctgagaaatgtttctaccggaatctaggagatccagttacctcagg
G-pig Win     14315 atccttcttgaatcctgagaaatgtttctaccggaatctaggagatccagttacctcagg

I-2 seq       14761 cttattccagttaaaaacttatctccgaatgattgagatggatgatttattcttacctttt
G-pig Win     14375 cttattccagttaaaaacttatctccgaatgattgagatggatgatttattcttacctttt

I-2 seq       14821 aattgcgaagaaccctgggaactgcactgccattgactttgtgctaaatcctagcggatt
G-pig Win     14435 aattgcgaagaaccctgggaactgcactgccattgactttgtgctaaatcctagcggatt
```

```
I-2 seq       14881 aaatgtccctgggtcgcaagacttaacttcatttctgcgccagattgtacgcaggaccat
G-pig Win     14495 aaatgtccctgggtcgcaagacttaacttcatttctgcgccagattgtacgcaggaccat

I-2 seq       14941 caccctaagtgcgaaaaacaaacttattaataccttatttcatgcgtcagctgacttcga
G-pig Win     14555 caccctaagtgcgaaaaacaaacttattaataccttatttcatgcgtcagctgacttcga

I-2 seq       15001 agacgaaatggtttgtaaatggctattatcatcaactcctgttatgagtcgttttgcggc
G-pig Win     14615 agacgaaatggtttgtaaatggctattatcatcaactcctgttatgagtcgttttgcggc

I-2 seq       15061 cgatatcttttcacgcacgccgagcgggaagcgattgcaaattctaggatacctggaagg
G-pig Win     14675 cgatatcttttcacgcacgccgagcgggaagcgattgcaaattctaggatacctggaagg

I-2 seq       15121 aacacgcacattattagcctctaagatcatcaacaataatacagagacaccggttttgga
G-pig Win     14735 aacacgcacattattagcctctaagatcatcaacaataatacagagacaccggttttgga

I-2 seq       15181 cagactgaggaaaataacattgcaaaggtggagcctatggtttagttatcttgatcattg
G-pig Win     14795 cagactgaggaaaataacattgcaaaggtggagcctatggtttagttatcttgatcattg

I-2 seq       15241 tgataatatcctggcggaggctttaacccaaataacttgcacagttgatttagcacagat
G-pig Win     14855 tgataatatcctggcggaggctttaacccaaataacttgcacagttgatttagcacagat

I-2 seq       15301 tctgagggaatattcatgggctcatattttagagggaagacctcttattggagccacact
G-pig Win     14915 tctgagggaatattcatgggctcatattttagagggaagacctcttattggagccacact

I-2 seq       15361 cccatgtatgattgagcaattcaaagtgttttggctgaaaccctacgaacaatgtccgca
G-pig Win     14975 cccatgtatgattgagcaattcaaagtgttttggctgaaaccctacgaacaatgtccgca

I-2 seq       15421 gtgttcaaatgcaaagcaaccaggtgggaaaccattcgtgtcagtggcagtcaagaaaca
G-pig Win     15035 gtgttcaaatgcaaagcaaccaggtgggaaaccattcgtgtcagtggcagtcaagaaaca

I-2 seq       15481 tattgttagtgcatggccgaacgcatcccgaataagctggactatcggggatggaatccc
G-pig Win     15095 tattgttagtgcatggccgaacgcatcccgaataagctggactatcggggatggaatccc

I-2 seq       15541 atacattggatcaaggacagaagataagataggacaacctgctattaaaccaaaatgtcc
G-pig Win     15155 atacattggatcaaggacagaagataagataggacaacctgctattaaaccaaaatgtcc

I-2 seq       15601 ttccgcagccttaagagaggccattgaattggcgtcccgtttaacatgggtaactcaagg
G-pig Win     15215 ttccgcagccttaagagaggccattgaattggcgtcccgtttaacatgggtaactcaagg

I-2 seq       15661 cagttcgaacagtgacttgctaataaaaccatttttggaagcacgagtaaatttaagtgt
G-pig Win     15275 cagttcgaacagtgacttgctaataaaaccatttttggaagcacgagtaaatttaagtgt

I-2 seq       15721 tcaagaaatacttcaaatgacccccttcacattactcaggaaatattgttcacaggtacaa
G-pig Win     15335 tcaagaaatacttcaaatgacccccttcacattactcaggaaatattgttcacaggtacaa

I-2 seq       15781 cgatcaatacagtcctcattctttcatggccaatcgtatgagtaattcagcaacgcgatt
G-pig Win     15395 cgatcaatacagtcctcattctttcatggccaatcgtatgagtaattcagcaacgcgatt

I-2 seq       15841 gattgtttctacaaacactttaggtgagttttcaggaggtggccagtctgcacgcgacag
G-pig Win     15455 gattgtttctacaaacactttaggtgagttttcaggaggtggccagtctgcacgcgacag

I-2 seq       15901 caatattattttccagaatgttataaattatgcagttgcactgttcgatattaaatttag
G-pig Win     15515 caatattattttccagaatgttataaattatgcagttgcactgttcgatattaaatttag

I-2 seq       15961 aaacactgaggctacagatatccaatataatcgtgctcaccttcatctaactaagtgttg
G-pig Win     15575 aaacactgaggctacagatatccaatataatcgtgctcaccttcatctaactaagtgttg

I-2 seq       16021 cacccgggaagtaccagctcagtatttaacatacacatctacattggatttagatttaac
G-pig Win     15635 cacccgggaagtaccagctcagtatttaacatacacatctacattggatttagatttaac

I-2 seq       16081 aagataccgagaaaacgaattgatttatgacagtaatcctctaaaaggaggactcaattg
G-pig Win     15695 aagataccgagaaaacgaattgatttatgacagtaatcctctaaaaggaggactcaattg
```

```
I-2 seq       16141 caatatctcattcgataatccatttttccaaggtaaacggctgaacattatagaagatga
G-pig Win     15755 caatatctcattcgataatccatttttccaaggtaaacggctgaacattatagaagatga

I-2 seq       16201 tcttattcgactgcctcacttatctggatgggagctagccaagaccatcatgcaatcaat
G-pig Win     15815 tcttattcgactgcctcacttatctggatgggagctagccaagaccatcatgcaatcaat

I-2 seq       16261 tatttcagatagcaacaattcatctacagacccaattagcagtggagaaacaagatcatt
G-pig Win     15875 tatttcagatagcaacaattcatctacagacccaattagcagtggagaaacaagatcatt

I-2 seq       16321 cactacccatttcttaacttatcccaagataggacttctgtacagttttggggcctttgt
G-pig Win     15935 cactacccatttcttaacttatcccaagataggacttctgtacagttttggggcctttgt

I-2 seq       16381 aagttattatcttggcaatacaattcttcggactaagaaattaacacttgacaatttttt
G-pig Win     15995 aagttattatcttggcaatacaattcttcggactaagaaattaacacttgacaatttttt

I-2 seq       16441 atattacttaactactcaaattcataatctaccacatcgctcattgcgaatacttaagcc
G-pig Win     16055 atattacttaactactcaaattcataatctaccacatcgctcattgcgaatacttaagcc

I-2 seq       16501 aacattcaaacatgcaagcgttatgtcacggttaatgagtattgatcctcatttttctat
G-pig Win     16115 aacattcaaacatgcaagcgttatgtcacggttaatgagtattgatcctcatttttctat

I-2 seq       16561 ttacataggcggtgctgcaggtgacagaggactctcagatgcggccaggttatttttgag
G-pig Win     16175 ttacataggcggtgctgcaggtgacagaggactctcagatgcggccaggttatttttgag

I-2 seq       16621 aacgtccatttcatcttttcttacatttgtaaaagaatggataattaatcgcggaacaat
G-pig Win     16235 aacgtccatttcatcttttcttacatttgtaaaagaatggataattaatcgcggaacaat

I-2 seq       16681 tgtccctttatggatagtatatccgctagagggtcaaaacccaacacctgtgaataattt
G-pig Win     16295 tgtccctttatggatagtatatccgctagagggtcaaaacccaacacctgtgaataattt

I-2 seq       16741 tctctatcagatcgtagaactgctggtgcatgattcatcaagacaacaggcttttaaaac
G-pig Win     16355 tctctatcagatcgtagaactgctggtgcatgattcatcaagacaacaggcttttaaaac

I-2 seq       16801 taccataagtgatcatgtacatcctcacgacaatcttgtttacacatgtaagagtacagc
G-pig Win     16415 taccataagtgatcatgtacatcctcacgacaatcttgtttacacatgtaagagtacagc

I-2 seq       16861 cagcaatttcttccatgcatcattggcgtactggaggagcagacacagaaacagcaaccg
G-pig Win     16475 cagcaatttcttccatgcatcattggcgtactggaggagcagacacagaaacagcaaccg

I-2 seq       16921 aaaatacttggcaagagactcttcaactggatcaagcacaaacaacagtgatggtcatat
G-pig Win     16535 aaaatacttggcaagagactcttcaactggatcaagcacaaacaacagtgatggtcatat

I-2 seq       16981 tgagagaagtcaagaacaaaccaccagagatccacatgatggcactgaacggaatctagt
G-pig Win     16595 tgagagaagtcaagaacaaaccaccagagatccacatgatggcactgaacggaatctagt

I-2 seq       17041 cctacaaatgagccatgaaataaaaagaacgacaattccacaagaaaacacgcaccaggg
G-pig Win     16655 cctacaaatgagccatgaaataaaaagaacgacaattccacaagaaaacacgcaccaggg

I-2 seq       17101 tccgtcgttccagtcctttctaagtgactctgcttgtggtacagcaaatccaaaactaaa
G-pig Win     16715 tccgtcgttccagtcctttctaagtgactctgcttgtggtacagcaaatccaaaactaaa

I-2 seq       17161 tttcgatcgatcgagacacaatgtgaaatttcaggatcataactcggcatccaagaggga
G-pig Win     16775 tttcgatcgatcgagacacaatgtgaaatttcaggatcataactcggcatccaagaggga

I-2 seq       17221 aggtcatcaaataatctcacaccgtctagtcctacctttctttacattatctcaagggac
G-pig Win     16835 aggtcatcaaataatctcacaccgtctagtcctacctttctttacattatctcaagggac

I-2 seq       17281 acgccaattaacgtcatccaatgagtcacaaacccaagacgagatatcaaagtacttacg
G-pig Win     16895 acgccaattaacgtcatccaatgagtcacaaacccaagacgagatatcaaagtacttacg

I-2 seq       17341 gcaattgagatccgtcattgataccacagtttattgtagatttaccggtatagtctcgtc
G-pig Win     16955 gcaattgagatccgtcattgataccacagtttattgtagatttaccggtatagtctcgtc
```

```
I-2 seq        17401 catgcattacaaacttgatgaggtcctttgggaaatagagagtttcaagtcggctgtgac
G-pig Win      17015 catgcattacaaacttgatgaggtcctttgggaaatagagagtttcaagtcggctgtgac

I-2 seq        17461 gctagcagagggagaaggtgctggtgccttactattgattcagaaataccaagttaagac
G-pig Win      17075 gctagcagagggagaaggtgctggtgccttactattgattcagaaataccaagttaagac

I-2 seq        17521 cttatttttcaacacgctagctactgagtccagtatagagtcagaaatagtatcaggaat
G-pig Win      17135 cttatttttcaacacgctagctactgagtccagtatagagtcagaaatagtatcaggaat

I-2 seq        17581 gactactcctaggatgcttctacctgttatgtcaaaattccataatgaccaaattgagat
G-pig Win      17195 gactactcctaggatgcttctacctgttatgtcaaaattccataatgaccaaattgagat

I-2 seq        17641 tattcttaacaactcagcaagccaaataacagacataacaaatcctacttggtttaaaga
G-pig Win      17255 tattcttaacaactcagcaagccaaataacagacataacaaatcctacttggtttaaaga

I-2 seq        17701 ccaaagagcaaggctacctaagcaagtcgaggttataaccatggatgcagagacaacaga
G-pig Win      17315 ccaaagagcaaggctacctaagcaagtcgaggttataaccatggatgcagagacaacaga

I-2 seq        17761 gaatataaacagatcgaaattgtacgaagctgtatataaattgatcttacaccatattga
G-pig Win      17375 gaatataaacagatcgaaattgtacgaagctgtatataaattgatcttacaccatattga

I-2 seq        17821 tcctagcgtattgaaagcagtggtccttaaagtctttctaagtgatactgagggtatgtt
G-pig Win      17435 tcctagcgtattgaaagcagtggtccttaaagtctttctaagtgatactgagggtatgtt

I-2 seq        17881 atggctaaatgataatttagccccgttttttgccactggttatttaattaagccaataac
G-pig Win      17495 atggctaaatgataatttagccccgttttttgccactggttatttaattaagccaataac

I-2 seq        17941 gtcaagtgctagatctagtgagtggtatctttgtctgacgaacttcttatcaactacacg
G-pig Win      17555 gtcaagtgctagatctagtgagtggtatctttgtctgacgaacttcttatcaactacacg

I-2 seq        18001 taagatgccacaccaaaaccatctcagttgtaaacaggtaatacttacggcattgcaact
G-pig Win      17615 taagatgccacaccaaaaccatctcagttgtaaacaggtaatacttacggcattgcaact

I-2 seq        18061 gcaaattcaacgaagcccatactggctaagtcatttaactcagtatgctgactgtgagtt
G-pig Win      17675 gcaaattcaacgaagcccatactggctaagtcatttaactcagtatgctgactgtgagtt

I-2 seq        18121 acatttaagttatatccgccttggttttccatcattagagaaagtactataccacaggta
G-pig Win      17735 acatttaagttatatccgccttggttttccatcattagagaaagtactataccacaggta

I-2 seq        18181 taacctcgtcgattcaaaaagaggtccactagtctctatcactcagcacttagcacatct
G-pig Win      17795 taacctcgtcgattcaaaaagaggtccactagtctctatcactcagcacttagcacatct

I-2 seq        18241 tagagcagagattcgagaattaactaatgattataatcaacagcgacaaagtcggactca
G-pig Win      17855 tagagcagagattcgagaattaactaatgattataatcaacagcgacaaagtcggactca

I-2 seq        18301 aacatatcactttattcgtactgcaaaaggacgaatcacaaaactagtcaatgattattt
G-pig Win      17915 aacatatcactttattcgtactgcaaaaggacgaatcacaaaactagtcaatgattattt

I-2 seq        18361 aaaattctttcttattgtgcaagcattaaaacataatgggacatggcaagctgagtttaa
G-pig Win      17975 aaaattctttcttattgtgcaagcattaaaacataatgggacatggcaagctgagtttaa

I-2 seq        18421 gaaattaccagagttgattagtgtgtgcaataggttctaccatattagagattgcaattg
G-pig Win      18035 gaaattaccagagttgattagtgtgtgcaataggttctaccatattagagattgcaattg

I-2 seq        18481 tgaagaacgtttcttagttcaaaccttatatttacatagaatgcaggattctgaagttaa
G-pig Win      18095 tgaagaacgtttcttagttcaaaccttatatttacatagaatgcaggattctgaagttaa

I-2 seq        18541 gcttatcgaaaggctgacagggcttctgagtttatttccggatggtctctacaggtttga
G-pig Win      18155 gcttatcgaaaggctgacagggcttctgagtttatttccggatggtctctacaggtttga

I-2 seq        18601 ttgaattaccgagcatagtatcctgatacttgcaaaggttggttattaacatacagatta
G-pig Win      18215 ttgaattaccgagcatagtatcctgatacttgcaaaggttggttattaacatacagatta
```

```
I-2 seq       18661 taaaaaactcataaattgctctcatacatcatattgatctaatctcaataaacaactatt
G-pig Win     18275 taaaaaactcataaattgctctcatacatcatattgatctaatctcaataaacaactatt

I-2 seq       18721 taaataacgaaaggagtccctatattatatactatatttagcctctctccctgcgtgata
G-pig Win     18335 taaataacgaaaggagtccctatattatatactatatttagcctctctccctgcgtgata

I-2 seq       18781 atcaaaaaattcacaatgcagcatgtgtgacatattactgccgcaatgaatttaacgcaa
G-pig Win     18395 atcaaaaaattcacaatgcagcatgtgtgacatattactgccgcaatgaatttaacgcaa

I-2 seq       18841 cataataaactctgcactctttataattaagctttaacgaaggtctgggctcatattgtt
G-pig Win     18455 cataataaactctgcactctttataattaagctttaacgaaggtctgggctcatattgtt

I-2 seq       18901 attgatataataatgttgtatcaatatcctgtcagatggaatagtgttttggttgataac
G-pig Win     18515 attgatataataatgttgtatcaatatcctgtcagatggaatagtgttttggttgataac

I-2 seq       18961 acaacttcttaaaacaaaattgatctttaagattaagtttttataattatcattacttt
G-pig Win     18575 acaacttcttaaaacaaaattgatctttaagattaagtttttataattatcattacttt

I-2 seq       19021 aatttgtcgttttaaaaacggtgatagccttaatctttgtgtaaaataagagattaggtg
G-pig Win     18635 aatttgtcgttttaaaaacggtgatagccttaatctttgtgtaaaataagagattaggtg

I-2 seq       19081 taataaccttaacatttttgtctagtaagctactatttcatacagaatgataaaattaaa
G-pig Win     18695 taataaccttaacatttttgtctagtaagctactatttcatacagaatgataaaattaaa

I-2 seq       19141 agaaaaggcaggactgtaaaatcagaaataccttctttacaatatagcagactagataat
G-pig Win     18755 agaaaaggcaggactgtaaaatcagaaataccttctttacaatatagcagactagataat

I-2 seq       19201 aatcttcgtgttaatgataattaagacattgaccacgctcatcagaaggctcgccagaat
G-pig Win     18815 aatcttcgtgttaatgataattaagacattgaccacgctcatcagaaggctcgccagaat

I-2 seq       19261 aaacgttgcaaaaaggattcctggaaaaatggtcgcacacaaaaatttaaaaataaatct
G-pig Win     18875 aaacgttgcaaaaaggattcctggaaaaatggtcgcacacaaaaatttaaaaataaatct

I-2 seq       19321 atttcttctttttgtgtgtccagggtcggcatggcatctccacctcctcgcggtccgac
G-pig Win     18935 atttcttctttttgtgtgtcca----------------------------------------

I-2 seq       19381 ctgggcatccgaaggaggacgtcgtccactcggatggctaagggagagcttctagacgta
G-pig Win     18958 ----------------------------------------------------------

I-2 seq       19441 cgctcgacgatccggctgctaacaaagcccgaaaggaagctgagttggctgctgccaccg
G-pig Win     18958 ----------------------------------------------------------

I-2 seq       19501 ctgagcaataactagcataacccctttggggcctctaaacgggtcttgaggggttttttgc
G-pig Win     18958 ----------------------------------------------------------

I-2 seq       19561 tgaaaggaggaactatatccggatcgagatcaattctgtgagcgtatggcaaacgaagga
G-pig Win     18958 ----------------------------------------------------------

I-2 seq       19621 aaaatagttatagtagccgcactcgatgggacatttcaacgtaaaccgtttaataatatt
G-pig Win     18958 ----------------------------------------------------------

I-2 seq       19681 ttgaatcttattccattatctgaaatggtggtaaaactaactgctgtgtgtatgaaatgc
G-pig Win     18958 ----------------------------------------------------------

I-2 seq       19741 tttaaggaggcttccttttctaaacgattgggtgaggaaaccgagatagaaataatagga
G-pig Win     18958 ----------------------------------------------------------

I-2 seq       19801 ggtaatgatatgtatcaatcggtgtgtagaaagtgttacatcgactcataatattatatt
G-pig Win     18958 ----------------------------------------------------------

I-2 seq       19861 ttttatctaaaaaactaaaaataaacattgattaaattttaatataatacttaaaaatgg
G-pig Win     18958 ----------------------------------------------------------
```

```
I-2 seq       19921 atgttgtgtcgttagataaaccgtttatgtattttgaggaaattgataatgagttagatt
G-pig Win     18958 ------------------------------------------------------------

I-2 seq       19981 acgaaccagaaagtgcaaatgaggtcgcaaaaaaactgccgtatcaaggacagttaaaac
G-pig Win     18958 ------------------------------------------------------------

I-2 seq       20041 tattactaggagaattatttttcttagtaagttacagcgacacggtatattagatggtg
G-pig Win     18958 ------------------------------------------------------------

I-2 seq       20101 ccaccgtagtgtatataggatctgctcccggtacacatatacgttatttgagagatcatt
G-pig Win     18958 ------------------------------------------------------------

I-2 seq       20161 tctataatttaggagtgatcatcaaatggatgctaattgacggccgccatcatgatccta
G-pig Win     18958 ------------------------------------------------------------

I-2 seq       20221 ttttaaatggattgcgtgatgtgactctagtgactcggttcgttgatgaggaatatctac
G-pig Win     18958 ------------------------------------------------------------

I-2 seq       20281 gatccatcaaaaaacaactgcatccttctaagattattttaatttctgatgtgagatcca
G-pig Win     18958 ------------------------------------------------------------

I-2 seq       20341 aacgaggaggaaatgaacctagtacggcggatttactaagtaattacgctctacaaaatg
G-pig Win     18958 ------------------------------------------------------------

I-2 seq       20401 tcatgattagtattttaaaccccgtggcgtctagtcttaaatggagatgcccgtttccag
G-pig Win     18958 ------------------------------------------------------------

I-2 seq       20461 atcaatggatcaaggacttttatatcccacacggtaataaaatgttacaacctttgctc
G-pig Win     18958 ------------------------------------------------------------

I-2 seq       20521 cttcatattcagggccgtcgttttacaacgtcgtgactgggaaaaccctggcgttaccca
G-pig Win     18958 ------------------------------------------------------------

I-2 seq       20581 acttaatcgccttgcagcacatccccctttcgccagctggcgtaatagcgaagaggcccg
G-pig Win     18958 ------------------------------------------------------------

I-2 seq       20641 caccgatcgcccttcccaacagttgcgcagcctgaatggcgaatggcgcgacgcgccctg
G-pig Win     18958 ------------------------------------------------------------

I-2 seq       20701 tagcggcgcattaagcgcggcgggtgtggtggttacgcgcagcgtgaccgctacacttgc
G-pig Win     18958 ------------------------------------------------------------

I-2 seq       20761 cagcgccctagcgcccgctcctttcgctttcttcccttcctttctcgccacgttcgccgg
G-pig Win     18958 ------------------------------------------------------------

I-2 seq       20821 ctttccccgtcaagctctaaatcgggggctcccttagggttccgatttagtgctttacg
G-pig Win     18958 ------------------------------------------------------------

I-2 seq       20881 gcacctcgaccccaaaaaacttgattagggtgatggttcacgtagtgggccatcgccctg
G-pig Win     18958 ------------------------------------------------------------

I-2 seq       20941 atagacggtttttcgccctttgacgttggagtccacgttctttaatagtggactcttgtt
G-pig Win     18958 ------------------------------------------------------------

I-2 seq       21001 ccaaactggaacaacactcaaccctatctcggtctattcttttgatttataagggatttt
G-pig Win     18958 ------------------------------------------------------------

I-2 seq       21061 gccgatttcggcctattggttaaaaaatgagctgatttaacaaaaatttaacgcgaattt
G-pig Win     18958 ------------------------------------------------------------

I-2 seq       21121 taacaaaatattaacgtttacaatttcccaggtggcacttttcggggaaatgtgcgcgga
G-pig Win     18958 ------------------------------------------------------------
```

```
I-2 seq       21181 acccctatttgtttatttttctaaatacattcaaatatgtatccgctcatgagacaataa
G-pig Win     18958 ------------------------------------------------------------

I-2 seq       21241 ccctgataaatgcttcaataatattgaaaaaggaagagtatgagtattcaacatttccgt
G-pig Win     18958 ------------------------------------------------------------

I-2 seq       21301 gtcgcccttattccctttttgcggcattttgccttcctgttttgctcacccagaaacg
G-pig Win     18958 ------------------------------------------------------------

I-2 seq       21361 ctggtgaaagtaaaagatgctgaagatcagttgggtgcacgagtgggttacatcgaactg
G-pig Win     18958 ------------------------------------------------------------

I-2 seq       21421 gatctcaacagcggtaagatccttgagagttttcgccccgaagaacgttttccaatgatg
G-pig Win     18958 ------------------------------------------------------------

I-2 seq       21481 agcacttttaaagttctgctatgtggcgcggtattatcccgtattgacgccgggcaagag
G-pig Win     18958 ------------------------------------------------------------

I-2 seq       21541 caactcggtcgccgcatacactattctcagaatgacttggttgagtactcaccagtcaca
G-pig Win     18958 ------------------------------------------------------------

I-2 seq       21601 gaaaagcatcttacggatggcatgacagtaagagaattatgcagtgctgccataaccatg
G-pig Win     18958 ------------------------------------------------------------

I-2 seq       21661 agtgataacactgcggccaacttacttctgacaacgatcggaggaccgaaggagctaacc
G-pig Win     18958 ------------------------------------------------------------

I-2 seq       21721 gctttttgcacaacatgggggatcatgtaactcgccttgatcgttgggaaccggagctg
G-pig Win     18958 ------------------------------------------------------------

I-2 seq       21781 aatgaagccataccaaacgacgagcgtgacaccacgatgcctgtagcaatggcaacaacg
G-pig Win     18958 ------------------------------------------------------------

I-2 seq       21841 ttgcgcaaactattaactggcgaactacttactctagcttcccggcaacaattaatagac
G-pig Win     18958 ------------------------------------------------------------

I-2 seq       21901 tggatggaggcggataaagttgcaggaccacttctgcgctcggcccttccggctggctgg
G-pig Win     18958 ------------------------------------------------------------

I-2 seq       21961 tttattgctgataaatctggagccggtgagcgtgggtctcgcggtatcattgcagcactg
G-pig Win     18958 ------------------------------------------------------------

I-2 seq       22021 gggccagatggtaagccctcccgtatcgtagttatctacacgacggggagtcaggcaact
G-pig Win     18958 ------------------------------------------------------------

I-2 seq       22081 atggatgaacgaaatagacagatcgctgagataggtgcctcactgattaagcattggtaa
G-pig Win     18958 ------------------------------------------------------------

I-2 seq       22141 ctgtcagaccaagtttactcatatatactttagattgatttaaaacttcatttttaattt
G-pig Win     18958 ------------------------------------------------------------

I-2 seq       22201 aaaaggatctaggtgaagatccttttttgataatctcatgaccaaaatcccttaacgtgag
G-pig Win     18958 ------------------------------------------------------------

I-2 seq       22261 ttttcgttccactgagcgtcagaccccgtagaaaagatcaaaggatcttcttgagatcct
G-pig Win     18958 ------------------------------------------------------------

I-2 seq       22321 ttttttctgcgcgtaatctgctgcttgcaaacaaaaaaaccaccgctaccagcggtggtt
G-pig Win     18958 ------------------------------------------------------------

I-2 seq       22381 tgtttgccggatcaagagctaccaactctttttccgaaggtaactggcttcagcagagcg
G-pig Win     18958 ------------------------------------------------------------
```

```
I-2 seq       22441 cagataccaaatactgtccttctagtgtagccgtagttaggccaccacttcaagaactct
G-pig Win     18958 ------------------------------------------------------------

I-2 seq       22501 gtagcaccgcctacatacctcgctctgctaatcctgttaccagtggctgctgccagtggc
G-pig Win     18958 ------------------------------------------------------------

I-2 seq       22561 gataagtcgtgtcttaccgggttggactcaagacgatagttaccggataaggcgcagcgg
G-pig Win     18958 ------------------------------------------------------------

I-2 seq       22621 tcgggctgaacgggggttcgtgcacacagcccagcttggagcgaacgacctacaccgaa
G-pig Win     18958 ------------------------------------------------------------

I-2 seq       22681 ctgagatacctacagcgtgagctatgagaaagcgccacgcttcccgaagggagaaaggcg
G-pig Win     18958 ------------------------------------------------------------

I-2 seq       22741 gacaggtatccggtaagcggcagggtcggaacaggagagcgcacgagggagcttccaggg
G-pig Win     18958 ------------------------------------------------------------

I-2 seq       22801 ggaaacgcctggtatctttatagtcctgtcgggtttcgccacctctgacttgagcgtcga
G-pig Win     18958 ------------------------------------------------------------

I-2 seq       22861 tttttgtgatgctcgtcaggggggcggagcctatggaaaaacgccagcaacgcggccttt
G-pig Win     18958 ------------------------------------------------------------

I-2 seq       22921 ttacggttcctggccttttgctggccttttgctcacatgttctttcctgcgttatcccct
G-pig Win     18958 ------------------------------------------------------------

I-2 seq       22981 gattctgtggataaccgtattaccgcctttgagtgagctgataccgctcgccgcagccga
G-pig Win     18958 ------------------------------------------------------------

I-2 seq       23041 acgaccgagcgcagcgagtcagtgagcgaggaagcggaagagcgcccaatacgcaaaccg
G-pig Win     18958 ------------------------------------------------------------

I-2 seq       23101 cctctccccgcgcgttggccgattcattaatgcagctggcacgacaggtttcccgactgg
G-pig Win     18958 ------------------------------------------------------------

I-2 seq       23161 aaagcgggcagtgagcgcaacgcaattaatgtgagttagctcactcattaggcaccccag
G-pig Win     18958 ------------------------------------------------------------

I-2 seq       23221 gctttacactttatgcttccggctcgtatgttgtgtggaattgtgagcggataacaattt
G-pig Win     18958 ------------------------------------------------------------

I-2 seq       23281 cacacaggaaacagctatgaccatgattacgcc
G-pig Win     18958 -----------------------------------
```