The Molecular Characterization and the Generation of a Reverse Genetics System for Kyasanur Forest Disease Virus

by

Bradley William Michael Cook

A Thesis submitted to the Faculty of Graduate Studies of

The University of Manitoba

in partial fulfilment of the requirements of the degree of

Master of Science

Department of Microbiology

University of Manitoba

Winnipeg, Manitoba, Canada

List of Abbreviations:

AHFV - Alkhurma Hemorrhagic Fever Virus

Amp – ampicillin

APOIV - Apoi Virus

ATP – adenosine tri-phosphate

BAC – bacterial artificial chromosome

BHK – Baby Hamster Kidney

BSA – bovine serum albumin

C1 – C-terminus fragment 1

C2 – C-terminus fragment 2

C - Capsid protein

cDNA – comlementary Deoxyribonucleic acid

CL – Containment Level

CO_2 – carbon dioxide

cHP - capsid hairpin

CNS - Central Nervous System

CPE – cytopathic effect

CS - complementary sequences

DENV1-4 - Dengue Virus

DIC - Disseminated Intravascular Coagulation (DIC)

DNA – Deoxyribonucleic acid

DTV - Deer Tick Virus

E - Envelope protein

EDTA - ethylenediaminetetraacetic acid

EM - Electron Microscopy

EMCV – Encephalomyocarditis Virus

ER - endoplasmic reticulum

FBS – fetal bovine serum

FP - fusion peptide

GGEV - Greek Goat Encephalitis Virus

GGYV - Gadgets Gully Virus

GMP - Guanosine mono-phosphate

GTP - Guanosine tri-phosphate

HBV- Hepatitis B Virus

HDV – Hepatitis Delta Virus

HIV - Human Immunodeficiency Virus

IFN – interferon

IRES – internal ribosome entry sequence

JEV - Japanese Encephalitis Virus

KADV - Kadam Virus

kDa - kilo Dalton

KFDV - Kyasanur Forest Disease Virus

KSIV - Karshi Virus

KUNV - Kunjin Virus

kV – kilo volts

LB – Luria-Bertani

LD - Lipid Droplets

LGTV - Langat Virus

LIV - Louping Ill Virus

M - Membrane protein

M-TB - Mammalian-Tick Borne

MAC - Membrane Attack Complex (MAC)

MB - Mosquito-Borne

MCS – multiple cloning site

MEAV - Meaban Virus

$MgCl_2$ – magnesium chloride

MMLV - Montana Myotis Leukoencephalitis Virus

MODV - Modoc Virus

mRNA - messenger RNA

MTPase - methyl-transferase

MVEV - Murray Valley Encephalitis Virus

N1 – N-terminus fragment 1

NaCl – sodium chloride

NKV - No Known Vector

NML – National Microbiology Laboratory

NS - Non-Structural proteins

NTP - nucleoside triphosphate

NTPase - nucleoside tri-phosphatase

OHFV$^{UVE/Li/Lin}$ - Omsk Hemorrhagic Fever Virus

PFU – plaque forming units

PHAC – Public Health Agency of Canada

POWV - Powassan Virus

prM - pre-Membrane protein

RBV - Rio Bravo Virus

RC - replication complex

RdRp - RNA-dependent RNA polymerase

RFV - Royal Farm Virus

RGS – reverse genetics system

RNA - Ribonucleic Acid

RNP – ribonucleoprotein complex

RPM – revolutions per minute

RT-PCR – reverse transcription polymerase chain reaction

RTPase - RNA tri-phosphatase

S-TB - Seabird-Tick Borne

SDS – sodium dodecyl sulfate

SDS-PAGE – sodium dodecyl sulfate-polyacrylamide gel electrophoresis

SL - stem-loop

SLEV - St. Louis Encephalitis Virus

SREV - Saumarez Reef Virus

ss - single-stranded

SSEV - Spanish Sheep Encephalomyelitis Virus

TAP – tobacco acid pyrophosphatase

TB - Tick-Borne

TBE – tris-borate ethylenediaminetetraacetic acid

TBE serocomplex - Tick-Borne Encephalitis serocomplex

$TBEV^{FE/Sib/EU}$ - Tick-Borne Encephalitis Virus

TSEV - Turkish Sheep Encephailitis Virus

TYUV - Tyuleniy Virus

UAR - Upstream AUG Region

UTR - Untranslated Region

WNV - West Nile Virus

YFV/YFV17D - Yellow Fever Virus

YOKV - Yokose Virus

Abstract:

Kyasanur Forest Disease Virus (KFDV) is a tick-borne, hemorrhagic fever-causing member of the *Flaviviridae*. With infections annually ranging from 50 to 1000 people in south-west India and the lack of effective treatments, a better understanding of this virus is needed. The development of a Reverse Genetics System for KFDV would provide the opportunity to address these issues in future studies. Using molecular techniques, the KFDV genome sequence was elucidated and the reverse genetics system was created. Utilizing this system live, infectious KFDV particles were produced from mammalian cell culture, thereby validating the success of the reverse genetics system. The implementation of this system will enable researchers to better study pathogenesis and disease progression, virus-host interplay, virion structure, genome replication and the emergence of effective therapeutics and vaccines.

Acknowledgements:

I would first like to acknowledge Dr. Steven Theriault and Dr. Deborah Court, for allowing me the opportunity to do the research that I enjoy and for being extremely patient and supportive. Thank you to my remaining committee members Dr. Teresa de Kievit and Dr. David Levin for their interest, knowledge and support. Mr. Jay Krishnan and Mr. Todd Cutts, thank you both for your endless knowledge, expertise and guidance. Diane Gordon, Bree-Ann Carruthers and the past members of the Applied Biosafety Research Program (ABRP) thanks for your continued encouragement. The National Microbiology Laboratory for the use of their facilities, especially the DNAcore department. The University of Manitoba and the department of Microbiology, with a special thank you to: Dr. Georg Hausner, for sequence analyses and Dr. Ivan Oresnik, for convincing me to pursue graduate work. Also thank you to Dr. Michael Holbrook for his advice.

I could not have done this without all of you!

Dedication:

My work is dedicated to my savior Jesus Christ and to the loving memories of my grandmother, mother and sister.

I would also like to dedicate this work to: My Kim Tran, for her endless support and to the rest of my close friends for their assistance.

List of Tables

List of Figures

Chapter 1: Introduction

1.1 Introduction to the Flaviviridae

Hemorrhagic Fever viruses are endemic in different and defined geographical regions throughout the world (Ebihara, Groseth et al. 2005). In general, these viruses are located in tropical and sub-tropical regions, ranging from North America to Africa to Asia. The majority of hemorrhagic fever-causing viruses are located in Africa and South America (Gubler 2001). To date, four viral families have been associated with hemorrhagic manifestations, *Arenaviridae*, *Filoviridae*, *Bunyaviridae* and *Flaviviridae* (Ebihara, Groseth et al. 2005). Notable viruses in these families include: Lassavirus, Ebolavirus, Marburgvirus, Crimean-Congo Hemorrhagic fever virus, Yellow fever virus, Dengue virus and Kyasanur Forest Disease Virus (KFDV) (Ebihara, Groseth et al. 2005; Uchil, Kumar et al. 2006).

Virologists have made attempts to classify virus species in the same taxonomic system currently used for classifying animals and bacteria. Viruses were left out of the Whittaker five kingdom and Woese three domain classification schemes, as viruses are acellular and require a host for metabolism. However, the major groups or taxonomic designations for viruses are the families and genera (Bauman 2009). Families are further divided into sub-families, genera, and species. Species of viruses can be split into strain, variant, clade, types and even sub-types, all of which have no specific taxonomic status (White and Fenner 1994). Discovery of viruses and differentiation between virus species can be performed by various methods. Some examples include serology, genotyping and restriction mapping. A higher taxon was given to three (four) viral families; *Filoviridae*, *Paramyxoviridae*, *Rhabdoviridae* and (*Bornaviridae*) are now classified as members of

the *Mononegavirales* viral order. The higher taxonomic status was granted due to international agreement, similarities in genome structure and replication strategies. Viral taxonomy is based on genome characteristics (informal grouping), followed by the Order or Family (formal grouping). Other similar groupings of virus families into viral Orders may occur in the near future (White and Fenner 1994; Fauquet, Mayo et al. 2005).

KFDV is a member of the *Flaviviridae* virus family, with no current viral Order (Fauquet, Mayo et al. 2005). Three genera are associated with this family: *Hepacivirus*, *Pestivirus* and *Flavivirus*. The *Flavivirus* genus is divided into three different groups: Tick-Borne (TB), Mosquito-Borne (MB) and No Known Vector (NKV) groups. The MB group includes viruses such as: Yellow Fever Virus, which is the type-strain, West Nile Virus and Dengue Virus. NKV viruses include: Entebba Bat Virus group (Entebba Bat Virus), Modoc Virus group (Modoc Virus), and Rio Bravo Virus group (Dakar Bat Virus). The TB viruses are divided into two sub-groups: Mammalian (M-TB) and Seabird (S-TB). Kadam Virus, Meaban Virus and Saumarez Reef Virus are placed within the sub-group of S-TB. M-TB includes: Langat Virus (LGV), Powassan Virus (POWV), Tick-Borne Encephalitis Virus (TBEV) and it contains the KFDV, Omsk Hemorrhagic Fever Virus (OHFV) and Alkhurma Hemorrhagic Fever Virus (AHFV), which are of interest to this study (Howley and Knipe 2007). It was determined that many of these viruses are antigenically similar; during the 1980s mouse polyclonal sera was used to examine various Flaviviruses. The viruses that were neutralized by these antibodies were defined as the "Tick-Borne Encephalitis serocomplex" viruses (Calisher, Karabatsos et al. 1989). Within this complex are the genetically-related TBEV, OHFV and KFDV (Pattnaik 2006). In contrast to the encephalitis causing TBEV, the OHFV, KFDV and AHFV (a

more recently discovered variant of KFDV) cause hemorrhagic fever (Gritsun, Frolova et al. 2003; Pattnaik 2006).

OHFV is transmitted via bite from infected ticks, namely *Dermacentor reticulates* and *Dermacentor marginatus*. OHFV has been isolated from both of these ticks, other arthropod vectors such as fleas, mosquitoes and biting midges, have been speculated to be carriers but this is not yet proven. OHFV has been reported in a unique region and this virus is currently endemic to Omsk, near Siberia, Russia. The first reported cases occurred in 1947 and its natural mammalian reservoir is the water vole. However, OHFV has been isolated from reptiles and amphibians as well (Lin, Li et al. 2003; Holbrook, Aronson et al. 2005). This virus uses the animal vector to survive the long winter months, i.e. it has an "overwintering" period. In addition, an infected tick remains so during all the aspects of its development (transstadial period) which helps with the persistence and survival of OHFV (Lin, Li et al. 2003).

Disease progression is characterized by an acute, biphasic illness. After infection the virus has an incubation period of 2-10 days; during this time the febrile period may be seen from 5-12 days. The second period may occur around 10-15 days post infection. Symptoms can include ocular suffusion, hemorrhagic fever manifestations, mainly in the nose, lungs and uterus, papulovesicular rash and capillary toxicosis. In comparison to KFDV and AHFV viral pathogenesis, OHFV very rarely causes encephalitis or neurological illnesses (Lin, Li et al. 2003; Holbrook, Aronson et al. 2005). Furthermore, the mortality rate of OHFV is relatively small, 0.4-2.5% compared to 25-30% for AHFV (Charrel, Zaki et al. 2005; Holbrook, Aronson et al. 2005; Charrel, Fagbo et al. 2007).

The AHFV was first isolated in Saudi Arabia in 1995. The virus was isolated from a patient who was ill with hemorrhagic fever manifestations. Preceding 2007, there were approximately 20 laboratory confirmed cases (virus isolations) reported in Saudi Arabia (Charrel, Fagbo et al. 2007). Despite this, there is the potential for newly acquired infections in the neighboring countries of Yemen and Oman. This may be due to an expanded radius of people, vector and reservoir species in the Arabian Peninsula (Charrel, Zaki et al. 2005; Charrel, Fagbo et al. 2007). AHFV has been associated with a sand-dwelling soft tick (*Ornithodoros savignyi)* that bites animals, mostly sheep, camels and humans. The case fatality rate is about 25-30%, which makes this the most deadly of all the Flaviviruses (Charrel, Zaki et al. 2005; Charrel, Fagbo et al. 2007). A study of common symptoms associated with 20 confirmed AHFV infections that spanned over a two year period (2001-2003), documented these symptoms: fever, headache, malaise, myalgia, chills, backache, nausea and vomiting. 55 percent of patients experienced hemorrhagic fever manifestations including epistaxis, petechiae, bleeding from gums, gastrointestinal hemorrhage and disseminated intravascular coagulation. 35 percent of those infected presented neurological disorders such as confusion, drowsiness, coma and encephalitis (Madani 2005). After the genome was sequenced, this virus was shown to be most closely related to KFDV. The sequence analysis was based on the envelope (E) protein and two Non-Structural (NS) proteins, NS3 and NS5, see section 1.4. AHFV was deemed to be a variant of KFDV (Charrel, Zaki et al. 2001; Charrel, Zaki et al. 2005). Using the NS3, NS5 and E protein sequences, the authors attempted to determine the evolutionary relationship of AHFV and KFDV. Comparisons with other Tick-Borne

Flaviviruses such as: LGV, POWV, DTV, OHFV and KFDV, using the same protein sequences, suggest a divergence of between 66-177 years ago (Charrel, Zaki et al. 2005).

The discovery of a new variant of KFDV called Nanjianyin virus was described in a February 2009 publication (Wang, Zhang et al. 2009). The patient infected with the unknown virus, identified as a Flavivirus through cross-reactivity with JEV antibodies, displayed symptoms similar to that of KFDV. The virus was isolated from the individual in Yunnan China in 1989 and subsequently lyophilized and stored. Researchers recently thawed this virus and performed molecular techniques to discern its genetic lineage (Wang, Zhang et al. 2009). The authors amplified the pre-membrane (prM), envelope (E) and RNA-dependent RNA polymerase (NS5) protein gene sequences and compared this with KFDV strain It P9605 polyprotein sequence (AY323490) (Grard, Moureau et al. 2007), AHFV strain 1176 and other Tick-Borne Encephalitis serocomplex viruses including: OHFV, POWV, LGV and TBEV using MegAlign software (DNA STAR) (Wang, Zhang et al. 2009). The results indicated that the prM gene of Nanjianyin virus was 99.6% similar to that of KFDV and showed 90.4% similarity with AHFV and 57.2-64.3% with that of the other Tick-Borne Encephalitis serocomplex viruses. The E protein gene sequence revealed a similarity of 99.8% with KFDV, 91.9% with the AHFV and less than 72% with the other TBE serocomplex members. The gene sequence encoding the NS5 protein, which is the most conserved gene within the *Flaviviridae,* gave a similarity of 99.6% with KFDV compared to 92.3% with AHFV and less than 77.6% with the TBE serocomplex viruses. Therefore, the sequencing data of Nanjianyin virus indicated that KFDV strain It P9605 (polyprotein AY323490) was the most similar of all the TBE serocomplex viruses tested, leading to the suggestion that Nanjianyin virus is a

variant of KFDV. The authors speculated that the residents in the Hengduan Mountain area seldom travel far. Thus, the individual from whom the Nanjianyin virus was isolated, most likely acquired this virus within this region. The authors attempted to further retrace the introduction of KFDV into the region, via previously infected migratory birds accompanied by KFDV-infected ticks (Wang, Zhang et al. 2009).

Hemorrhagic fever producing properties of KFDV, OHFV and AHFV have not been attributed to any type of genetic lineage within their phylogenetic cluster. A study by Grard et al. in 2007 (Grard, Moureau et al. 2007), describes lineages of many Flaviviruses using complete open reading frame or polyprotein sequences. The lineages of the hemorrhagic fever-causing viruses are genetically associated with non-hemorrhagic fever-causing viruses, for example: POWV is clustered near KFDV and LGV-OHFV-TBEV/LIV are also linked together as seen in Figure 1.1.1 (Grard, Moureau et al. 2007).

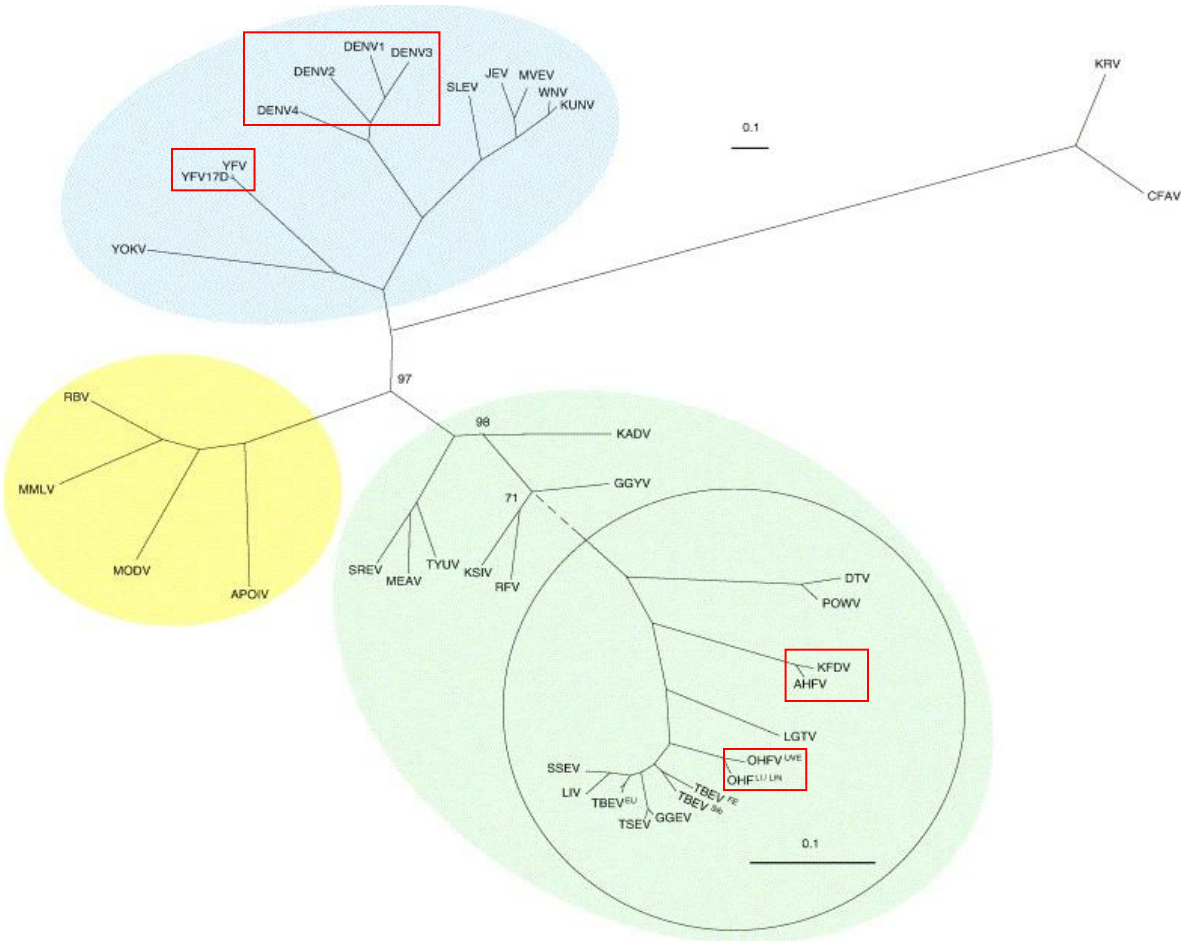


Figure 1.1.1: Phylogenetic tree based on complete polyprotein sequences from hemorrhagic and non-hemorrhagic fever causing Flaviviruses.

The Tick-Borne (TB), Mosquito-Borne (MB) and No Known Vector (NKV) Flavivirus groups are shown in green, blue and yellow respectively. Red boxes indicate the hemorrhagic fever causing Flaviviruses. The abbreviations are as follows: Yokose Virus (YOKV), Yellow Fever Virus (YFV/YFV17D), Dengue Virus (DENV1-4), St. Louis Encephalitis Virus (SLEV), Japanese Encephalitis Virus (JEV), Murray Valley Encephalitis Virus (MVEV), West Nile Virus (WNV), Kunjin Virus (KUNV), Rio Bravo Virus (RBV), Montana Myotis Leukoencephalitis Virus (MMLV), Modoc Virus (MODV), Apoi Virus (APOIV), Saumarez Reef Virus (SREV), Meaban Virus (MEAV), Tyuleniy Virus (TYUV), Karshi Virus (KSIV), Royal Farm Virus (RFV), Gadgets Gully Virus (GGYV), Kadam Virus (KADV), Deer Tick Virus (DTV), Powassan Virus (POWV), Kyasanur Forest Disease Virus (KFDV), Alkhurma Hemorrhagic Fever Virus (AHFV), Langat Virus (LGTV), Omsk Hemorrhagic Fever Virus ($OHFV^{UVE/Li/Lin}$), Tick-Borne

Encephalitis Virus ($TBEV^{FE/Sib/EU}$), Greek Goat Encephalitis Virus (GGEV), Turkish Sheep Encephailitis Virus (TSEV), Louping Ill Virus (LIV) and Spanish Sheep Encephalomyelitis Virus (SSEV)(Grard, Moureau et al. 2007). Figure referenced with permission granted from Elsevier Limited.

Since the TBE serocomplex has a subset of hemorrhagic fever-causing viruses, Lin et al. (Lin, Li et al. 2003) attempted to discern differences in neurotropic and hemorrhagic properties within the serocomplex. While focusing on the E protein, three amino acid residues were discovered to be similar in the hemorrhagic fever causing TB viruses when compared to non-hemorrhagic fever causing TBE serocomplex viruses (Lin, Li et al. 2003). Two of the three residues, E-457 and E-489, are in the stem-anchor region of the E protein, see section 1.4. The authors focused their studies on the third residue (E-76), as its location is on the portion of the protein that is required for virus-host receptor fusion, near the fusion peptide which is located in the second domain of the E protein of hemorrhagic fever-causing viruses. The 76th amino acid residue in the E protein (E-76) is an alanine instead of the threonine residue as in the other encephalitic viruses. Using computer software, reconstructions of the E protein featuring the replacement of the threonine with an alanine causes a destabilization in the E protein, as a hydroxyl group in the threonine residue is lost. With this loss of stability of the E protein dimer, the fusion peptide is sequestered in the dimer. Possible pre-mature exposure of the fusion peptide may be responsible for the occurrence of "accidental" fusions of hemorrhagic fever causing TB viruses. The result may be a change in tissue tropism and perhaps hemorrhagic fever properties in those infected (Lin, Li et al. 2003). Tropism is the ability of viruses to infect certain tissues preferentially and is governed by three main factors: susceptibility (amount receptors for virus entry), permissivity (host factors aiding in virus infectivity) and accessibility (viral interaction with host) (Flint, Enquist et al. 2004). Perhaps the "accidental" fusions alter the typical tropism of the TBE serocomplex viruses from neurotropic viruses infecting central nervous system tissue, to pantropic

viruses infecting multiple tissues and cell types (Lin, Li et al. 2003; Flint, Enquist et al. 2004). This theory is entirely possible since the exact cellular receptors for Flaviviruses are not entirely known or are poorly characterized (Howley and Knipe 2007; Hanley and Weaver 2010). However, the same threonine to alanine mutation appeared in an isolate of TBEV (strain D1283 Far Eastern Subtype) studied by Hayasaka et al. (Hayasaka, Ivanov et al. 2001) which was not analyzed by Lin et al (Lin, Li et al. 2003). Using a mouse model, infection with this virus led to neuro-pathogenic symptoms and no hemorrhagic fever manifestations. Therefore, hemorrhagic fever manifestations must be more complex than a single residue difference, between these viruses and non-hemorrhagic fever causing viruses (Hayasaka, Ivanov et al. 2001; Grard, Moureau et al. 2007). Animal models are used to study many viruses that infect humans. The type of model used to mimic human pathology may depend on the pathogens' ability to infect the animal tested, for example monkeys, ferrets, mice and guinea pigs (Flint, Enquist et al. 2009). Not all viruses will be able to infect, propagate, transmit and cause virulence in certain models, abolishing the relevance of the model for viral pathogenesis in humans (Flint, Enquist et al. 2009). An example of virus adaptation can be seen with guinea pigs and adult mice. Ebola Virus-Zaire subtype causes lethal hemorrhagic fever in human and non-human primates, but is non-lethal in guinea pigs and adult mice. Sequential passaging of the virus from adult mouse to adult mouse, using viruses isolated from the spleen and liver. The mouse-adapted virus through repeated passaging, was able to be lethal in adult mice from 5-16 weeks old. The characteristics of infection and titres were similar to that of primates, leading to a successful small animal model of Ebola Virus-Zaire (Bray, Davis et al. 1999). Adaptation to small animal models may not be a concern for the viruses

within the Tick-Borne Encephalitis serocomplex. The study by Hayasaka et al. in 2001 (Hayasaka, Ivanov et al. 2001), did not describe the need for TBEV adaptation to their mouse model. Concurrently in 2005, Holbrook et al. (Holbrook, Aronson et al. 2005) described the first use of a small animal model for a hemorrhagic fever-causing Flavivirus (OHFV), without the need for host adaptation. The opposite can be said for YFV and small animal models (Holbrook, Aronson et al. 2005). Lethal infections in laboratory settings have been described for KFDV. In small animals such as sucking mice, weaned mice and suckling hamsters, KFDV was lethal when injected through the intracerebral route (Richman, Whitley et al. 2009).

1.2 KFDV Virus Epidemiology

In 1957 an epizootic (an epidemic in animal populations)(Artsob and Lindsay 2008) event occurred within a population of monkeys, in a localized region of southwest India. This area is the Kyasanur Forest of Shimoga district, Karnataka state, India (Pattnaik 2006). At the time, this unknown virus had very high infectivity rates and virulence within humans and primates. Thus this virus was named Kyasanur Forest Disease Virus (KFDV) (Dandawate, Desai et al. 1994). The typical monkeys that acquire the virus are the black-faced langur (*Presbytis entellus*) and red-faced bonnet (*Macaca radiate*) (Pattnaik 2006). Shortly after the outbreak, the first human cases were observed. Epizootics continued throughout the Shimoga district almost every year since the first KFDV discovery in 1957. Specifically, the 1964-1973 year period had the highest rates of death. It is thought that the Shimoga district is now endemic for KFDV (Pattnaik 2006). Human infections are presumed to parallel those of the primates; reported human

cases range from 50 to over 1,000 infections per year (Venugopal, Gritsun et al. 1994; Goodman, Dennis et al. 2005; Pattnaik 2006).

Neutralizing antibodies against KFDV have been detected in humans outside of the Shimoga district in the Gujrat and West Bangal states. Oddly, people positive for KFDV antibodies were most prevalent in the Andaman and Nicobar Islands, off the east coast of India, refer to Figure 1.2.1. Some researchers are considering this area to be a potentially "silent focus" of KFDV and are monitoring the area closely (Pattnaik 2006). KFDV-specific antibodies can arise from an infection or a simple exposure to antigen(s) from KFDV. The viruses' abilities to cause infections can be negatively impacted as virus-specific antibodies can block receptor binding or cause virions to aggregate together (Flint, Enquist et al. 2004).

Clinical symptoms of KFD can be seen in two stages, *i.e.* it has a biphasic course of illness. After the incubation period of approximately 3-8 days, "flu-like" symptoms develop. This is called a febrile or "with fever period" and symptoms include: fever, headache, vomiting, severe back and extremity pain and prostration. The febrile period can last for one to two weeks and convalescence may begin ending the infection (Goodman, Dennis et al. 2005; Pattnaik 2006; Grard, Moureau et al. 2007; Office of Laboratory Safety 2008). However, on average 55% of patients may suffer more extreme and severe complications, in a second phase of the illness. In 35% of afflicted individuals, such abnormalities may include Central Nervous System (CNS) disorders such as: stiff neck, abnormal reflexes and inflammation. Another 55% of patients suffer the second phase illness course and may succumb to severe ailments such as: meningoencephalitis, conjunctival inflammation, bradycardia, hemorrhagic fever

manifestations (hemorrhagic pneumonitis, gastrointestinal bleeding), coma and possibly death (Adhikari Prabha, Prabhu et al. 1993; Howley and Knipe 2007). Of these severe symptoms, meningoencephalitis is the most common and coma and bronchopneumonia are the least common (Goodman, Dennis et al. 2005; Pattnaik 2006; Grard, Moureau et al. 2007; Office of Laboratory Safety 2008). The main contributor to hemorrhagic fever is believed to be "disseminated intravascular coagulation (DIC)" (Adhikari Prabha, Prabhu et al. 1993; Pattnaik 2006). Congealing or clotting of blood within the blood vessels can lead to shock and of course, "multi-organ system failure". What makes KFDV so dangerous is its high morbidity but, fortunately a relatively low mortality rate of from 2-10% (Pattnaik 2006; Office of Laboratory Safety 2008). This is despite the fact that up to 54% of patients may develop encephalitis or hemorrhagic fever in the second phase of illness (Adhikari Prabha, Prabhu et al. 1993). Monkeys display the same types of symptoms as humans but, have a much higher mortality rate of approximately 85% (Dobler).

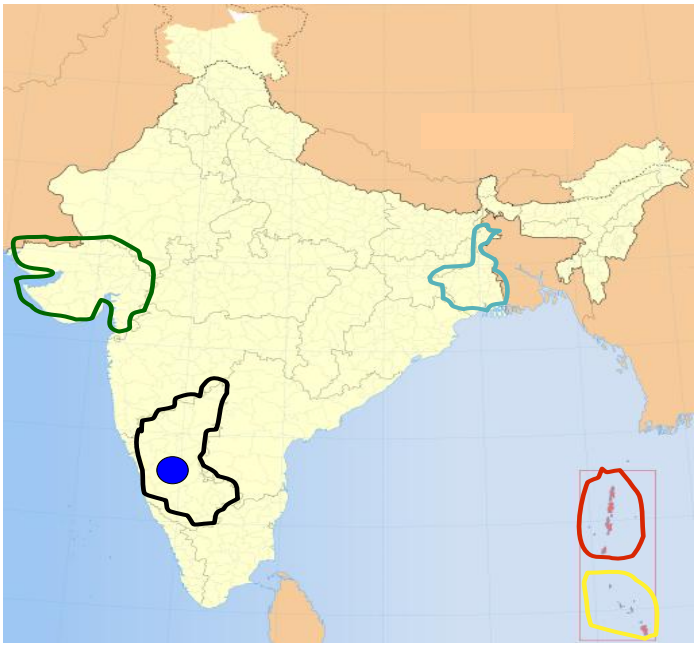


Figure 1.2.1: India.

Shimoga district: dark blue circle, Karnataka state: black line, Gujarat state: green line, West Bengal state: light blue line, Andaman islands: red line, Nicobar islands: yellow line. Adapted from (Pattnaik 2006; PlaneMad/Wikipedia). Figure referenced with permission granted from PlaneMad/Wikipedia.

1.3 KFDV Virus Transmission

1.3.1 Arthropods as Vectors for Virus Transmission

When considering disease transmission, arthropods are definitely public health threats and should not be underestimated. Mosquitoes and ticks are very diverse in the potential human pathogens they may harbor. The disease-causing organisms that can be transmitted to humans via arthropods, include bacteria (*Rickettsiae*, *Spirochetes*, etc.), protozoans (*Anaplasma*, *Ehrlichia*, etc.), and viruses (*Nairoviruses*, *Coltiviruses* and *Flavivruses*). Zooantroponoses, have been described as the human tick-borne infections, in which animal diseases become acquired by humans. Humans are considered "dead-end" or incidental hosts (Goodman, Dennis et al. 2005) because the virus will not replicate to high enough titres in the blood for an un-infected arthropod to take up the pathogen from a human host (Artsob and Lindsay 2008). The incidental nature of these tick-borne diseases can be seen, when we consider the types of exposures by which humans can contract such infections. Besides a bite from an infected tick, there are a few other methods in which humans can come in contact with tick-borne diseases. Q fever, and tularemia can be acquired from direct contact, ingestion and aerosol exposure to wastes from animals previously infected by tick vectors, including rodents, rabbits and beavers. In hospital settings, Crimean-Congo Hemorrhagic Fever is a possible nosocomial disease, with infection occurring via aerosols or via direct contact with mucous, blood and/or saliva. Blood transfusions can contribute to Babesiosis, mostly in endemic areas and Colorado Tick Fever; however there is only one reported case (Goodman, Dennis et al. 2005). Maternal transmission of *Flaviviruses* can also occur

from an infected mother to progeny, when ovarioles are infected (Goodman, Dennis et al. 2005; Howley and Knipe 2007).

1.3.2 Tick Feeding Anatomy and Patterns in Virus Transmission

In order to understand the feeding and therefore disease transmission aspect, one must understand the tick mouthparts and how they function. Collectively, the mouthparts shown in Figure 1.3.1, the palps, chelicerae, hypostome and posterior basis capituli are called the Capitulum (Goodman, Dennis et al. 2005).

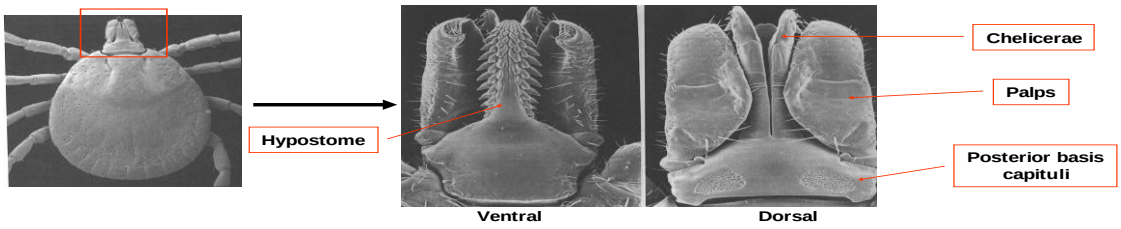


Figure 1.3.1: Structural feeding components from hard (*Ixodidae, spp.)* ticks (Goodman, Dennis et al. 2005). Figure referenced with permission granted from ASM Press.

The soft ticks have a similar capitulum structure but, the palps are shorter and a "hood-like" structure covers the capitulum when not in use (Goodman, Dennis et al. 2005). Looking at the capitulum structure from the outside in, the palps are described as "leg-like" structures that detect host chemicals, using chemosensory sensilla (hair-like epithelial sensors). In between the palps are the chelicerae, which protrude farther than the palps. These resemble appendages that can move and contain sharp teeth on the exterior. This feature resembles a saw and is responsible for cutting through the host's skin in preparation for attachment and feeding. The actual feeding aspect is enabled by the hypostome. This appendage is like a large tooth with two functions: a channel for blood intake and escape for saliva into the wound and surrounding tissues, and an anchor for attachment to begin the feeding process. The last part of the capitulum is the posterior basal capituli which serves as the bottom part of the mouth. Being flexible, this component allows for up and down movement of the capitulum (Goodman, Dennis et al. 2005).

Distribution of Flaviviruses is primarily by hard and soft ticks, although, tick and animal vectors for each virus species have not been identified. These combinations of tick and animal vectors are usually characteristic of a certain geographical region because the respective virus species is localized within that specific location (Goodman, Dennis et al. 2005; Howley and Knipe 2007). Such a dynamic relationship between vector, reservoir and virus, can lead to spreading or localization of certain pathogens. As an example, POWV originated in Eurasia and has its own combinations of vectors and reservoirs. POWV was thought to be recently introduced into North America, as its first isolation in the Western hemisphere occurred in 1958 (Howley and Knipe 2007; Hinten, Beckett et

al. 2008). POWV and its variant DTV (Deer Tick Virus) are widespread in South-eastern Canada and the North-eastern United States and have their own vectors and reservoirs (Goodman, Dennis et al. 2005; Hinten, Beckett et al. 2008). Additionally, POWV is currently the only known member of the TBE serocomplex present in North America (Howley and Knipe 2007).

Ticks are considered the most successful arthropods and their existence can be dated back to 200 million years ago. A tick may survive for approximately one year without food. Their only nutritional source is blood from feeding on small and large mammals such as cattle, monkeys and even humans. Two families or types of ticks exist, either "soft" or *Argasidae* and "hard" or *Ixodidae*. What separates the types is the presence of a "scutum" which is a hard-shell on the dorsal body (back-side) of the tick. The hard or *Ixodidae* family, have the most species implicated with disease transmission: *Dermacentor spp., Amblyomma spp., Haemaphylsalis spp., Hyalomma spp.* and *Rhipicephalus spp*. The *Ixodidae* ticks have three distinct stages of their life cycle, with each stage being dictated by blood meals. Throughout the tick life cycle, larvae, nymphs and adults look the same except, progressively smaller between molts and larvae have three pairs of legs not four. Beginning with the larval stage, once the egg hatches, tick larvae will feed, detach and seek shelter to molt into the next aspect of the life cycle. *Ixodidae* ticks have an interesting method of attaching to their unsuspecting prey in order to feed. While laying low in grasses, they will extend their legs outwards. The legs have tiny barbs that will snag when a potential host grazes past and the tick can begin with feeding. The stage-dependant size differences in the ticks, only slightly modifies their attack. The larvae usually remain low on vegetation in order to attach to smaller animal

prey, the larger nymph and adult ticks can climb higher to access larger animals. After feeding and subsequent molting, the nymph will repeat the cycle of feeding, hiding and molting, allowing the tick to progress into the adult phase of their life cycle. The adult stage is where feeding, mating and egg laying will occur.

Establishment and the maintenance of pathogens between tick and vectors can be described as involving three steps. A tick must become infected by feeding on a host that has viremia high enough to transfer to the tick during feeding. Ticks, once infective, can transmit the pathogen to another host or to other ticks during a co-feeding event. In the latter case, the uninfected tick must be feeding in the same vicinity as the infective tick (Goodman, Dennis et al. 2005; Howley and Knipe 2007). Finally, there must be an adequate number of susceptible reservoirs for infection and vectors for delivery of the pathogen, in order for the tick-reservoir cycle to be maintained (Goodman, Dennis et al. 2005). Three types of transmission help to maintain pathogen levels. Horizontal transmission is defined by repeated cycles of tick-reservoir transmission and the animals that aid in maintenance may also be called amplifying hosts (Goodman, Dennis et al. 2005; Artsob and Lindsay 2008). Once infected, ticks may remain so for the remainder of their lives, despite molting after a blood meal. Thus, in each stage of tick development the pathogen remains. This is referred to as "transstadial transmission" (Goodman, Dennis et al. 2005). Vertical transmission from mother to progeny, termed "transovarial transmission" can also occur, if the ovarioles of the female tick are infected (Goodman, Dennis et al. 2005; Howley and Knipe 2007). Tangential transmission is defined by the breaking of the pathogen's maintenance cycle (Goodman, Dennis et al. 2005), in which reservoirs are unable to become infected or are non-susceptible hosts. In more detail, the

host either cannot become infected or cannot sustain adequate levels of viremia to infect other arthropods. These types of hosts are synonymously termed: Zooprophylactic, incidental and/or dead-end hosts (Goodman, Dennis et al. 2005; Artsob and Lindsay 2008). The ability of a pathogen to be transmitted to a new host, vector or reservoir, can be altered due to differences in climate. In wet tropical climates the complete life cycle may occur many times in a single year. And in more temperate to sub-arctic climates, the cycle can take two or more years to complete. In India, where KFDV is present, the climate of Karnataka's endemic region has alternating wet and dry seasons (Goodman, Dennis et al. 2005; Pattnaik 2006).

1.3.3 Ticks and Transmission of KFDV, OHFV and AHFV

Kyasanur Forest Disease Virus (KFDV) is one of many "arboviruses". Arboviruses are defined as, "arthropod-borne viruses". All viruses that fall into this category are able to be transmitted to and from vertebrate hosts and arthropods requiring blood meals, also known as "hematophagus arthropods" (Artsob and Lindsay 2008). In 1898-1901, Yellow Fever Virus, a Flavivirus, was the first arbovirus identified. This was based on the hypothesis of a Cuban physician, Dr. Carlos Finlay, who believed that Yellow Fever Virus was transmitted to humans by mosquitoes and not acquired from filth (Artsob and Lindsay 2008). In order for Arboviruses to be sustained in an environment, three criteria must be met. Firstly, horizontal transmission with a plentiful amount of vertebrate hosts or "amplifying hosts" or reservoirs for the virus is required. Such vertebrate hosts can include small to large mammals, birds and reptiles. Secondly, amplification of the arbovirus must reach high enough titres, in order to be acquired by

the vector (tick or mosquito). Lastly, the virus must be able to replicate in the vector. This replication must lead to transmission through the saliva or from leakage of a previous blood meal into the new host or reservoir (Artsob and Lindsay 2008).

Kyasanur Forest Disease Virus has been successfully isolated from 16 different Ixodes ticks, including species from the *Haemaphysalis* and *Dermacento*r genera (Pattnaik 2006). However, in a study of an *Argasidae* (soft-tick) family, KFDV was successfully isolated from *Ornithodoros chiropterphila* (Bhat and Goverdhan 1973). May this be a 17th type of KDFV-transmitting tick? (Charrel, Fagbo et al. 2007). Bhat et al. in 1973 had discovered that *Ornithodoros crossi* was a capable vector of KFDV during the nymphal phase. Despite the maintenance of the virus through molting to the adult phase; however, transmission rates declined in the adult ticks (Bhat and Goverdhan 1973). Similarly, another soft-tick, *Ornithodoros tholozani,* has been documented as a secondary vector for KFDV transmission (Goodman, Dennis et al. 2005). The ramifications of such a discovery mean that the hard ticks (*Ixodidae*) may not be the only type of vector for KFDV. Perhaps the soft ticks (*Argasidae*) may need to be further investigated for their potential KFDV associations. Several animal reservoirs have been suggested for KFDV including birds, shrews, field mice, forest rats, Indian crested porcupines, bats, giant flying squirrels, black-naped hares (Goodman, Dennis et al. 2005; Pattnaik 2006). These potential reservoirs are typically infected during the dry season (October-December), as tick eggs hatch from and drop from the underside of tree leaves. During the second part of the dry season (January-May), after the fed larvae molt into nymphs, the same reservoir animals are again subject to feeding. The next molting event allows for the maturation of adult ticks. It is in this wet/monsoon season (June-September) that the ticks

feed on larger animal reservoirs, like cattle, and begin to procreate and females lay their eggs. The constant cycles help to sustain KFDV in this forested locale. Even though most of these animals are reservoirs and have no symptoms, neutralizing antibodies can be found in their blood. Monkeys are always present in these forested areas and can become infected at any stage of the ticks' life cycle. However, monkeys are dead-end hosts of KFDV and may succumb to their illness. Similarly, humans are also dead-end hosts and are susceptible to KFDV. Humans visit the forested areas frequently for rice paddy harvesting and wood collection during the months of November-December and December-May, respectively. It is during these dry season months (November-May) that the highest incidences of KFDV infections occur in humans (Pattnaik 2006).

Besides maintaining the tick-reservoir cycle of KFDV, animals may contribute to human transmission via bite, consumption of infected animal products (milk and meat) or contact with infected blood (Howley and Knipe 2007). Infection through the consumption of infected milk has been demonstrated with many tick-borne Flaviviruses and primates including: OHFV, LIV, POWV, AHFV and KFDV (Goodman, Dennis et al. 2005; Pattnaik 2006; Howley and Knipe 2007).

1.4 Flavivirus Structure

Morphologically, Flaviviruses are quite complex in the structure of their virion, otherwise known as an infectious particle, which is capable of initiating a new infection. When the virus is a mature infectious particle, it is characterized as a dense icosahedral nucleocapsid (~35-50 nm) that houses the single-stranded (ss) RNA genome of positive-sensed polarity and is surrounded by a lipid bilayer envelope carrying two surface

proteins. The M (Membrane) and E (Envelope) proteins are associated with each other in tightly-packed dimers that lay parallel to the envelope. With the use of electron microscopy (EM), the virion was shown to be smooth-looking and has a “golf-ball-like” appearance (Kuhn, Zhang et al. 2002; Gehrke, Ecker et al. 2003). Unlike other viruses, like HIV (Human Immunodeficiency Virus) and Influenza virus, there are no spike proteins or protein projections (Kuhn, Zhang et al. 2002; Howley and Knipe 2007).

Flaviviruses have a positive-stranded RNA genome of about 10-11 kb in length. Since the genome is positive-stranded, it can be translated immediately as it is equivalent to messenger RNA (mRNA). Translation of the genome creates a long polyprotein that is then co and post-translationally cleaved into 10 proteins by cellular and viral proteases. The 5' end of the genome encodes for three structural proteins and seven non-structural proteins follow to the 3' end. The polyprotein is organized in the following way: C-prM-E-NS1-NS2A-NS2B-NS3-NS4A-NS4B-NS5, Figure 1.5.3 (Harris, Holden et al. 2006; Howley and Knipe 2007). The structural proteins include the C (capsid) protein, which is 11 kDa and is the only protein that makes up the nucleocapsid. The other two structural proteins are the only two envelope-associated proteins. prM (pre-membrane) is a glycosylated precursor-protein of about 27 kDa, which is then cleaved during the assembly/release stage by a host cellular protease called furin. Cleavage provides an unglycosylated M (membrane) protein of approximately 8 kDa. The second membrane protein is the glycosylated 54-kDa, E (envelope) protein. The protein is composed of three domains (I, II and III), the second domain contains the fusion peptide (FP), visually represented in Figures: 4B and 4C. The third domain is thought to be responsible for

virus attachment and entrance via a fusion mechanism that involves the FP (Gehrke, Ecker et al. 2003; Harris, Holden et al. 2006; Howley and Knipe 2007).

There are seven non-structural proteins, NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5. The NS1 protein is described as a cell-associated or non-virion associated glycoprotein, which can form dimers and become secreted into the extracellular environment. The exact roles of this protein still remain speculative. Some believe that it aids in viral pathogenesis and replication (Howley and Knipe 2007). Others believe, when secreted, NS1 may promote antibody production and direct complement-mediated lysis to those cells infected with Flaviviruses aiding in viral spread and disease severity. When secreted, NS1 binds to the virus-infected cell and reduces the formation of the Membrane Attack Complex (MAC). In addition, NS1 may bind to the complement system regulator complement factor H. This can negatively regulate the complement system and inhibit lysis of virus-infected cells (Krishna, Rangappa et al. 2009). NS1 has also been implicated in increased complement-mediated anapylatoxin release. This may result in more severe disease progression of DV, specifically Dengue Hemorrhagic Fever symptoms (Avirutnan, Mehlhop et al. 2008). Another thought is that this protein can mimic human molecules (Howley and Knipe 2007)**.** Host protein mimicry by NS1 may be a reason for Dengue Hemorrhagic Fever disease severity in DV infections. A study by Chang et al. 2002 (Chang, Shyu et al. 2002), demonstrated that NS1 of DV shares structural similarity with an RGD peptide motif which is important for matrix-integrin cell adhesion. Antibodies generated against NS1 were shown to be cross-reactive against a RGD peptide and fibrinogen (Falconar 1997) and therefore, may negatively impact

matrix proteins in blood vessels and blood clotting (Falconar 1997; Chang, Shyu et al. 2002).

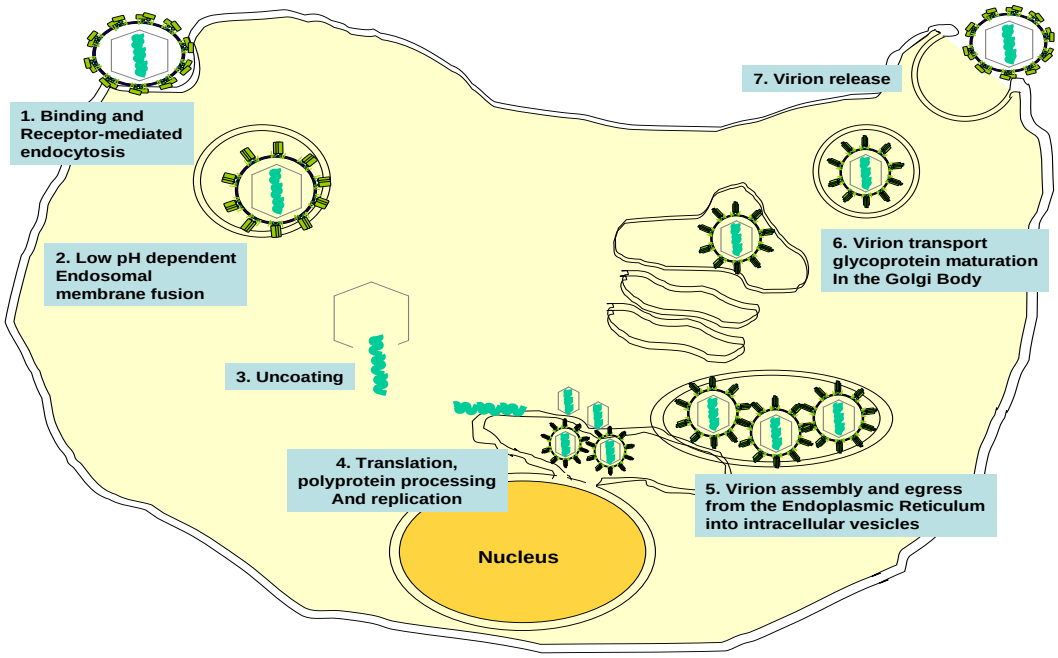
The NS2A protein is a hydrophobic protein that is cleaved from NS1 by an unknown enzyme, which may be present in the endoplasmic reticulum. Mutational analyses have shown that NS2A is critical for viral assembly and can block interferon (IFN) responses (Pastorino, Peyrefitte et al. 2006; Leung, Pijlman et al. 2008). NS2B forms a complex with NS3, to form the Viral Serine Protease. The role of NS2B is to be a co-factor for NS3 and it anchors the complex, using hydrophobic amino acids, into the membrane of the endoplasmic reticulum (Schrauf, Schlick et al. 2008). NS3 is a viral serine protease, which has homology with trypsin (Schrauf, Schlick et al. 2008). The N-terminus of NS3 is responsible for the cleavage of NS2A/NS2B, NS2B/NS3, NS3/NS4A and NS4B/NS5 protein junctions, as long as it has the NS2B co-factor to form the functional viral-serine protease (Pastorino, Peyrefitte et al. 2006). Within the polyprotein, the NS2B-NS3 junction and the C-terminus of NS3 sites undergo an auto-catalytic event to free the protease and allow for its activity (Luo, Xu et al. 2008). Previous studies have compared the viral serine protease target sites in Dengue virus and West Nile virus polyprotein. These have a similar sequence motif and are highly conserved for all Flaviviruses (Schrauf, Schlick et al. 2008). The conserved sequence motif is denoted by Lysine-Arginine (K, R) or Arginine-Arginine-(Glycine, Serine, Alanine or Threonine) (R, R, X, where X is G, S A or T). However, at the C proteins' C-terminal cleavage site, the sequence is Lysine / Arginine-Arginine-Glycine-Lysine-Arginine-Arginine-Serine, in which, the two underlined sites may be recognition sites for the NS2B-NS3 complex (Grard, Moureau et al. 2007; Schrauf, Schlick et al. 2008). Other known activities of NS3

are on the C-terminus of the protein: an RNA helicase/NTPase (nucleoside tri-phosphatase) activity for RNA template unwinding by hydrolysis, and an RTPase (RNA tri-phosphatase) action, involved in addition of a cap on the 5' end of the genomic RNA.

NS4A and NS4B are both hydrophobic proteins and have proposed roles in IFN inhibition, with NS4B being the stronger antagonist. Both proteins may be involved in the replication cycle. NS4A can perform membrane re-arrangements, once cleaved from the polyprotein by a currently unknown enzyme. NS4B is associated with the double-stranded RNA portion of the replication cycle. NS5 has two vital roles for Flaviviruses. The carboxyl terminus of NS5 is an RNA-dependent RNA polymerase essential for viral genome replication. In addition, the amino terminus of NS5 is an MTPase (methyl-transferase). Methyl-transferase activity is responsible for the methylation on the 5' RNA cap at two separate sites (Gehrke, Ecker et al. 2003; Howley and Knipe 2007). The amino terminus has also been implicated as a guanyltransferase. This activity results in the transfer of a Guanosine mono-phosphate (GMP) moiety to the RNA and providing the RNA cap (Issur, Geiss et al. 2009).

Flanking the coding region of the Flavivirus genome are the 5' UTR and the 3' UTR. For individual Flaviviruses, the 5' UTR is essentially similar in secondary structure but, different in sequence. 5' UTRs range between 95-132 bases long, with MB viruses having shorter 5' UTRs than in TB viruses (Proutski, Gould et al. 1997; Villordo and Gamarnik 2009). However, the 3' UTR is between 400-700 nucleotides in length but, quite different in sequence for MB and TB subgroups (Howley and Knipe 2007).

1.5 Flavivirus Life Cycle



.

Figure 1.5.1: Flavivirus life cycle. Adapted from (Howley and Knipe 2007; Samsa, Mondotte et al. 2009; Mahy and Van Regenmortel 2010).

1.5.1 Viral entry

The Flavivirus life cycle begins with attachment and fusion of the E (envelope) protein with an unknown cellular receptor(s) and the virion enters the host cell in endosomes, through a process termed receptor-mediated endocytosis (Howley and Knipe 2007; LaFemina 2009; Hanley and Weaver 2010).

Once the virus is introduced into the host cell, fusion of viral and endosomal membranes occur by a pH-dependent mechanism. Consequently, the nucleocapsid is released into the cytoplasm (Harris, Holden et al. 2006). Class II membrane fusion proteins are responsible for membrane-membrane fusion while in endosomes to allow Flaviviruses genomic RNA to enter the host's cytoplasm. Flavivirus-endosome membrane fusion occurs in a pH-dependant process, at an acidic pH of around 5.4 (Stiasny, Kossl et al. 2007). Attempts have been made to assess the necessity of acidic pH for fusion by Stiasny et al. in 2007 (Stiasny, Kossl et al. 2007). Tick-Borne Encephalitis Virus (TBEV) particles and liposomes were incubated at an acidic pH of 5.4 and an alkaline pH of 10.0 to see if this would allow for fusion. Viewing the mixtures with an electron microscope revealed binding of E protein to the liposomal membrane at both pH levels. However, fusion and an "electron dense particle" being inserted into the liposome were seen, only at pH 5.4. The authors believe the "electron dense particle" is the TBEV nucleocapsid (Stiasny, Kossl et al. 2007). Another study attempted to see if the Class II fusion proteins required a certain lipid composition in the host membrane for fusion. Lipid composition made a difference only at 4^{o}C; in the absence of cholesterol in the target membrane fusion did not occur. This result is not relevant to a real life infection, as the optimal conditions require temperatures of at least 37 ^{o}C. The authors do suggest that

Flaviviruses require cholesterol in the target membrane for fusion (Stiasny, Koessl et al. 2003). Futhermore, cholesterol may enhance the ability of virus-membrane fusion (Kielian 2006). Lipid rafts or microdomains with cholesterol have been suggested to be required for fusion mechanisms (fusion, post fusion stability and membrane curvature) for other viral families such as Herpesviruses, Filoviruses and Coronaviruses (Teissier and Pecheur 2007).

Membrane-membrane fusion inside the endosome is obtained through the fusion peptide (FP), which is sequestered in the E protein dimer (Kielian 2006). Figure 1.5.2.A depicts the mature virion with the E protein as a dimer that is responsible for both the binding to a receptor and fusion properties. On the virion surface and lying parallel to the lipid envelope, the E protein forms a network or "glycoprotein cage" of 90 E protein dimers surrounding the icosahedral nucleocapsid (Howley and Knipe 2007). A drop in pH causes the dimers to dissociate into monomers that now have FP exposed called pre-fusion or monomeric intermediate (Kielian 2006; Stiasny, Kossl et al. 2007). The FPs, are "spike-like" projections which make contact with the host membrane. Homotrimerization occurs from low pH and cholesterol forming the "post-fusion" trimer conformation. Then five to six homotrimers interact to form a ring-like structure on the target membrane with the FPs inserted (Kielian 2006). For fusion to occur domain III of E must re-orient towards the FP on domain II, to drive the viral-membrane fusion event, demonstrated in Figure 1.5.2.C (Kielian 2006; Stiasny, Kossl et al. 2007). Following fusion the nucleocapsid enters the cytoplasm through the "fusion pore", where uncoating begins to release the genomic RNA, as summarized in Figure 1.5.2.D (Howley and Knipe 2007; Stiasny, Kossl et al. 2007; Fritz, Stiasny et al. 2008). The fusion mechanism has been

attributed to histidine residues in the E protein. The acidic conditions will cause protonation of Histidine 323 which acts as a bridge between domains I and II of the E protein. Protonated Histidine 323 is proposed to break a salt bridge between Glutamate 373 (domain III) and Arginine 9 (domain I) and a subsequent salt bridge with Histidine 323 and Glutamate 373 is created. This allows for exposure of the FP and permits formation of the post-fusion trimer (Fritz, Stiasny et al. 2008).

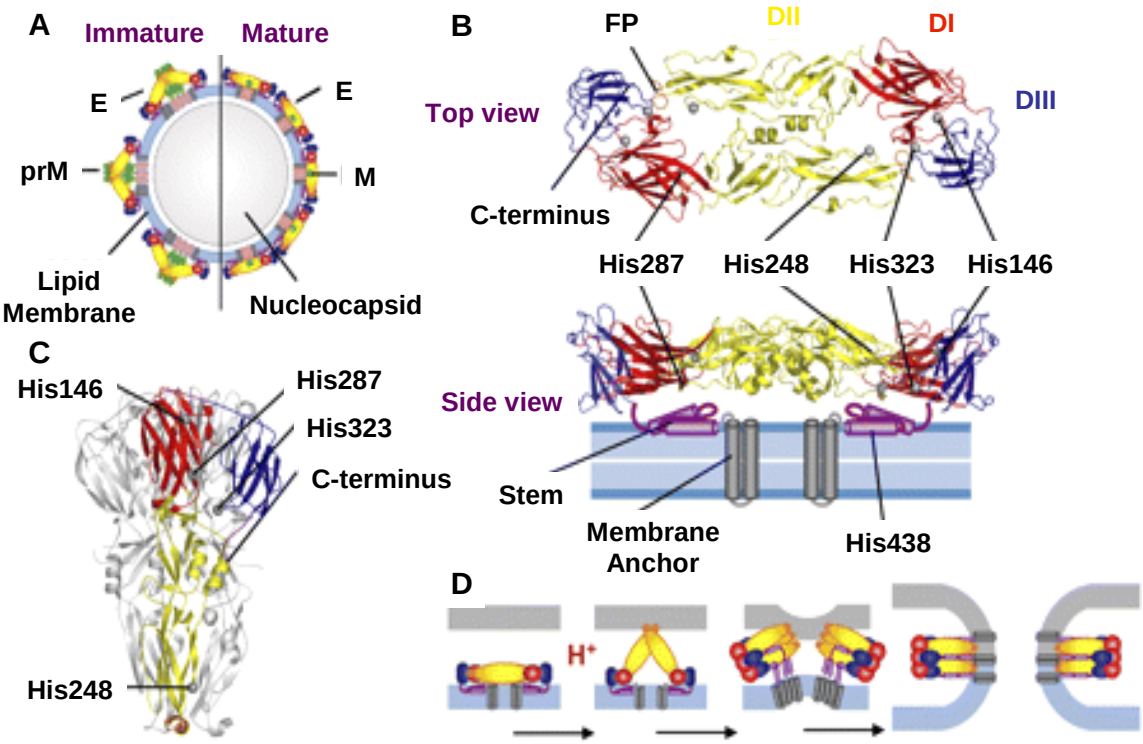


Figure 1.5.2: Flavivirus pH-dependent fusion components and mechanism.

A) Immature virion with prM protein covering the fusion peptide. Mature virion with fusion peptide is hidden in between domain I and III of the neighboring E protein. B) E protein dimer with histidine residue His323 as the pH indicator and the other residues which are required for the structural integrity of the post-fusion trimer. C) Post-fusion trimer with fusion peptide inserted into the host membrane. D) Flavivirus fusion mechanism. Nucleocapsid: off-white, Lipid membrane / envelope: light blue, E protein domains: I, II and III are red, yellow and dark blue respectively, Fusion peptide: orange, Stem of E protein: purple, Transmembrane domains in E protein: dark grey and Host membrane: light gray (Fritz, Stiasny et al. 2008). Figure referenced with permission granted from Karin Stiasny.

1.5.2 Viral translation and replication

The positive-stranded RNA genome serves as a template for genome replication and translation of viral proteins. Replication occurs within cytoplasmic replication complexes (RC) near perinuclear membranes and translation occurs nearby at the rough endoplasmic reticulum (ER). There are two important events that occur during replication, using the positive-strand RNA as a template for negative-strand RNA synthesis (anti-genomic RNA) and negative-strand RNA being replicated back into positive-strand RNA. This is done in an asymmetric manner, meaning positive-stranded RNA is generated in a higher proportion than that of negative-stranded RNA. Importantly, the negative-stranded anti-genome is vital to these viruses in their life cycle. Anti-genomic RNA serves as the precursor for positive RNA synthesis, which is needed for genomes for progeny viruses and for translation to generate viral proteins (Howley and Knipe 2007; Villordo and Gamarnik 2009).

Flavivirus translation and replication are thought to be coupled. The RC is further defined to include NS3 and NS2A bound to NS5, which assemble during translation of the genome. In addition, other NS proteins and possibly some host factors may also be part of the RC complex (Howley and Knipe 2007; Villordo and Gamarnik 2009). Another study indicates that the RC consists of the NS1, NS3, NS5, NS2A and NS4A proteins (Khromykh, Sedlak et al. 2000). Despite the apparent coupling of translation and RNA synthesis, there is coordination of each event so that they do not interfere with each other. Translation happens first, followed by the formation of the RC and subsequent genome replication (Harris, Holden et al. 2006). A short hairpin region that is conserved within

MB and TB Flaviviruses located in the C gene provides a pausing for ribosomes to allow for an authentic AUG (start codon) recognition. This hairpin, designated cHP, is slightly downstream (12-16 nucleotides) of the authentic AUG (start codon) due to the secondary structure. As the ribosomes scan the RNA it pauses at the hairpin. Unwinding of the hairpin allows the recognition of the authentic start codon, instead of another downstream start codon that would truncate the polyprotein. Thus, the cHP may prevent a "poor initiation context" or leaky ribosome scanning (Davies and Kaufman 1992; Harris, Holden et al. 2006). Once translation is completed the RC complex binds to ~100 conserved nucleotides on the 3' UTR, termed the 3' SL (stem loop) and is shuttled to cellular membranes or perinuclear membranes, which are both proposed to be sites for replication (Harris, Holden et al. 2006; Howley and Knipe 2007; Villordo and Gamarnik 2009).

The 5' and 3' UTR regions are considered to be very important for RNA synthesis, translation and the regulation of both events, which was found experimentally with West Nile Virus, Dengue Virus, Japanese Encephalitis Virus and Kunjin Virus (Harris, Holden et al. 2006). Mutations in both 5' and 3' UTR of Dengue Virus have led to reductions in viral replication, nerovirulence and in viremia for cell culture, mice and monkeys, respectively (Harris, Holden et al. 2006). Translation and replication occur in different directions: translation being a 5' to 3' movement of the ribosome, and replication involving 3' to 5' movement of the RC. Translation precedes replication, as cyclization of the UTRs occurs through conserved complementary sequences (CS). Since these processes are in opposite ends of the genome, the collision of translational machinery and RC would be possible. It has been suggested that some viral and cellular factors utilize

the 5' and 3' UTRs to control such a catastrophe (Harris, Holden et al. 2006). In addition to the 3' SL, the CSs have two sets of hybridization sequences (Harris, Holden et al. 2006). The 5' CS is located in the C gene about 40 nucleotides past the start codon of the polyprotein. The binding partner, 3' CS, is slightly upstream of the 3' SL and mutations in either of these binding partners crippled replication but, replication was rescued with restoration of a new set of base pairing sequences, different from the CS sequences. This study demonstrated that cyclization was vital for replication and was not sequence-dependent, for Kunjin virus, West Nile Virus, Dengue Virus and Yellow Fever viruses (Kofler, Hoenninger et al. 2006; Villordo and Gamarnik 2009). A second set of complementary sequences was also found to be necessary for replication in DV and WNV: the 5' UAR (Upstream AUG Region) and 3' UAR within the stem-region of the 3' SL (Villordo and Gamarnik 2009).

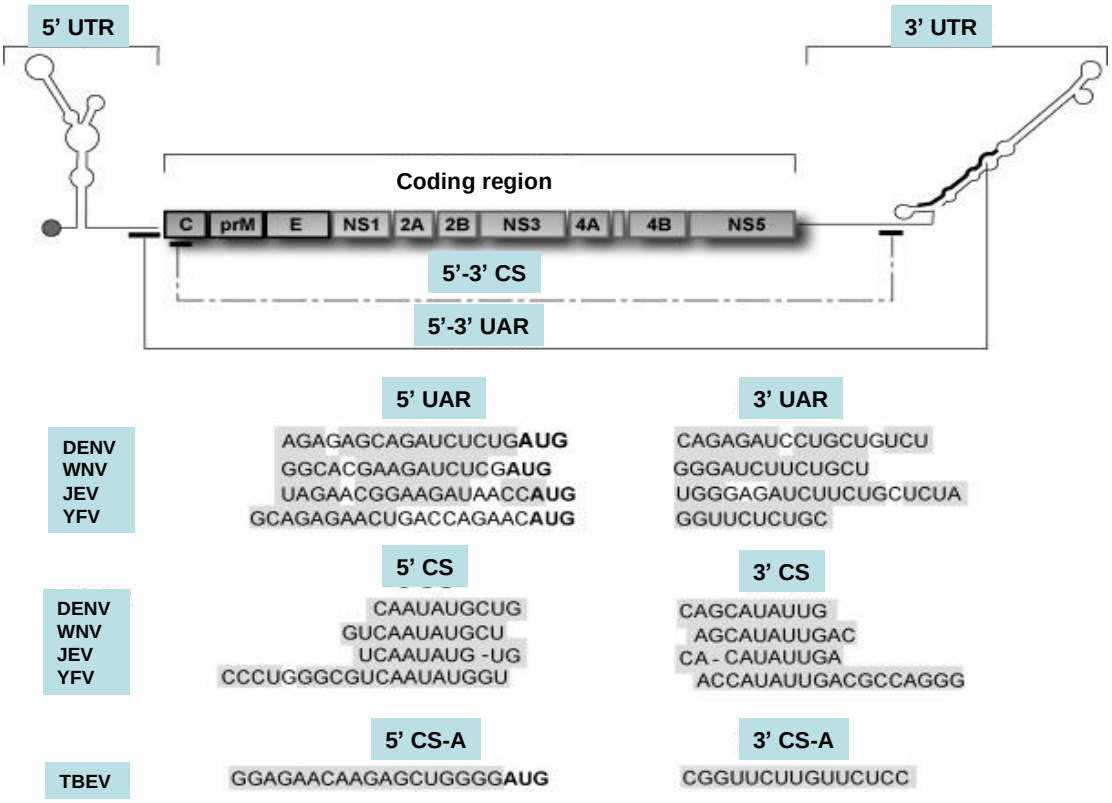


Figure 1.5.3: Flavivirus genome featuring the 5' and 3' UAR and CS sequences in both MB and TB virus groups.

DENV (Dengue Virus), WNV (West Nile Virus), JEV (Japanese Encephalitis), YFV (Yellow Fever Virus), TBEV (Tick-Borne Encephalitis Virus) and the authentic start codon of the polyprotein is indicated in bold (**AUG**) (Villordo and Gamarnik 2009). Figure referenced with permission granted from Elsevier Limited.

An exception to 5' CS and 3' CS binding is found in the Tick-Borne Flaviviruses, which do contain these sequences, but they are slightly different in sequence and are not functional. In the Tick-Borne Encephalitis Virus (TBEV) genome, both 5' and 3' sequences are complementary inverted repeats, called 5'-CS-A and 3'-CS-A. The 5' UTR contains the 5'-CS-A and within the 3' SL of the 3' UTR is the 3'-CS-A, similar in function to the 5' and 3' UAR regions (Kofler, Hoenninger et al. 2006). The CS-A regions are highly conserved in Tick-Borne but not in Mosquito-Borne Flaviviruses. However, the CS motifs for the Mosquito-Borne group have similar positions in the Tick-Borne group and they are designated 5'-CS-B and 3'-CS-B. Unlike the CS-A regions, the CS-B regions were both found to be not essential for replication, but further studies are needed to define their exact role(s) (Kofler, Hoenninger et al. 2006).

Further addition to the model for replication and the role of genome cyclization in Flaviviruses has been proposed by Villordo et al. in 2009 (Villordo and Gamarnik 2009). For the Dengue Virus system, it is proposed that NS5 binds to the 5' UTR at the SLA (Stem Loop A, containing a promoter region). Genome cyclization (5' – 3' UAR and 5' – 3' CS in Mosquito-Borne viruses and 5'-3' CS-A and possibly, 5'-3' CS-B in Tick-Borne viruses) allows for replication from the 3' UTR to the 5' UTR generating a negative-stranded RNA, depicted in Figure 1.5.4. Cyclization is thought to be necessary for positive-strand synthesis from the negative-strand template. However, this still remains to be elucidated (Harris, Holden et al. 2006; Villordo and Gamarnik 2009).

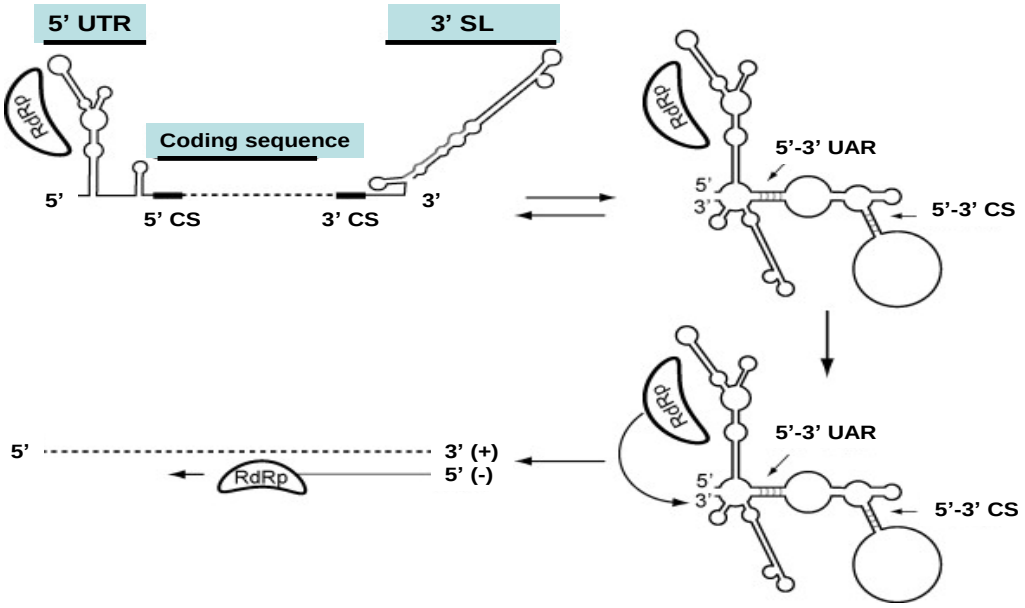


Figure 1.5.4: Proposed replication model for DV.

The RNA dependent RNA polymerase (RdRp) binds to the SLA region of the 5' UTR and begins RNA synthesis from the 3' UTR through to the 5' UTR (Villordo and Gamarnik 2009). Figure referenced with permission granted from Elsevier Limited.

NS5 binding to the SLA is thought to bring the polymerase-promoter complex to the initiation of replication site within the base of the 3' SL. Unwinding of the 3' SL base region is a requirement for initiation, perhaps this unwinding is aided by the helicase/NTPase activity of NS3 (Villordo and Gamarnik 2009). Using the DV genome, the SLA was determined to be the promoter element for the RdRp. They divided the DV genome into two segments, one with the 5' UTR-RNA and RNA-3' UTR and quantitatively determining RNA synthesis. The 5' UTR was found to promote RNA synthesis from the DV NS5 (RdRp). They also performed mutational studies within the 5' UTR region and assessed that the SLA region promoted a 10-fold increase in RNA replication (Filomatori, Lodeiro et al. 2006).

This functional study suggests that the promoter region for the NS5 polymerase is located within the 5' UTRs SLA or SL1. Additionally, the SLAs for all Flaviviruses could possibly be interchangeable (Villordo and Gamarnik 2009). Cyclization of the complementary ends of viral genomes generates what have been called, "panhandle-like" structures and the process is similar to some negative-strand virus replication models in the *Bunyaviridae*, *Arenaviridae* and *Orthomyxoviridae* (Villordo and Gamarnik 2009).

Since RNA replication and methylation of the 5' RNA cap are carried out by the same protein, NS5, it has been postulated that replication and capping may be coupled (Dong, Ray et al. 2007). In general, capping of RNA happens on cellular transcripts (post-translational modification) and viral genomes. In both cases the cap adds stability to the transcript and is used for ribosome recognition of the transcript (Kapp and Lorsch 2004; Howley and Knipe 2007). A cap is a methylated guanosine nucleoside, meaning a guanine (nitrogenous base) is attached to a ribose sugar at the first labeled carbon (C1) of

the ribose sugar (Bauman 2009). There are three types of caps described: Type 0, Type 1 and Type 2. The types are distinguished by the number of methylation events. Type 0 caps are generated by the removal a phosphate group of the first nucleotide on the pre-mRNA by an RNA triphosphatase (RTPase) and the addition of a guanosine monophosphate (GMP) by an RNA guanyltransferase to generate a 5'-5' triphosphate. RNA guanine-methyltransferase adds a methyl group to the N-7 position of the base. The type 0 cap is designated as m^7GpppN, where m = methylation on position 7, G = guanine base and pppN = the first nucleoside triphosphate (NTP) of the RNA transcript (Dong, Ray et al. 2007). Type 1 and type 2 caps are defined as having the first and second NTP methylated on the 2' OH of the ribose sugar. Type 1 and type 2 caps are denoted as m^7GpppNm and m^7GpppNmNm, respectively (Dong, Ray et al. 2007). Flavivirus RNAs have a Type 1 cap, whose addition is facilitated by two viral proteins, NS3 and NS5. The first step of capping is performed by the NS3 protein. In addition to the helicase and serine protease activities, it has an RTPase function that removes one of the three phosphates on the 5' end of the nascent RNA genome. The second step involves the addition of a guanosine mono-phosphate (GMP) cap from guanosine tri-phosphate (GTP) nucleotide (Li, Clum et al. 1999; Yon, Teramoto et al. 2005; Dong, Ray et al. 2007; Howley and Knipe 2007). Recently, the NS5 protein has been demonstrated to have guanylyltransferase activity (Issur, Geiss et al. 2009). Two methylation events are performed by the virally-encoded MTPase (Methyl-transferase). This occurs first on N-7 and then on the 2'-OH on the ribose of the guanosine cap structure, creating a type 1 cap structure. In a study by Dong et al. 2007, four different Flaviviruses were studied: West Nile Virus, Dengue Virus-1, Yellow Fever Virus and Powassan Virus, to show that

methylation is dependent on the 5' RNA sequence (Dong, Ray et al. 2007). This situation is that for cellular mRNA, in which capping works in a sequence-independent manner. Furthermore, Dong et al. 2007, demonstrated that blocking of capping by NS5 prevented replication of West Nile Virus in cell culture. There was no affect on the cell culture in which the virus was tested. With that established, inhibition of virus cap methylation events may be a future direction for antiviral therapies against Flaviviruses (Dong, Ray et al. 2007).

1.5.3 Virus assembly and release

Packaging of nascent genomes into the capsids to generate progeny viruses also occurs within the replication complex (Howley and Knipe 2007; Denison 2008; LaFemina 2009). After replication and translation, the assembly phase is characterized by the association of nascent genomic RNA with the C (capsid) protein by an unknown mechanism (Samsa, Mondotte et al. 2009; Villordo and Gamarnik 2009). There are two different proposals for this unknown mechanism using ER-derived organelles called lipid droplets (LD). These organelles have been shown to become surrounded by the C protein in DENV virus-infected cells. At some later point in the infection, these LDs transport to the ER and interact with the nascent viral RNA to begin virus assembly. Another suggestion is that nucleocapsid formation occurs on the surface of the LDs preceding virion morphogenesis (Samsa, Mondotte et al. 2009). During budding through the ER lumen, the lipid bilayer, prM and E proteins are acquired and KFDV is transported out of the plasma membrane through vesicle packets or double membrane vesicles. Exocytosis from the host cell occurs as the packets pass through the Golgi apparatus (Harris, Holden et al. 2006; Denison 2008; LaFemina 2009; Villordo and Gamarnik 2009). Within the

secretory pathway, an important cleavage of prM to M by furin occurs before the viruses are released into the extracellular environment (Kuhn, Zhang et al. 2002). Uncleaved "immature" viruses have a rough and jagged appearance resulting from 60 prM-E monomers projections clustered in triplets. Cleavage of the pr component of M defines the virus as "mature" and M-E homodimers re-orient parallel to the lipid bilayer into 90 head to tail dimers, implying the dimers are infectious in susceptible cells (Howley and Knipe 2007; Villordo and Gamarnik 2009). The FP on the E protein is covered by pr; this pr cap-like structure is thought to prevent premature fusion with the host membrane, until the virus is released from the host (Perera and Kuhn 2008).

1.6 Reverse Genetic Systems (RGS)

1.6.1. Infectious Clone Systems for Analyzing Viral Replication

The first infectious clone systems for RNA viruses was developed for the positive-stranded Poliovirus, and paved the way for other positive-stranded viruses (Ebihara, Groseth et al. 2005). The first Flaviviruses infectious clone system generated infectious Yellow Fever viruses (YFV) entirely from cloned cDNA in 1989 (Rice, Grakoui et al. 1989). Briefly, this system transcribed infectious Yellow Fever Virus RNA from a cDNA template, *in vitro* and the RNA was then transfected into mammalian cells. After transfection, the mammalian host cells would allow the infectious RNA to act as a natural YFV infection. The cDNA template of YFV was from the digestion of three fragments cloned separately into λ (Lambda) phage vectors. All three fragments were ligated together using unique restriction sites and the resulting full-length YF cDNA genome was then transcribed using SP6 DNA-dependent RNA polymerase and

transfected into confluent BHK (Baby Hamster Kidney) cell culture monolayers. The authors also added a methylated cap analog, mimicking a natural Flavivirus cap, into the transcription reaction. Results indicated that the infectious clone system gave titres equivalent to the original YF17D virus and that capping had improved infectivity of the RNA. Infectivity was calculated as plaque forming units (PFU) per microgram of transfected RNA. Capped RNA gave 110 PFU / μg versus the un-capped RNA, which produced <10 PFU / μg (Rice, Grakoui et al. 1989). Two years later, a similar experiment using Dengue Virus type 4 was performed (Lai, Zhao et al. 1991) and the salient points are summarized in Figure 1.6.1. The authors combined two fragments, using a unique restriction site. The 5' and 3' fragments, when ligated together, generated the full-length DV-4 cDNA genome. Two silent mutations were incorporated in the cDNA to generate a unique restriction site. Similar to the previous YFV infectious clone system, transcription was performed with SP6 DNA-dependent RNA polymerase and a methylated cap analogue. The following transfection procedure was performed with both the mutated RNA and wild type DV RNA separately, as a control. To confirm the progeny viruses were from the transcribed cDNA clone, progeny virus genomic RNA was extracted and RT-PCR (Reverse Transcription Polymerase Chain Reaction) was performed. Subsequent digestion of the unique restriction site created by the silent mutations, confirmed that infectious RNA was generated from the cDNA construct. Furthermore, Western blotting confirmed the presence of viral proteins: prM, E, NS1 and NS3 from the progeny viruses following transfection with RNA (Lai, Zhao et al. 1991). Once the first infectious clone system was successful, many more clone systems for Flaviviruses soon followed including Japanese Encephalitis Virus (Sumiyoshi, Hoke et al. 1992; Yun, Kim et al.

2003)), Dengue Virus-2 (Kapoor, Zhang et al. 1995; Kinney, Butrapet et al. 1997; Polo, Ketner et al. 1997; Gualano, Pryor et al. 1998)), Kunjin Virus (Khromykh and Westaway 1994)) and Tick-Borne Encephalitis Virus (Mandl, Ecker et al. 1997)). Problems with stability of the full-length genomes in cDNA format in *Escherichia coli* cloning strains were reported for some viruses including Yellow Fever, Japanese Encephalitis Virus and Dengue virus-2 strain New Guinea C (Hurrelbrink, Nestorowicz et al. 1999). The genetic instability gave rise to non-sense mutations in the cDNA and also gave low DNA yield during harvesting from *E. coli* (Yun, Kim et al. 2003). Rice et al. 1989 reported poor vector harvesting and deleterious effects were seen in their *E. coli* host, which was used for plasmid propagation. The authors believe that negative effects on the bacterial host could be attributed to an homologous peptide in the Flaviviruses polyprotein which was not defined by the authors (Rice, Grakoui et al. 1989). Methods to deal with this instability include changing bacterial hosts for cloning and using different vectors to express the cDNA (Rice, Grakoui et al. 1989; Hurrelbrink, Nestorowicz et al. 1999; Yun, Kim et al. 2003). One successful cDNA vector was a Bacterial Artificial Chromosome (BAC) vector (Yun, Kim et al. 2003). For an infectious clone system developed for Japanese Encephalitis Virus (JEV), switching to a BAC kept the construct stable for over 180 generations in *E. coli.* Using pBeloBAC11 (derivative named pBAC/SV), virus rescue studies gave titres as high as ~2.0 x 10^5 PFU / ml at 24 hours and 2.3 x 10^6 PFU/ml 48 hours post infection, once the RNA was transcribed *in vitro* and tranfected into BHK cell culture monolayers. Of note, these authors also linearized their cDNA construct with a unique restriction site, downstream of the 3' UTR. Treatment with mung bean nuclease removed the overhangs from the cohesive restriction enzyme. Using this

treatment, non-viral nucleotides from run-off transcription would be eliminated. Non-viral nucleotides can negatively impact the infectivity of RNA generated *in vitro* from cDNA constructs (Lai, Zhao et al. 1991; Yun, Kim et al. 2003). For JEV, titres increased to 5.6 x 10^6 PFU/ml at 24 hr in comparison to the previous pBAC/SV rescue titre (2.0 x 10^5). Conversely, there was no significant increase in titre at 48 hrs with and without the RNA transcription modifications (2.3 x 10^6 *vs* 2.4 x 10^6 PFU/ml) (Yun, Kim et al. 2003).

To achieve successful infection, viral RNA transcripts or RNA genomes must associate with their viral-encoded proteins, like viral replicase. Therefore, transcripts generated from the system have to mimic the viral RNA as closely as possible (Boyer and Haenni 1994). In addition to run-off transcription, a HDV (Hepatitis Delta Virus) ribozyme may be added. The HDV virus is known as a "defective satellite virus" which is present during Hepatitis B Virus (HBV) infections. HDV is a circular negative-sense, single-stranded RNA sub-virus that uses self-cleavage and ligation abilities for replication, via a double rolling circle mechanism (White and Fenner 1994). This ribozyme has autolytic activity and will cleave the RNA before the first nucleotide of its sequence. Previous work suggested that cleavage was from an acid-base catalysis mechanism. A trans-esterification reaction between the ribozymes' first nucleotide and the previous nucleotide will occur, resulting in a 3' cyclical phosphate group and a 5' hydroxyl group (Ciesiolka, Wrzesinski et al. 2001). Such a reaction is thought to involve a cytosine in the ribozyme sequence acting as a base, which de-protonates the 2' hydroxyl group. Attack of the 5'-3' phospho-diester bond creates the 3' cyclical-phosphate group. A hydrated metal cation, usually magnesium, donates a proton to the 5' hydroxyl group. The reverse has also been suggested, in which magnesium hydroxide is the base

and the cytosine is the acid, generating the 3' cyclical phosphate and 5' hydroxyl group, respectively (Ciesiolka, Wrzesinski et al. 2001; Fedoruk-Wyszomirska, Giel-Pietraszuk et al. 2009). Therefore, if the ribozyme is placed downstream of the viral sequence, then non-viral residues can be removed by the RNA itself. An additional benefit is that the 3' RNA has a cyclic phosphate group, which may increase protection from cellular exonucleases (Kawaoka 2004).

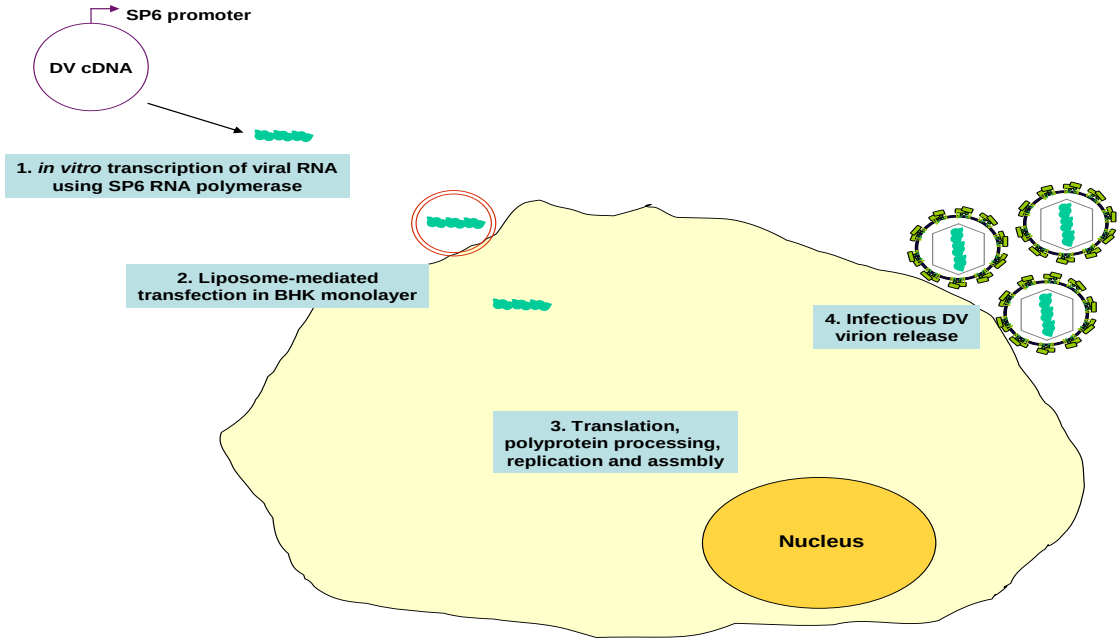


Figure 1.6.1: Infectious clone system for DV. Adapted from Lai et al. 1991 (Lai, Zhao et al. 1991).

1.6.2. Reverse Genetic Systems for analyzing infectious disease

The first infectious cDNA clone systems for negative-stranded viruses occurred following the positive-stranded systems were generated and include the virus familes: *Filovridae*, *Rhabdoviridae*, *Paramyxoviridae* and *Orthomyxoviridae*. Negative-stranded viruses need to be transcribed into their positive-stranded (antigenome) in order to be translated (Ebihara, Groseth et al. 2005). Their nucleocapsids contain viral proteins required for replication, i.e. Filoviruses infectious particles include their RNA-dependent RNA polymerase (Volchkov, Volchkova et al. 2001). Ultimately, the rescue of live and infectious RNA virus particles can be achieved, entirely from cloned cDNA and using mammalian cell culture. Establishment of such systems has been focused on the *Mononegavirales* viral order, featuring negative-stranded RNA virus families (Kawaoka 2004). In using the *Mononegavirales*, a number of families with negative-strand RNA genomes were used as a model for the RGS, an important functional unit that is necessary for replication is the ribonucleoprotein complex (RNP) (Kawaoka 2004) (Figure 1.6.2). RNP complexes are believed to bind to a promoter site on the 3' end of the RNA genome. After binding, the RNA dependent RNA polymerase (RdRp) will drive the replication of the genome into a positive-strand form. Positive-strand RNA is the equivalent to mRNA, which is important in generating the viral proteins and also more negative-strand RNA to serve as nascent viral genomes. Packaging, assembly and release of progeny viruses from the infected cell can now be achieved. Therefore, the RGS can mimic the natural course of infection, allowing researchers to manipulate the viruses in cDNA format. The RGS system takes advantage of two important replication (RNP) and attachment (envelope proteins) functional units (Kawaoka 2004).

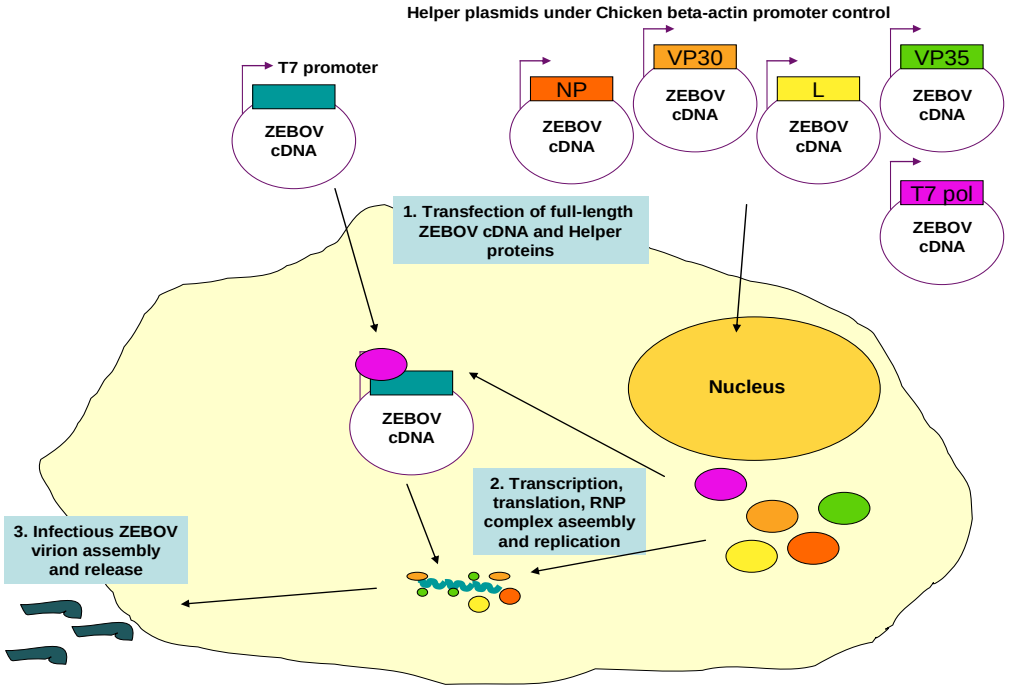


Figure 1.6.2: Reverse Genetics System for Ebola Virus-Zaire (ZEBOV). Adapted from (Theriault, Groseth et al. 2004; Ebihara, Groseth et al. 2005).

Since the RNP complex is so vital to replication, it is of utmost importance to understand its components. The Nucleoprotein (N) is responsible for encapsidation of the nascent genomes and begins assembly of new virus particles. Phosphoprotein (P) is vital for the encapsidation of RNA by N, as P acts as a chaperone. The RdRp, encoded by the L gene, is the catalytic protein for transcription of the genome into an anti-genomic RNA. RNP complexes are necessary for the replication of all negative-stranded RNA viruses. Apparently, replication and encapsidation by RNP complexes are linked. The second functional unit in the RC is the envelope protein(s). Attachment and an initial round of infection are facilitated by the envelope protein(s).

1.6.3. Generation of a RGS for KFDV

In comparison to the reverse genetics systems for *Mononegavirales,* a system for KFDV could also require an RNP complex or in the case of *Flaviviridae* the RC complex. Such a requirement would counter the idea that Flaviviruses require translation to occur first before replication (Harris, Holden et al. 2006; Richman, Whitley et al. 2009). This would indicate that the RC of Flaviviruses is within the virion and not utilized once it is generated after translation of the genome. With that established, this RGS for KFDV was designed to follow the principles of the negative-stranded viral rescue systems. As with *Mononegavirales,* after transfection of mammalian cells with the RGS, viral proteins forming the viral RC complex would be made first. The RC complex will drive replication to generate the anti-genomic (negative-strand) KFD RNA (Kawaoka 2004). To our knowledge, the RC of KFDV still remains to be fully elucidated. We hypothesize that the capsid (C) protein will act as the N protein, NS5 (RdRp) will function as does L and finally the NS3 / NS2B will provide viral serine

protease, nucleoside triphosphatase (capping) and helicase activities. With these proteins making up the proposed RC complex, when produced from the helper plasmids, replication and efficient rescue of KFD virions should occur. However, when the positive-stranded RNA genome enters the cell, translation into the polyprotein begins, which will generate the RC. From here the RC will replicate the genome inside of intracellular, convoluted or vesicle packets (Harris, Holden et al. 2006; Villordo and Gamarnik 2009). With that in mind, the helper proteins may not be necessary for rescue of virus particles in this positive-stranded model because the RC complex will assemble from the polyprotein following cleavage. In principal for negative-stranded virus models, a successful rescue requires a minimum of three criteria to be met: firstly, production of viral anti-genomic RNA must be successful. This negative-strand of RNA will be used as a precursor to synthesize viral proteins and genomic RNA. Secondly, expression of the minimal viral proteins that can form the viral RC complex is necessary. The purpose of this complex is to drive the transcription or replication of the anti-genomic RNA into genomic RNA. Third, an *in vivo* system must be used to support the rescue attempts. This cellular environment can be obtained by using cell culture and or animal models (Kawaoka 2004).

The efficiency or yield of virus rescue using the reverse genetics system can be affected by the promoter used for the RNA polymerase. This is important as different promoters have different strengths and binding abilities, which could affect yield and length of transcription, for both rescue plasmids and the reverse genetics construct (Boyer and Haenni 1994). The Kyasanur Forest Disease virus system will utilize RNA polymerase II within the nucleus of the mammalian host cell.

The KFDV reverse genetics system will contain two types of mammalian expression vectors: pTM1 (Moss, Elroy-Stein et al. 1990)(Figure 1.6.3.A) for the KFD genome expression and pCAGGS-MCS (Niwa, Yamamura et al. 1991) (Figure 1.6.3.B), for the biosynthesis of the helper proteins and the formation of the viral replication complex. Benefits of the pCAGGS expression are: a strong chicken beta-actin promoter and stable and efficient transcription in the nucleus by RNA Polymerase II (Neumann, Feldmann et al. 2002; Ebihara, Groseth et al. 2005). Mammalian promoters can be used in any mammalian expression system but in this system, the actin promoter is isolated from chickens because it is stronger than the human equivalent. In eukayotic cells, actin is a major component of the cytoskeleton and is transcribed by RNA polymerase II (Niwa, Yamamura et al. 1991; Bauman 2009). Preceding the promoter is a Human Cytomegalovirus (CMV) immediate-early promoter enhancer. With the enhancer, high level expression of downstream cloned genes will be very efficient and allow polymerases to be more stable at the promoter (Clontech). The reverse genetics construct will be comprised of the full-length cDNA genome of KFDV cloned in the pTM1 mammalian expression vector (Moss, Elroy-Stein et al. 1990). There are several features associated with this vector that are beneficial for many studies. First, a bacteriophage-based transcription system will be used to allow high efficiency, specific and regulated transcription. Bacteriophage T7 DNA-dependent RNA polymerase is a single-subunit structure polymerase that does not need co-factors. T7 RNA pol is very functional in eukaryotes, has been used in many expression systems and if desired, a nuclear localization signal (NLS) can be added to allow for nuclear activity (Kawaoka 2004). This enzyme has high transcription rates, high processivity and is functional in the

cytoplasm (Moss, Elroy-Stein et al. 1990). The T7 RNA polymerase has been used for RGS systems for viruses that replicate strictly in the mammalian cytoplasm (Volchkov, Volchkova et al. 2001; Neumann, Feldmann et al. 2002). T7 RNA pol will be expressed under control of the chicken beta-actin promoter of the pCAGGS-MCS vector and will drive transcription of the KFDV cDNA genome in the cytoplasm. This system will be similar to the natural cytoplasmic replication cycle of Flaviviruses (Howley and Knipe 2007).

In mammalian systems, mRNA is more stable and translation is more efficient with a capped transcript (Kapp and Lorsch 2004) but another approach is to use a cap-independent translation system. The pTM1 vector has been previously modified with an Internal Ribosome Entry Sequence (IRES) to allow cap-independent translation. Encephalomyocarditis Virus (EMCV) of the *Picornaviridae* family does not have a cap on its RNA genome. However, what it does have is an internal ribosome entry site (IRES) on the 5' UTR, which facilitates cap-independent translation. The addition of the EMCV 5' UTR increases the expression rates of pTM1 transcripts by 5-10 fold. Therefore, capping of transcripts from this vector is not needed for efficient translation (Moss, Elroy-Stein et al. 1990). The Multiple Cloning Site (MCS) has an ATG or start codon provided by the first restriction site, *Nco*1 (Moss, Elroy-Stein et al. 1990). For KFDV, our system used the natural start codon at the start of the C protein. Therefore, the *Nco*1 and the ATG were removed by basic cloning procedures described in section 2.5.

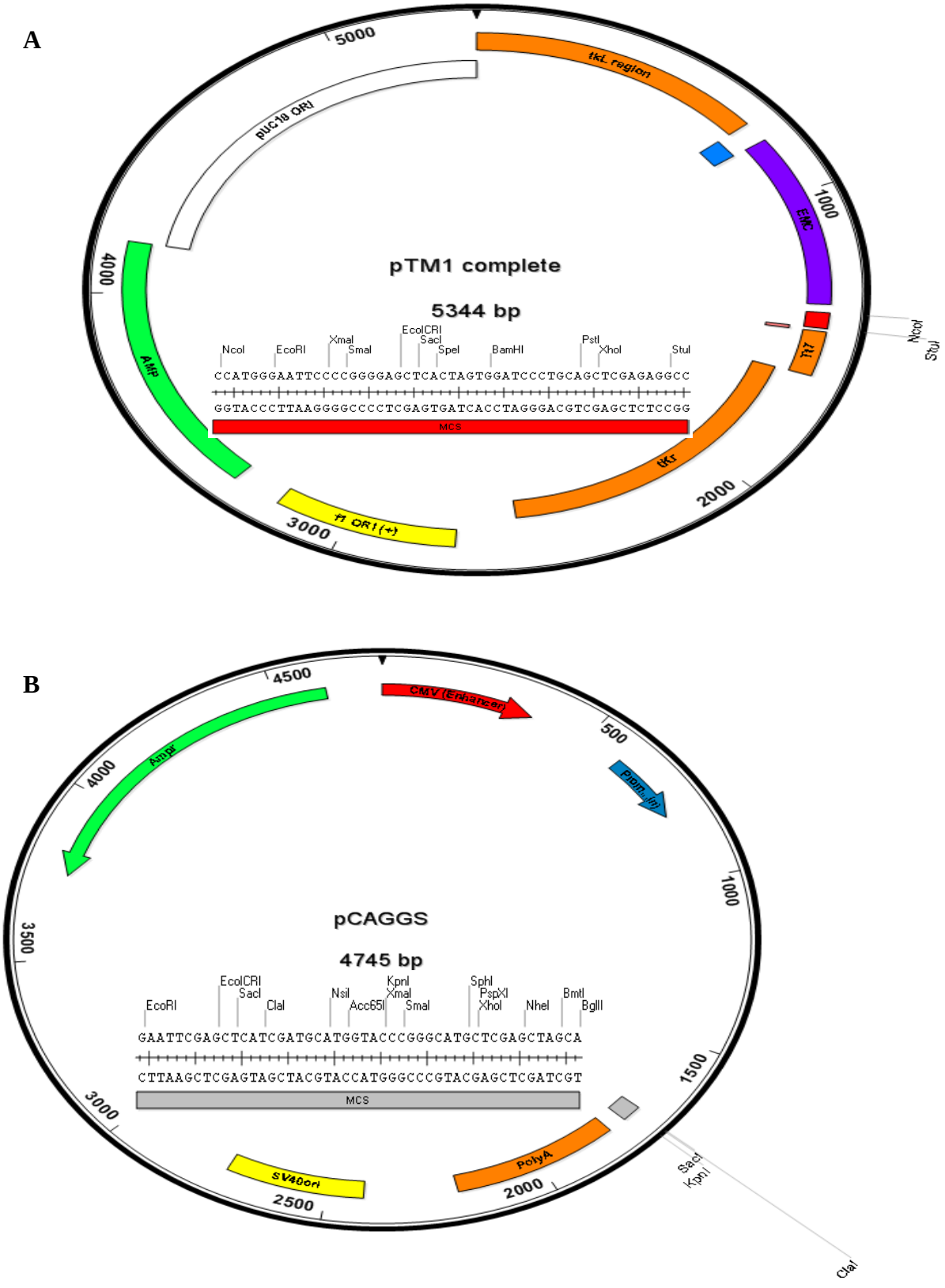


Figure 1.6.3: Vectors for the generation of the KFDV RGS.

A) The pTM1 mammalian expression vector, showing the *Nco*1 and *Stu*1 sites of the MCS. MCS: red, IRES: purple, T7 promoter: blue, Ampicillin (Amp) resistance gene: green. B) The pCAGGS-MCS mammalian expression vector, showing the *Sca*1, *Cla*1 and *Kpn*1 of the MCS. MCS: silver, CMV-enhancer: red, Chicken-beta actin promoter: blue, Amp resistance gene: green. Adapted from (Moss, Elroy-Stein et al. 1990; Niwa, Yamamura et al. 1991) and created using the SeqBuilder program (DNASTAR, Madison, WI., USA).

When trying to rescue KFDV, a problem that must be avoided is the possible hybridization of anti-genomic RNA with the mRNA of the helper plasmids. This can occur as mRNA is positive-strand and anti-genomic RNA will be in a negative-stranded orientation. Consequently, hybridization could occur, which will prevent RC assembly and halt the rescue attempt. Additionally, this partial double-stranded RNA (dsRNA), will trigger the antiviral state provided by Interferons (IFN), which may also block rescue attempts (Kawaoka 2004). IFNs are cytokines that exert strong responses to viral infections. The antiviral state will be induced for both virus-infected cells and surrounding cells by blocking viral and infected target cell replication and protein synthesis. Combined with cell growth arrest and induction of apoptosis, IFN will limit the spread of viruses to neighboring cells (Goodbourn, Didcock et al. 2000; Kawaoka 2004). However, our system should avoid this as the rescue proteins will be transcribed and translated before transcription of viral anti-genomic RNA. This will occur because the T7 RNA polymerase needs to be expressed first, in order for the cDNA genome of KFDV to be transcribed. The RC complex proteins will also be expressed from the helper plasmids at the same time as T7 RNA pol and before the cDNA is transcribed. Hybridization may still occur but, the original timing of transcription of the helper proteins in the nucleus and KFDV genome in the cytoplasm (requiring T7 RNA pol), may limit the amount of hybridization.

1.7. Research goals

Development of a reverse genetics system (RGS) will enable researchers to study KFDV and supply a useful application for other positive-stranded RNA viruses. Such a system will allow researchers to study the biological properties and mechanisms involved in viral life cycle characteristics and pathogenesis through genetic manipulations. A better understanding of the complexities of the interactions between viruses and their hosts and will inevitably, propel vaccine strategies using both cell culture and animal models. However, in order to create the RGS system, the full-length genomic sequence of KFDV must be elucidated and constructs of the full genome and helper proteins must be generated. Finally, successful rescue of live and infectious KFDV particles will validate the RGS system. The set of helper proteins consisting of each of the ten viral-encoding proteins of KFDV will enable complementation studies and understanding of the roles of each protein in pathogenesis and each aspect of the virus life cycle including: adsorption, uncoating, replication, assembly and egress. Additionally, positive rescues with just the cDNA genome will confirm the positive-stranded genome as the only requirement for genome replication. This would be in accordance with data from the *in vitro* infectious clone systems previously described for Flaviviruses (Rice, Grakoui et al. 1989; Lai, Zhao et al. 1991; Pekosz, He et al. 1999), with the hypothesis that the nucleocapsid only contains the viral genome encapsidated by the C protein (Gehrke, Ecker et al. 2003; Howley and Knipe 2007). Conversely, such a result will be unlike the negative-stranded virus models that requires a viral RNP complex for successful rescue of infectious virus particles (Pekosz, He et al. 1999; Kawaoka 2004).

The specific goals of this research are: to clone and sequence the KFDV genome, create subgenomic clones for the RGS and test the accomplishment of the RGS in cell culture.

Chapter 2: Materials and Methods

2.1. KFD infection, RNA harvest and reverse transcription

The Containment Level-4 (CL-4) and Containment Level 3+ (CL-3+) aspects of this research were performed by qualified employees of the Public Health Agency of Canada (PHAC), Todd Cutts and Steven Theriault. KFDV was propagated in VeroE6 cells within a CL-4 laboratory at an MOI of 0.1. Infected cells were then harvested at 72 hours post infection when 70-80% CPE was established (Theriault, Groseth et al. 2004). The supernatant from the harvest was concentrated on a 20% sucrose cushion at 50,000 g and extracted using TRIzolLS reagent kit (Invitrogen, Burlington, Ont., Canada), according to the manufacturer's protocol. Samples were removed from CL-4 for further processing and analysis. Under CL-3+ conditions, the single stranded (ss) cDNA was generated using Thermoscript (Invitrogen) in a two-step reverse transcription reaction. The first-step of the single-stranded cDNA synthesis began with approximately 500 ng total RNA, divided into 17 separate reactions for reverse transcription. With the RNA as a template, single-stranded cDNA (ss cDNA) was generated with each tube containing an anti-sense primer, refer to Appendix A.1 and Figure 3.1.1.A. Thus, KFD1as to KFD13as and F1as to F4as primers were in the 17 tubes. The cDNA amplification was achieved with reverse transcriptase for 3 hours at 52°C. Since the KFD RNA is infectious, a subsequent incubation with RNaseH enzyme (New England Biolabs, Pickering, Ont.,

Canada) was performed at 37°C for 30 minutes. The ss cDNA was removed from the CL3+ laboratory before beginning the second step of the reverse transcription reaction within a CL-2 laboratory as described in section 2.2.

The 3' UTR region was reverse transcribed in CL-3+ conditions beginning with the addition of a RNA adapter. The adapter was a 5'P-dT(25)-dd(quencher)3' (New England BioLabs), Appendix A.1, was ligated to the carboxyl terminal coding region of the KFDV RNA facilitated by T4 RNA ligase for 30 minutes at 37^{o}C. Amplification with an antisense adapter primer generated a ss cDNA, from reverse transcription for 3 hours at 52°C. The second reverse transcription step was then performed in a CL-2 laboratory as described in section 2.2.

The 5'UTR cDNA was amplified following the ligation of the 5' UTR to the 3' UTR. This RNA ligation was made possible by exchanging the 5' Cap with a phosphate group, using Tobacco Acid Pyrophosphatase (TAP), incubated for 30 minutes at 37 ^{o}C (Epicentre Biotechnologies, Markham, Ont., Canada). The ligation of the 5' phosphate and 3' OH with T4 RNA ligase I (New England BioLabs) for 30 minutes at 37 ^{o}C, resulting in the circularization of the KFDV genome. Subsequent cDNA synthesis of the 5' fragment was achieved using reverse transcriptase for 3 hours at 52°C. Second step or PCR amplification was performed in a CL-2 laboratory as described in section 2.2.

2.2. PCR reactions for all four fragments for sequencing and RGS construction, 3' UTR-adapter cDNA, 5' UTR-ligated cDNA product and all 10 KFDV rescue proteins from the cDNA template

All PCR reactions were performed according to Figure 3.1.1.B and Appendix A.1, in a total volume of 50 µl, using iproof High-Fidelity DNA polymerase (Bio-Rad Laboratories, Mississauga, Ont., Canada): 10 µl 5X iproof HF buffer (1X final concentration), 1 µl dNTP mixture (200 µM each nucleotide final concentration), 0.5 µl 100 µM sense primer and 0.5 µl 100 µM anti-sense primer (1 µM final concentration each), 1 µl iproof HF DNA polymerase (2 Units / µl final concentration), 1 µl DNA template (diluted to 50-100 nm / µl final concentration) and 36 µl sterile water. The control reaction had the same components, except the 1 µl template was not added and the resulting volume was compensated with sterile water (Sambrook and Russell 2001).

The cycling conditions began with an initial denaturation period of 98^{o}C for 1 minute, repeated cycles (28 cycles) of denaturation, hybridization and extension: 98^{o}C for 20 seconds, 56-64^{o}C for 20 seconds and 72^{o}C for 20 seconds per kb of target DNA, respectively. The final extension at 72^{o}C for 2 minutes completes the PCR reaction. This was then followed by a 4^{o}C indefinite pause (Sambrook and Russell 2001). Refer to Appendix A.1, for primer names and respective sequences.

Following PCR reactions for all ten KFDV proteins, the four KFDV fragments, 3' UTR and 5' UTR amplicons for both sequencing and RGS construction were gel isolated and purified. Most gels were 1% agarose; however, for smaller products the gels were

1.5% and 2% agarose. The 1% agarose gels were made with one gram agarose powder and 100 millilitres of 1X TBE (Tris-Borate EDTA) buffer (Sambrook, Fritsch et al. 1989). DNA samples and 1 kb+ DNA ladder (New England BioLabs) were mixed with 1X loading dye, consisting of SDS, ficoll, EDTA and bromophenol blue (Sambrook, Fritsch et al. 1989). Gel electrophoresis was at 120 volts for periods of 20 minutes to 1 hour, depending on the separation of bands sizes necessary.

2.3. Colony-PCR screening reactions for all four fragments for sequencing and RGS construction, 3′ UTR-adapter, 5′ UTR-ligated product and all 10 KFDV rescue proteins from cDNA template

After the prospective clones were created and transformed into *E. coli*, as explained in section 2.6., the ensuing transformants (bacterial colonies) were subjected to PCR screening. All of the colony-PCR screens were accomplished in 30 µl total volume with Go *Taq* green master mix (ProMega, Napean, Ont., Canada), 250 nM of each sense and anti-sense primer (Appendix A.1) and a single colony was added to each reaction and the residual cells on the loop were streaked onto an LB-Amp plate for future plasmid harvesting of the positive clones, as explained in section 2.7.

The thermocycler program began with a ten minute denaturation phase at 95^{o}C followed by repeated cycles (40) of: denaturation, hybridization and extension at 95^{o}C for 30 seconds, 56-64^{o}C for 30 seconds and 72^{o}C for one minute per kb. A final extension for five minutes at 72^{o}C completed the PCR program.

2.4. Restriction digestion of DNA for cloning and clone confirmation

Reactions used restriction enzymes only from New England Biolabs. Typical reactions were in 20 µl total volumes for all PCR inserts and cloning vectors. 2 µl 10X NEB buffer (buffer was dependent on the enzyme(s) being used), 1 µl of each restriction enzyme (~10 units / µl), 1 µg cloning vector, clones from Mini and Maxi preparations (volume dependent on the DNA vector concentration), were added and the mix was brought up to up to 20 µl with sterile water. However, adaptations were made accordingly: PCR inserts, once gel isolated and purified were added up to 20 µl. Bovine Serum Albumin (BSA) was added if recommended by the enzyme manufacturer and enzyme amounts were accounted for in the mixture. Digestions were usually performed at 37^{o}C or at other temperatures, as indicated by the manufacturer for approximately one hour.

2.5. Ligation reaction for sequencing and RGS cloning

PCR inserts from amplification of KFDV fragments one, two, three, four, 5' UTR and 3' UTR cDNA, were blunt-end cloned utilizing Zero Blunt TOPO PCR Cloning Kit for Sequencing (Invitrogen) which allows for sequencing of DNA inserts into the provided pCR4blunt-TOPO vector. These ligations were set up in accordance with the manufacturer's protocol in a total volume of 6 µl: 4 µl PCR product, 1 µl TOPO vector (pCR4) and 1 µl salt (1.2 M NaCl and 0.06 M $MgCl_2$) solution (for chemical transformations) or 1 µl of sterile water (for electroporation transformations). The addition of the salt solution was recommended by the manufacturer for increased

efficiency of transformations with chemically competent *E. coli*. The ligation reactions were mixed gently and allowed to incubate 10-30 minutes at room temperature (21-23^{o}C).

The ligations for RGS constructs were achieved with T4 DNA ligase (New England BioLabs). Typical reactions involved the addition of the DNA vector to DNA insert in a molar ratio of 1:2.5 in 20 µl total reaction volumes: 12 µl insert, 5 µl vector, 2 µl 10X T4 DNA ligase buffer (100 mM ATP added) and 1 µl T4 DNA ligase. Control ligation reactions had 12 µl of sterile water, instead of DNA insert. Incubation times were either 18^{o}C for 4 hours or 16^{o}C overnight (~15-18 hours) (Sambrook and Russell 2001). The MCS of pTM1 was removed by digestion with *Nco*1 and *Stu*1 restriction enzymes and was replaced with the DNA adapter, Appendix A.1. The adapter contained the *Nco*1 and *Stu*1 sites, along with two required restriction sites, *Age*1 and *Kas*1, both are not present in the vector or in the KFDV genome.

2.6. Transformation of *Escherichia coli*

Transformation by chemical or electroporation methods was completed with Chemically-Competent or Electro-Competent Top10 *E. coli* (Invitrogen) for most clones developed in the study. The exception is the final KFDV RGS and KFDV RGS-ribo (RGS with the ribozyme addition) clones in which, XL10-Gold Ultra-Competent *E. coli* (Stratagene, Cedar Creek, TX, USA) was used in chemical transformations. Transformation via chemical methods followed the manufacturer's protocols for their respective competent cells. For the Top10 cells by Invitrogen, the protocol included:

thawing chemical-competent cells from -70°C on ice, adding 3-4 µl of ligation mixture to ~50-100 µl of completely thawed competent cells, mixing gently and incubating 15-30 minutes on ice. The heat shock step lasted for 30 seconds at 42°C and the tubes were immediately transferred to ice. The recovery phase was aided by the addition of 250 µl SOC medium (Sambrook, Fritsch et al. 1989), with incubation for 1 hour at 37°C and shaking at 280 rpm. After recovery, the mixture was spread plated (200 µl and 20 µl) onto LB (Luria-Bertani) medium plates supplemented with 100 µg/ml of Amp. Transformants were generated after an overnight incubation at 37°C for approximately 12-16 hours. Transformation using the XL10-Gold chemically-competent cells followed the same protocol as above with minor alterations suggested by the manufacturer. During the thawing process, 4 µl of XL10-Gold β-Mercaptoethanol mix was added to the cells and they were incubated on ice for 10 minutes, with swirling every 2 minutes. After the heat shock step, the cells were incubated on ice for two minutes and 900 µl of SOC medium was then added. The Stratagene protocol recommends NZY broth as the recovery medium; however, the addition of SOC medium did not appear to adversely affect the transformants after incubation.

Electroporation of Electro-Competent Top10 cells began with thawing the cells on ice immediately upon removal from -70°C. Once completely thawed, 3-4 µl of ligation was added and gently mixed. Then the contents were added to a pre-chilled cuvette on ice. The pulse was set to 2.5 kV, according to the program on the electroporator (Bio-Rad). Immediately after, 900 µl of SOC medium was added and the cells were allowed to recover for 1 hour at 37°C with shaking at 280 rpm. Following the recovery period, 200

μl and 20 μl aliquots were spread plated on LB-Amp (100 μg/ml) plates and incubated overnight (12-16 hours) at 37^{o}C.

2.7. Plasmid DNA harvesting and purification via alkaline lysis procedure

The confirmation of generated clones required plasmid isolations from the transformed *E. coli*. This involved using mini prep and maxi prep kits (Qiagen, Mississauga, Ont., Canada) following the manufacturer's protocols on the colony-PCR positive samples, explained in section 2.3. Quantification of plasmid isolations was performed by a spectrophotometer (Implen NanoPhotometer). All clones testing positive by colony-PCR (section 2.3) and restriction digestion (section 2.4) were DNA sequenced by the NML DNA core facility using internal primer sets outlined in Figure 3.2 and Appendix A.1.

2.8. Media and antibiotic preparations

All of the constructs were made in vectors encoding β-lactamase, which is the enzyme that degrades the antibiotic Amp. Therefore, plates for culturing transformants, broth cultures for plasmid isolation and purification were supplemented with Amp at a final concentration of 100 μg / ml. Mini and Maxi preps were grown in 5-10 ml and 100-300 ml of LB-Lenox broth, respectively. The LB-Amp media plates and LB-Lenox broth were both kindly prepared and provided by the NML media department.

2.9. KFDV rescue

For each rescue, either Vero E6 (African Green Monkey Kidney) or BHK-21 (Baby Hamster Kidney) tissue culture monolayers were split into 4-6 well plates with the intent of having ~80% confluency, 24 hours before the intended rescue. The following day, each transfection utilized nine wells of either Vero E6 or BHK-21 tissue culture monolayers. The transfections were set up the same in terms of media and transfection reagent but included different combinations of plasmids, outlined in Table 3.1. According to an adapted protocol, kindly provided by Dr. Theriault, 100 µl of serum-free media Opti-MEM (Gibco, Burlington, Ont., Canada) was added to nine sterile tubes, a 2 µl aliquot of TransIT-LT1 (Mirus, Ottawa, Ont., Canada) transfection reagent, was then added to each tube and allowed to incubate for 5 minutes at room temperature. 1 µg of each DNA plasmid construct was added in the appropriate order and was followed by a 30 minute incubation step at room temperature. A subsequent 900 µl addition of Opti-MEM completed the preparation of the DNA complexes for transfection. The growth medium from the nine wells (three wells per plate) was removed and the transfection mixtures were then applied to their respective wells. The tissue culture plates were transferred in the CL-4 suites and subject to an overnight incubation at 37^{o}C with 5% CO_2. After incubation but, before 24 hours the transfection mixture for each well was replaced with 4 ml of Dulbecco's Modified Eagle Medium (DMEM) (Gibco) supplemented with 2% Fetal Bovine Serum (FBS) (Gibco) and 1% Penicillin-Streptomycin (PenStrep)(Gibco). The plates were re-incubated for 4 days under the same conditions in CL-4. Blind passaging immediately followed the 4 day incubation, which requires 9-T 25 tissue culture flasks (Vero E6 or BHK-21), prepared for 80% confluency

after 24 hours incubation. Blind passaging involved the removal of the T 25 media and the addition of 2 ml media from one of the nine tranfected wells from the 6 well plates. Incubation with the infectious 2 ml of medium lasted for 30 to 60 minutes at 37^{o}C and 5% CO_2 and then 5 ml of DMEM supplemented with 2% FBS and 1% PenStrep was added. Confirmation of a successful rescue was performed after 14 days of incubation at 37^{o}C and 5% CO_2 or until ~70-80% CPE was noticed.

2.10. Confirmation of KFDV rescue attempts with RT-PCR

Once CPE was seen between 7-14 days after the blind passage, tissue culture media was collected and monolayers were harvested by scraping the cells off of the bottom of each tissue culture flask. Both samples were inactivated by addition of RLT buffer from the TRIzolLS reagent kit (Invitrogen) and safely removed from the CL-4 suite. While in CL-3+ laboratory conditions, the cell monlayers were homogenized using the QIAshredder kit (Qiagen) and RNA was harvested with RNeasy kit (Qiagen). Verification of the presence of RNA from KFDV was assessed using Qiagen OneStep RT-PCR (Qiagen) in 50 µl total reactions: 10.0 µl 5X Qiagen OneStep RT-PCR buffer, 2.0 µl dNTP mix (400 µM each nucleotide final concentration), 0.3 µl 100 µM KFD12s (or KFD7s)(600 nM final concentration), 0.3 µl 100 µM KFD12as (or KFD7as)(600 nM final concentration), 2 µl RT-PCR enzyme mix, 5 µl isolated template RNA and 30.4 µl sterile water. For the negative control, the template was replaced with 5 µl of sterile water. These contents were mixed and subjected to the following cycling conditions: 50^{o}C for 30 minutes, 95 ^{o}C 15 minutes, 28 cycles of denaturation, annealing and

extension (95 ^{o}C for 1 minute, 58 ^{o}C for 1 minute and 72 ^{o}C for 1 minute, respectively), a final extension of 72 ^{o}C for 10 minutes. The completed reactions were treated with RNAse H enzyme (New England Biolabs) for 20 minutes at room temperature and then run on a 1% agarose gel for 45 minutes at 100 volts.

Chapter 3: Results

3.1. KFDV genome sequencing

The KFDV polyprotein sequence was divided into four overlapping fragments, Figure 3.1.1.A, based on a previously published polyprotein sequence (AY323490). Primers were synthesized to amplify each fragment, 27 sense and anti-sense internal primer pairs were used for the sequencing of the internal regions of each fragment, Appendix A.1. The cDNAs were reverse transcribed from the KFDV RNA genome inside of a CL-3+ facility. RT-PCR was performed with each anti-sense primer for all four fragments and 13 internal primers, Figure 3.1.1.A.

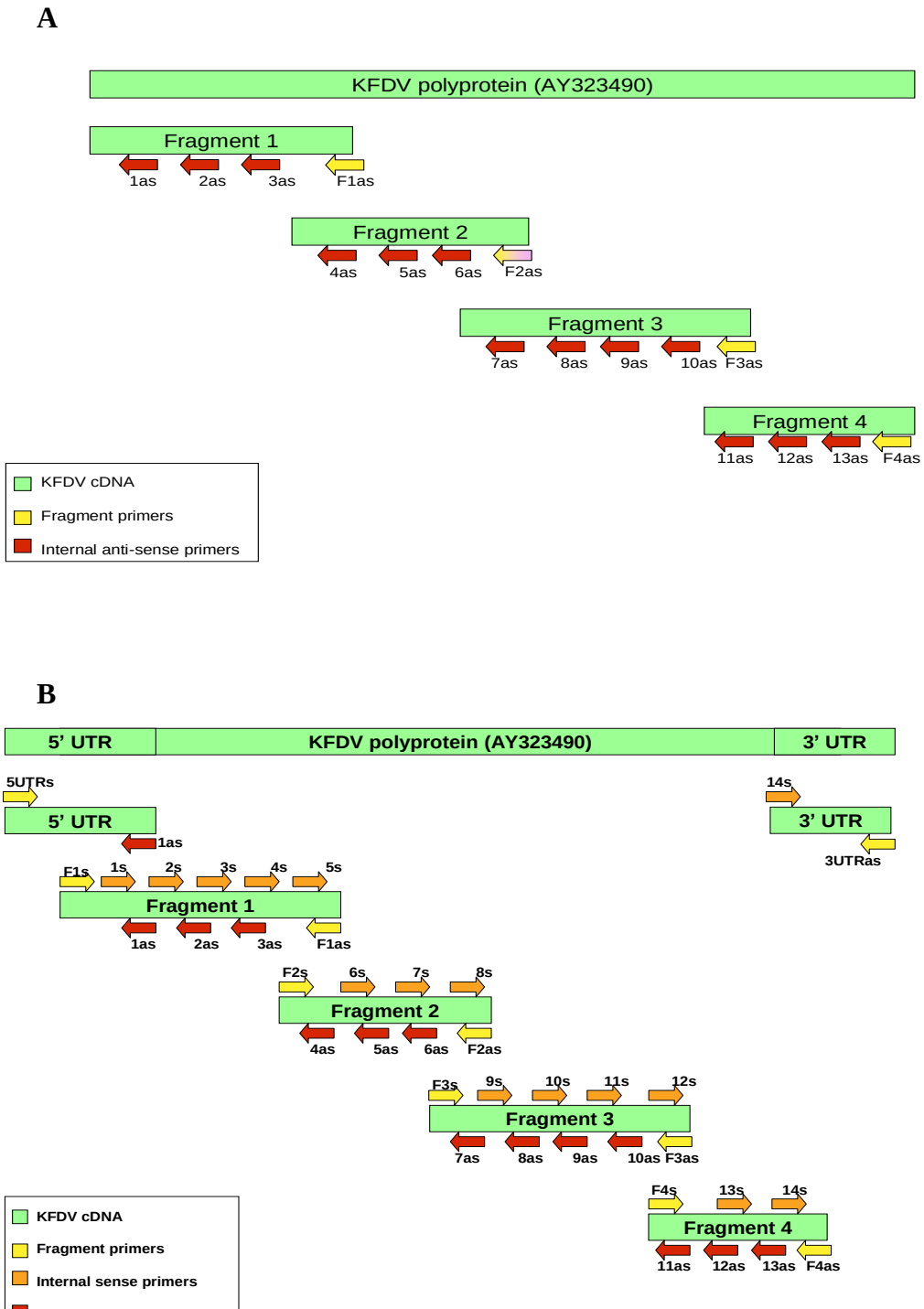


Figure 3.1.1: A) KFDV cDNA generation using RT-PCR and antisense primers based on published KFDV polyprotein sequence (AY323490). Each amplified cDNA was titled according to the antisense primer that was used, ie. KFD9as-cDNA template. B) Strategy of KFDV full-length genome sequencing.

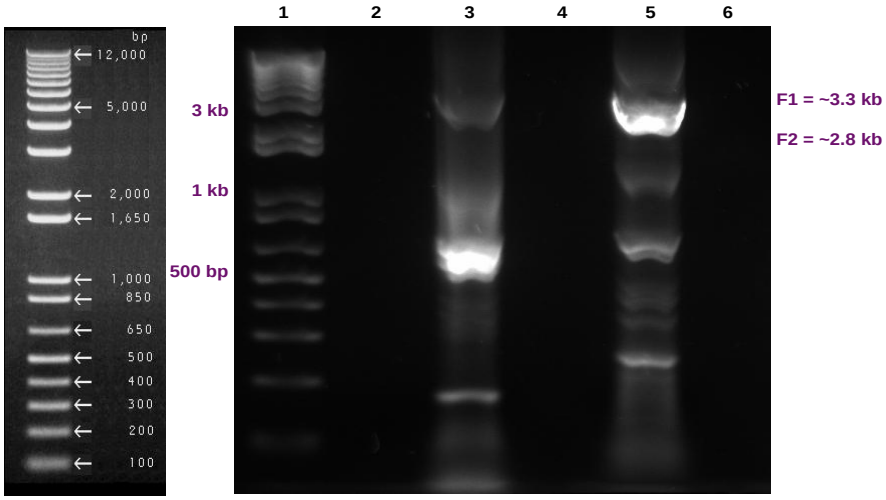


Figure 3.1.2: PCR amplification of the first and second fragments of KFDV.

Lane 1: 1 kb+ DNA ladder (Invitrogen), Lane 2: Sterile water and loading dye, Lane 3: Fragment one, Lane 4: PCR negative control for fragment one, Lane 5: Fragment two, Lane 6: PCR negative control of fragment two, PCR products were analyzed on a 1% agarose gel.

Amplification of fragments one and two were performed using: KFD5as derived cDNA template, Figure 3.1.1.A., with KFDF1s / KFDF1as, Figure 3.1.1.B., and KFD8as cDNA template with KFDF2s / KFDF2as, respectively. The resulting bands, as shown in Figure 3.1.2., of ~3.3 kb and ~2.8 kb were blunt-end cloned into pCR4-TOPO sequencing vectors. Confirmation of these clones was executed with PCR using KFD1s / KFD3as (~2.1 kb) for segment one and KFD5s / KFD6as (~1.2 kb) for segment two, these primer sets are internal to fragments one and two. The clones were propagated and harvested from their *E. coli* host Top10 strains. The clones were verified by DNA sequencing.

The full-length fragment 3 amplification using KFD3s / KFD3as was unsuccessful on many attempts using different primers, cDNA templates and annealing temperatures. Consequently, three overlapping pieces of fragment 3 were generated and designated N1 (N-terminus 1), C1 (C-terminus 1) and C2 (C-terminus 2). The N1 portion was first amplified with KFDF3s / KFD9as on the KFD9as-cDNA generated template. Processing of the resulting ~ 1.7 kb gel isolated DNA and blunt-end cloning into pCR4, immediately followed. PCR screening with KFD8s / KFD8as gave a ~650 bp band. Sequencing of this clone validated the N1 sequence. The C1 terminus was created via KFD9s / KFDF3as on KFDF3as-cDNA template. An amplicon of ~2.5 kb was gel isolated, purified and cloned into pCR4-TOPO. Screening with KFD10s / KFD10as affirmed the clone. Sequencing later demonstrated an unexpected result at the latter end of the fragment. A mispriming event may have occurred as the sequence was in agreement with AY323490 from KFD9s (or nucleotide position 6253) to slightly past KFD10as, at nucleotide position 8046. Position 8047 is the 3' sequence of the KFDF3as primer, followed by the pCR4 vector sequence. Thus ~790 bp of segment three was

unaccounted for. Due to this event, a region to overlap C1 and complete the fragment was amplified and designated C2. The C2 portion of the third fragment of KFDV was generated by amplification of the KFD12as cDNA generated template and KFD11s / KFDF3as, giving a ~1.0 kb band. Colony-PCR using the same primers as in the PCR reaction, followed by sequencing confirmed this F3 (C2)-pCR4-TOPO clone. Figure 3.1.3, displays the combination of N1, C1 and C2, which embodies the full third fragment, PCR amplified using F3s/F3-pTM1as which was used in the RGS development, Figure 3.2.1.

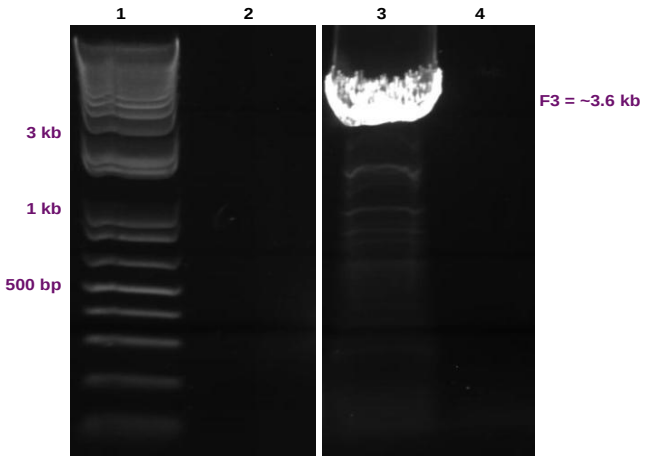


Figure 3.1.3: PCR amplification of the third fragment of KFDV.

Lane 1: 1 kb+ DNA ladder (Invitrogen), Lane 2: Sterile water and loading dye, Lane 3: Fragment three, Lane 4: PCR negative control of fragment three, on a 1% agarose gel. Insignificant lanes were removed from the image.

Amplification of the ~1.7 kb fourth fragment was enabled using KFDF4s / F4as-pTM1, shown in Figure 3.1.4. Gel purification, ligation into pCR4-TOPO, transformation and colony-PCR screening with KFDF4s / F4as-pTM1 gave confidence in the success of the cloning. Subsequently, sequencing provided the final sequence of the polyprotein region of the KFDV genome.

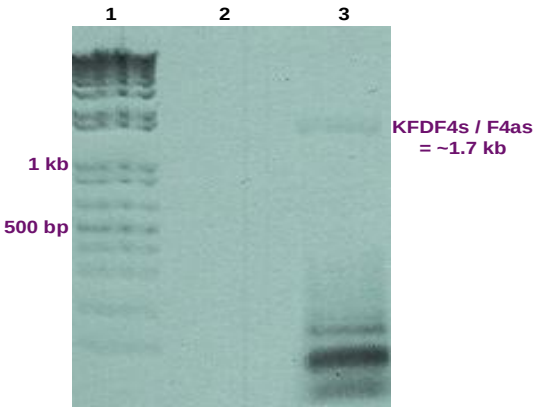


Figure 3.1.4: PCR amplification of fourth fragment of KFDV.

Lane 1: 1 kb+ DNA ladder (Invitrogen), Lane 2: Sterile water and loading dye, Lane 3: Fragment four, on a 1% agarose gel.

The sequence determination of the 3' UTR began with the addition of an oligo-T adapter to the 3' end of the KFDV RNA genome. Synthesis of the cDNA and subsequent PCR amplification with internal primer 14s and Oligo-Tas, provided a product of approximately 592 bp. After blunt-end cloning and transformation, PCR of one colony screening with KFD14s / F4-pTM1as generated a ~211 bp product, as indicated by Figure 3.1.5. Sequencing confirmed the region of interest as part of fragment four (87 bp), 3' UTR (392 bp) and the adapter region (113 bp).

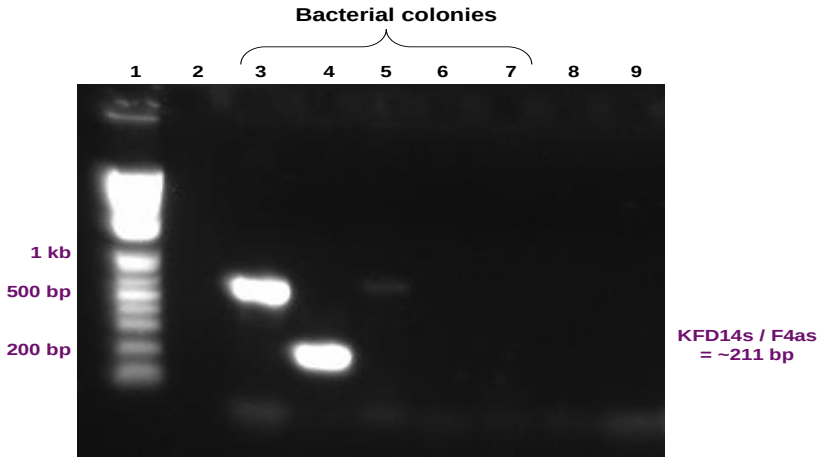


Figure 3.1.5: Colony-PCR screen of KFDV 3' UTR.

Lane 1: 1 kb+ DNA ladder (Invitrogen), Lane 2: Sterile water, Lanes 3-7: PCR products from five isolated colonies, Lanes 8 and 9: PCR negative control of 3' UTR, on a 1% agarose gel. Lane 4 has the desired PCR product.

Elucidation of the 3'UTR, allowed the determination of the 5' UTR sequence with increased certainty because ligation of the unknown 5'UTR region with the known 3' UTR region allowed for amplification using internal sequencing primers: 14s and 1as. The resulting ~1.2 kb PCR product was a combination of the part of fragment four (87 bp), 3' UTR (392 bp), 5' UTR (131 bp) and a portion of fragment one (595 bp). Cloning into pCR4-TOPO and colony-PCR screening with KFD14s / KFD1as demonstrated a band of ~1.2 kb, visualized in Figure 3.1.6. Sequencing and digestion with *Bsa*1 confirmed this construct. The discovery of the 131 bp 5' UTR sequence concluded the elucidation of the full-genome characterization of KFDV.

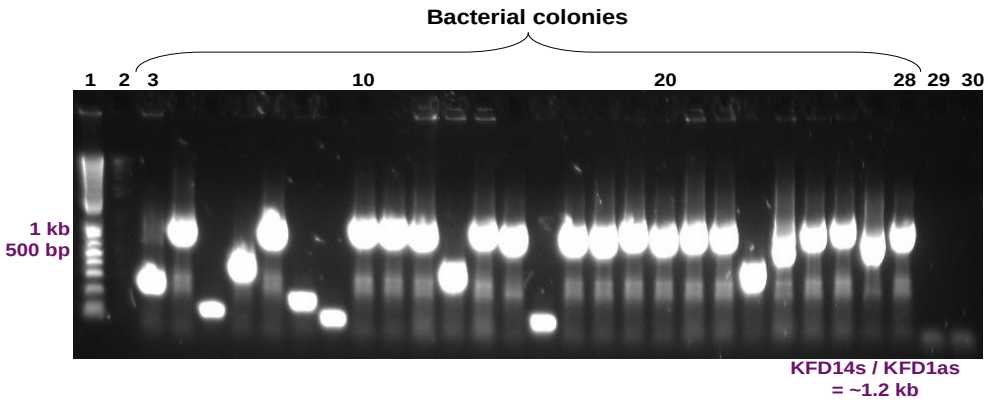


Figure 3.1.6: Colony-PCR screen of KFDV 5' UTR.

Lane 1: 1 kb+ DNA ladder (Invitrogen), Lane 2: Sterile water, Lanes 3-28: PCR products generated from 26 isolated colonies, Lanes 29 and 30: PCR negative control of 5' UTR, on a 1% agarose gel.

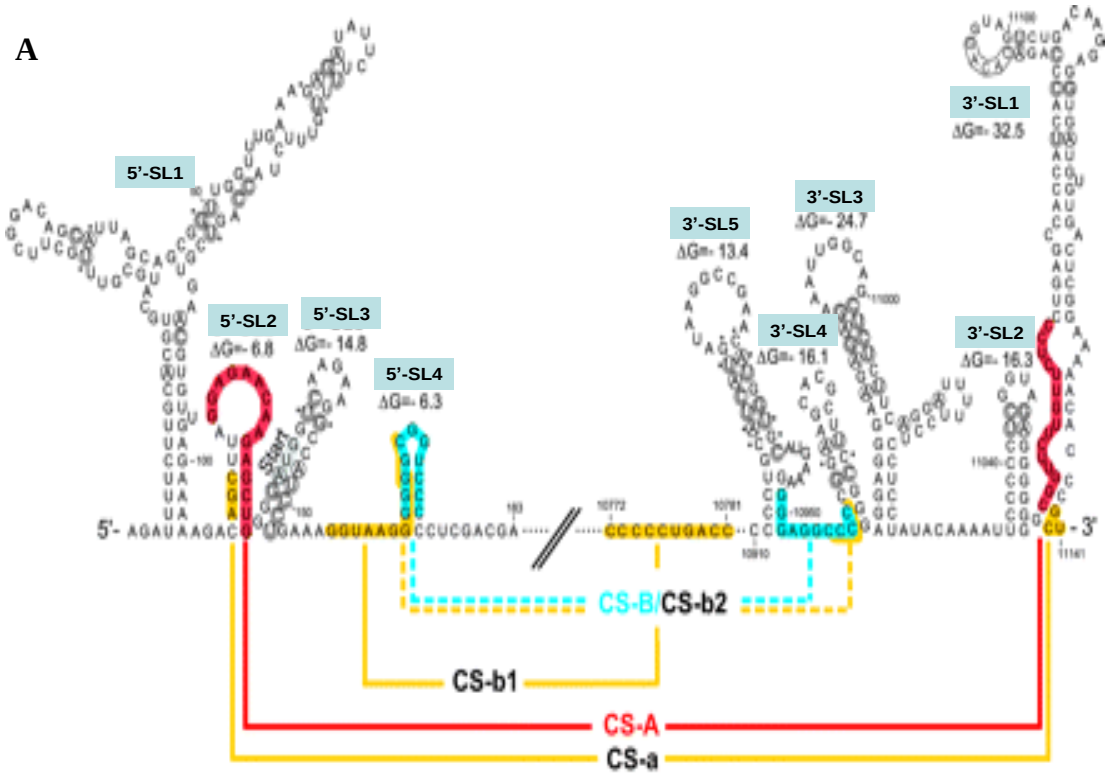
The sequencing of the KFDV genome produced the first annotated full-length KFDV virus sequence (Appendix A.2), which is 10,774 bp in length. During comparisons with the published polyprotein sequence (AY323490) of 10,376 bp in length, sequencing exposed ten points of difference and the resulting amino acid changes are bracketed. In total there are five transitions and four transversions and one deletion: the transitions were at nucleotide positions: C 3489 T (T 1163 S), C 4962 T, A 6559 G (T 2187 A), G 6910 A (A 2304 T) and A 10, 029 G i.e. Position 3489 differed between AY323490 and our strain. Transversions were noted at positions: C 3488 G (T 11663 S), T 5528 A (V 1843 E), C 9147 A and C 9150 G, the T 10, 309 – was deleted in our strain. However, this deletion was downstream of the NS5 gene sequence and thus outside of the open reading frame of the KFDV polypeptide. At the amino acid level four missense mutations were identified within the coding region of KFDV. Of these the most significant for the overall protein structures may be seen in the valine (V) to glutamic acid (E) in position 1843. The properties of these residues are somewhat different in valine being hydrophobic and glutamic acid being acidic.

Stem-loop structures were demonstrated in TBEV-neudoerfl strain in 2006 (Kofler, Hoenninger et al. 2006). TB flaviviruses and MB flaviviruses utilize stem-loop structures to aid in replication and translation of their genomes (Harris, Holden et al. 2006; Villordo and Gamarnik 2009). KFDV being a member of the TB flaviviruses appears to be no different with respect to possessing secondary structures in the 5' and 3' UTR regions. Compared to the secondary structures of TBEV-Neudoerfl as described by (Kofler, Hoenninger et al. 2006), Figure 3.1.7.A, the structures were found for KFDVs 5'

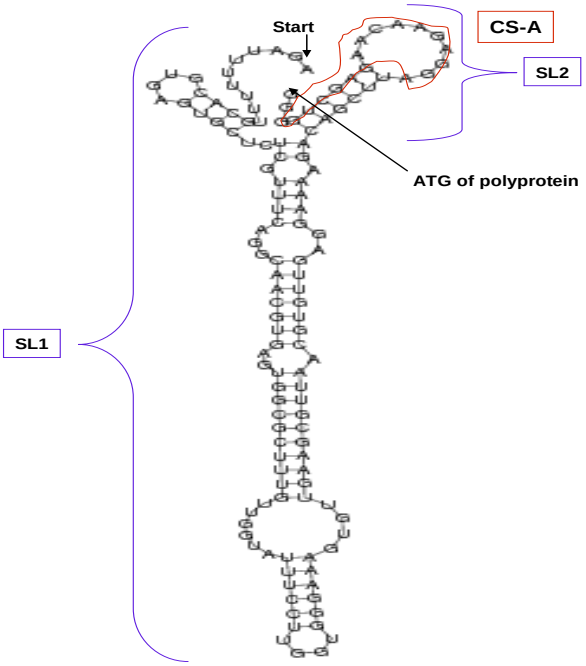
and 3' UTR regions, Figure 3.1.7.B and C. The RNA structures were elucidated using the RNAfold WebServer program (Gruber, Lorenz et al. 2008). The parameters were set according to the programs default settings of 37^{o}C and ionic conditions. The RNA structures are generated using minimum free energy predictions, based on Watson-Crick base-pairing and individual nearest-neighbor-hydrogen bond (INN-HB) calculations. The INN-HB gives free energy contributions and the smallest amount of free energy is governs the secondary structure that is produced. Stem-loop structure stabilities are sequence dependent but, the most stable loops should have the lowest free energy (Mathews, Sabina et al. 1999). The significant 5' SL, specifically the SLA or SL1 is thought to be the RNA polymerase promoter and the second stem-loop is located upstream of the polyprotein start codon (Villordo and Gamarnik 2009). For the most part, KFDV shares the 5' SL (or 5' SL1 (Kofler, Hoenninger et al. 2006)) sequence with TBEV. Alignment of the 5' UTR regions indicates a percent sequence similarity of 70.8 percent (Appendix A.3). In looking at the alignment output file the main locations of sequence divergence is in the SL1 region from bases 55-85 (data not shown). Perhaps this is the reason for the different secondary structures produced. Additionally, such sequence diversity could be expected since Tick-Borne Flaviviruses have been reported to have different sequences within the 5' SL1 region (Kofler, Hoenninger et al. 2006). Furthermore the side stem loop of many Tick-Borne Flaviviruses may vary due a "hypervariable region" (Villordo and Gamarnik 2009). These may result in alternate folding patterns by the computer-based algorithm and final structures are somewhat speculative. The second stem-loop (SL2) is in the same position and 5' CS-A region is conserved in both viruses, Figures 3.1.7.A, 3.1.7.B and 1.5.3.

The stem loops of the 3' SL, specifically the positions of 3' SL1, SL2, SL3, SL4 and SL5 of KFDV are similar to those of the 3' UTR of TBEV. Within the SL1 loop, the 3' CS-A is conserved for both viruses (Figure 3.1.7.A and C). Since, the 5' CS-A and 3' CS-A are thought to be vital for replication of the viral RNA (Villordo and Gamarnik 2009), conservation of both sequences indicate that the 5' and 3' UTRs of KFDV are indeed authentic. When TBEV-Neudoerfl (GenBank accession number: U27495) was folded using RNAfold, the results mimicked those of Figure 3.1.7.A (data not shown). The sequence percent similarity was 73.7% (Appendix A.3). However in contrast to the 5' UTRs the sequences difference was more spread throughout in the 3' UTR. This may be the reason why the RNA structure is more similar for this region than in the 5' UTR.

A


5'-SL1
5'-SL2
ΔG= -6.8
5'-SL3
ΔG= -14.8
5'-SL4
ΔG= -6.3
3'-SL1
ΔG= -32.5
3'-SL3
ΔG= -24.7
3'-SL5
ΔG= -13.4
3'-SL4
ΔG= -16.1
3'-SL2
ΔG= -16.3
Start
5'- AGAUAAGAC
GGAAAGGUAAGG
CCUCGACGA
CCCCCUGACC
AUAUACAAAAUUG
CU -3'
CS-B/CS-b2
CS-b1
CS-A
CS-a

B


Start
CS-A
SL2
ATG of polyprotein
SL1

C

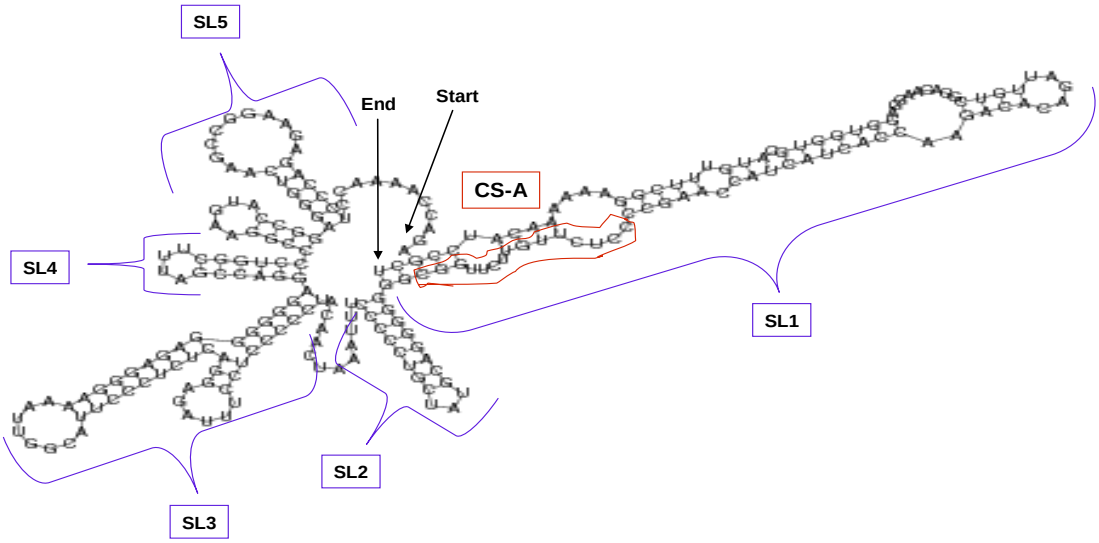


Figure 3.1.7: RNA Secondary Structures of TBEV and KFDV.

A) Secondary structures from TBEV-Neudoerfl strain produced from Vienna RNA package (Kofler, Hoenninger et al. 2006). B) Secondary structures from KFDV 5' UTR created from RNAfold WebServer (Gruber, Lorenz et al. 2008). C) KFDV 3' UTR secondary structure of the last 232 bp of the genome (10, 543-10, 774 bases) based on comparison with TBEV-Neudoerfl strain in A) (10, 910-11, 141 bases). This structure was generated from RNAfold WebServer (Gruber, Lorenz et al. 2008). Figure referenced with permission granted from American Society for Microbiology (ASM).

To investigate the relationships among other Flaviviruses with full-length sequence data, phylogenetic analysis was performed using the DAMBE program (Xia and Xie 2001), ClustalW alignment and Maximum Likelihood (ML) approaches. The ML method describes evolutionary relationships based on the aligned data set and maximizes the chances that the observed information is the best estimate of the data set (Baxevanis and Ouelette 2001). As expected, the TB and MB flaviviruses do cluster together and KFDV clusters with the viruses associated with the TBE serocomplex (Calisher, Karabatsos et al. 1989). AHFV and KFDV form a clade and KFDV is genetically distinct from all of the viruses analyzed. These results are consistent with previous phylogenetic analyses performed with the E, NS5 and prM-E proteins (Lin, Li et al. 2003; Wang, Zhang et al. 2009). When repeated with other MB flaviviruses as the outgroup, the clusters remained similar for each tree (data not shown). Therefore, the tree with YFV as the outgroup is shown in Figure 3.1.8.

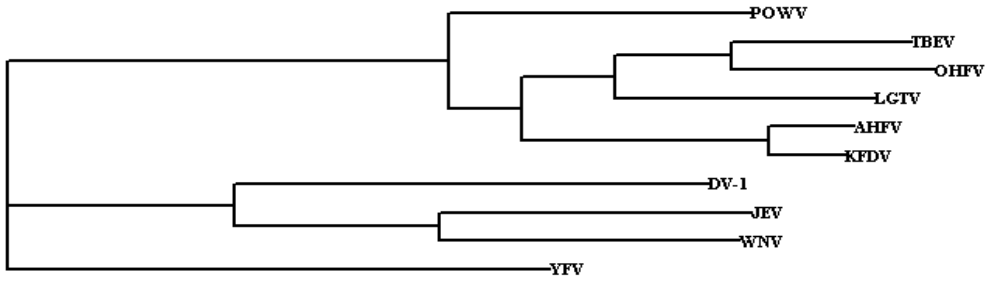


Figure 3.1.8: Phylogenetic analysis using the Maximum Likelihood (ML) method, with Yellow Fever Virus (YFV) as the "outgroup".

The Pubmed "Unique Identifier Numbers" (UID) are as follows: 10160 = Japanese Encephalitis Virus (JEV), 10217 = Dengue Virus-1 (DV-1), 10413 = Yellow Fever Virus (YFV), 17333 = Omsk Hemorrhagic Fever Virus (OHFV), 16744 = Alkhurma Hemorrhagic Fever Virus (AHFV), 16179 = Langat Virus (LGTV), 10598 = Tick-Borne Encephalitis Virus-Neudoerfl (TBEV), 16176 = Powassan Virus (POWV) and 21528 = West Nile Virus-NY99 (WNV).

3.2. RGS Construction

Expression the KFDV genome was designed for the eukaryotic expression vector, pTM1. Sequencing of the full-length genome of KFDV did not reveal unique restriction sites that correlated with the multiple cloning site (MCS) of pTM1. Thus, pTM1 was modified with the addition of an adapter sequence, as described in section 2.5 and appendix 3.2. The new plasmid was designated pTM1-Adapter and is 5.3 kb in size and was verified by digestion with *Xba*1 and *Kas*1 and sequencing. In the original pTM1 *Xba*1 cuts once and *Kas*1 is not present. However, since *Kas*1 is present in the adapter sequence, digestion generated ~1.3-kb and ~3.9-kb fragments.

Construction of the RGS was performed in a step-wise manner building the entire genome, fragment by fragment using unique restriction sites in between each section. For this purpose, each segment was PCR amplified from their pCR4-TOPO sequencing vector template. The primer sets were the same in sequence, except some had the addition of a restriction site on the 5' end in order to build the full length cDNA construct as shown in Figure 3.2.1.

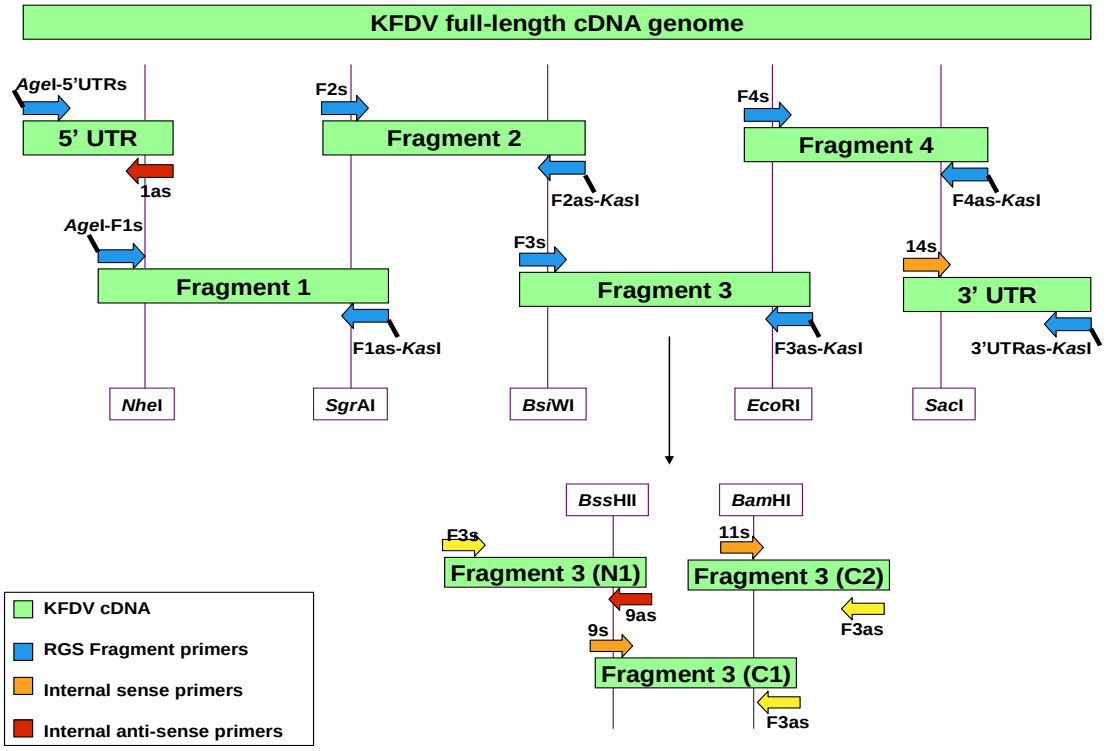


Figure 3.2.1: RGS cloning strategy of the KFDV full-length genome.

Fragment one was amplified using KFDF1-pTM1s with a 5' *Age*1 restriction site and KFDF1-pTM1as containing a 5' *Kas*1 site. This ~3.3-kb segment was cloned into pTM1-Adapter aided by the added restriction sites. Colony-PCR with KFD1s and KFD3as confirmed the cloning as a ~2.1-kb product was produced. Subsequent digestion with *Eco*R1 and *Eag*1, produced two fragments of ~5.0 kb and ~3.5 kb. The pTM1-Adapter had two different fragments of ~3.5 kb and ~1.7 kb and served as a negative control. Sequencing of F1-pTM1-Adapter later confirmed this initial RGS clone.

A unique restriction site, *Sgr*A1, within fragments one and two was utilized for the next cloning. The second piece was amplified with KFDF2s and KFDF2-pTM1as (*Kas*1 added), giving a PCR product of ~2.8 kb. After gel purification and colony-PCR screening using KFD3s and KFD5as, the clone was grown up and the plasmid was purified. Initial verification was employed by digestion with *Hin*d111, F2-pTM1-adapter and F1-pTM1-adapter generated three (~5.5 kb, ~4.4 kb and ~1.0 kb) and two (~7.6 kb and ~1.0 kb) sections. Final verification was enabled by sequencing of the F2-pTM1-Adapter clone.

Addition of fragment three involved a few cloning steps to assemble the full fragment, summarized in Figure 3.2.1. The C1 and C2 components were PCR amplified using: KFD9s / KFDC1as and KFD11s / KFDF3-pTM1as, respectively. These three fragments of ~1.7 kb, ~1.8 kb and ~1.0 kb were gel isolated and purified. Digestion of a unique *Bam*H1 site between the C1 and C2 portions, allowed for their ligation. Amplification by PCR using KFD9s / KFDF3-pTM1as generated a band of ~2.6 kb. After gel isolation and purification, the newly designated C-C was blunt-end cloned into pCR4-TOPO. Screening by colony-PCR of C-C-pCR4-TOPO with KFD11s / KFD11as

gave a band of ~800 bp. Sequencing and digestion with *Bsa*1 confirmed the clone. Sub-cloning of N1 from the N1-pCR4-TOPO vector made for sequencing into C-C-pCR4-TOPO created the full fragment. Facilitation of the sub-cloning involved *Not*1 (from pCR4) and *Bss*H11 (N1 and C1 overlap). Colony-PCR screening with KFD8s / KFD10as demonstrated a ~2.4 kb amplicon. Digestion with *Eag*1 and sequencing confirmed the construct. With fragment three fully assembled, sub-cloning into the RGS was possible as fragment three contained the *Kas*1 site. Digestion of KFDF2-pTM1 and F3-pCR4-TOPO with *Kas*1, along with a unique site, *Bsi*W1, gave ~3.6 kb (F3) and 10.7 kb (5'UTR-F1-F2-pTM1) fragments. Processing, ligation, transformation and colony-PCR screening with the KFD8s / KFD8as primer set (~650 bp), gave confidence for the successful subcloning procedure. The clone was substantiated by sequencing and digestion with the *Pst*1 restriction enzyme.

The 5 'UTR region was amplified with a 5' *Age*1 restriction site added. Digestion with *Age*1 and a common *Nhe*1 inside of fragment one, were utilized for cloning. Following the ligation and transformation steps, colony-PCR screening with 5'UTR-pTM1s / KFD1as gave a ~735 bp band. This 5' UTR-F1-F3-pTM1 (~14.3 kb) construct was confirmed by sequencing and digestion with *Kas*1 and *Xba*1.

The fourth fragment and the 3' UTR were added during the same cloning step. The previous cloning of fragment four and 3' UTR in their sequencing vectors provided an opportunity to PCR amplify both as one fragment. Fragment four and 3' UTR were amplified using their respective primers: KFDF4s / KFDF4-pTM1as and KFD14s / KFD3UTRas. These two bands of ~ 1.7 kb and ~480 bp, were then digested with a unique site, *Sca*1 and ligated together using a one to one ratio. The ligated product was

then cloned into pCR4-TOPO and sequencing was performed to assure that no PCR-based mutations had been acquired. After sequencing, PCR amplification with KFDF4s / KFD3UTRas resulted in a ~1.9 kb band, which was isolated on an agarose gel and purified. Preparation for cloning into the RGS, included digestion of the PCR-acquired *Kas*1 and the unique-overlapping *Eco*R1 site between fragments three and four. After ligation, the *E. coli* Top10 cloning strain was switched to *E. coli* XL10-Gold, which is a cloning strain designed for larger plasmids, with the ability to produce high yields during plasmid harvesting. Confidence in the successful cloning procedure was gained by colony-PCR screening with KFD12s / KFD12as; this gave a band of ~534 bp. Verification of the clone and now full-length KFDV genome was supported by sequencing and digestion of the ~16.1 kb product with *Pst*1, giving four bands (Figure 3.2.2: 10.3, 3.4, 2.8 and 0.4 kb). This full-length cDNA genome clone was designated "KFDV RGS".

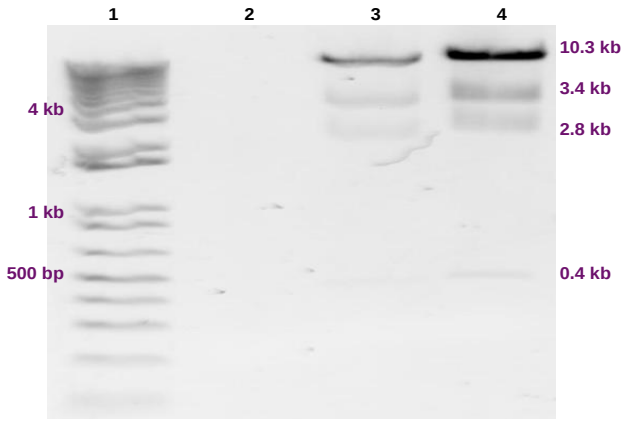


Figure 3.2.2: Restriction digestion of KFDV RGS with *Pst*1.

Lane 1: 1 kb+ DNA ladder (Invitrogen), Lane 2: Sterile water and loading dye, Lane 3: 5 µl KFDV RGS, Lane 4: 12 µl KFDV RGS, on a 1% agarose gel.

Despite PCR amplification of each fragment comprising the full-length KFDV cDNA genome from the sequencing vectors, only one PCR-induced mutation had occurred. This base-pair replacement C 978 T was within fragment one of the KFDV RGS clone and was corrected back into a cytosine deoxyribonucleotide by sub-cloning. The unique restriction sites *Nhe*1 and *Sgr*A1, within F1-pCR4-TOPO sequencing vector and the RGS were utilized. The ~2.7-kb mutated segment was replaced by the same size piece of DNA, sequencing of the construct, assured that the clone was in agreement with the sequenced KFDV genome.

The final aspect of the KFDV RGS cloning strategy was the addition of the HDV ribozyme, which needed a few preparation steps before hand, summarized in Figure 3.2.4. In anticipation of the HDV ribozyme sequence located at the immediate boundary of the full-length KFDV genome, the fourth fragment / 3' UTR was subcloned out of the RGS and inserted into a pTM1 vector. Cloning was verified by restriction digestion using *Eco*R1 and *Pac*1. The KFDV-F4 subclone produced two bands (~1935 and 5290 bp) and the pTM1 control also had two bands (54 and 5290 bp) (data not shown). The HDV sequence was amplified from a previously made clone (RGS for Ebola-Zaire) which was used as the PCR template (kindly provided by Dr. S. Theriault). Amplification was executed with the addition of a *Kas*1 and *Pac*1 restriction site on the sense and anti-sense primers, respectively. Gel electrophoresis on a 2% agarose gel revealed a PCR product of ~103 bp, shown in Figure 3.2.3.

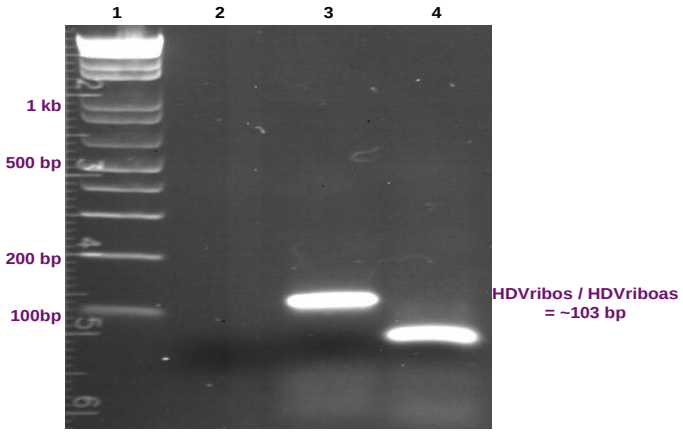


Figure 3.2.3: PCR amplification of the HDV ribozyme.

Lane 1: 1 kb+ DNA ladder (Invitrogen), Lane 2: Sterile water and loading dye, Lane 3: HDV ribozyme, Lane 4: PCR negative control of HDV ribozyme, on a 2% agarose gel.

This ribozyme fragment was extracted, purifed and cloned into the KFDV-F4 subclone using the added enzyme cut sites. Digestion with *Kas*1 and *Pac*1 gave two bands for both KFDV-F4-ribo subclone and KFDV-F4 subclone (control): ~7, 200 and 98 bp; ~7, 200 and 24 bp, seen on a 2% agarose gel. The KFDV-F4-ribo subclone was sequenced to authenticate the ribozymes' sequence. Original cloning of the KFDV cDNA genome into the pTM1-Adapter was facilitated by *Kas*1, since the ribozyme cleaves its immediate 5' end and was cloned into the RGS via *Kas*1 on its 5' end. This may result in the *Kas*1 recognition site becoming incorporated into the KFDV RNA when expressed by T7 RNA polymerase. Any non-viral nucleotides may inhibit efficient rescue of KFDV, much like the situation in previous RGS systems including: *Rhabdoviridae, Paramyxoviridae, Filoviridae, Bunyaviridae* and *Arenaviridae* (Kawaoka 2004). A deletion of the *Kas*1 recognition site (5' GGCGCC 3') would fuse the 3' UTR with the ribozyme and allow for authentic viral RNA to be transcribed. Circular mutagenesis re-created the KFDV-F4-ribo subclone clone without the *Kas*1 site. Sequencing of the clone demonstrated no PCR-induced mutations and a successful deletion of the *Kas*1 site. The last step involved subcloning of the F4-3' UTR-ribozyme into the RGS using *Eco*R1 and *Pac*1 restriction sites, Figure 3.2.4. Colony-PCR screening revealed the successful subcloning. Digestion using *Sfi*1 and *Pac*1 gave two bands (~281 and ~15, 879 bp) and the RGS control showed two bands as well (~216 and ~15, 879 bp), (Figure 3.2.5). The larger bands were un-distinguishable on the gel due to the high gel percentage which favours separation of smaller bands. However, the main focus was the smaller bands which indicated the addition of the ribozyme. Sequencing of the entire clone gave the final verification of the ~16, 160 bp KFDV RGS-ribo clone.

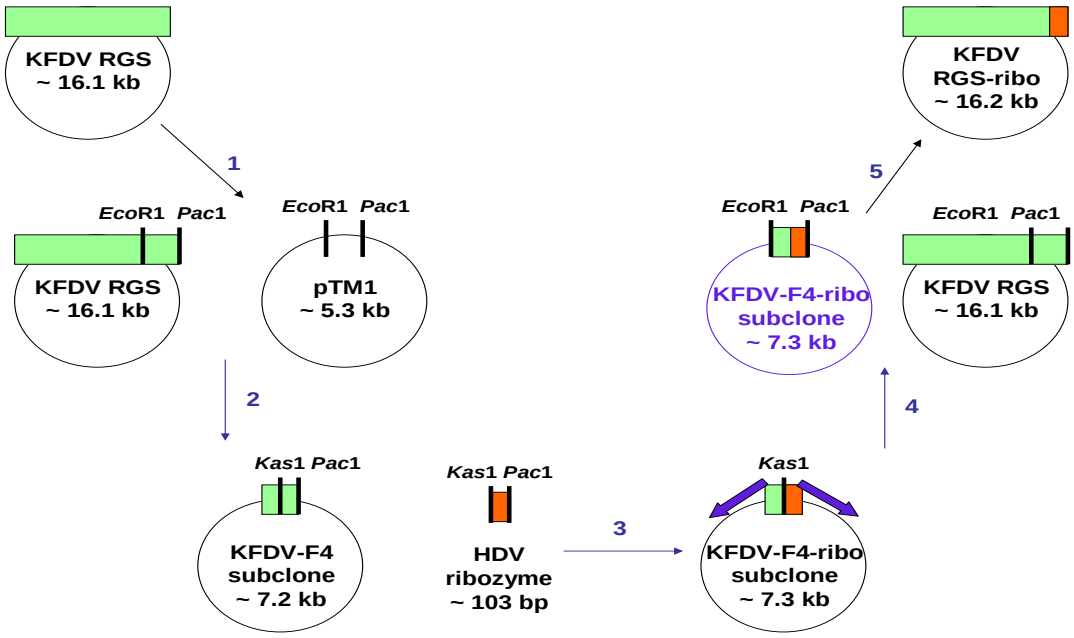


Figure 3.2.4: The strategy for creating KFDV RGS-ribo.

1) Subcloning of KFDV fragment 4-3' UTR into pTM1 vector. 2) Cloning of PCR amplified HDV ribozyme into the KFDV-F4 subclone. 3) Site-directed mutagenesis to remove the *Kas*1 restriction site and regenerate the KFDV F4-ribo subclone. 4) Subcloning of the fragment 4-3' UTR-HDV ribo segment into the KFDV RGS. 5) The final product, KFDV RGS-ribo.

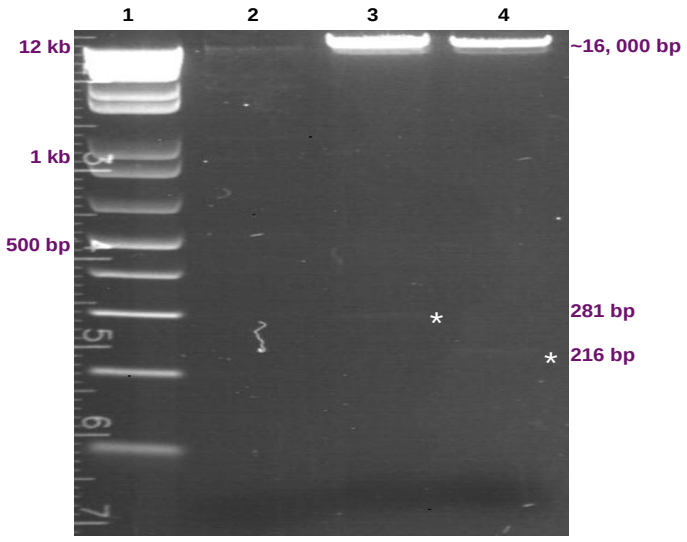


Figure 3.2.5: Restriction digestion of KFDV RGS-ribo with *Sfi*1 and *Pac*1.

Lane 1: 1 kb+ DNA ladder (Invitrogen), Lane 2: Sterile water and loading dye, Lane 3: KFDV RGS-ribo, Lane 4: KFDV RGS control, on a 2% agarose gel. The 281 and 216 bp bands are indicated by the asterisks on the right side of each band.

3.3. KFDV helper proteins

Construction of the ten virally-encoded protein clones of KFDV began, first with the mapping of each gene based on previously proposed protease recognition sites (Grard, Moureau et al. 2007). Amplification of each gene was performed using primer sets synthesized with restriction sites on the 5' ends from the KFDV cDNA template. Accompanying the restriction sites was the addition of either a start (ATG) or stop (TAA) codon added according to the sense or anti-sense oligonucleotides. Eight of the ten genes: C (~306 bp), prM (~510 bp), E (~1506 bp), NS1 (~1068 bp), NS2B (~399 bp), NS3 (~1881 bp), NS4A (~396 bp) and NS4B (~774 bp), were amplified and cloned into the eukaryotic expression vector pCAGGS-MCS, with *Sac*1 and *Kpn*1. The last two genes: NS2A (~705 bp) and NS5 (~2727 bp), were amplified and cloned using *Cla*1 and *Kpn*1, as *Sac*1 had cut sites within these genes. All ten clones were confirmed by digestion in which each of the entire genes cloned in were cut out of the vector, as shown in Figure 3.3.1. Sequencing confirmed the helper protein clones and their respective start and stop codon additions. The coding regions of each protein will be in frame as each protein is translated into a polyprotein, as in the native Flavivirus genome. The helper proteins will offer another tool for studying KFDV and its life cycle including: the requirement of a positive stranded RC complex and for complementation studies to assess the functional roles of the viral proteins.

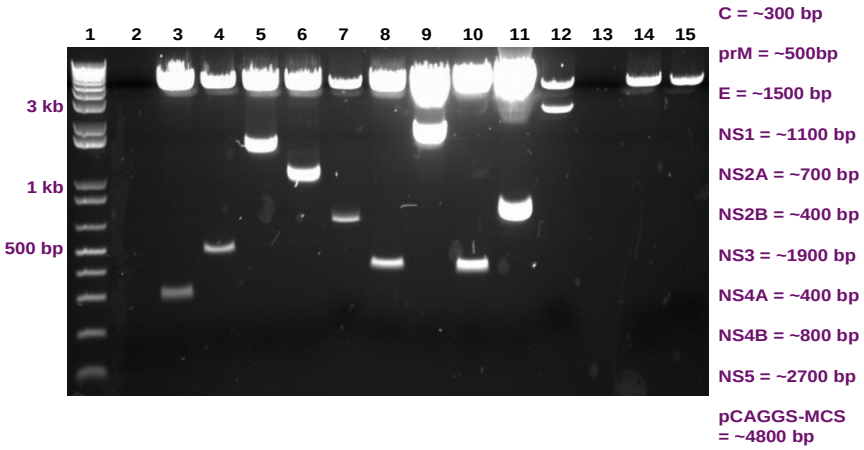


Figure 3.3.1: Restriction digestion of the KFDV genes that encode the helper proteins cloned into pCAGGS-MCS with *Sac*1 and *Kpn*1 for C, prM, E, NS1, NS2B, NS3, NS4A and NS4B. The NS2A and NS5 proteins were digested with *Cla*1 and *Kpn*1.

Lane 1: 1 kb+ DNA ladder (Invitrogen), Lane 2: Sterile water and loading dye, Lane 3: C protein, Lane 4: prM, Lane 5: E, Lane 6: NS1, Lane 7: NS2A, Lane 8: NS2B, Lane 9: NS3, Lane 10: NS4A, Lane 11: NS4B, Lane 12: NS5, Lane 13: Space, Lane 14: pCAGGS-MCS control digested with *Sac*1 and *Kpn*1, Lane 15: pCAGGS-MCS control digested with *Cla*1 and *Kpn*1, on a 1% agarose gel.

3.4. KFDV Rescue Attempts

Our first attempts to rescue infectious KFDV particles entirely from cloned cDNA used Vero E6 mammalian cell culture, the KFDV RGS construct, a T7 RNA polymerase clone and different combinations of the rescue proteins as listed in Table 3.1 according to the transfection procedure outlined in section 2.9.

Table 3.1: Description of the Kyasanur Forest Disease Virus rescue

Wells	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9 (-)**
Support plasmids (1µg)	NS3/NS5 / NS2A/T7 RNApol	NS3/NS5 / NS2A/C/ T7 RNApol	NS1/NS2 A/ NS3/NS4 A/ NS5/T7 RNApol	NS1/NS2 A/ NS3/NS4 A/ NS5/C/T 7 RNApol	NS1/NS2 A/ NS3/NS4 A/ NS4B/NS 5/ T7 RNApol	NS1/NS2 A/ NS3/NS4 A/ NS4B/NS 5/ C/T7 RNApol	All 10 virus proteins/ T7 RNApol	T7 RNApol	-
RGS construct (1µg)	cDNA-pTM1	cDNA-pTM1	cDNA-pTM1	cDNA-pTM1	cDNA-pTM1	cDNA-pTM1	cDNA-pTM1	cDNA-pTM1	-
First Rescue Results (+/-)	+	+	+	+	+	+	-	+	-
Second Rescue Results (+/-)	+	+	+	+	+	+	+	+	-

After 4-5 days, any virions present, will have egressed into the extracellular environment. Blind passage in CL-4 will remove these infectious KFDV particles from each well and they will be used to infect a fresh monolayer of Vero E6 cells in T-25 culture flasks.

Incubation for up to 14 days may be necessary to see any signs of cytopathic effect (CPE) (Theriault, Groseth et al. 2004). Approximately 60 percent CPE was seen, up to and after 14 days post infection (data not shown due to technical limitation in CL-4). After CPE were seen, which each T-25 flask was processed for RNA harvesting from the culture media and cell monolayer. Confirmation of the KFDV infection was obtained by RT-PCR. Positive results for infectious KFDV particles were obtained by RT-PCR using primers within the NS5 protein, KFD12s / KFD12as, which yielded an amplicon of approximately 534 bp, shown in Figure 3.4.1. The results indicate that KFDV was successfully rescued from cell monolayer samples one through six and from the media or supernatants from samples one, four, five, six and eight.

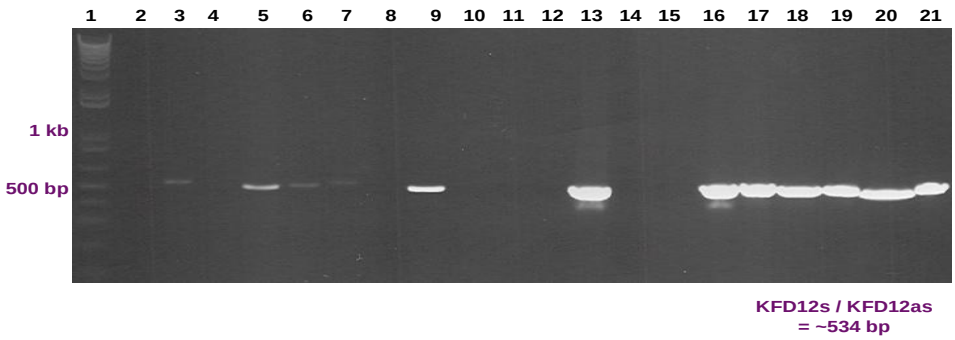


Figure 3.4.1: RT-PCR amplification of tissue culture supernatants and cell monolayers to confirm KFDV rescue.

Lane 1: 1 kb+ DNA ladder (Invitrogen),Lane 2: Sterile water and loading dye, Lanes 3-9: tissue culture supernatant samples 1, 2, 4, 5, 6, 7 and 8 respectively, Lane 10: negative control: tissue culture supernatant sample 9, Lane 11: Space, Lane 12: RT-PCR negative control, Lane 13: PCR positive control (KFDV RGS template), Lanes 14 and 15: Space, Lanes 16-21: Cell monolayer samples 1, 2, 3, 4, 5 and 6 respectively, on a 1% agarose gel. Tissue culture sample 3, cell monlayer samples 7, 8 and 9 were lost during sample processing and are not included on the gel.

A second rescue attempt involved BHK-21 (Baby Hamster Kidney) mammalian cells, instead of the previous Vero E6 cell culture. The switch in cell lines was based on the thought that different cell cultures and virus species, may result in variable visual CPE detection (Richman, Whitley et al. 2009). Perhaps more pronounced CPE may be noticed in the new cell line. Additionally the BHK-21 cell type was recommended by a colleague Dr. Michael Holbrook (personal communication). In contrast to the first rescue the KFDV cDNA genome had the HDV ribozyme added to the RGS construct (KFD RGS-ribo), thus two rescue parameters were altered. The transfection procedure was set up as described in section 2.9 and each well was transfected according to Table 3.1. After blind passaging, the processing of the supernatant (cell culture medium) was performed 14 days past the blind passaging of the culture media containing KFDV virus particles.

The CPE was more pronounced in the second rescue and developed faster than in the first rescue. Some samples developed CPE within six-seven days after blind passaging (data not shown). Positive wells were again determined by RT-PCR with primers KFD7s and KFD7as giving a product of approximately 793 bp. The results are shown in Figure 3.4.2.

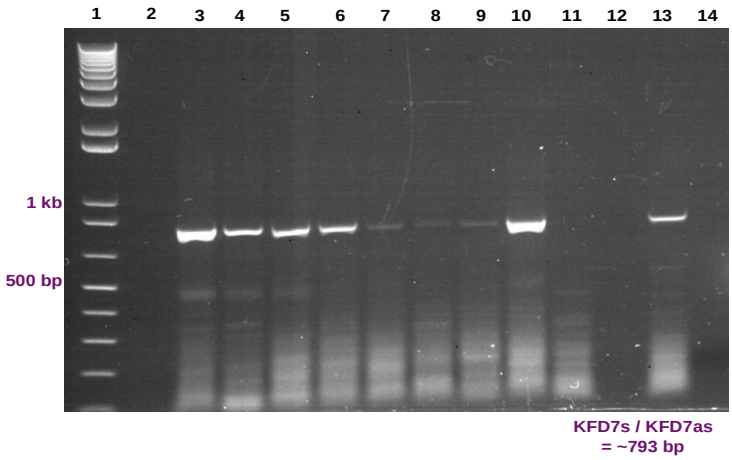


Figure 3.4.2: RT-PCR amplification of tissue culture supernatants to confirm KFDV rescue.

Lane 1: 1 kb+ DNA ladder (Invitrogen), Lane 2: Sterile water and loading dye, Lanes 3-10: tissue culture supernatant samples 1, 2, 4, 5, 6, 7 and 8 respectively, Lane 11: negative control: tissue culture supernatant sample 9, Lane 12: Space, Lane 13: PCR positive control (KFDV RGS-ribo template) and Lane 14: RT-PCR negative control on a 1% agarose gel.

Chapter 4: Discussion

Emerging infectious diseases, like KFDV, are an ever increasing threat to the global population. Hemorrhagic fever viruses have case fatality rates from 5-90% (Ebihara, Groseth et al. 2005) and may increase their continental distribution through travel and global shipment (Peters 2006). Additionally, Dengue Virus (DV), much like KFDV needs a forest cycle for virus maintenance. DV is now able to replicate to high enough titres in humans, termed the "Urban cycle", such that the forest cycle is no longer needed (Hanley and Weaver 2010). Therefore, the establishment of RGS for many emerging infectious viruses, allows a stable system to study these deadly pathogens. Furthermore, viral-host interplay, virus biology and the generation of potential therapeutics and/or vaccines, are possible for negative and positive-stranded viruses by researchers (Kawaoka 2004; Ebihara, Groseth et al. 2005).

The KFDV RGS was based on a previous model for an Ebola virus RGS using a transcription mechanism for rescue attempts in the eukaryotic host's cytoplasm (Theriault, Groseth et al. 2004). The transfection of the full-length genome complement of KFDV is under control of the bacteriophage T7 promoter and different combinations of the putative RC complex proteins of KFDV along with T7 RNA polymerase, under control of the eukaryotic chicken β-actin promoter was applied. The results from the first rescue indicated that only the full-length genome needed to be transcribed for virion formation. This primarily suggests that a RC complex is not necessary for rescue of KFDV using this system. Conversely KFDVs negative-stranded counterpart, Ebola virus, needs its RC complex equivalent (RNP complex) for successful genome replication and subsequent virion formation (Kawaoka 2004). Our result is consistent with numerous

reports that Flaviviruses genetic material undergoes translation first, followed by replication using the RC complex proteins (Harris, Holden et al. 2006; Howley and Knipe 2007; Villordo and Gamarnik 2009). Thus, KFDV is no exception and the addition of the possible RC complex proteins in rescue attempts is unnecessary. The Flavivirus virion is composed of structural proteins and does not bring any non-structural proteins with the virion into the infected cell. Therefore unlike, the negative-strand viruses, no RC/RNP is brought in with the nucleocapsid when it enters the cytoplasm (White and Fenner 1994; Kawaoka 2004; Richman, Whitley et al. 2009). So the current results agree with the understanding of Flavivirus biology.

KFDV rescues were initially assessed by the production of CPE; further confirmation by molecular methods soon followed. CPE can be described as morphological changes that occur during viral infection in cell culture. Such alterations in cellular morphology can be visualized by a light microscope in many different formations: shrinking / rounding of cells, loss of adherence, cellular lysis and syncytia formation, just to name a few (Howley and Knipe 2007; Richman, Whitley et al. 2009). The first and second rescue attempts displayed CPE which was more pronounced in the latter. This had featured a new cell type (BHK-21) and the addition of a Hepatitis Delta Virus (HDV) ribozyme. Despite the fact that Flaviviruses can show CPE in Vero E6 and BHK-21 cells, the variability in CPE production by different Flavivrus species and cell types has been documented (Richman, Whitley et al. 2009). Therefore, the thought of switching to the BHK-21 cell line seemed reasonable and the system did display more pronounced CPE. The addition of the HDV ribozyme has been previously used for other RGS systems. The auto-catalytic RNA, when placed immediately following the KFDV

RNA genome, will precisely cleave from the KFDV RNA. This will avoid any non-viral ribonucleosides from the KFDV RNA, which may inhibit viral replication (Ebihara, Groseth et al. 2005). Rescues of KFDV in both attempts demonstrate that all transfections worked successfully (Table 3.1, Figures 3.4.1 and 3.4.2). However, the increasing amounts of helper plasmids appear to be inhibiting the production or release of virions. Transfected wells (Table 3.1., 4-7, Figure 3.4.1., lanes 5-8; wells: 2-7 in Figure 3.4.2. lanes 4-9), show decreases in viral RNA based on RT-PCR reactions. However, this RNA level was assessed from the tissue culture medium. Conversely, when the cell culture was subjected to RT-PCR, the signal had similar intensities for wells 1-6. Such a result would indicate an inability to produce intact infectious particles or their limited ability to egress from the infected cells. Perhaps the result is due to an escalation in helper protein content in the latter transfected wells. Furthermore, tranfections using just the KFDV RGS-ribo and T7 RNA-pCAGGS-MCS clones (well 8) appeared to be the most successful for rescuing KFDV infectious particles (Figure 3.4.1., lane 9 and Figure 3.4.2., lane 10). The RT-PCR signals for the extracellular virions present in the cell culture medium were the strongest, especially when compared to the other transfected wells.

The production of KFDV virions demonstrates that the RNA polymerase II-dependent RGS system was successful. The KFDV genome appears to be authentic, at least with respect to the requirements for translation and the replication aspects of the Flavivirus life cycle. The evidence of genome authenticity, besides successful KFDV rescue, is further exemplified by the sequence analysis. The cyclization sequences in the 5' and 3' UTR (CS-A regions) were present in KFDV, Figure 3.1.7. B and C. The stem-loops and the CS-A sequences were very similar to those documented for TBEV-

Neudoerfl strain, Figure 3.1.7.A and (Kofler, Hoenninger et al. 2006). Phylogenetic analysis of the KFDV full-length genome was compared with other Mosquito-Borne (MB) and Tick-Borne (TB) flaviviruses. KFDV is clustered with AHFV within the TB flavivirus cluster and the MB flavivirus form their own cluster (Figure 3.1.8). This result agrees with previously reported phylogenetic analysis for Flaviviruses (Lin, Li et al. 2003; Howley and Knipe 2007).

Chapter 5: Conclusion

5.1 KFDV Conclusions

The KFDV isolate, in summation, is very similar to the published polyprotein sequence. Based on sequencing data, there were ten base pair polymorphisms and four subsequent amino acid residue differences. Secondary structures in the UTR regions of the RNA genome were determined to be similar to that of TBEV-Neudoerfl strain; both are in the same antigenic complex. Phylogenic analysis of full-length Flavivirus genomes dictates that, KFDV forms a distinct clade within the TBE serocomplex with AHFV (a KFDV variant). Furthermore, infectious KFDV particles were rescued entirely from cloned cDNA using the RGS system.

5.2 Future Work

Some of the future applications of reverse genetics can include the understanding of adsorption, replication, maturation and release of viruses in a real-time infection. All of theses aspects of virology could be facilitated by various cell culture and animal

models. The generation of helper proteins can be used in complementation studies, in which different genes in the KFDV genome could be deleted and assayed for their importance during the viral life cycle. A study by Jones et al. in 2005 (Jones, Patkar et al. 2005), using YFV, incorporated gene deletions and assessed the ability for YFV replicons to assemble virus particles without the corresponding protein products: C, NS1 and NS3. Removal of the C protein completely stopped packaging; the lack of the NS3 protein impaired assembly as well. The Jones study concluded that, despite the apparent lack of involvement of the NS1 protein in assembly, all of the NS proteins are likely involved in assembly, at some level (Jones, Patkar et al. 2005; Leung, Pijlman et al. 2008). This is important because it is thought that replication is coupled with virus assembly, showing dual roles of the NS proteins (Jones, Patkar et al. 2005; Harris, Holden et al. 2006; Leung, Pijlman et al. 2008). Researchers can even go as far as studying functional relatedness of viral proteins from closely and distantly-related viruses. Theriault et al. 2004 (Theriault, Groseth et al. 2004), replaced different combinations of viral proteins involved in replication of Ebola Virus-Zaire (ZEBOV). The authors elucidated that in general, proteins from the related Ebola Virus-Reston (REBOV) had a better rescue efficiency than did any combinations from Marburg Virus-Lake Victoria strain (MARV) (Theriault, Groseth et al. 2004). In relation to KFDV, a similar study could be performed. Functional replication studies with KFDV and other closely-related tick-borne viruses such as: Omsk Hemorrhagic Fever Virus (OHFV), Alkhurma Hemorrhagic Fever Virus (AHFV) and/or Tick-Borne Encephalitis Virus (TBEV) and with more distantly-related viruses could include: Yellow Fever Virus (YF), Dengue Fever Virus (DV), Japanese Encephalitis Virus (JEV) can be employed with the RGS. Another area of study can be

similar to a study by Cristea et al. 2006 (Cristea, Carroll et al. 2006). In this study, these authors looked at the replication complex of Sindbis Virus from the *Alphaviridae*. Using a Green Fluorescent Protein (GFP) fusion with a known protein necessary for replication in Sindbis Virus, once fluorescence was detected, cells were harvested at two hour intervals. Cryogenic lysis was used and the protein-replication complexes were processed using Sodium-Dodecyl-Sulfate Polyacrylamide Gel Electrophoresis (SDS-PAGE). From here, the researchers could determine which proteins, viral and/or host, formed the viral replication complex and their location inside infected cells (Cristea, Carroll et al. 2006).

Vaccine development is a very attractive area of study from a public health perspective. With that said, there are many types of methods for vaccine development. Of course, all could be applicable to KFDV; examples of such methods can be seen with Influenza viruses. Using reverse genetics "6:2 reassortments" are used to combat the seasonal flu. How this works is antigenic drift and antigenic shift will alter the current hemagglutinin (HA) and neuraminidase (NA) and reverse genetics can use the new or this season's HA/NA to replace the old or last season's HA/NA. Simply put, each year the Influenza Virus keeps its regular backbone of "6 segments" and just the determinants of infection (HA/NA) "2 segments" are updated, all using reverse genetics to cause virus reassortments (Kawaoka 2004). However, KFDV is different than Influenza Virus because the genome is non-segmented. Despite the genome differences, the system can still be utilized. Attenuation of viruses is another type vaccination approach. Attenuation can be achieved by performing single-gene mutations causing temperature-sensitive (ts) viruses. These mutations can give a restricted growth characteristic *in vitro* at 37^{o}C and hopefully allow the body to mount a strong immune response, when the virus is present at

human physiological body temperature (Kawaoka 2004). Some viruses are able to antagonize and avoid the immune response (IR) in the cell and/or body. This gives the pathogen a competitive advantage in establishing an infection and increasing chances of survival by delaying the innate IR (Goodbourn, Didcock et al. 2000; Kawaoka 2004). From an attenuation stand point, determining which part of the gene involved in this antagonism can be exploited by reverse genetics. Once a virus is deficient in abolishing the IR, the body will be primed for a strong response when challenged with a wild-type (wt) virus. Chances of the body clearing the invading organism would be increased (Kawaoka 2004). Of course with any technology, public health concerns must be addressed; reverse genetics systems are no exception. The ability to generate recombinant viruses with various phenotypes can be beneficial, in terms of vaccination strategies and understanding viruses as a whole. But, this technology, in some instances, may be detrimental as undesired phenotypes can be created. This may adversely affect public health. For example, a study by Kobasa et al. 2004 (Kobasa, Takada et al. 2004), re-created the 1918 Influenza Virus, commonly known as the "Spanish flu". While the focus of the study was to determine the reasons for the high pathogenicity and case fatality ratio of the 1918 virus, it showed how simple a revival of such a deadly virus can be using reverse genetics. However, with public and governmental regulation and sanctions on an international scale, this concern hopefully, can be avoided. Only in a research setting should this ingenuity be used to continue to broaden our knowledge on pathogens concerning public health (Kawaoka 2004; Kobasa, Takada et al. 2004).

Futhermore, RGS systems can be used for determination of biological properties of various Flaviviruses within a Containment Level-2 (CL-2) facility (Gehrke, Ecker et

al. 2003; Yoshii and Holbrook 2009). Virus-like particles (VLPs) generated from cDNA clones of TBEV through chinese hamster ovary cells (CHO), will resemble the natural infectious version. Two separate clones will drive the formation of the VLPs: a cDNA clone with deletions of the prM and E genes in the entire genome of TBEV; another plasmid encoding the TBEV deleted genes, after transfection would recombine into the genome of the cell line. The constitutively expressed surface proteins, prM-E, within the cell line will allow for the TBEV particles to assemble and be released. Once these virions are passaged into a fresh cell line without the integrated prM and E genes, no cytopathic effect (CPE) will be seen. As these virions will carry a genome with the two deleted genes, they will be unable to form functional virions, unless a second vector carrying the deleted genes is supplied. These VLPs will never be able to become infectious and initiate new rounds of infection (Gehrke, Ecker et al. 2003). A similar experiment was performed with Omsk Hemorrhagic Fever Virus by Yoshii and Holbrook in 2009. Reporter genes can be cloned in place of the structural genes, analyzing VLP characteristics during an infection is possible with this system (Yoshii and Holbrook 2009). An important point to note is the possibility of recombination or reversion events. No spontaneous reversion occurred in this study (Yoshii and Holbrook 2009) and recombination of Flaviviruses in natural settings has never been reported (Yoshii and Holbrook 2009).

To our knowledge, this *in vivo* Reverse Genetics System (RGS) is the first of its kind for Flaviviruses. In addition, this RGS may pave the way for research opportunities of the other Flaviviruses which are enormous threats to public health. A new vaccination approach may also be studied, as new cases of KDFV infection appear to be increasing

despite routine vaccination (Pattnaik 2006). The threat of the urban cycle in Dengue Virus (DV) (Hanley and Weaver 2010) infected areas adds an increased risk to people. DV would be using humans as reservoir for virus amplification, which is significant because a DV virus infection can be life-threatening. An added concern is the introduction and maintenance of the first documented TB flavivirus in North America, Powassan virus (POWV) (Howley and Knipe 2007; Hinten, Beckett et al. 2008). Consequently, if such events became possible for KFDV, the at risk population would increase and KFDV could become a global health concern. Therefore, the importance of better understanding pathogenesis and virus-host relationships, at the molecular level, is essential for the development of effective therapeutics. Thus, the above describes the importance of the establishment of the RGS for KFDV.

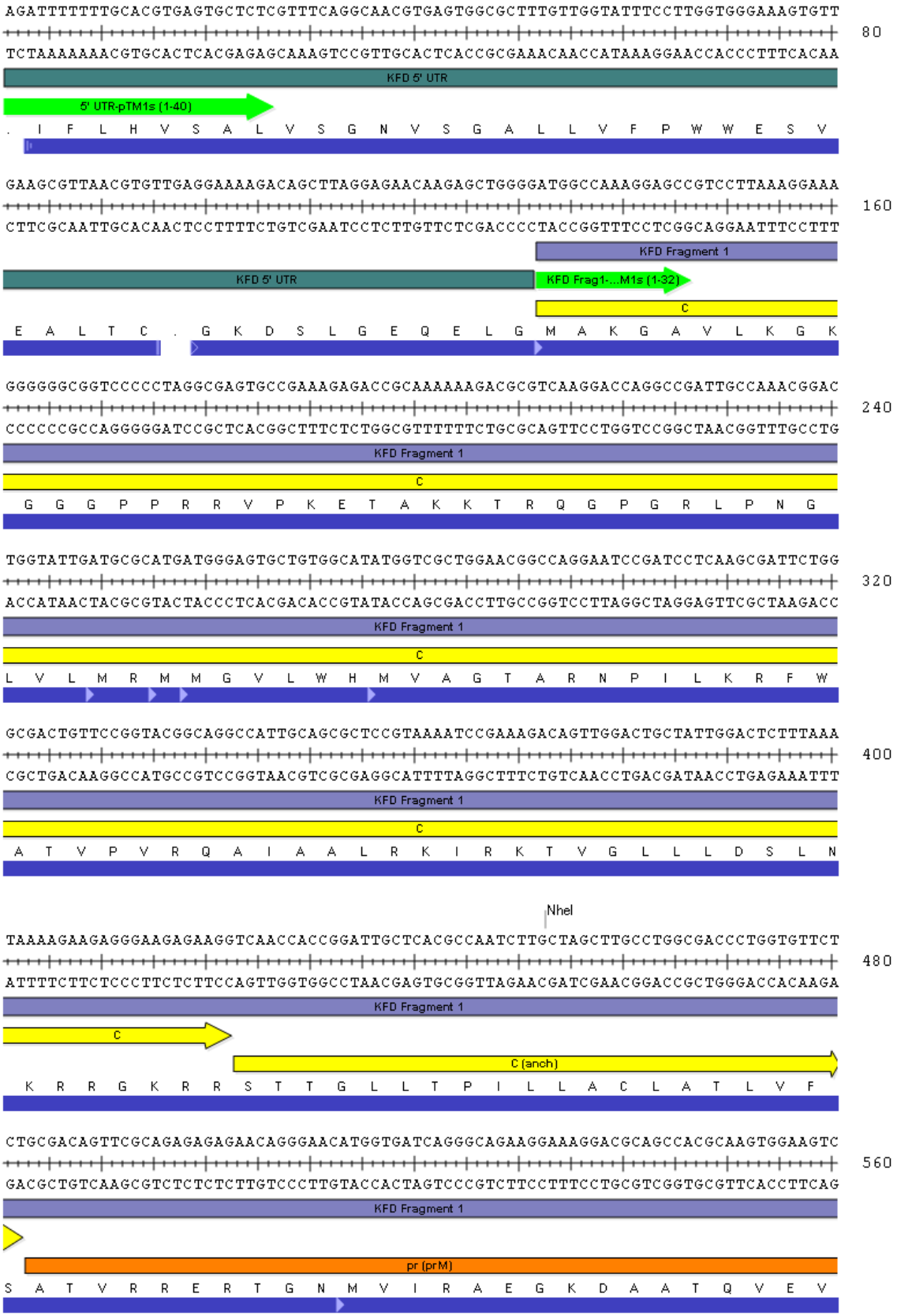
Appendix:

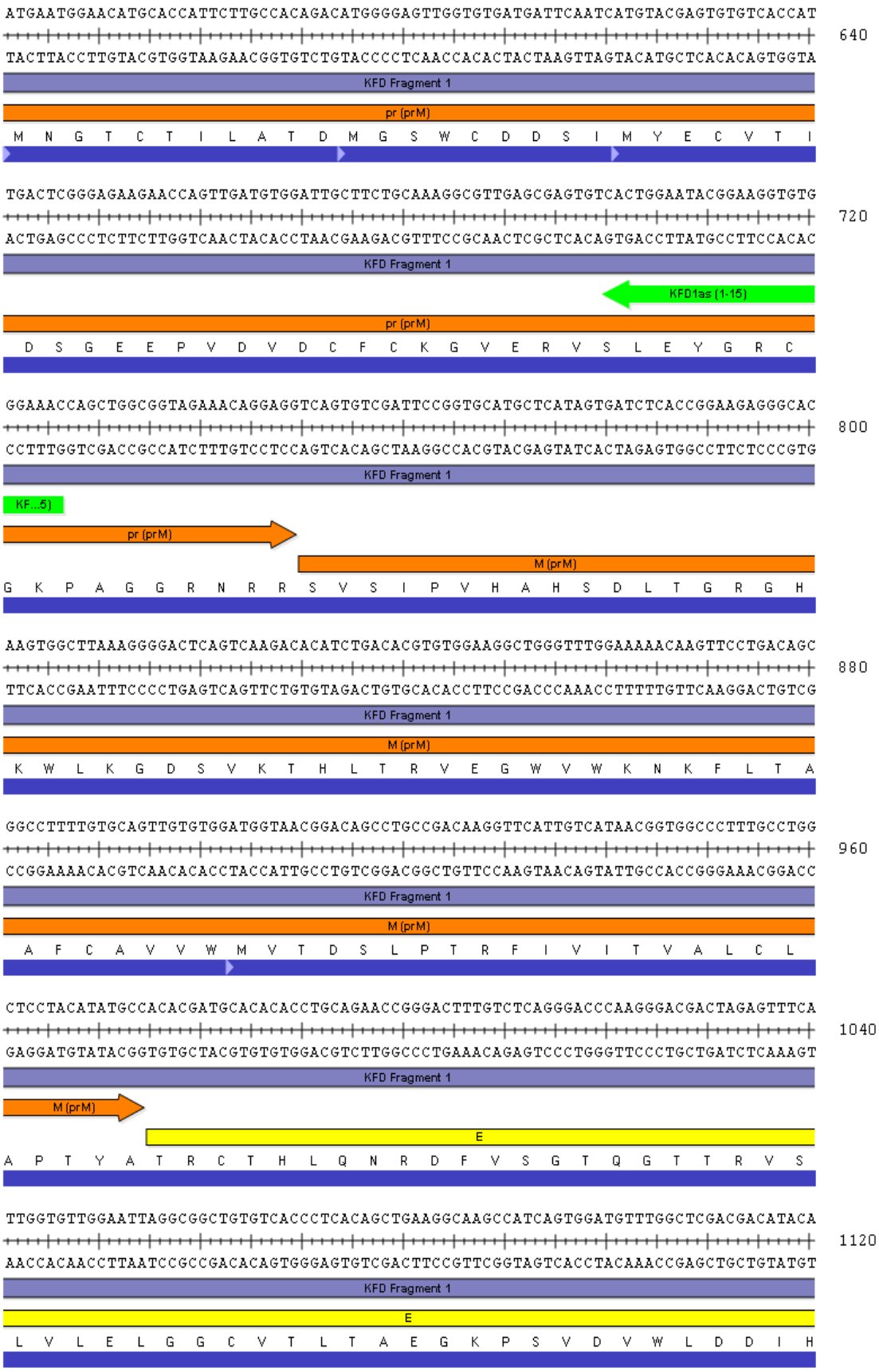
Table A.1: The Primer sequences used for KFDV genome sequencing and for the generation of the RGS for KFDV

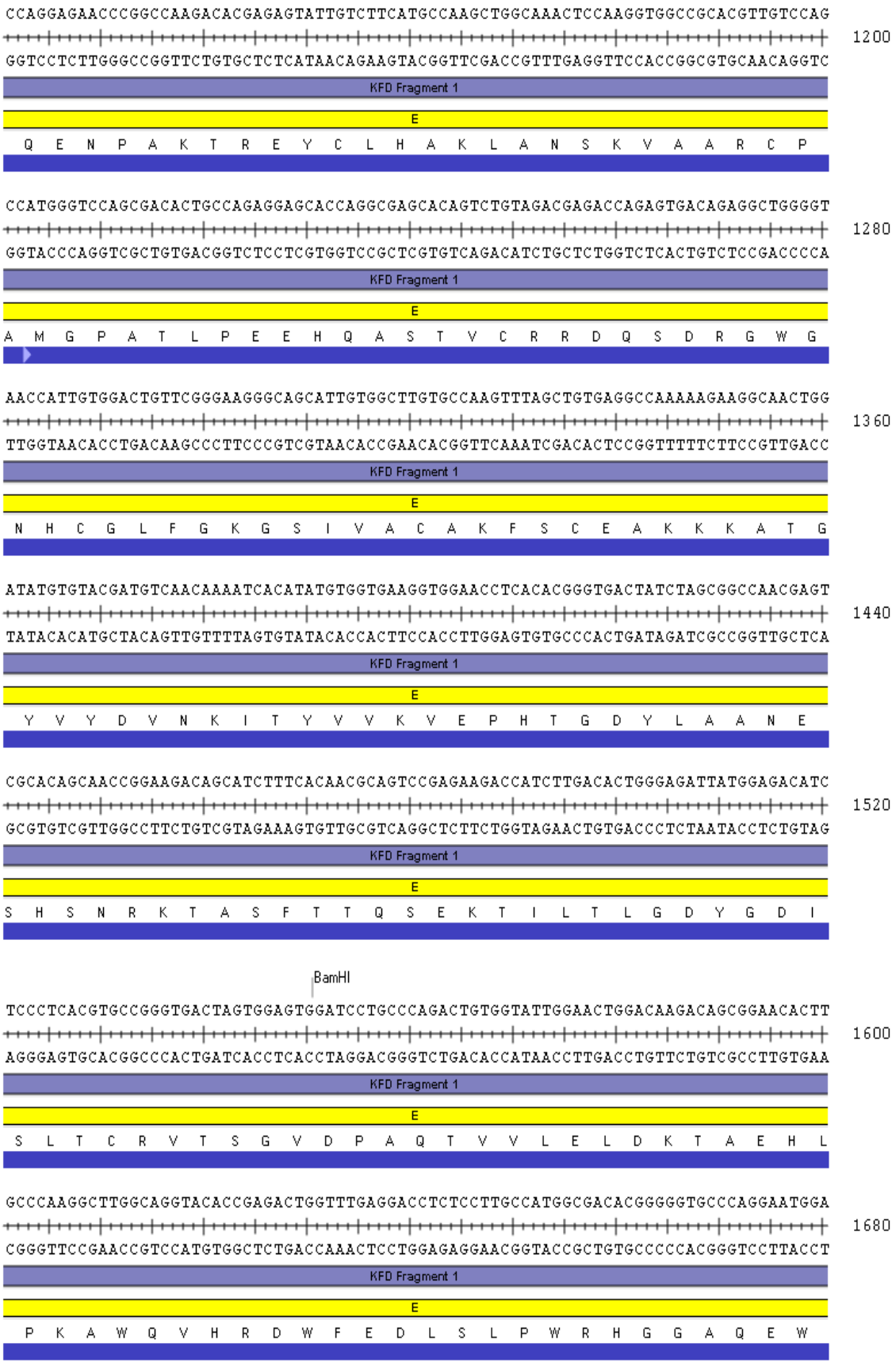
Position	Name	Sequence	Restriction Site / Notes
KFD 1-1	KFD1s	GCGGTCCCCCTAGGCGAGTGC	
KFD 1-2	KFD2s	CCCTTTGCCTGGCTCCTACATA	
KFD 1-3	KFD3s	ACCTGCGATGTGGGCCTTGAAAA	
KFD 1-4	KFD4s	TCTGTCCTTGCGGCATCTCTAAAA	
KFD 1-5	KFD5s	CCTGGGACTGCGGTTAAGATTGAC	
KFD 1-6	KFD6s	TTGGAGGCGAGGGAGTGGGTCTTA	
KFD 1-7	KFD7s	GGGGTATGGAGCCAAGGGAGTTC	
KFD 1-8	KFD8s	CGGAGAGTGGAGAGTGGGCTTTGA	
KFD 1-9	KFD9s	TTGCAGCCGGTGTGGAAGGACT	
KFD 1-10	KFD10s	GCAGAGGGCGCACAGGGTAT	
KFD 1-11	KFD11s	GCAGAGGCGGGTGGTCATACTACG	
KFD 1-12	KFD12s	TGCGGGAAAGAAAACACCACGACT	
KFD 1-13	KFD13s	AAATGACATGGCCAAGGTGAGAAA	
KFD 1-14	KFD14s	GGCCCCGAGAGGTATGCTGATTAT	
KFD 1-15	KFD1as	GTTTCCCACACCTTCCGTATTCCAGTG	
KFD 1-16	KFD2as	CTCGGACTGCGTTGTGAAAGAT	
KFD 1-17	KFD3as	ATGCCGCCTACCGAACCGAAGTC	
KFD 1-18	KFD4as	CACTCCCGTTGCATTCCTGT	
KFD 1-19	KFD5as	CCAGCACGAGGAGCAAGAAATAGG	
KFD 1-20	KFD6as	ACCAACCCTGCAAGAAGCCATAAC	
KFD 1-21	KFD7as	ATGCCCTCTTGCCGCGATACTGTG	
KFD 1-22	KFD8as	AGGTCCAGCTCCGATCCAACACAT	
KFD 1-23	KFD9as	CACCACTTGCCACTGAACTG	
KFD 1-24	KFD10as	ATGGCCTTTCCCTCCTATCGTGTA	
KFD 1-25	KFD11as	AAAGGCCGTTGTGTCCGTCAT	
KFD 1-26	KFD12as	GTTCGGTGTTCGCCTTCCAGATAA	
KFD 1-27	KFD13as	CCTGCGGGACCCAATCAAT	
KFD 1-28	RNA3Adapter	TTTTTTTTTTTTTTTTTTTTTTTTT3-dT-Q	
KFD 1-29	Adapter primer	AAAAAAAAAAAAAAAAAAAAAAAAA	
KFD 1-30	pTM1 adaptor-s	gtacCCATGGACCGGTgactagcaGGCGCCAGGCCTgtac	NcoI-AgeI-linker-KasI-StuI
KFD 1-31	pTM1 adaptor-as	gtacAGGCCTGGCGCCtgctagtcACCGGTCCATGGgtac	NcoI-AgeI-linker-KasI-StuI rev
KFD 1-32	KFD Frag1-pTM1s	GTACaccggtATGGCCAAAGGAGCC	Add 5' AgeI (Adaptor)
KFD 1-33	KFD Frag1-pTM1as	GTACggcgccCCTTCCCGCTTTCCG	Add 3' KasI (Adaptor)
KFD 1-34	KFD Frag2-pTM1s	TGGCAGAGTTTGGGATGGGAATGA	
KFD 1-35	KFD Frag2-pTM1as	GTACggcgccAACCGCCCCACCTTTTG	Add 3' KasI (Adaptor)

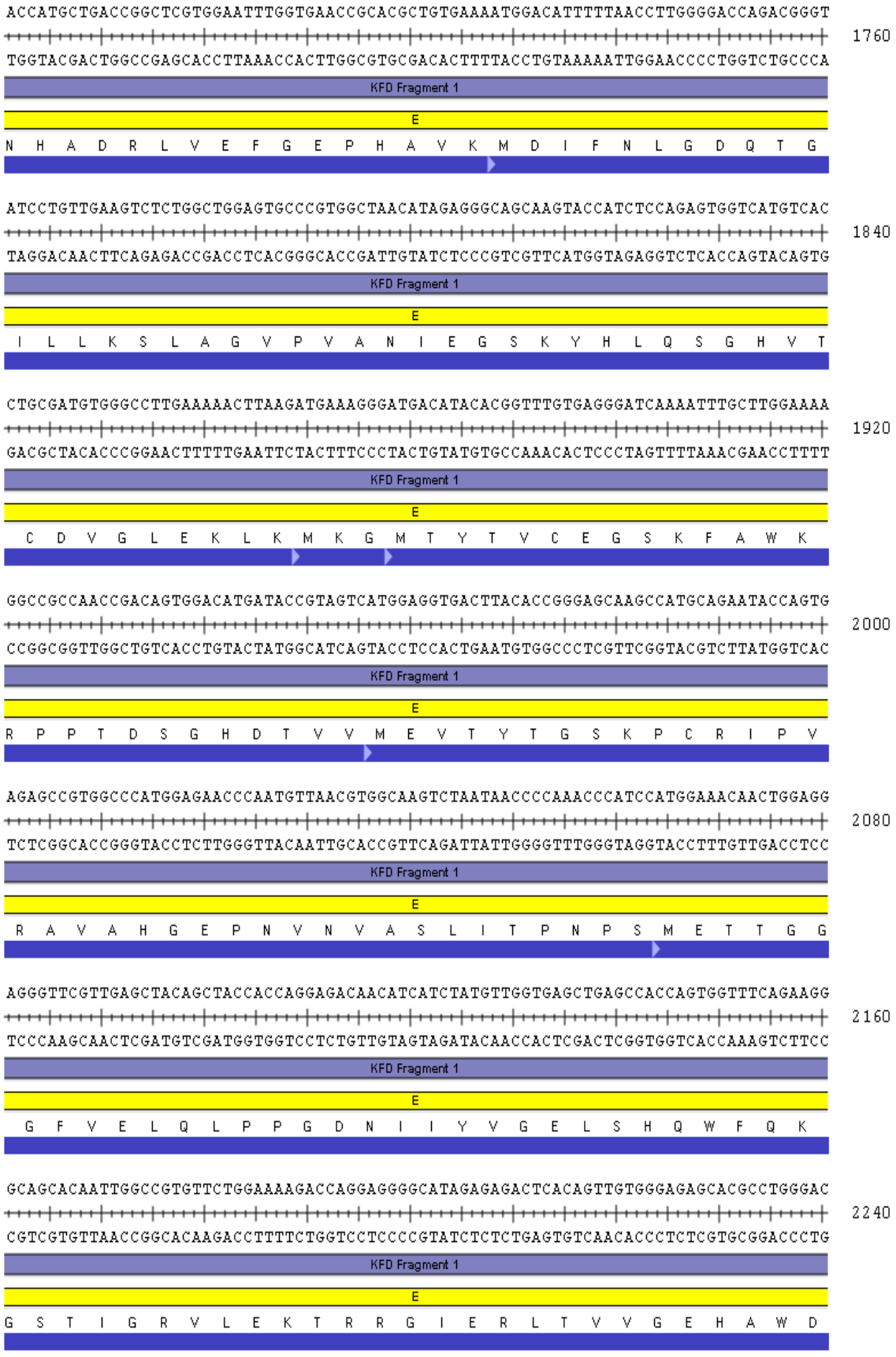
KFD 1-36	KFD Frag3-pTM1s	CCATGCCACGTACGTCCACAGAAG	
KFD 1-37	KFD Frag3-pTM1as	GTACggcgccCAAAACTCGGGGTCCTCC	Add 3' KasI (Adaptor)
KFD 1-38	KFD Frag4-pTM1s	TGCGGGAAAGAAAACACCACGACT	
KFD 1-39	KFD Frag4-pTM1as	GTACggcgccCAGCTTCCCGTCACCAAATG	Add 3' KasI (Adaptor)
KFD 1-40	KFD 5' UTR-pTM1s	GTACaccggtAGATTTTTTTGCACGTGAGTGCTCTC	Add 5' AgeI (Adaptor)
KFD 1-41	KFD 3' UTR-pTM1as	GTACggcgccAGCGGATGTTTTTTCCGAAAC	Add 3' KasI (Adaptor)
KFD 1-42	KFDvpVirCs	GTACgagctcATGGCCAAAGGAGC	Add SacI
KFD 1-43	KFDvpVirCas	GTACggtaccTTACCTTCTCTTCCCTCTTC	Add KpnI
KFD 1-44	KFDvpprMs	GTACgagctcATGGCGACAGTTCGCAGAGAGAGAACAGG	Add SacI
KFD 1-45	KFDvpprMas	GTACggtaccTTAGGCATATGTAGGAGCCAGGCAAAGGG	Add KpnI
KFD 1-46	KFDvpEs	GTACgagctcATGACACGATGCACACACCTG	Add SacI
KFD 1-47	KFDvpEas	GTACggtaccTTAAGCACCAACCCCCAGTGTCATTGTC	Add KpnI
KFD 1-48	KFDvpNS1s	GTACgagctcATGGATATGGGCTGTGCAATTGATGCTAACAGGATG	Add SacI
KFD 1-49	KFDvpNS1as	GTACggtaccTTAGTCAGCCAGCACCATCGACCTCACCAGA	Add KpnI
KFD 1-50	KFDvpNS2As	GTACatcgatATGAACGGAGCCATGCTAAGTGAAGGTGGAGTC	Add ClaI
KFD 1-51	KFDvpNS2Aas	GTACggtaccTTACCGTCTGTTCCTTCGCTCGGCAAGCT	Add KpnI
KFD 1-52	KFDvpNS2Bs	GTACgagctcATGTCTTTCAGTGAACCACTGACGGTGGTG	Add SacI
KFD 1-53	KFDvpNS2Bas	GTACggtaccTTACCTTCGCCCCGAGCCCAGC	Add KpnI
KFD 1-54	KFDvpNS3s	GTACgagctcATGTCTGAACTTGTCTTCTCTGGA	Add SacI
KFD 1-55	KFDvpNS3as	GTACggtaccTTAGCGACGCCC	Add KpnI
KFD 1-56	KFDvpNS4As	GTACgagctcATGAGTGTTGGTGATGTTTTGGGTGGTCTGGC	Add SacI
KFD 1-57	KFDvpNS4Aas	GTACggtaccTTATCGCTGTTTTCCCGGCTCTGGCTG	Add KpnI
KFD 1-58	KFDvpNS4Bs	GTACgagctcATGAACGAAATGGGCATGTTGGACAAGACAA	Add SacI
KFD 1-59	KFDvpNS4Bas	GTACggtaccTTATCTGCGTGTTCCAGTTGTCCTAAGCCAC	Add KpnI
KFD 1-60	KFDvpNS5s	GTACatcgatATGGGAGGTGCCGAGGGAGAGACTCTTGG	Add ClaI
KFD 1-61	KFDvpNS5as	GTACggtaccTTAGATGATATTGCTTTCCAACTTCAGGTCCCAGTG	Add KpnI
KFD 1-62	pCAGGSs	TGCTGGTTGTTGTGCTGTCTCATCATTTTG	Sequencing Upstream MCS
KFD 1-63	pCAGGSas	ATGTCCCCATAATTTTTGGCAGAGGGA	Sequencing Downstream MCS
KFD 1-64	pTM1MCSs	GACGTGGTTTTCCTTTGAAAAACACGATAATA	Sequencing Upstream MCS
KFD 1-65	pTM1MCSas	TAGCAGCCGGATCGTCGACTTAATTAATTAGG	Sequencing Downstream MCS

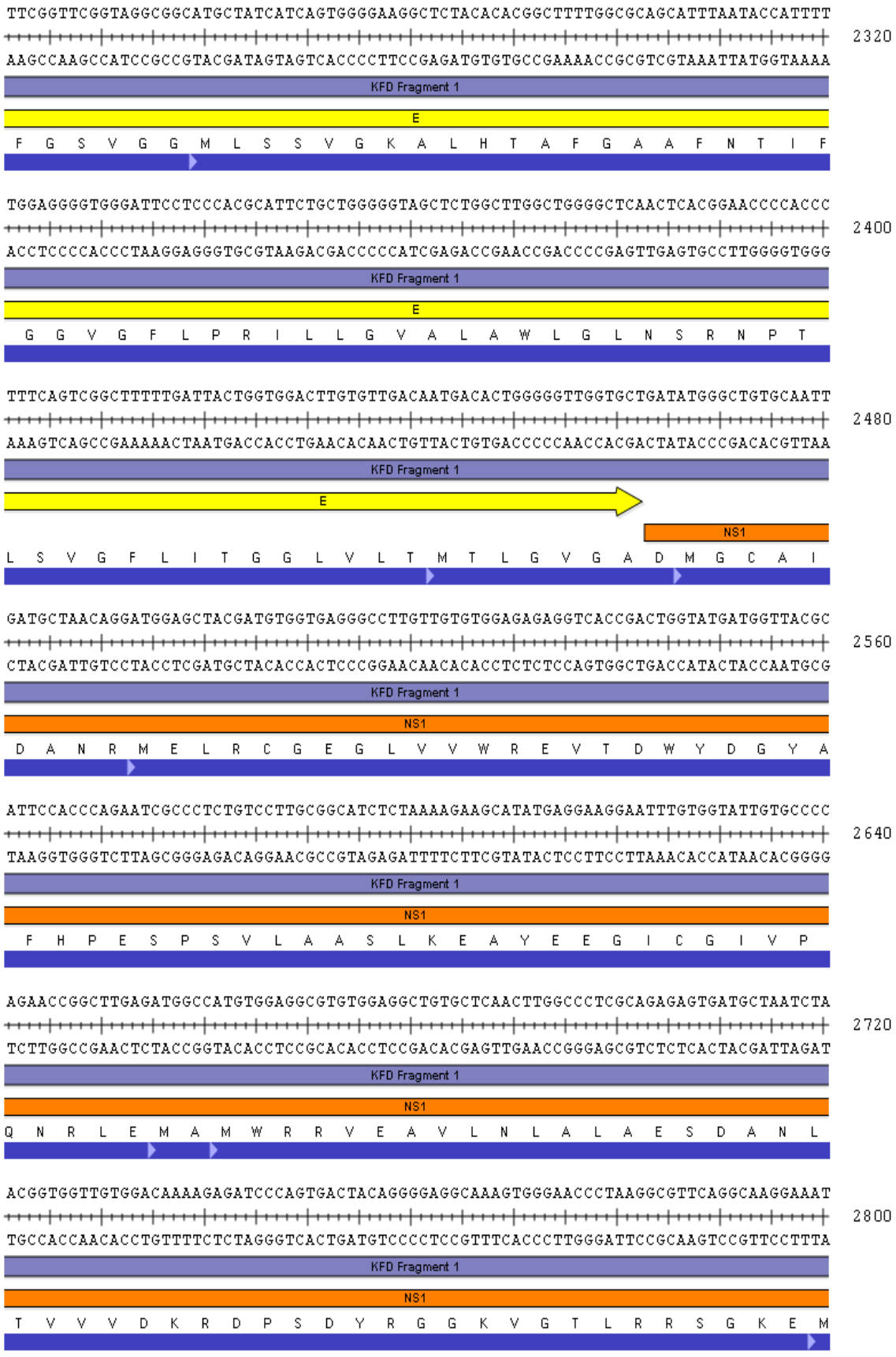
KFD 1-66	KFD mutF1as	CCGGTTCTGCAGGTGTGTGCATCG	Colony-PCR mutation in fragment 1
KFD 1-67	HDVKFDs	GTACggcgccGGGTCGGCAT	Add KasI
KFD 1-68	HDVKFDas	GTACttaattaaTTAGCTCTCCCTTAGCCATCCGAGT	Add PacI
KFD 1-69	mpHDVKFDas	CATGCCGACCCAGCGGATGTT	Delete KasI (site-directed mutagenesis)

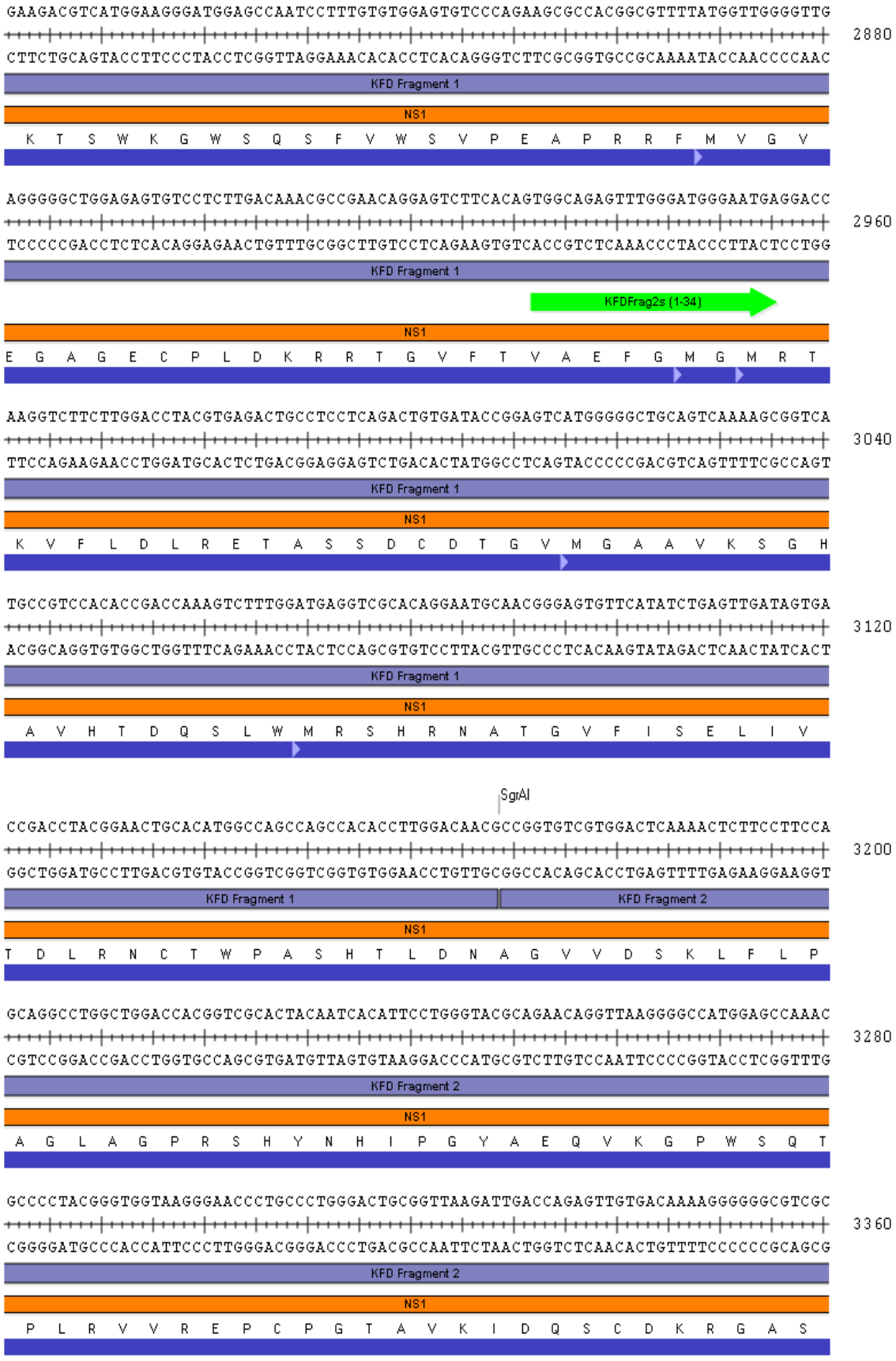

AGATTTTTTTGCACGTGAGTGCTCTCGTTTCAGGCAACGTGAGTGGCGCTTTGTTGGTATTTCCTTGGTGGGAAAGTGTT
80
TCTAAAAAAAACGTGCACTCACGAGAGCAAAGTCCGTTGCACTCACCGCGAAACAACCATAAAGGAACCACCCTTTCACAA
KFD 5' UTR
5' UTR-pTM1s (1-40)
. I F L H V S A L V S G N V S G A L L V F P W W E S V
GAAGCGTTAACGTGTTGAGGAAAAGACAGCTTAGGAGAACAAGAGCTGGGGATGGCCAAAGGAGCCGTCCTTAAAGGAAA
160
CTTCGCAATTGCACAACTCCTTTTCTGTCGAATCCTCTTGTTCTCGACCCCTACCGGTTTCCTCGGCAGGAATTTCCTTT
KFD Fragment 1
KFD 5' UTR
KFD Frag1-...M1s (1-32)
C
E A L T C . G K D S L G E Q E L G M A K G A V L K G K
GGGGGGCGGTCCCCCTAGGCGAGTGCCGAAAGAGACCGCAAAAAAGACGCGTCAAGGACCAGGCCGATTGCCAAACGGAC
240
CCCCCCGCCAGGGGGATCCGCTCACGGCTTTCTCTGGCGTTTTTTCTGCGCAGTTCCTGGTCCGGCTAACGGTTTGCCTG
KFD Fragment 1
C
G G G P P R R V P K E T A K K T R Q G P G R L P N G
TGGTATTGATGCGCATGATGGGAGTGCTGTGGCATATGGTCGCTGGAACGGCCAGGAATCCGATCCTCAAGCGATTCTGG
320
ACCATAACTACGCGTACTACCCTCACGACACCGTATACCAGCGACCTTGCCGGTCCTTAGGCTAGGAGTTCGCTAAGACC
KFD Fragment 1
C
L V L M R M M G V L W H M V A G T A R N P I L K R F W
GCGACTGTTCCGGTACGGCAGGCCATTGCAGCGCTCCGTAAAATCCGAAAGACAGTTGGACTGCTATTGGACTCTTTAAA
400
CGCTGACAAGGCCATGCCGTCCGGTAACGTCGCGAGGCATTTTAGGCTTTCTGTCAACCTGACGATAACCTGAGAAATTT
KFD Fragment 1
C
A T V P V R Q A I A A L R K I R K T V G L L L D S L N
NheI
TAAAAGAAGAGGGAAGAGAAGGTCAACCACCGGATTGCTCACGCCAATCTTGCTAGCTTGCCTGGCGACCCTGGTGTTCT
480
ATTTTCTTCTCCCTTCTCTTCCAGTTGGTGGCCTAACGAGTGCGGTTAGAACGATCGAACGGACCGCTGGGACCACAAGA
KFD Fragment 1
C
C (anch)
K R R G K R R S T T G L L T P I L L A C L A T L V F
CTGCGACAGTTCGCAGAGAGAGAACAGGGAACATGGTGATCAGGGCAGAAGGAAAGGACGCAGCCACGCAAGTGGAAGTC
560
GACGCTGTCAAGCGTCTCTCTCTTGTCCCTTGTACCACTAGTCCCGTCTTCCTTTCCTGCGTCGGTGCGTTCACCTTCAG
KFD Fragment 1
pr (prM)
S A T V R R E R T G N M V I R A E G K D A A T Q V E V

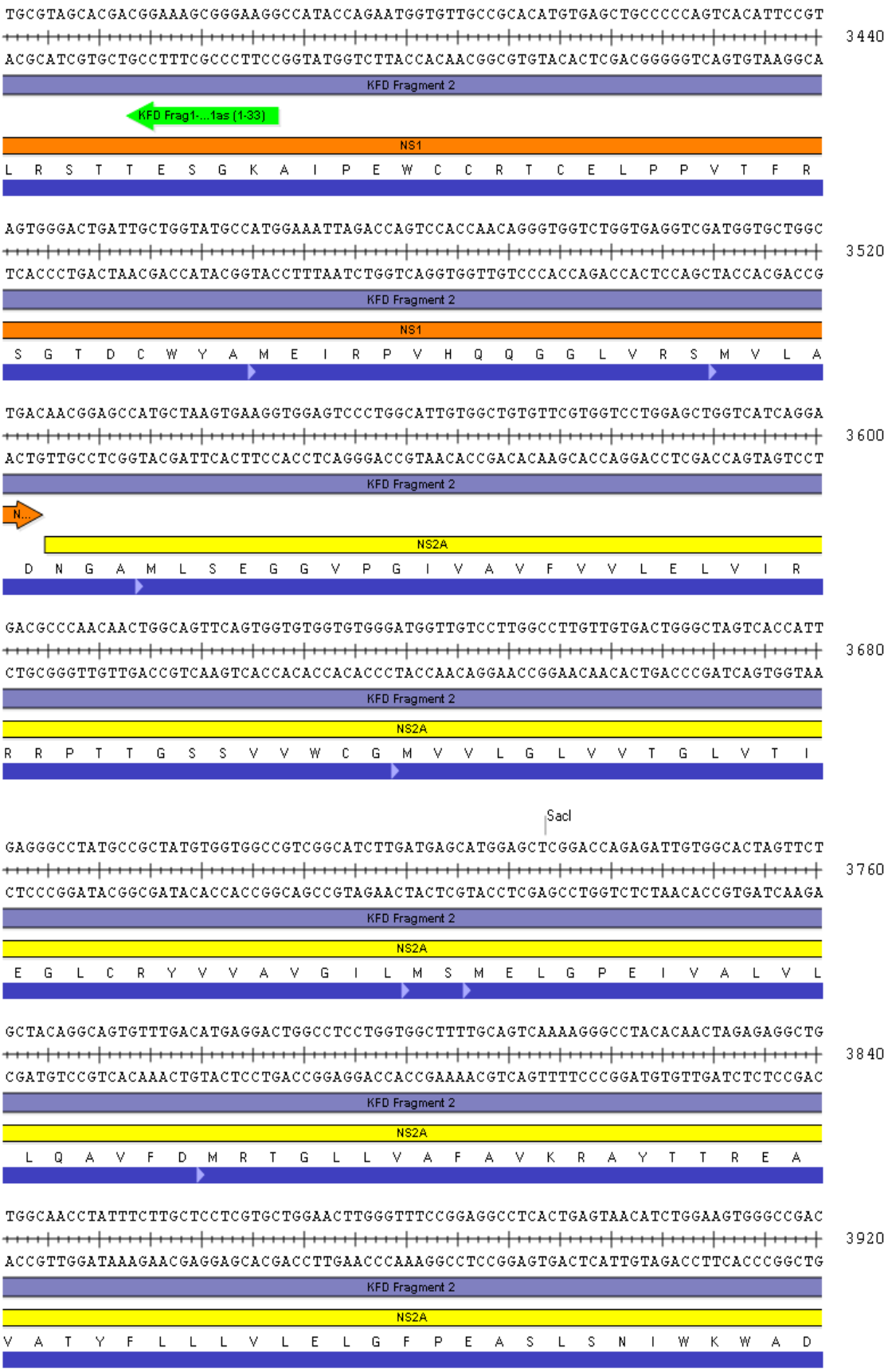

ATGAATGGAACATGCACCATTCTTGCCACAGACATGGGGAGTTGGTGTGATGATTCAATCATGTACGAGTGTGTCACCAT
640
TACTTACCTTGTACGTGGTAAGAACGGTGTCTGTACCCCTCAACCACACTACTAAGTTAGTACATGCTCACACAGTGGTA
KFD Fragment 1
pr (prM)
M N G T C T I L A T D M G S W C D D S I M Y E C V T I
TGACTCGGGAGAAGAACCAGTTGATGTGGATTGCTTCTGCAAAGGCGTTGAGCGAGTGTCACTGGAATACGGAAGGTGTG
720
ACTGAGCCCTCTTCTTGGTCAACTACACCTAACGAAGACGTTTCCGCAACTCGCTCACAGTGACCTTATGCCTTCCACAC
KFD Fragment 1
KFD1as (1-15)
pr (prM)
D S G E E P V D V D C F C K G V E R V S L E Y G R C
GGAAACCAGCTGGCGGTAGAAACAGGAGGTCAGTGTCGATTCCGGTGCATGCTCATAGTGATCTCACCGGAAGAGGGCAC
800
CCTTTGGTCGACCGCCATCTTTGTCCTCCAGTCACAGCTAAGGCCACGTACGAGTATCACTAGAGTGGCCTTCTCCCGTG
KFD Fragment 1
KF...5)
pr (prM)
M (prM)
G K P A G G R N R R S V S I P V H A H S D L T G R G H
AAGTGGCTTAAAGGGGACTCAGTCAAGACACATCTGACACGTGTGGAAGGCTGGGTTTGGAAAAACAAGTTCCTGACAGC
880
TTCACCGAATTTCCCCTGAGTCAGTTCTGTGTAGACTGTGCACACCTTCCGACCCAAACCTTTTTGTTCAAGGACTGTCG
KFD Fragment 1
M (prM)
K W L K G D S V K T H L T R V E G W V W K N K F L T A
GGCCTTTTGTGCAGTTGTGTGGATGGTAACGGACAGCCTGCCGACAAGGTTCATTGTCATAACGGTGGCCCTTTGCCTGG
960
CCGGAAAACACGTCAACACACCTACCATTGCCTGTCGGACGGCTGTTCCAAGTAACAGTATTGCCACCGGGAAACGGACC
KFD Fragment 1
M (prM)
A F C A V V W M V T D S L P T R F I V I T V A L C L
CTCCTACATATGCCACACGATGCACACACCTGCAGAACCGGGACTTTGTCTCAGGGACCCAAGGGACGACTAGAGTTTCA
1040
GAGGATGTATACGGTGTGCTACGTGTGTGGACGTCTTGGCCCTGAAACAGAGTCCCTGGGTTCCCTGCTGATCTCAAAGT
KFD Fragment 1
M (prM)
E
A P T Y A T R C T H L Q N R D F V S G T Q G T T R V S
TTGGTGTTGGAATTAGGCGGCTGTGTCACCCTCACAGCTGAAGGCAAGCCATCAGTGGATGTTTGGCTCGACGACATACA
1120
AACCACAACCTTAATCCGCCGACACAGTGGGAGTGTCGACTTCCGTTCGGTAGTCACCTACAAACCGAGCTGCTGTATGT
KFD Fragment 1
E
L V L E L G G C V T L T A E G K P S V D V W L D D I H

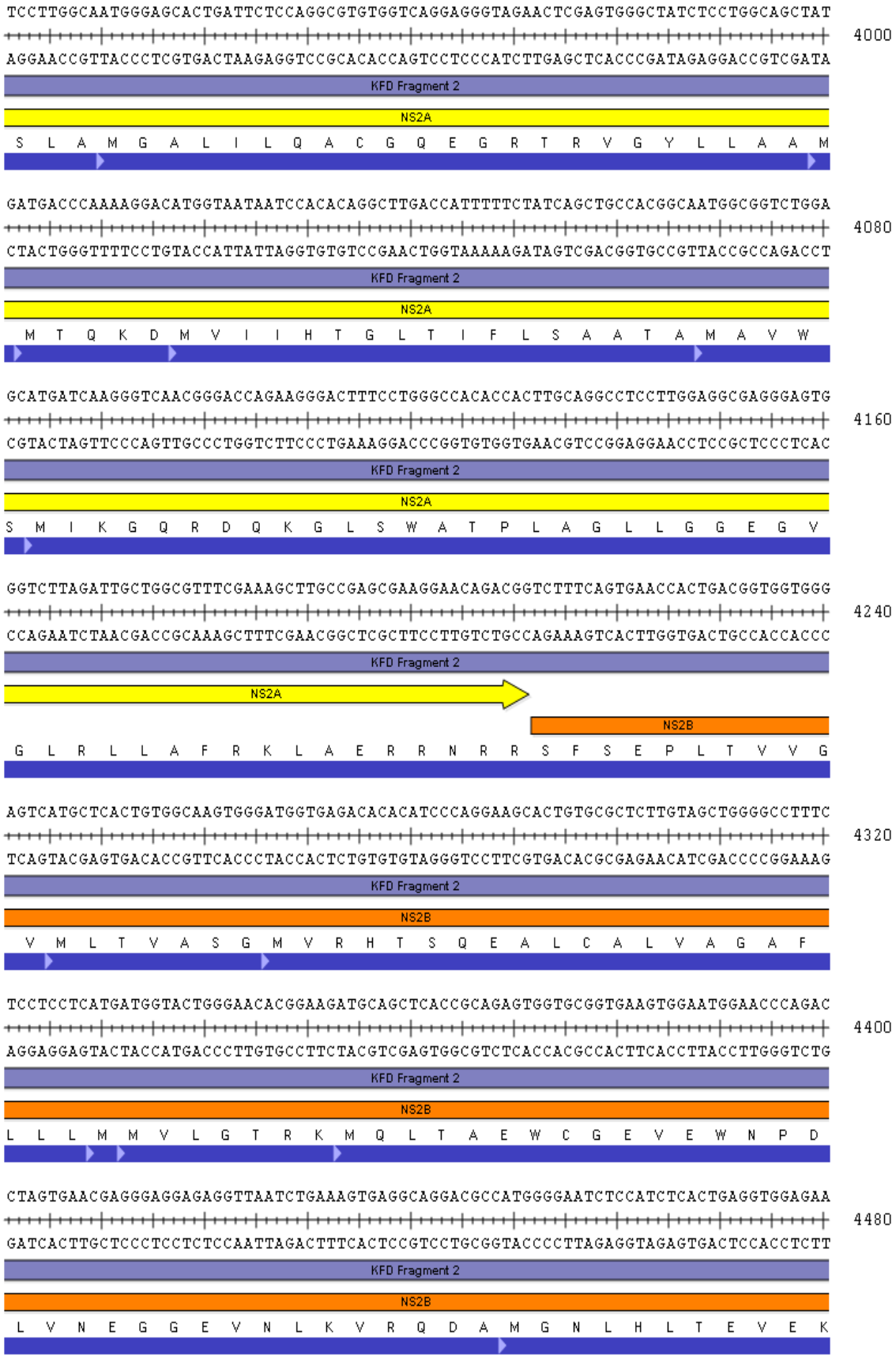

CCAGGAGAACCCGGCCAAGACACGAGAGTATTGTCTTCATGCCAAGCTGGCAAACTCCAAGGTGGCCGCACGTTGTCCAG
1200
GGTCCTCTTGGGCCGGTTCTGTGCTCTCATAACAGAAGTACGGTTCGACCGTTTGAGGTTCCACCGGCGTGCAACAGGTC
KFD Fragment 1
E
Q E N P A K T R E Y C L H A K L A N S K V A A R C P
CCATGGGTCCAGCGACACTGCCAGAGGAGCACCAGGCGAGCACAGTCTGTAGACGAGACCAGAGTGACAGAGGCTGGGGT
1280
GGTACCCAGGTCGCTGTGACGGTCTCCTCGTGGTCCGCTCGTGTCAGACATCTGCTCTGGTCTCACTGTCTCCGACCCCA
KFD Fragment 1
E
A M G P A T L P E E H Q A S T V C R R D Q S D R G W G
AACCATTGTGGACTGTTCGGGAAGGGCAGCATTGTGGCTTGTGCCAAGTTTAGCTGTGAGGCCAAAAAGAAGGCAACTGG
1360
TTGGTAACACCTGACAAGCCCTTCCCGTCGTAACACCGAACACGGTTCAAATCGACACTCCGGTTTTTCTTCCGTTGACC
KFD Fragment 1
E
N H C G L F G K G S I V A C A K F S C E A K K K A T G
ATATGTGTACGATGTCAACAAAATCACATATGTGGTGAAGGTGGAACCTCACACGGGTGACTATCTAGCGGCCAACGAGT
1440
TATACACATGCTACAGTTGTTTTAGTGTATACACCACTTCCACCTTGGAGTGTGCCCACTGATAGATCGCCGGTTGCTCA
KFD Fragment 1
E
Y V Y D V N K I T Y V V K V E P H T G D Y L A A N E
CGCACAGCAACCGGAAGACAGCATCTTTCACAACGCAGTCCGAGAAGACCATCTTGACACTGGGAGATTATGGAGACATC
1520
GCGTGTCGTTGGCCTTCTGTCGTAGAAAGTGTTGCGTCAGGCTCTTCTGGTAGAACTGTGACCCTCTAATACCTCTGTAG
KFD Fragment 1
E
S H S N R K T A S F T T Q S E K T I L T L G D Y G D I
BamHI
TCCCTCACGTGCCGGGTGACTAGTGGAGTGGATCCTGCCCAGACTGTGGTATTGGAACTGGACAAGACAGCGGAACACTT
1600
AGGGAGTGCACGGCCCACTGATCACCTCACCTAGGACGGGTCTGACACCATAACCTTGACCTGTTCTGTCGCCTTGTGAA
KFD Fragment 1
E
S L T C R V T S G V D P A Q T V V L E L D K T A E H L
GCCCAAGGCTTGGCAGGTACACCGAGACTGGTTTGAGGACCTCTCCTTGCCATGGCGACACGGGGGTGCCCAGGAATGGA
1680
CGGGTTCCGAACCGTCCATGTGGCTCTGACCAAACTCCTGGAGAGGAACGGTACCGCTGTGCCCCCACGGGTCCTTACCT
KFD Fragment 1
E
P K A W Q V H R D W F E D L S L P W R H G G A Q E W

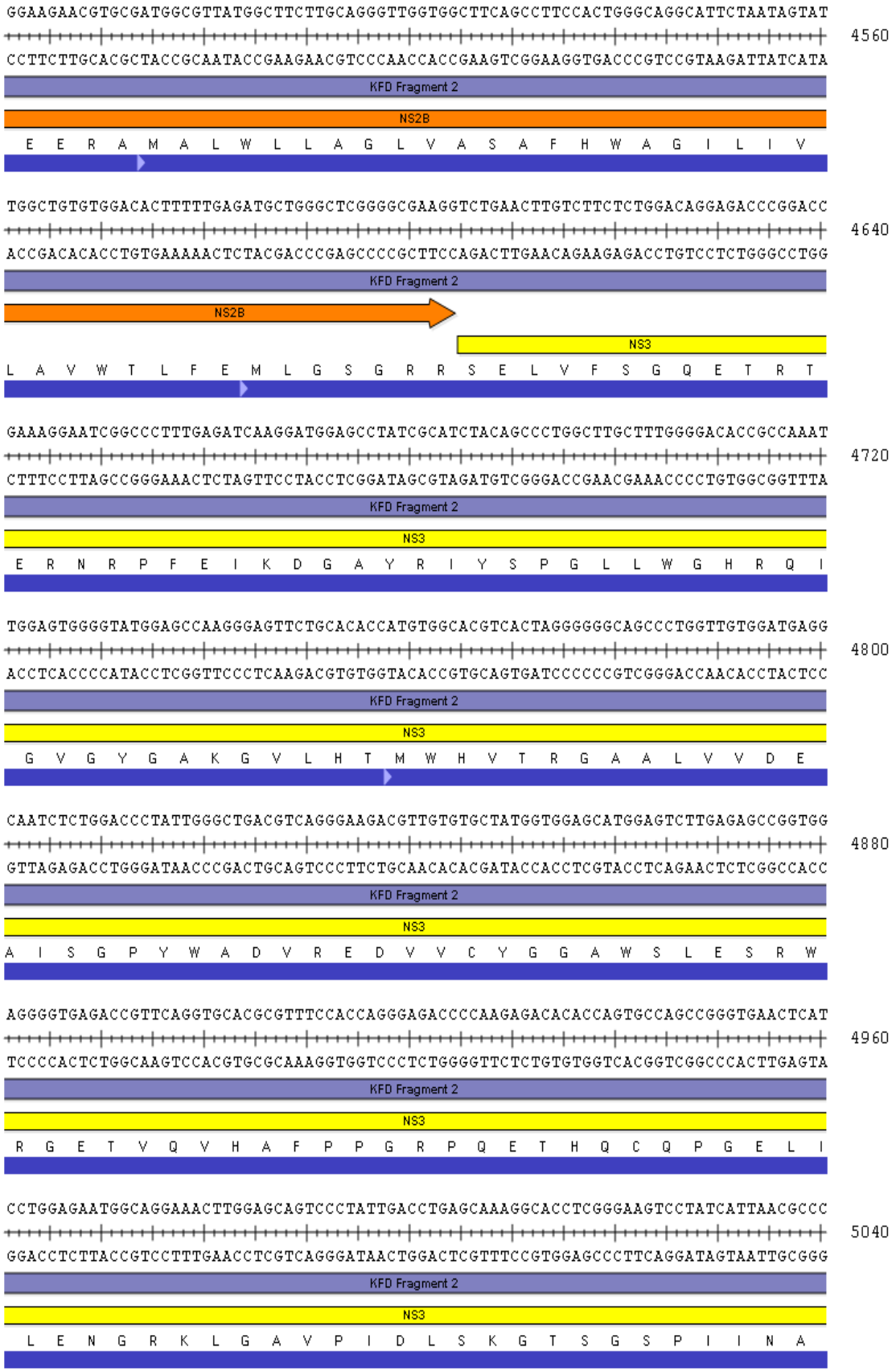

ACCATGCTGACCGGCTCGTGGAATTTGGTGAACCGCACGCTGTGAAAATGGACATTTTTAACCTTGGGGACCAGACGGGT
1760
TGGTACGACTGGCCGAGCACCTTAAACCACTTGGCGTGCGACACTTTTACCTGTAAAAATTGGAACCCCTGGTCTGCCCA
KFD Fragment 1
E
N H A D R L V E F G E P H A V K M D I F N L G D Q T G
ATCCTGTTGAAGTCTCTGGCTGGAGTGCCCGTGGCTAACATAGAGGGCAGCAAGTACCATCTCCAGAGTGGTCATGTCAC
1840
TAGGACAACTTCAGAGACCGACCTCACGGGCACCGATTGTATCTCCCGTCGTTCATGGTAGAGGTCTCACCAGTACAGTG
KFD Fragment 1
E
I L L K S L A G V P V A N I E G S K Y H L Q S G H V T
CTGCGATGTGGGCCTTGAAAAACTTAAGATGAAAGGGATGACATACACGGTTTGTGAGGGATCAAAATTTGCTTGGAAAA
1920
GACGCTACACCCGGAACTTTTTGAATTCTACTTTCCCTACTGTATGTGCCAAACACTCCCTAGTTTTAAACGAACCTTTT
KFD Fragment 1
E
C D V G L E K L K M K G M T Y T V C E G S K F A W K
GGCCGCCAACCGACAGTGGACATGATACCGTAGTCATGGAGGTGACTTACACCGGGAGCAAGCCATGCAGAATACCAGTG
2000
CCGGCGGTTGGCTGTCACCTGTACTATGGCATCAGTACCTCCACTGAATGTGGCCCTCGTTCGGTACGTCTTATGGTCAC
KFD Fragment 1
E
R P P T D S G H D T V V M E V T Y T G S K P C R I P V
AGAGCCGTGGCCCATGGAGAACCCAATGTTAACGTGGCAAGTCTAATAACCCCAAACCCATCCATGGAAACAACTGGAGG
2080
TCTCGGCACCGGGTACCTCTTGGGTTACAATTGCACCGTTCAGATTATTGGGGTTTGGGTAGGTACCTTTGTTGACCTCC
KFD Fragment 1
E
R A V A H G E P N V N V A S L I T P N P S M E T T G G
AGGGTTCGTTGAGCTACAGCTACCACCAGGAGACAACATCATCTATGTTGGTGAGCTGAGCCACCAGTGGTTTCAGAAGG
2160
TCCCAAGCAACTCGATGTCGATGGTGGTCCTCTGTTGTAGTAGATACAACCACTCGACTCGGTGGTCACCAAAGTCTTCC
KFD Fragment 1
E
G F V E L Q L P P G D N I I Y V G E L S H Q W F Q K
GCAGCACAATTGGCCGTGTTCTGGAAAAGACCAGGAGGGGCATAGAGAGACTCACAGTTGTGGGAGAGCACGCCTGGGAC
2240
CGTCGTGTTAACCGGCACAAGACCTTTTCTGGTCCTCCCCGTATCTCTCTGAGTGTCAACACCCTCTCGTGCGGACCCTG
KFD Fragment 1
E
G S T I G R V L E K T R R G I E R L T V V G E H A W D

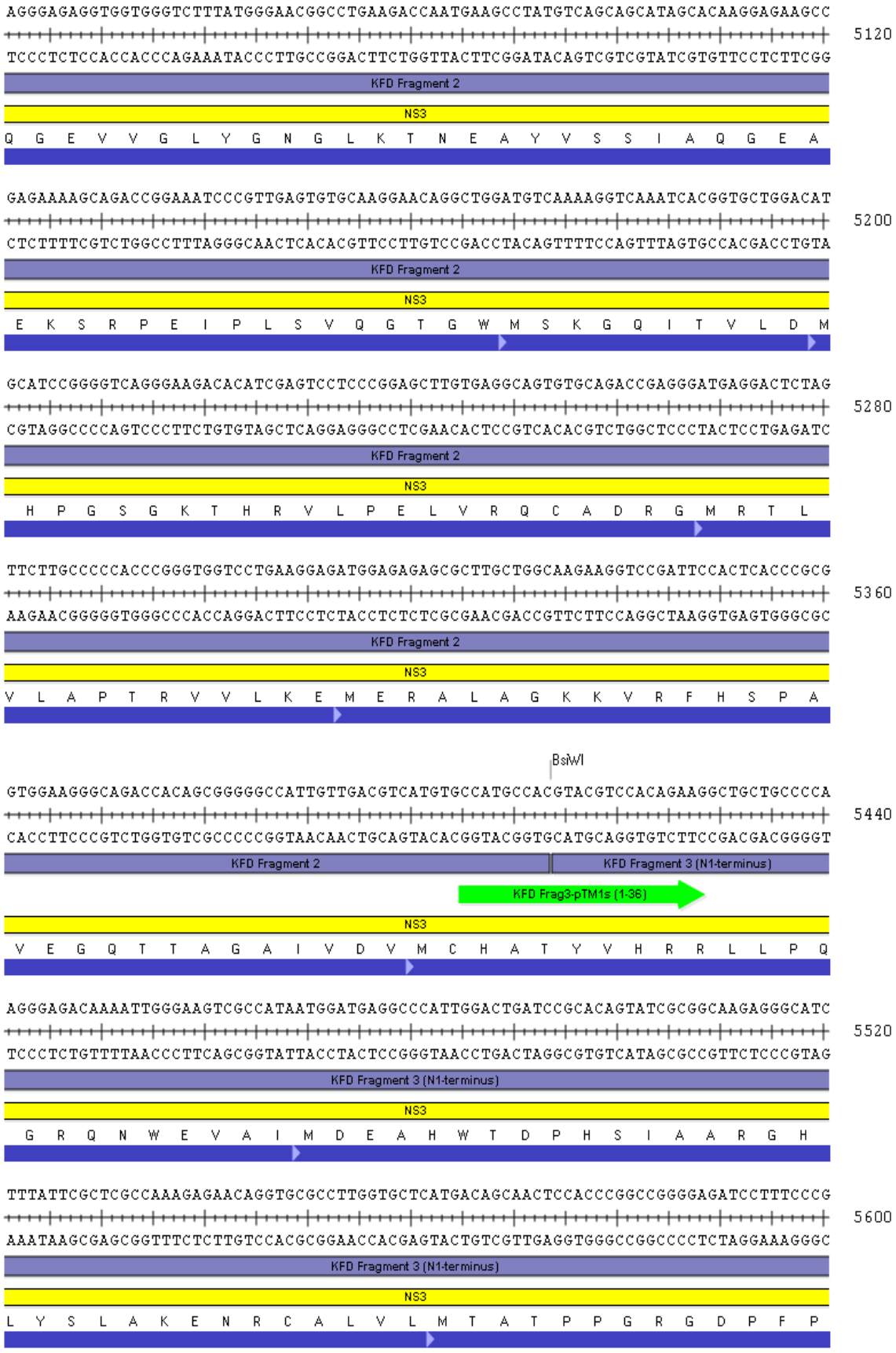

TTCGGTTCGGTAGGCGGCATGCTATCATCAGTGGGGAAGGCTCTACACACGGCTTTTGGCGCAGCATTTAATACCATTTT
2320
AAGCCAAGCCATCCGCCGTACGATAGTAGTCACCCCTTCCGAGATGTGTGCCGAAAACCGCGTCGTAAATTATGGTAAAA
KFD Fragment 1
E
F G S V G G M L S S V G K A L H T A F G A A F N T I F
TGGAGGGGTGGGATTCCTCCCACGCATTCTGCTGGGGGTAGCTCTGGCTTGGCTGGGGCTCAACTCACGGAACCCCACCC
2400
ACCTCCCCACCCTAAGGAGGGTGCGTAAGACGACCCCCATCGAGACCGAACCGACCCCGAGTTGAGTGCCTTGGGGTGGG
KFD Fragment 1
E
G G V G F L P R I L L G V A L A W L G L N S R N P T
TTTCAGTCGGCTTTTTGATTACTGGTGGACTTGTGTTGACAATGACACTGGGGGTTGGTGCTGATATGGGCTGTGCAATT
2480
AAAGTCAGCCGAAAAACTAATGACCACCTGAACACAACTGTTACTGTGACCCCCAACCACGACTATACCCGACACGTTAA
KFD Fragment 1
E
NS1
L S V G F L I T G G L V L T M T L G V G A D M G C A I
GATGCTAACAGGATGGAGCTACGATGTGGTGAGGGCCTTGTTGTGTGGAGAGAGGTCACCGACTGGTATGATGGTTACGC
2560
CTACGATTGTCCTACCTCGATGCTACACCACTCCCGGAACAACACACCTCTCTCCAGTGGCTGACCATACTACCAATGCG
KFD Fragment 1
NS1
D A N R M E L R C G E G L V V W R E V T D W Y D G Y A
ATTCCACCCAGAATCGCCCTCTGTCCTTGCGGCATCTCTAAAAGAAGCATATGAGGAAGGAATTTGTGGTATTGTGCCCC
2640
TAAGGTGGGTCTTAGCGGGAGACAGGAACGCCGTAGAGATTTTCTTCGTATACTCCTTCCTTAAACACCATAACACGGGG
KFD Fragment 1
NS1
F H P E S P S V L A A S L K E A Y E E G I C G I V P
AGAACCGGCTTGAGATGGCCATGTGGAGGCGTGTGGAGGCTGTGCTCAACTTGGCCCTCGCAGAGAGTGATGCTAATCTA
2720
TCTTGGCCGAACTCTACCGGTACACCTCCGCACACCTCCGACACGAGTTGAACCGGGAGCGTCTCTCACTACGATTAGAT
KFD Fragment 1
NS1
Q N R L E M A M W R R V E A V L N L A L A E S D A N L
ACGGTGGTTGTGGACAAAAGAGATCCCAGTGACTACAGGGGAGGCAAAGTGGGAACCCTAAGGCGTTCAGGCAAGGAAAT
2800
TGCCACCAACACCTGTTTTCTCTAGGGTCACTGATGTCCCCTCCGTTTCACCCTTGGGATTCCGCAAGTCCGTTCCTTTA
KFD Fragment 1
NS1
T V V V D K R D P S D Y R G G K V G T L R R S G K E M

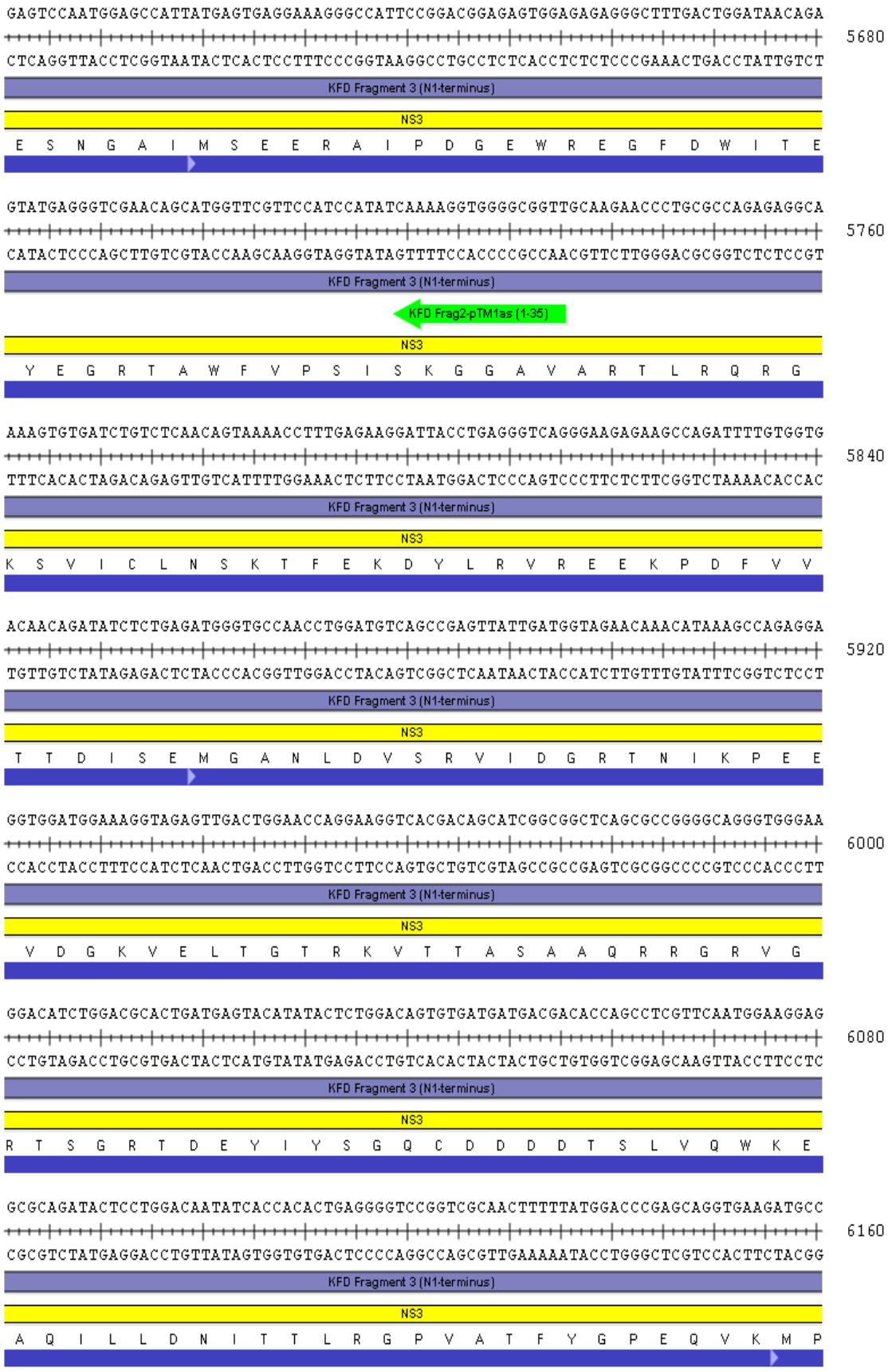

GAAGACGTCATGGAAGGGATGGAGCCAATCCTTTGTGTGGAGTGTCCCAGAAGCGCCACGGCGTTTTATGGTTGGGGTTG
2880
CTTCTGCAGTACCTTCCCTACCTCGGTTAGGAAACACACCTCACAGGGTCTTCGCGGTGCCGCAAAATACCAACCCCAAC
KFD Fragment 1
NS1
K T S W K G W S Q S F V W S V P E A P R R F M V G V
AGGGGGCTGGAGAGTGTCCTCTTGACAAACGCCGAACAGGAGTCTTCACAGTGGCAGAGTTTGGGATGGGAATGAGGACC
2960
TCCCCCGACCTCTCACAGGAGAACTGTTTGCGGCTTGTCCTCAGAAGTGTCACCGTCTCAAACCCTACCCTTACTCCTGG
KFD Fragment 1
KFDFrag2s (1-34)
NS1
E G A G E C P L D K R R T G V F T V A E F G M G M R T
AAGGTCTTCTTGGACCTACGTGAGACTGCCTCCTCAGACTGTGATACCGGAGTCATGGGGGCTGCAGTCAAAAGCGGTCA
3040
TTCCAGAAGAACCTGGATGCACTCTGACGGAGGAGTCTGACACTATGGCCTCAGTACCCCCGACGTCAGTTTTCGCCAGT
KFD Fragment 1
NS1
K V F L D L R E T A S S D C D T G V M G A A V K S G H
TGCCGTCCACACCGACCAAAGTCTTTGGATGAGGTCGCACAGGAATGCAACGGGAGTGTTCATATCTGAGTTGATAGTGA
3120
ACGGCAGGTGTGGCTGGTTTCAGAAACCTACTCCAGCGTGTCCTTACGTTGCCCTCACAAGTATAGACTCAACTATCACT
KFD Fragment 1
NS1
A V H T D Q S L W M R S H R N A T G V F I S E L I V
SgrAI
CCGACCTACGGAACTGCACATGGCCAGCCAGCCACACCTTGGACAACGCCGGTGTCGTGGACTCAAAACTCTTCCTTCCA
3200
GGCTGGATGCCTTGACGTGTACCGGTCGGTCGGTGTGGAACCTGTTGCGGCCACAGCACCTGAGTTTTGAGAAGGAAGGT
KFD Fragment 1
KFD Fragment 2
NS1
T D L R N C T W P A S H T L D N A G V V D S K L F L P
GCAGGCCTGGCTGGACCACGGTCGCACTACAATCACATTCCTGGGTACGCAGAACAGGTTAAGGGGCCATGGAGCCAAAC
3280
CGTCCGGACCGACCTGGTGCCAGCGTGATGTTAGTGTAAGGACCCATGCGTCTTGTCCAATTCCCCGGTACCTCGGTTTG
KFD Fragment 2
NS1
A G L A G P R S H Y N H I P G Y A E Q V K G P W S Q T
GCCCCTACGGGTGGTAAGGGAACCCTGCCCTGGGACTGCGGTTAAGATTGACCAGAGTTGTGACAAAAGGGGGGCGTCGC
3360
CGGGGATGCCCACCATTCCCTTGGGACGGGACCCTGACGCCAATTCTAACTGGTCTCAACACTGTTTTCCCCCCGCAGCG
KFD Fragment 2
NS1
P L R V V R E P C P G T A V K I D Q S C D K R G A S

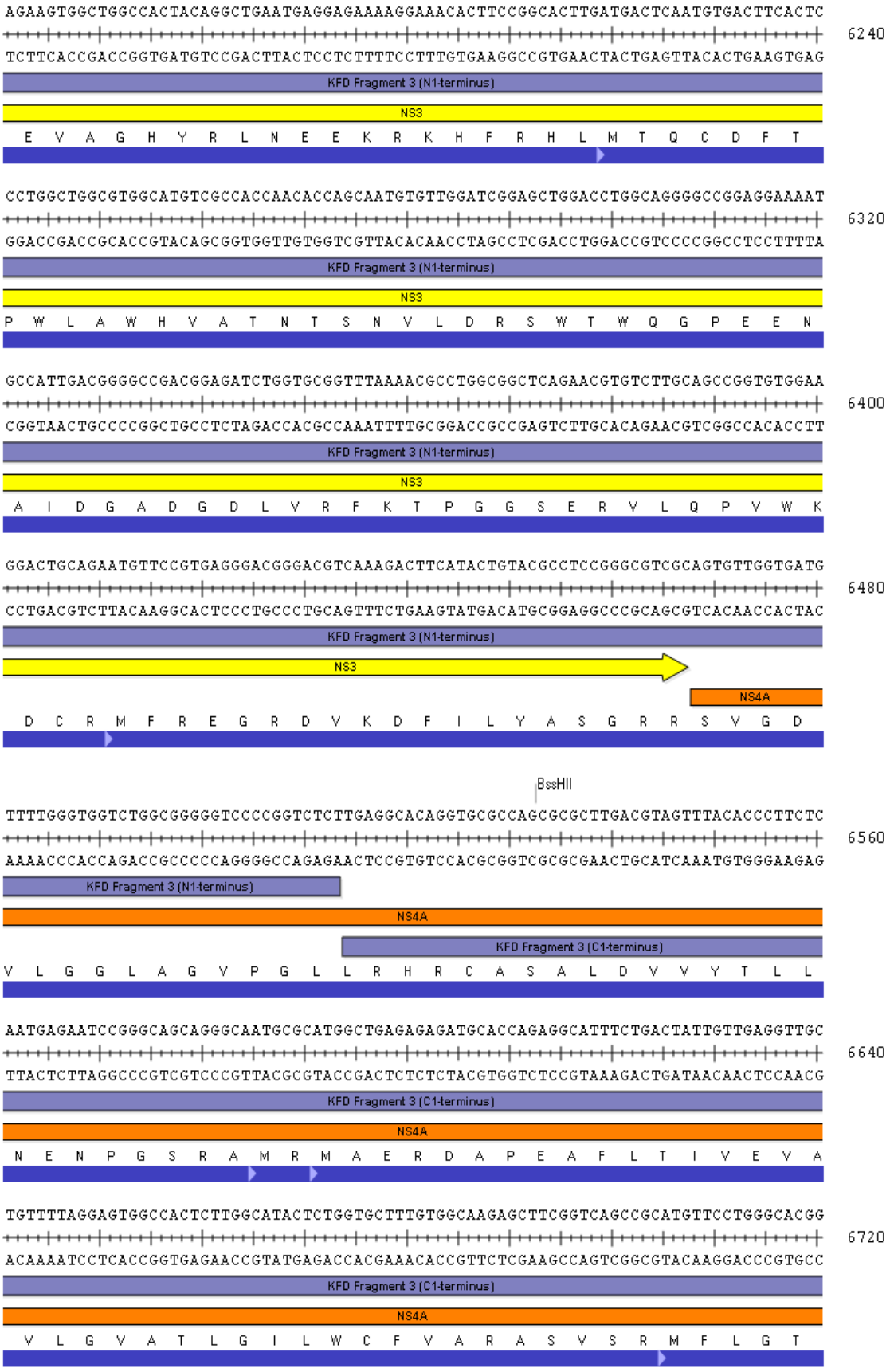

TGCGTAGCACGACGGAAAGCGGGAAGGCCATACCAGAATGGTGTTGCCGCACATGTGAGCTGCCCCCAGTCACATTCCGT
3440
ACGCATCGTGCTGCCTTTCGCCCTTCCGGTATGGTCTTACCACAACGGCGTGTACACTCGACGGGGGTCAGTGTAAGGCA
KFD Fragment 2
KFD Frag1-...1as (1-33)
NS1
L R S T T E S G K A I P E W C C R T C E L P P V T F R
AGTGGGACTGATTGCTGGTATGCCATGGAAATTAGACCAGTCCACCAACAGGGTGGTCTGGTGAGGTCGATGGTGCTGGC
3520
TCACCCTGACTAACGACCATACGGTACCTTTAATCTGGTCAGGTGGTTGTCCCACCAGACCACTCCAGCTACCACGACCG
KFD Fragment 2
NS1
S G T D C W Y A M E I R P V H Q Q G G L V R S M V L A
TGACAACGGAGCCATGCTAAGTGAAGGTGGAGTCCCTGGCATTGTGGCTGTGTTCGTGGTCCTGGAGCTGGTCATCAGGA
3600
ACTGTTGCCTCGGTACGATTCACTTCCACCTCAGGGACCGTAACACCGACACAAGCACCAGGACCTCGACCAGTAGTCCT
KFD Fragment 2
N...
NS2A
D N G A M L S E G G V P G I V A V F V V L E L V I R
GACGCCCAACAACTGGCAGTTCAGTGGTGTGGTGTGGGATGGTTGTCCTTGGCCTTGTTGTGACTGGGCTAGTCACCATT
3680
CTGCGGGTTGTTGACCGTCAAGTCACCACACCACACCCTACCAACAGGAACCGGAACAACACTGACCCGATCAGTGGTAA
KFD Fragment 2
NS2A
R R P T T G S S V V W C G M V V L G L V V T G L V T I
SacI
GAGGGCCTATGCCGCTATGTGGTGGCCGTCGGCATCTTGATGAGCATGGAGCTCGGACCAGAGATTGTGGCACTAGTTCT
3760
CTCCCGGATACGGCGATACACCACCGGCAGCCGTAGAACTACTCGTACCTCGAGCCTGGTCTCTAACACCGTGATCAAGA
KFD Fragment 2
NS2A
E G L C R Y V V A V G I L M S M E L G P E I V A L V L
GCTACAGGCAGTGTTTGACATGAGGACTGGCCTCCTGGTGGCTTTTGCAGTCAAAAGGGCCTACACAACTAGAGAGGCTG
3840
CGATGTCCGTCACAAACTGTACTCCTGACCGGAGGACCACCGAAAACGTCAGTTTTCCCGGATGTGTTGATCTCTCCGAC
KFD Fragment 2
NS2A
L Q A V F D M R T G L L V A F A V K R A Y T T R E A
TGGCAACCTATTTCTTGCTCCTCGTGCTGGAACTTGGGTTTCCGGAGGCCTCACTGAGTAACATCTGGAAGTGGGCCGAC
3920
ACCGTTGGATAAAGAACGAGGAGCACGACCTTGAACCCAAAGGCCTCCGGAGTGACTCATTGTAGACCTTCACCCGGCTG
KFD Fragment 2
NS2A
V A T Y F L L L V L E L G F P E A S L S N I W K W A D

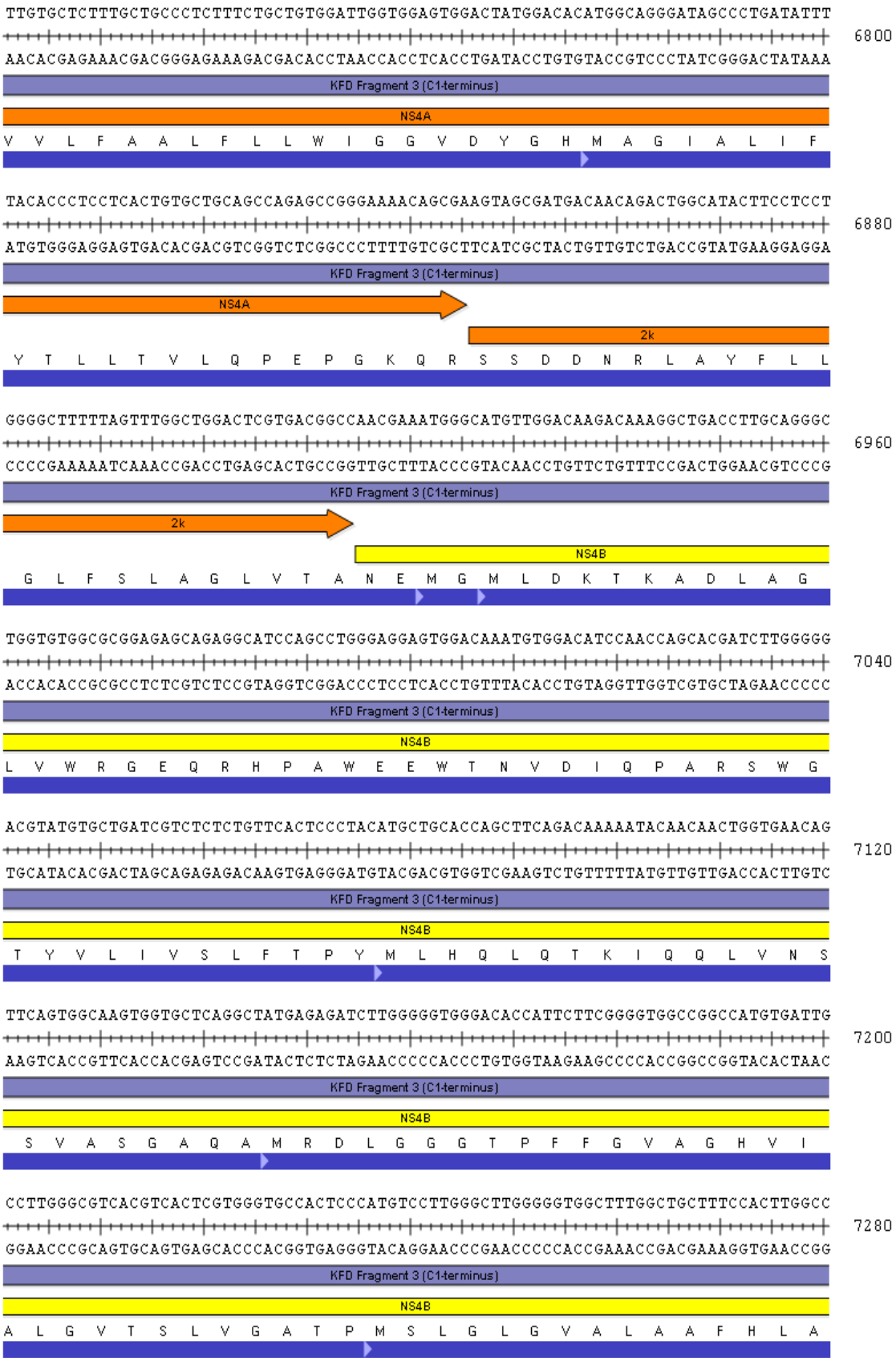

TCCTTGGCAATGGGAGCACTGATTCTCCAGGCGTGTGGTCAGGAGGGTAGAACTCGAGTGGGCTATCTCCTGGCAGCTAT
4000
AGGAACCGTTACCCTCGTGACTAAGAGGTCCGCACACCAGTCCTCCCATCTTGAGCTCACCCGATAGAGGACCGTCGATA
KFD Fragment 2
NS2A
S L A M G A L I L Q A C G Q E G R T R V G Y L L A A M
GATGACCCAAAAGGACATGGTAATAATCCACACAGGCTTGACCATTTTTCTATCAGCTGCCACGGCAATGGCGGTCTGGA
4080
CTACTGGGTTTTCCTGTACCATTATTAGGTGTGTCCGAACTGGTAAAAAGATAGTCGACGGTGCCGTTACCGCCAGACCT
KFD Fragment 2
NS2A
M T Q K D M V I I H T G L T I F L S A A T A M A V W
GCATGATCAAGGGTCAACGGGACCAGAAGGGACTTTCCTGGGCCACACCACTTGCAGGCCTCCTTGGAGGCGAGGGAGTG
4160
CGTACTAGTTCCCAGTTGCCCTGGTCTTCCCTGAAAGGACCCGGTGTGGTGAACGTCCGGAGGAACCTCCGCTCCCTCAC
KFD Fragment 2
NS2A
S M I K G Q R D Q K G L S W A T P L A G L L G G E G V
GGTCTTAGATTGCTGGCGTTTCGAAAGCTTGCCGAGCGAAGGAACAGACGGTCTTTCAGTGAACCACTGACGGTGGTGGG
4240
CCAGAATCTAACGACCGCAAAGCTTTCGAACGGCTCGCTTCCTTGTCTGCCAGAAAGTCACTTGGTGACTGCCACCACCC
KFD Fragment 2
NS2A
NS2B
G L R L L A F R K L A E R R N R R S F S E P L T V V G
AGTCATGCTCACTGTGGCAAGTGGGATGGTGAGACACACATCCCAGGAAGCACTGTGCGCTCTTGTAGCTGGGGCCTTTC
4320
TCAGTACGAGTGACACCGTTCACCCTACCACTCTGTGTGTAGGGTCCTTCGTGACACGCGAGAACATCGACCCCGGAAAG
KFD Fragment 2
NS2B
V M L T V A S G M V R H T S Q E A L C A L V A G A F
TCCTCCTCATGATGGTACTGGGAACACGGAAGATGCAGCTCACCGCAGAGTGGTGCGGTGAAGTGGAATGGAACCCAGAC
4400
AGGAGGAGTACTACCATGACCCTTGTGCCTTCTACGTCGAGTGGCGTCTCACCACGCCACTTCACCTTACCTTGGGTCTG
KFD Fragment 2
NS2B
L L L M M V L G T R K M Q L T A E W C G E V E W N P D
CTAGTGAACGAGGGAGGAGAGGTTAATCTGAAAGTGAGGCAGGACGCCATGGGGAATCTCCATCTCACTGAGGTGGAGAA
4480
GATCACTTGCTCCCTCCTCTCCAATTAGACTTTCACTCCGTCCTGCGGTACCCCTTAGAGGTAGAGTGACTCCACCTCTT
KFD Fragment 2
NS2B
L V N E G G E V N L K V R Q D A M G N L H L T E V E K

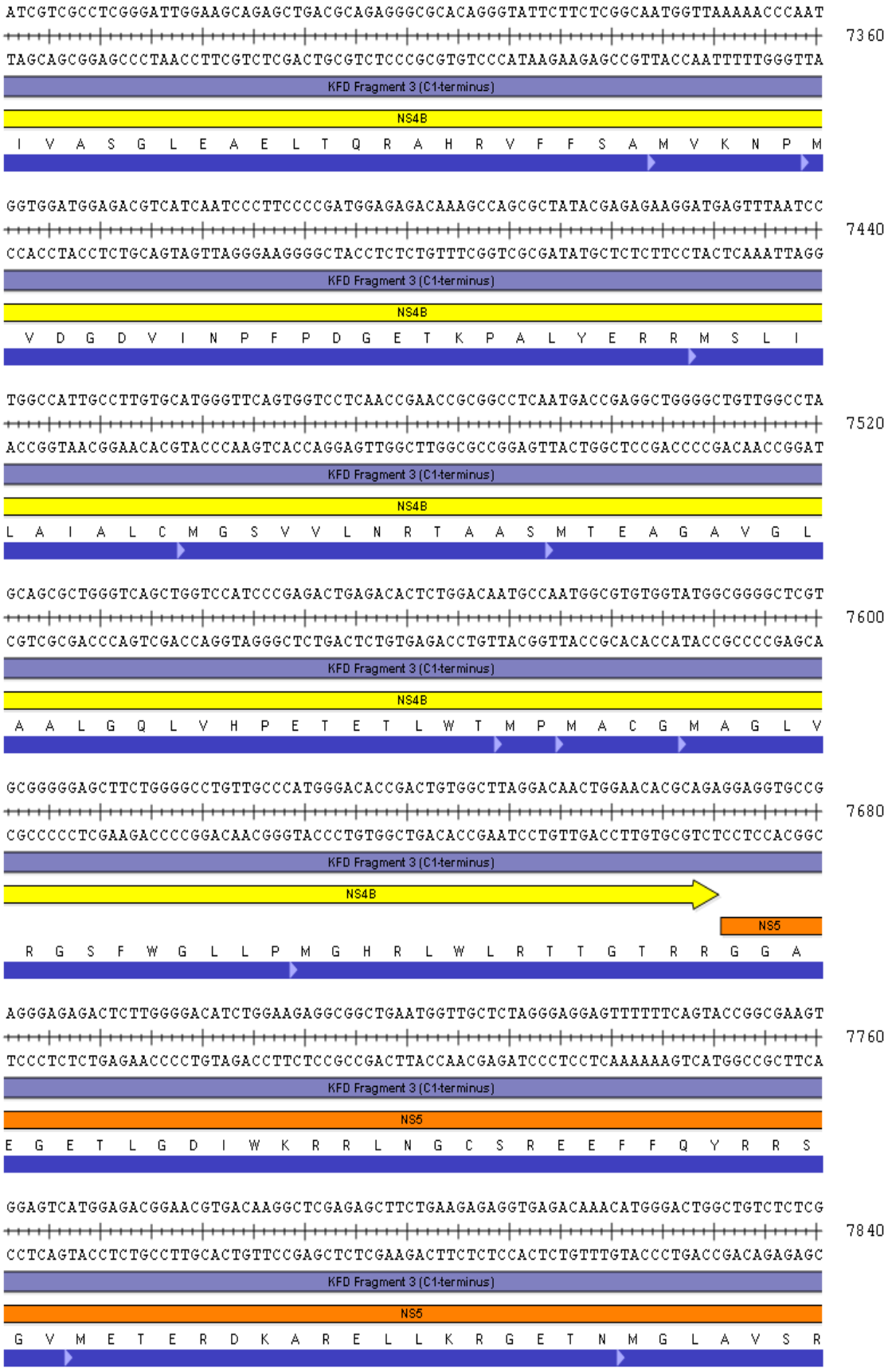

GGAAGAACGTGCGATGGCGTTATGGCTTCTTGCAGGGTTGGTGGCTTCAGCCTTCCACTGGGCAGGCATTCTAATAGTAT
4560
CCTTCTTGCACGCTACCGCAATACCGAAGAACGTCCCAACCACCGAAGTCGGAAGGTGACCCGTCCGTAAGATTATCATA
KFD Fragment 2
NS2B
E E R A M A L W L L A G L V A S A F H W A G I L I V
TGGCTGTGTGGACACTTTTTGAGATGCTGGGCTCGGGGCGAAGGTCTGAACTTGTCTTCTCTGGACAGGAGACCCGGACC
4640
ACCGACACACCTGTGAAAAACTCTACGACCCGAGCCCCGCTTCCAGACTTGAACAGAAGAGACCTGTCCTCTGGGCCTGG
KFD Fragment 2
NS2B
NS3
L A V W T L F E M L G S G R R S E L V F S G Q E T R T
GAAAGGAATCGGCCCTTTGAGATCAAGGATGGAGCCTATCGCATCTACAGCCCTGGCTTGCTTTGGGGACACCGCCAAAT
4720
CTTTCCTTAGCCGGGAAACTCTAGTTCCTACCTCGGATAGCGTAGATGTCGGGACCGAACGAAACCCCTGTGGCGGTTTA
KFD Fragment 2
NS3
E R N R P F E I K D G A Y R I Y S P G L L W G H R Q I
TGGAGTGGGGTATGGAGCCAAGGGAGTTCTGCACACCATGTGGCACGTCACTAGGGGGGCAGCCCTGGTTGTGGATGAGG
4800
ACCTCACCCCATACCTCGGTTCCCTCAAGACGTGTGGTACACCGTGCAGTGATCCCCCCGTCGGGACCAACACCTACTCC
KFD Fragment 2
NS3
G V G Y G A K G V L H T M W H V T R G A A L V V D E
CAATCTCTGGACCCTATTGGGCTGACGTCAGGGAAGACGTTGTGTGCTATGGTGGAGCATGGAGTCTTGAGAGCCGGTGG
4880
GTTAGAGACCTGGGATAACCCGACTGCAGTCCCTTCTGCAACACACGATACCACCTCGTACCTCAGAACTCTCGGCCACC
KFD Fragment 2
NS3
A I S G P Y W A D V R E D V V C Y G G A W S L E S R W
AGGGGTGAGACCGTTCAGGTGCACGCGTTTCCACCAGGGAGACCCCAAGAGACACACCAGTGCCAGCCGGGTGAACTCAT
4960
TCCCCACTCTGGCAAGTCCACGTGCGCAAAGGTGGTCCCTCTGGGGTTCTCTGTGTGGTCACGGTCGGCCCACTTGAGTA
KFD Fragment 2
NS3
R G E T V Q V H A F P P G R P Q E T H Q C Q P G E L I
CCTGGAGAATGGCAGGAAACTTGGAGCAGTCCCTATTGACCTGAGCAAAGGCACCTCGGGAAGTCCTATCATTAACGCCC
5040
GGACCTCTTACCGTCCTTTGAACCTCGTCAGGGATAACTGGACTCGTTTCCGTGGAGCCCTTCAGGATAGTAATTGCGGG
KFD Fragment 2
NS3
L E N G R K L G A V P I D L S K G T S G S P I I N A

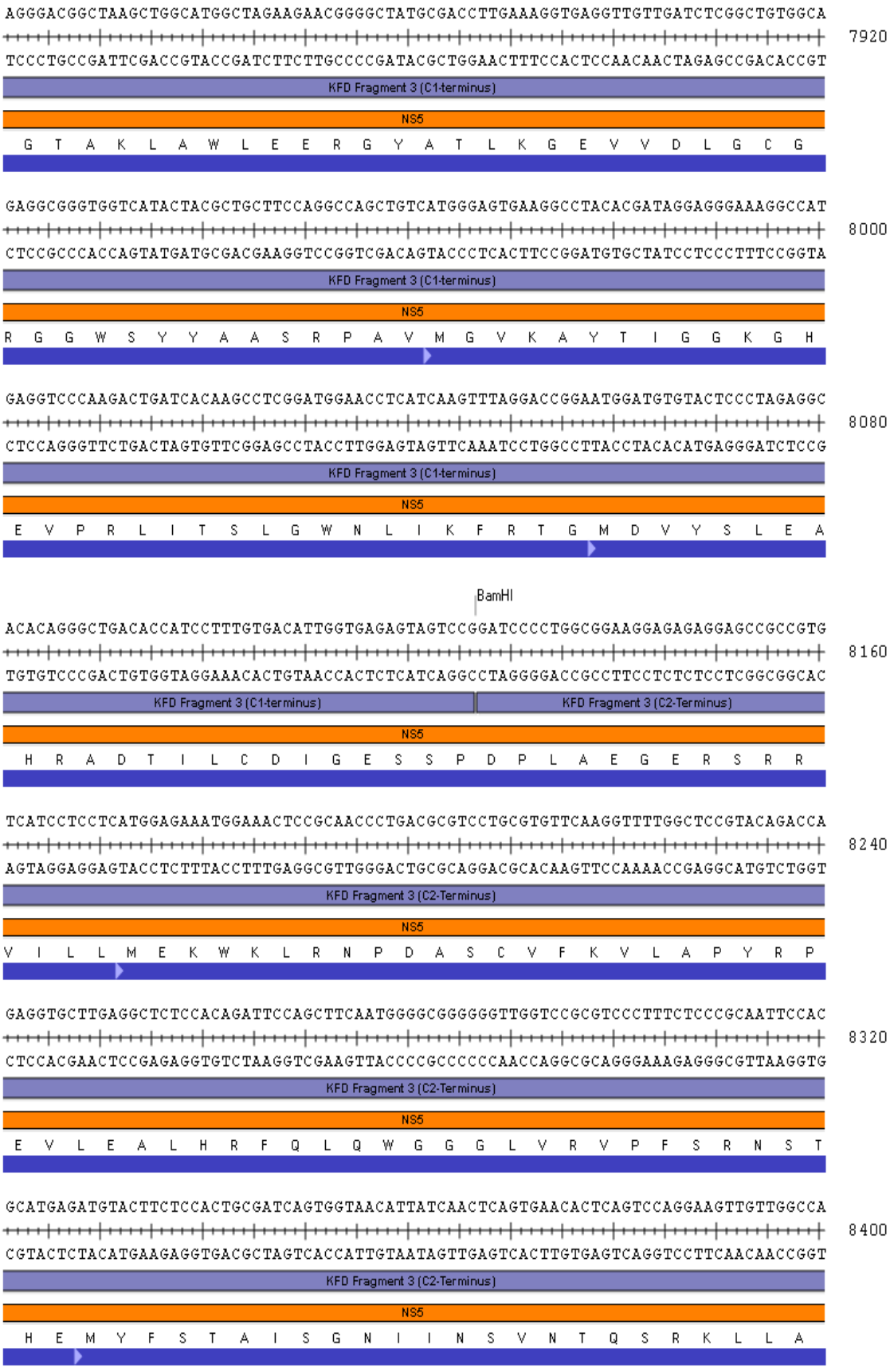

AGGGAGAGGTGGTGGGTCTTTATGGGAACGGCCTGAAGACCAATGAAGCCTATGTCAGCAGCATAGCACAAGGAGAAGCC
5120
TCCCTCTCCACCACCCAGAAATACCCTTGCCGGACTTCTGGTTACTTCGGATACAGTCGTCGTATCGTGTTCCTCTTCGG
KFD Fragment 2
NS3
Q G E V V G L Y G N G L K T N E A Y V S S I A Q G E A
GAGAAAAGCAGACCGGAAATCCCGTTGAGTGTGCAAGGAACAGGCTGGATGTCAAAAGGTCAAATCACGGTGCTGGACAT
5200
CTCTTTTCGTCTGGCCTTTAGGGCAACTCACACGTTCCTTGTCCGACCTACAGTTTTCCAGTTTAGTGCCACGACCTGTA
KFD Fragment 2
NS3
E K S R P E I P L S V Q G T G W M S K G Q I T V L D M
GCATCCGGGGTCAGGGAAGACACATCGAGTCCTCCCGGAGCTTGTGAGGCAGTGTGCAGACCGAGGGATGAGGACTCTAG
5280
CGTAGGCCCCAGTCCCTTCTGTGTAGCTCAGGAGGGCCTCGAACACTCCGTCACACGTCTGGCTCCCTACTCCTGAGATC
KFD Fragment 2
NS3
H P G S G K T H R V L P E L V R Q C A D R G M R T L
TTCTTGCCCCCACCCGGGTGGTCCTGAAGGAGATGGAGAGAGCGCTTGCTGGCAAGAAGGTCCGATTCCACTCACCCGCG
5360
AAGAACGGGGGTGGGCCCACCAGGACTTCCTCTACCTCTCTCGCGAACGACCGTTCTTCCAGGCTAAGGTGAGTGGGCGC
KFD Fragment 2
NS3
V L A P T R V V L K E M E R A L A G K K V R F H S P A
BsiWI
GTGGAAGGGCAGACCACAGCGGGGGCCATTGTTGACGTCATGTGCCATGCCACGTACGTCCACAGAAGGCTGCTGCCCCA
5440
CACCTTCCCGTCTGGTGTCGCCCCCGGTAACAACTGCAGTACACGGTACGGTGCATGCAGGTGTCTTCCGACGACGGGGT
KFD Fragment 2
KFD Fragment 3 (N1-terminus)
KFD Frag3-pTM1s (1-36)
NS3
V E G Q T T A G A I V D V M C H A T Y V H R R L L P Q
AGGGAGACAAAATTGGGAAGTCGCCATAATGGATGAGGCCCATTGGACTGATCCGCACAGTATCGCGGCAAGAGGGCATC
5520
TCCCTCTGTTTTAACCCTTCAGCGGTATTACCTACTCCGGGTAACCTGACTAGGCGTGTCATAGCGCCGTTCTCCCGTAG
KFD Fragment 3 (N1-terminus)
NS3
G R Q N W E V A I M D E A H W T D P H S I A A R G H
TTTATTCGCTCGCCAAAGAGAACAGGTGCGCCTTGGTGCTCATGACAGCAACTCCACCCGGCCGGGGAGATCCTTTCCCG
5600
AAATAAGCGAGCGGTTTCTCTTGTCCACGCGGAACCACGAGTACTGTCGTTGAGGTGGGCCGGCCCCTCTAGGAAAGGGC
KFD Fragment 3 (N1-terminus)
NS3
L Y S L A K E N R C A L V L M T A T P P G R G D P F P

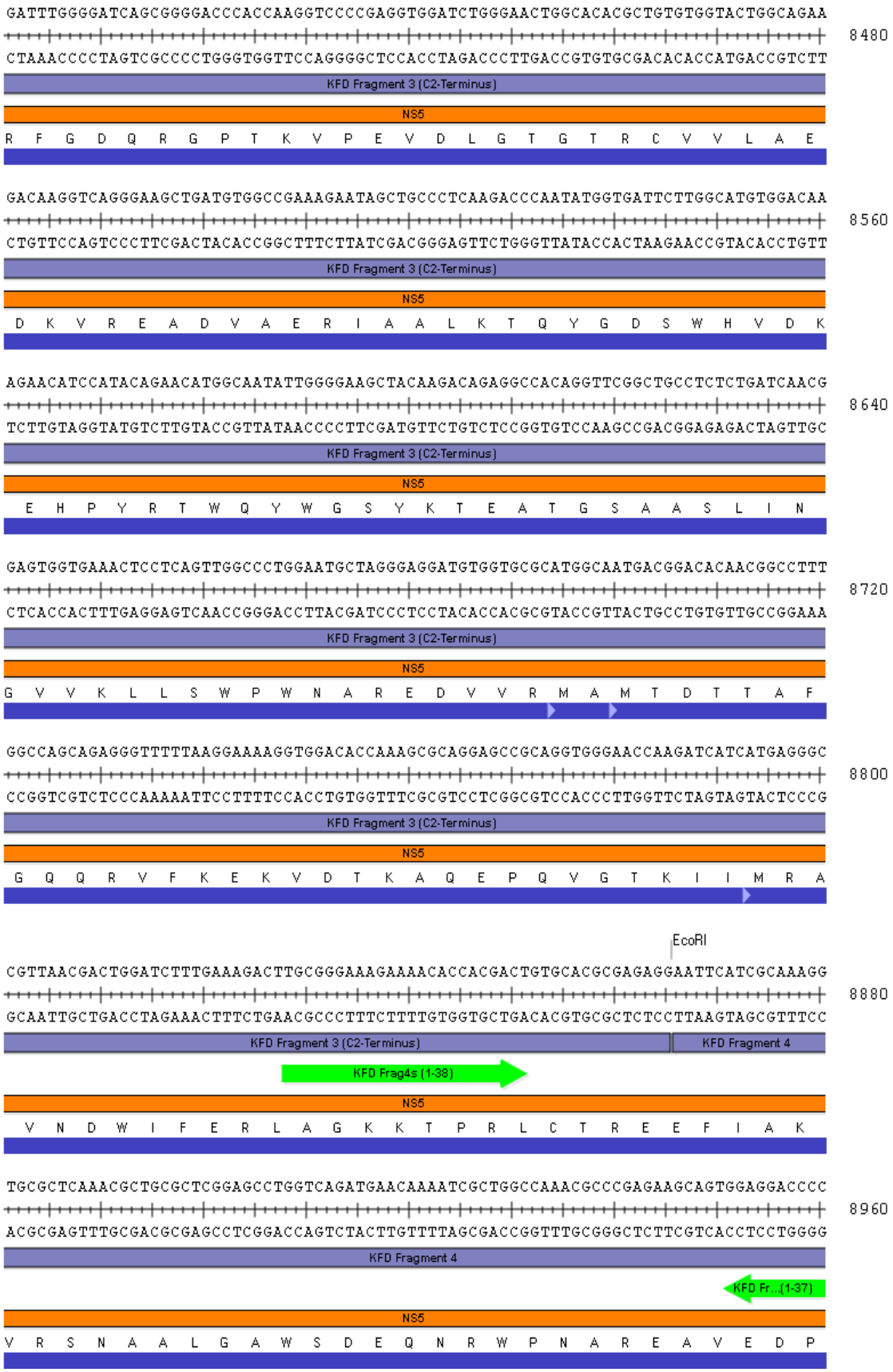

GAGTCCAATGGAGCCATTATGAGTGAGGAAAGGGCCATTCCGGACGGAGAGTGGAGAGAGGGCTTTGACTGGATAACAGA
5680
CTCAGGTTACCTCGGTAATACTCACTCCTTTCCCGGTAAGGCCTGCCTCTCACCTCTCTCCCGAAACTGACCTATTGTCT
KFD Fragment 3 (N1-terminus)
NS3
E S N G A I M S E E R A I P D G E W R E G F D W I T E
GTATGAGGGTCGAACAGCATGGTTCGTTCCATCCATATCAAAAGGTGGGGCGGTTGCAAGAACCCTGCGCCAGAGAGGCA
5760
CATACTCCCAGCTTGTCGTACCAAGCAAGGTAGGTATAGTTTTCCACCCCGCCAACGTTCTTGGGACGCGGTCTCTCCGT
KFD Fragment 3 (N1-terminus)
KFD Frag2-pTM1as (1-35)
NS3
Y E G R T A W F V P S I S K G G A V A R T L R Q R G
AAAGTGTGATCTGTCTCAACAGTAAAACCTTTGAGAAGGATTACCTGAGGGTCAGGGAAGAGAAGCCAGATTTTGTGGTG
5840
TTTCACACTAGACAGAGTTGTCATTTTGGAAACTCTTCCTAATGGACTCCCAGTCCCTTCTCTTCGGTCTAAAACACCAC
KFD Fragment 3 (N1-terminus)
NS3
K S V I C L N S K T F E K D Y L R V R E E K P D F V V
ACAACAGATATCTCTGAGATGGGTGCCAACCTGGATGTCAGCCGAGTTATTGATGGTAGAACAAACATAAAGCCAGAGGA
5920
TGTTGTCTATAGAGACTCTACCCACGGTTGGACCTACAGTCGGCTCAATAACTACCATCTTGTTTGTATTTCGGTCTCCT
KFD Fragment 3 (N1-terminus)
NS3
T T D I S E M G A N L D V S R V I D G R T N I K P E E
GGTGGATGGAAAGGTAGAGTTGACTGGAACCAGGAAGGTCACGACAGCATCGGCGGCTCAGCGCCGGGGCAGGGTGGGAA
6000
CCACCTACCTTTCCATCTCAACTGACCTTGGTCCTTCCAGTGCTGTCGTAGCCGCCGAGTCGCGGCCCCGTCCCACCCTT
KFD Fragment 3 (N1-terminus)
NS3
V D G K V E L T G T R K V T T A S A A Q R R G R V G
GGACATCTGGACGCACTGATGAGTACATATACTCTGGACAGTGTGATGATGACGACACCAGCCTCGTTCAATGGAAGGAG
6080
CCTGTAGACCTGCGTGACTACTCATGTATATGAGACCTGTCACACTACTACTGCTGTGGTCGGAGCAAGTTACCTTCCTC
KFD Fragment 3 (N1-terminus)
NS3
R T S G R T D E Y I Y S G Q C D D D D T S L V Q W K E
GCGCAGATACTCCTGGACAATATCACCACACTGAGGGGTCCGGTCGCAACTTTTTATGGACCCGAGCAGGTGAAGATGCC
6160
CGCGTCTATGAGGACCTGTTATAGTGGTGTGACTCCCCAGGCCAGCGTTGAAAAATACCTGGGCTCGTCCACTTCTACGG
KFD Fragment 3 (N1-terminus)
NS3
A Q I L L D N I T T L R G P V A T F Y G P E Q V K M P

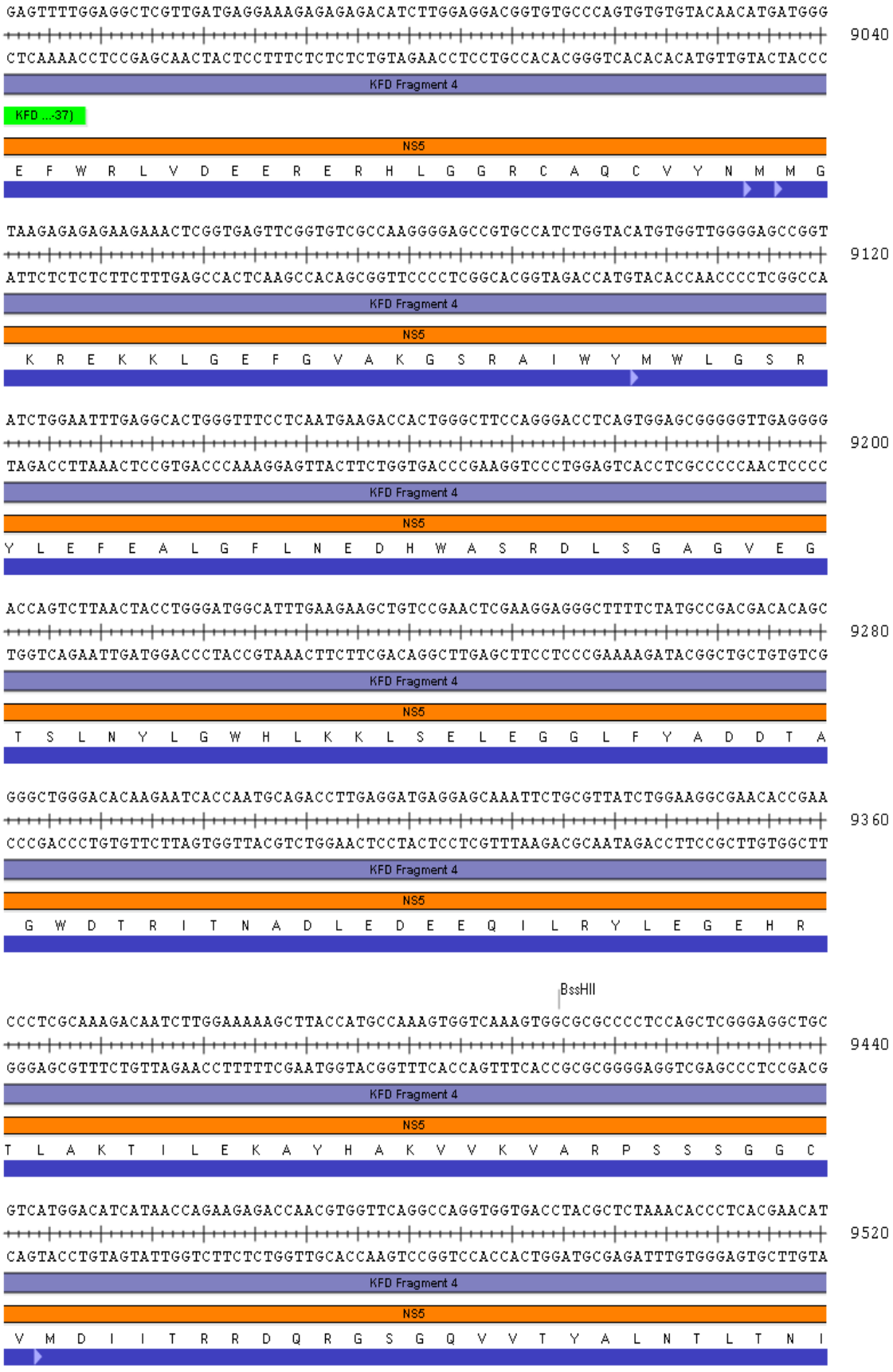

AGAAGTGGCTGGCCACTACAGGCTGAATGAGGAGAAAAGGAAACACTTCCGGCACTTGATGACTCAATGTGACTTCACTC
6240
TCTTCACCGACCGGTGATGTCCGACTTACTCCTCTTTTCCTTTGTGAAGGCCGTGAACTACTGAGTTACACTGAAGTGAG
KFD Fragment 3 (N1-terminus)
NS3
E V A G H Y R L N E E K R K H F R H L M T Q C D F T
CCTGGCTGGCGTGGCATGTCGCCACCAACACCAGCAATGTGTTGGATCGGAGCTGGACCTGGCAGGGGCCGGAGGAAAAT
6320
GGACCGACCGCACCGTACAGCGGTGGTTGTGGTCGTTACACAACCTAGCCTCGACCTGGACCGTCCCCGGCCTCCTTTTA
KFD Fragment 3 (N1-terminus)
NS3
P W L A W H V A T N T S N V L D R S W T W Q G P E E N
GCCATTGACGGGGCCGACGGAGATCTGGTGCGGTTTAAAACGCCTGGCGGCTCAGAACGTGTCTTGCAGCCGGTGTGGAA
6400
CGGTAACTGCCCCGGCTGCCTCTAGACCACGCCAAATTTTGCGGACCGCCGAGTCTTGCACAGAACGTCGGCCACACCTT
KFD Fragment 3 (N1-terminus)
NS3
A I D G A D G D L V R F K T P G G S E R V L Q P V W K
GGACTGCAGAATGTTCCGTGAGGGACGGGACGTCAAAGACTTCATACTGTACGCCTCCGGGCGTCGCAGTGTTGGTGATG
6480
CCTGACGTCTTACAAGGCACTCCCTGCCCTGCAGTTTCTGAAGTATGACATGCGGAGGCCCGCAGCGTCACAACCACTAC
KFD Fragment 3 (N1-terminus)
NS3
NS4A
D C R M F R E G R D V K D F I L Y A S G R R S V G D
BssHII
TTTTGGGTGGTCTGGCGGGGGTCCCCGGTCTCTTGAGGCACAGGTGCGCCAGCGCGCTTGACGTAGTTTACACCCTTCTC
6560
AAAACCCACCAGACCGCCCCCAGGGGCCAGAGAACTCCGTGTCCACGCGGTCGCGCGAACTGCATCAAATGTGGGAAGAG
KFD Fragment 3 (N1-terminus)
NS4A
KFD Fragment 3 (C1-terminus)
V L G G L A G V P G L L R H R C A S A L D V V Y T L L
AATGAGAATCCGGGCAGCAGGGCAATGCGCATGGCTGAGAGAGATGCACCAGAGGCATTTCTGACTATTGTTGAGGTTGC
6640
TTACTCTTAGGCCCGTCGTCCCGTTACGCGTACCGACTCTCTCTACGTGGTCTCCGTAAAGACTGATAACAACTCCAACG
KFD Fragment 3 (C1-terminus)
NS4A
N E N P G S R A M R M A E R D A P E A F L T I V E V A
TGTTTTAGGAGTGGCCACTCTTGGCATACTCTGGTGCTTTGTGGCAAGAGCTTCGGTCAGCCGCATGTTCCTGGGCACGG
6720
ACAAAATCCTCACCGGTGAGAACCGTATGAGACCACGAAACACCGTTCTCGAAGCCAGTCGGCGTACAAGGACCCGTGCC
KFD Fragment 3 (C1-terminus)
NS4A
V L G V A T L G I L W C F V A R A S V S R M F L G T

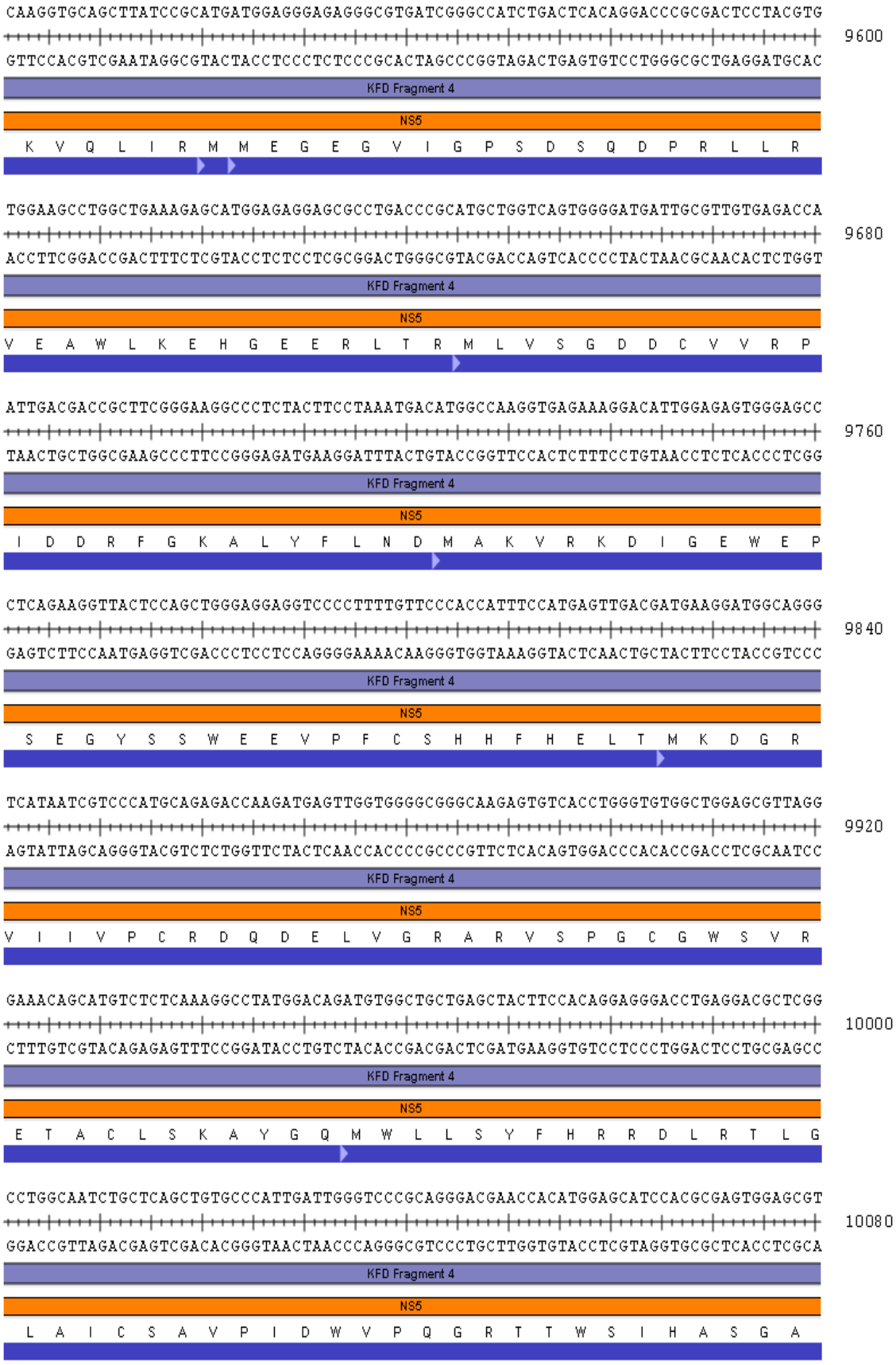

TTGTGCTCTTTGCTGCCCTCTTTCTGCTGTGGATTGGTGGAGTGGACTATGGACACATGGCAGGGATAGCCCTGATATTT
6800
AACACGAGAAACGACGGGAGAAAGACGACACCTAACCACCTCACCTGATACCTGTGTACCGTCCCTATCGGGACTATAAA
KFD Fragment 3 (C1-terminus)
NS4A
V V L F A A L F L L W I G G V D Y G H M A G I A L I F
TACACCCTCCTCACTGTGCTGCAGCCAGAGCCGGGAAAACAGCGAAGTAGCGATGACAACAGACTGGCATACTTCCTCCT
6880
ATGTGGGAGGAGTGACACGACGTCGGTCTCGGCCCTTTTGTCGCTTCATCGCTACTGTTGTCTGACCGTATGAAGGAGGA
KFD Fragment 3 (C1-terminus)
NS4A
2k
Y T L L T V L Q P E P G K Q R S S D D N R L A Y F L L
GGGGCTTTTTAGTTTGGCTGGACTCGTGACGGCCAACGAAATGGGCATGTTGGACAAGACAAAGGCTGACCTTGCAGGGC
6960
CCCCGAAAAAATCAAACCGACCTGAGCACTGCCGGTTGCTTTACCCGTACAACCTGTTCTGTTTCCGACTGGAACGTCCCG
KFD Fragment 3 (C1-terminus)
2k
NS4B
G L F S L A G L V T A N E M G M L D K T K A D L A G
TGGTGTGGCGCGGAGAGCAGAGGCATCCAGCCTGGGAGGAGTGGACAAATGTGGACATCCAACCAGCACGATCTTGGGGG
7040
ACCACACCGCGCCTCTCGTCTCCGTAGGTCGGACCCTCCTCACCTGTTTACACCTGTAGGTTGGTCGTGCTAGAACCCCC
KFD Fragment 3 (C1-terminus)
NS4B
L V W R G E Q R H P A W E E W T N V D I Q P A R S W G
ACGTATGTGCTGATCGTCTCTCTGTTCACTCCCTACATGCTGCACCAGCTTCAGACAAAAATACAACAACTGGTGAACAG
7120
TGCATACACGACTAGCAGAGAGACAAGTGAGGGATGTACGACGTGGTCGAAGTCTGTTTTTATGTTGTTGACCACTTGTC
KFD Fragment 3 (C1-terminus)
NS4B
T Y V L I V S L F T P Y M L H Q L Q T K I Q Q L V N S
TTCAGTGGCAAGTGGTGCTCAGGCTATGAGAGATCTTGGGGGTGGGACACCATTCTTCGGGGTGGCCGGCCATGTGATTG
7200
AAGTCACCGTTCACCACGAGTCCGATACTCTCTAGAACCCCCACCCTGTGGTAAGAAGCCCCACCGGCCGGTACACTAAC
KFD Fragment 3 (C1-terminus)
NS4B
S V A S G A Q A M R D L G G G T P F F G V A G H V I
CCTTGGGCGTCACGTCACTCGTGGGTGCCACTCCCATGTCCTTGGGCTTGGGGGTGGCTTTGGCTGCTTTCCACTTGGCC
7280
GGAACCCGCAGTGCAGTGAGCACCCACGGTGAGGGTACAGGAACCCGAACCCCCACCGAAACCGACGAAAGGTGAACCGG
KFD Fragment 3 (C1-terminus)
NS4B
A L G V T S L V G A T P M S L G L G V A L A A F H L A

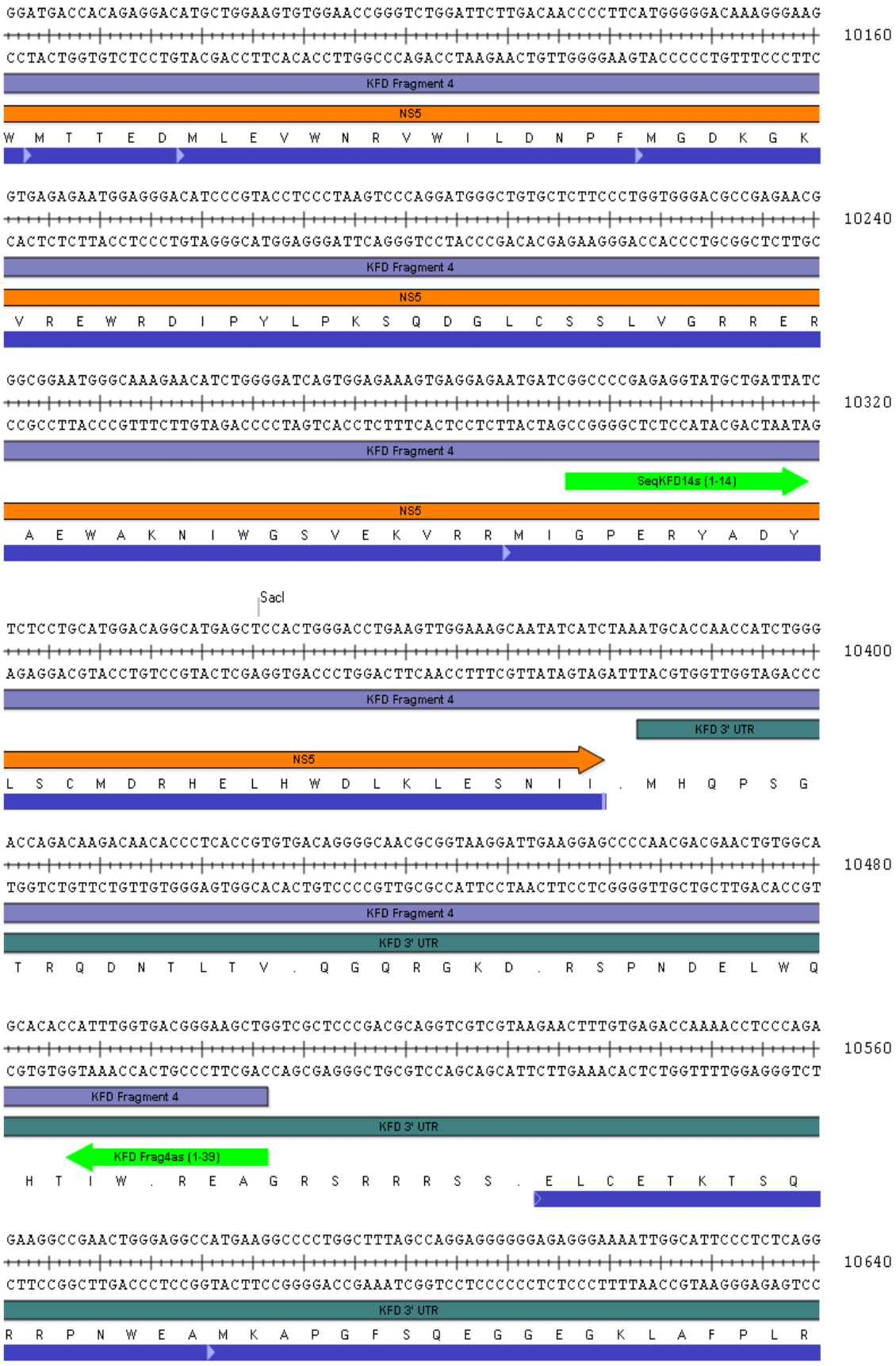

ATCGTCGCCTCGGGATTGGAAGCAGAGCTGACGCAGAGGGCGCACAGGGTATTCTTCTCGGCAATGGTTAAAAACCCAAT
7360
TAGCAGCGGAGCCCTAACCTTCGTCTCGACTGCGTCTCCCGCGTGTCCCATAAGAAGAGCCGTTACCAATTTTTGGGTTA
KFD Fragment 3 (C1-terminus)
NS4B
I V A S G L E A E L T Q R A H R V F F S A M V K N P M
GGTGGATGGAGACGTCATCAATCCCTTCCCCGATGGAGAGACAAAGCCAGCGCTATACGAGAGAAGGATGAGTTTAATCC
7440
CCACCTACCTCTGCAGTAGTTAGGGAAGGGGCTACCTCTCTGTTTCGGTCGCGATATGCTCTCTTCCTACTCAAATTAGG
KFD Fragment 3 (C1-terminus)
NS4B
V D G D V I N P F P D G E T K P A L Y E R R M S L I
TGGCCATTGCCTTGTGCATGGGTTCAGTGGTCCTCAACCGAACCGCGGCCTCAATGACCGAGGCTGGGGCTGTTGGCCTA
7520
ACCGGTAACGGAACACGTACCCAAGTCACCAGGAGTTGGCTTGGCGCCGGAGTTACTGGCTCCGACCCCGACAACCGGAT
KFD Fragment 3 (C1-terminus)
NS4B
L A I A L C M G S V V L N R T A A S M T E A G A V G L
GCAGCGCTGGGTCAGCTGGTCCATCCCGAGACTGAGACACTCTGGACAATGCCAATGGCGTGTGGTATGGCGGGGCTCGT
7600
CGTCGCGACCCAGTCGACCAGGTAGGGCTCTGACTCTGTGAGACCTGTTACGGTTACCGCACACCATACCGCCCCGAGCA
KFD Fragment 3 (C1-terminus)
NS4B
A A L G Q L V H P E T E T L W T M P M A C G M A G L V
GCGGGGGAGCTTCTGGGGCCTGTTGCCCATGGGACACCGACTGTGGCTTAGGACAACTGGAACACGCAGAGGAGGTGCCG
7680
CGCCCCCTCGAAGACCCCGGACAACGGGTACCCTGTGGCTGACACCGAATCCTGTTGACCTTGTGCGTCTCCTCCACGGC
KFD Fragment 3 (C1-terminus)
NS4B
NS5
R G S F W G L L P M G H R L W L R T T G T R R G G A
AGGGAGAGACTCTTGGGGACATCTGGAAGAGGCGGCTGAATGGTTGCTCTAGGGAGGAGTTTTTTCAGTACCGGCGAAGT
7760
TCCCTCTCTGAGAACCCCTGTAGACCTTCTCCGCCGACTTACCAACGAGATCCCTCCTCAAAAAAGTCATGGCCGCTTCA
KFD Fragment 3 (C1-terminus)
NS5
E G E T L G D I W K R R L N G C S R E E F F Q Y R R S
GGAGTCATGGAGACGGAACGTGACAAGGCTCGAGAGCTTCTGAAGAGAGGTGAGACAAACATGGGACTGGCTGTCTCTCG
7840
CCTCAGTACCTCTGCCTTGCACTGTTCCGAGCTCTCGAAGACTTCTCTCCACTCTGTTTGTACCCTGACCGACAGAGAGC
KFD Fragment 3 (C1-terminus)
NS5
G V M E T E R D K A R E L L K R G E T N M G L A V S R


AGGGACGGCTAAGCTGGCATGGCTAGAAGAACGGGGCTATGCGACCTTGAAAGGTGAGGTTGTTGATCTCGGCTGTGGCA
7920
TCCCTGCCGATTCGACCGTACCGATCTTCTTGCCCCGATACGCTGGAACTTTCCACTCCAACAACTAGAGCCGACACCGT
KFD Fragment 3 (C1-terminus)
NS5
G T A K L A W L E E R G Y A T L K G E V V D L G C G
GAGGCGGGTGGTCATACTACGCTGCTTCCAGGCCAGCTGTCATGGGAGTGAAGGCCTACACGATAGGAGGGAAAGGCCAT
8000
CTCCGCCCACCAGTATGATGCGACGAAGGTCCGGTCGACAGTACCCTCACTTCCGGATGTGCTATCCTCCCTTTCCGGTA
KFD Fragment 3 (C1-terminus)
NS5
R G G W S Y Y A A S R P A V M G V K A Y T I G G K G H
GAGGTCCCAAGACTGATCACAAGCCTCGGATGGAACCTCATCAAGTTTAGGACCGGAATGGATGTGTACTCCCTAGAGGC
8080
CTCCAGGGTTCTGACTAGTGTTCGGAGCCTACCTTGGAGTAGTTCAAATCCTGGCCTTACCTACACATGAGGGATCTCCG
KFD Fragment 3 (C1-terminus)
NS5
E V P R L I T S L G W N L I K F R T G M D V Y S L E A
BamHI
ACACAGGGCTGACACCATCCTTTGTGACATTGGTGAGAGTAGTCCGGATCCCCTGGCGGAAGGAGAGAGGAGCCGCCGTG
8160
TGTGTCCCGACTGTGGTAGGAAACACTGTAACCACTCTCATCAGGCCTAGGGGACCGCCTTCCTCTCTCCTCGGCGGCAC
KFD Fragment 3 (C1-terminus)
KFD Fragment 3 (C2-Terminus)
NS5
H R A D T I L C D I G E S S P D P L A E G E R S R R
TCATCCTCCTCATGGAGAAATGGAAACTCCGCAACCCTGACGCGTCCTGCGTGTTCAAGGTTTTGGCTCCGTACAGACCA
8240
AGTAGGAGGAGTACCTCTTTACCTTTGAGGCGTTGGGACTGCGCAGGACGCACAAGTTCCAAAACCGAGGCATGTCTGGT
KFD Fragment 3 (C2-Terminus)
NS5
V I L L M E K W K L R N P D A S C V F K V L A P Y R P
GAGGTGCTTGAGGCTCTCCACAGATTCCAGCTTCAATGGGGCGGGGGGGTTGGTCCGCGTCCCTTTCTCCCGCAATTCCAC
8320
CTCCACGAACTCCGAGAGGTGTCTAAGGTCGAAGTTACCCCGCCCCCCAACCAGGCGCAGGGAAAGAGGGCGTTAAGGTG
KFD Fragment 3 (C2-Terminus)
NS5
E V L E A L H R F Q L Q W G G G L V R V P F S R N S T
GCATGAGATGTACTTCTCCACTGCGATCAGTGGTAACATTATCAACTCAGTGAACACTCAGTCCAGGAAGTTGTTGGCCA
8400
CGTACTCTACATGAAGAGGTGACGCTAGTCACCATTGTAATAGTTGAGTCACTTGTGAGTCAGGTCCTTCAACAACCGGT
KFD Fragment 3 (C2-Terminus)
NS5
H E M Y F S T A I S G N I I N S V N T Q S R K L L A


GATTTGGGGATCAGCGGGGACCCACCAAGGTCCCCGAGGTGGATCTGGGAACTGGCACACGCTGTGTGGTACTGGCAGAA
8480
CTAAACCCCTAGTCGCCCCTGGGTGGTTCCAGGGGCTCCACCTAGACCCTTGACCGTGTGCGACACACCATGACCGTCTT
KFD Fragment 3 (C2-Terminus)
NS5
R F G D Q R G P T K V P E V D L G T G T R C V V L A E
GACAAGGTCAGGGAAGCTGATGTGGCCGAAAGAATAGCTGCCCTCAAGACCCAATATGGTGATTCTTGGCATGTGGACAA
8560
CTGTTCCAGTCCCTTCGACTACACCGGCTTTCTTATCGACGGGAGTTCTGGGTTATACCACTAAGAACCGTACACCTGTT
KFD Fragment 3 (C2-Terminus)
NS5
D K V R E A D V A E R I A A L K T Q Y G D S W H V D K
AGAACATCCATACAGAACATGGCAATATTGGGGAAGCTACAAGACAGAGGCCACAGGTTCGGCTGCCTCTCTGATCAACG
8640
TCTTGTAGGTATGTCTTGTACCGTTATAACCCCTTCGATGTTCTGTCTCCGGTGTCCAAGCCGACGGAGAGACTAGTTGC
KFD Fragment 3 (C2-Terminus)
NS5
E H P Y R T W Q Y W G S Y K T E A T G S A A S L I N
GAGTGGTGAAACTCCTCAGTTGGCCCTGGAATGCTAGGGAGGATGTGGTGCGCATGGCAATGACGGACACAACGGCCTTT
8720
CTCACCACTTTGAGGAGTCAACCGGGACCTTACGATCCCTCCTACACCACGCGTACCGTTACTGCCTGTGTTGCCGGAAA
KFD Fragment 3 (C2-Terminus)
NS5
G V V K L L S W P W N A R E D V V R M A M T D T T A F
GGCCAGCAGAGGGTTTTTAAGGAAAAGGTGGACACCAAAGCGCAGGAGCCGCAGGTGGGAACCAAGATCATCATGAGGGC
8800
CCGGTCGTCTCCCAAAAATTCCTTTTCCACCTGTGGTTTCGCGTCCTCGGCGTCCACCCTTGGTTCTAGTAGTACTCCCG
KFD Fragment 3 (C2-Terminus)
NS5
G Q Q R V F K E K V D T K A Q E P Q V G T K I I M R A
EcoRI
CGTTAACGACTGGATCTTTGAAAGACTTGCGGGAAAGAAAACACCACGACTGTGCACGCGAGAGGAATTCATCGCAAAGG
8880
GCAATTGCTGACCTAGAAACTTTCTGAACGCCCTTTCTTTTGTGGTGCTGACACGTGCGCTCTCCTTAAGTAGCGTTTCC
KFD Fragment 3 (C2-Terminus)
KFD Fragment 4
KFD Frag4s (1-38)
NS5
V N D W I F E R L A G K K T P R L C T R E E F I A K
TGCGCTCAAACGCTGCGCTCGGAGCCTGGTCAGATGAACAAAATCGCTGGCCAAACGCCCGAGAAGCAGTGGAGGACCCC
8960
ACGCGAGTTTGCGACGCGAGCCTCGGACCAGTCTACTTGTTTTAGCGACCGGTTTGCGGGCTCTTCGTCACCTCCTGGGG
KFD Fragment 4
KFD Fr...(1-37)
NS5
V R S N A A L G A W S D E Q N R W P N A R E A V E D P


GAGTTTTGGAGGCTCGTTGATGAGGAAAGAGAGAGACATCTTGGAGGACGGTGTGCCCAGTGTGTGTACAACATGATGGG
9040
CTCAAAACCTCCGAGCAACTACTCCTTTCTCTCTCTGTAGAACCTCCTGCCACACGGGTCACACACATGTTGTACTACCC
KFD Fragment 4
KFD ...-37)
NS5
E F W R L V D E E R E R H L G G R C A Q C V Y N M M G
TAAGAGAGAGAAGAAACTCGGTGAGTTCGGTGTCGCCAAGGGGAGCCGTGCCATCTGGTACATGTGGTTGGGGAGCCGGT
9120
ATTCTCTCTCTTCTTTGAGCCACTCAAGCCACAGCGGTTCCCCTCGGCACGGTAGACCATGTACACCAACCCCTCGGCCA
KFD Fragment 4
NS5
K R E K K L G E F G V A K G S R A I W Y M W L G S R
ATCTGGAATTTGAGGCACTGGGTTTCCTCAATGAAGACCACTGGGCTTCCAGGGACCTCAGTGGAGCGGGGGTTGAGGGG
9200
TAGACCTTAAACTCCGTGACCCAAAGGAGTTACTTCTGGTGACCCGAAGGTCCCTGGAGTCACCTCGCCCCCAACTCCCC
KFD Fragment 4
NS5
Y L E F E A L G F L N E D H W A S R D L S G A G V E G
ACCAGTCTTAACTACCTGGGATGGCATTTGAAGAAGCTGTCCGAACTCGAAGGAGGGCTTTTCTATGCCGACGACACAGC
9280
TGGTCAGAATTGATGGACCCTACCGTAAACTTCTTCGACAGGCTTGAGCTTCCTCCCGAAAAGATACGGCTGCTGTGTCG
KFD Fragment 4
NS5
T S L N Y L G W H L K K L S E L E G G L F Y A D D T A
GGGCTGGGACACAAGAATCACCAATGCAGACCTTGAGGATGAGGAGCAAATTCTGCGTTATCTGGAAGGCGAACACCGAA
9360
CCCGACCCTGTGTTCTTAGTGGTTACGTCTGGAACTCCTACTCCTCGTTTAAGACGCAATAGACCTTCCGCTTGTGGCTT
KFD Fragment 4
NS5
G W D T R I T N A D L E D E E Q I L R Y L E G E H R
BssHII
CCCTCGCAAAGACAATCTTGGAAAAAGCTTACCATGCCAAAGTGGTCAAAGTGGCGCGCCCCTCCAGCTCGGGAGGCTGC
9440
GGGAGCGTTTCTGTTAGAACCTTTTTCGAATGGTACGGTTTCACCAGTTTCACCGCGCGGGGAGGTCGAGCCCTCCGACG
KFD Fragment 4
NS5
T L A K T I L E K A Y H A K V V K V A R P S S S G G C
GTCATGGACATCATAACCAGAAGAGACCAACGTGGTTCAGGCCAGGTGGTGACCTACGCTCTAAACACCCTCACGAACAT
9520
CAGTACCTGTAGTATTGGTCTTCTCTGGTTGCACCAAGTCCGGTCCACCACTGGATGCGAGATTTGTGGGAGTGCTTGTA
KFD Fragment 4
NS5
V M D I I T R R D Q R G S G Q V V T Y A L N T L T N I


CAAGGTGCAGCTTATCCGCATGATGGAGGGAGAGGGCGTGATCGGGCCATCTGACTCACAGGACCCGCGACTCCTACGTG
9600
GTTCCACGTCGAATAGGCGTACTACCTCCCTCTCCCGCACTAGCCCGGTAGACTGAGTGTCCTGGGCGCTGAGGATGCAC
KFD Fragment 4
NS5
K V Q L I R M M E G E G V I G P S D S Q D P R L L R
TGGAAGCCTGGCTGAAAGAGCATGGAGAGGAGCGCCTGACCCGCATGCTGGTCAGTGGGGATGATTGCGTTGTGAGACCA
9680
ACCTTCGGACCGACTTTCTCGTACCTCTCCTCGCGGACTGGGCGTACGACCAGTCACCCCTACTAACGCAACACTCTGGT
KFD Fragment 4
NS5
V E A W L K E H G E E R L T R M L V S G D D C V V R P
ATTGACGACCGCTTCGGGAAGGCCCTCTACTTCCTAAATGACATGGCCAAGGTGAGAAAGGACATTGGAGAGTGGGAGCC
9760
TAACTGCTGGCGAAGCCCTTCCGGGAGATGAAGGATTTACTGTACCGGTTCCACTCTTTCCTGTAACCTCTCACCCTCGG
KFD Fragment 4
NS5
I D D R F G K A L Y F L N D M A K V R K D I G E W E P
CTCAGAAGGTTACTCCAGCTGGGAGGAGGTCCCCTTTTGTTCCCACCATTTCCATGAGTTGACGATGAAGGATGGCAGGG
9840
GAGTCTTCCAATGAGGTCGACCCTCCTCCAGGGGAAAACAAGGGTGGTAAAGGTACTCAACTGCTACTTCCTACCGTCCC
KFD Fragment 4
NS5
S E G Y S S W E E V P F C S H H F H E L T M K D G R
TCATAATCGTCCCATGCAGAGACCAAGATGAGTTGGTGGGGCGGGCAAGAGTGTCACCTGGGTGTGGCTGGAGCGTTAGG
9920
AGTATTAGCAGGGTACGTCTCTGGTTCTACTCAACCACCCCGCCCGTTCTCACAGTGGACCCACACCGACCTCGCAATCC
KFD Fragment 4
NS5
V I I V P C R D Q D E L V G R A R V S P G C G W S V R
GAAACAGCATGTCTCTCAAAGGCCTATGGACAGATGTGGCTGCTGAGCTACTTCCACAGGAGGGACCTGAGGACGCTCGG
10000
CTTTGTCGTACAGAGAGTTTCCGGATACCTGTCTACACCGACGACTCGATGAAGGTGTCCTCCCTGGACTCCTGCGAGCC
KFD Fragment 4
NS5
E T A C L S K A Y G Q M W L L S Y F H R R D L R T L G
CCTGGCAATCTGCTCAGCTGTGCCCATTGATTGGGTCCCGCAGGGACGAACCACATGGAGCATCCACGCGAGTGGAGCGT
10080
GGACCGTTAGACGAGTCGACACGGGTAACTAACCCAGGGCGTCCCTGCTTGGTGTACCTCGTAGGTGCGCTCACCTCGCA
KFD Fragment 4
NS5
L A I C S A V P I D W V P Q G R T T W S I H A S G A


GGATGACCACAGAGGACATGCTGGAAGTGTGGAACCGGGTCTGGATTCTTGACAACCCCTTCATGGGGGACAAAGGGAAG
10160
CCTACTGGTGTCTCCTGTACGACCTTCACACCTTGGCCCAGACCTAAGAACTGTTGGGGAAGTACCCCCTGTTTCCCTTC
KFD Fragment 4
NS5
W M T T E D M L E V W N R V W I L D N P F M G D K G K
GTGAGAGAATGGAGGGACATCCCGTACCTCCCTAAGTCCCAGGATGGGCTGTGCTCTTCCCTGGTGGGACGCCGAGAACG
10240
CACTCTCTTACCTCCCTGTAGGGCATGGAGGGATTCAGGGTCCTACCCGACACGAGAAGGGACCACCCTGCGGCTCTTGC
KFD Fragment 4
NS5
V R E W R D I P Y L P K S Q D G L C S S L V G R R E R
GGCGGAATGGGCAAAGAACATCTGGGGATCAGTGGAGAAAGTGAGGAGAATGATCGGCCCCGAGAGGTATGCTGATTATC
10320
CCGCCTTACCCGTTTCTTGTAGACCCCTAGTCACCTCTTTCACTCCTCTTACTAGCCGGGGCTCTCCATACGACTAATAG
KFD Fragment 4
SeqKFD14s (1-14)
NS5
A E W A K N I W G S V E K V R R M I G P E R Y A D Y
SacI
TCTCCTGCATGGACAGGCATGAGCTCCACTGGGACCTGAAGTTGGAAAGCAATATCATCTAAATGCACCAACCATCTGGG
10400
AGAGGACGTACCTGTCCGTACTCGAGGTGACCCTGGACTTCAACCTTTCGTTATAGTAGATTTACGTGGTTGGTAGACCC
KFD Fragment 4
KFD 3' UTR
NS5
L S C M D R H E L H W D L K L E S N I I . M H Q P S G
ACCAGACAAGACAACACCCTCACCGTGTGACAGGGGCAACGCGGTAAGGATTGAAGGAGCCCCAACGACGAACTGTGGCA
10480
TGGTCTGTTCTGTTGTGGGAGTGGCACACTGTCCCCGTTGCGCCATTCCTAACTTCCTCGGGGTTGCTGCTTGACACCGT
KFD Fragment 4
KFD 3' UTR
T R Q D N T L T V . Q G Q R G K D . R S P N D E L W Q
GCACACCATTTGGTGACGGGAAGCTGGTCGCTCCCGACGCAGGTCGTCGTAAGAACTTTGTGAGACCAAAACCTCCCAGA
10560
CGTGTGGTAAACCACTGCCCTTCGACCAGCGAGGGCTGCGTCCAGCAGCATTCTTGAAACACTCTGGTTTTGGAGGGTCT
KFD Fragment 4
KFD 3' UTR
KFD Frag4as (1-39)
H T I W . R E A G R S R R R S S . E L C E T K T S Q
GAAGGCCGAACTGGGAGGCCATGAAGGCCCCTGGCTTTAGCCAGGAGGGGGGAGAGGGAAAATTGGCATTCCCTCTCAGG
10640
CTTCCGGCTTGACCCTCCGGTACTTCCGGGGACCGAAATCGGTCCTCCCCCCTCTCCCTTTTAACCGTAAGGGAGAGTCC
KFD 3' UTR
R R P N W E A M K A P G F S Q E G G E G K L A F P L R

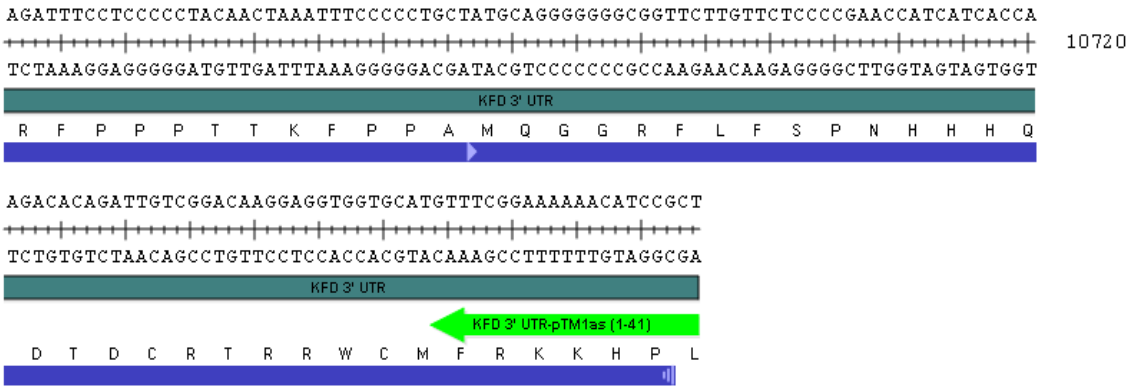

AGATTTCCTCCCCCTACAACTAAATTTCCCCCTGCTATGCAGGGGGGGCGGTTCTTGTTCTCCCCGAACCATCATCACCA
10720
TCTAAAGGAGGGGGATGTTGATTTAAAGGGGGACGATACGTCCCCCCCGCCAAGAACAAGAGGGGCTTGGTAGTAGTGGT
KFD 3' UTR
R F P P P T T K F P P A M Q G G R F L F S P N H H H Q
AGACACAGATTGTCGGACAAGGAGGTGGTGCATGTTTCGGAAAAAACATCCGCT
TCTGTGTCTAACAGCCTGTTCCTCCACCACGTACAAAGCCTTTTTTGTAGGCGA
KFD 3' UTR
KFD 3' UTR-pTM1as (1-41)
D T D C R T R R W C M F R K K H P L

Figure A.2: KFDV genome in cDNA format.

Primers: green, 5' and 3'UTRs: teal, C/E/NS2A/NS3/NS4B genes: yellow, prM/NS1/NS2B/NS4A/NS5 genes: orange, open reading frame: blue. The genome was constructed using DNAstar. Annotated genes were from published KFDV polyprotein sequence (AY323490) (Grard, Moureau et al. 2007).

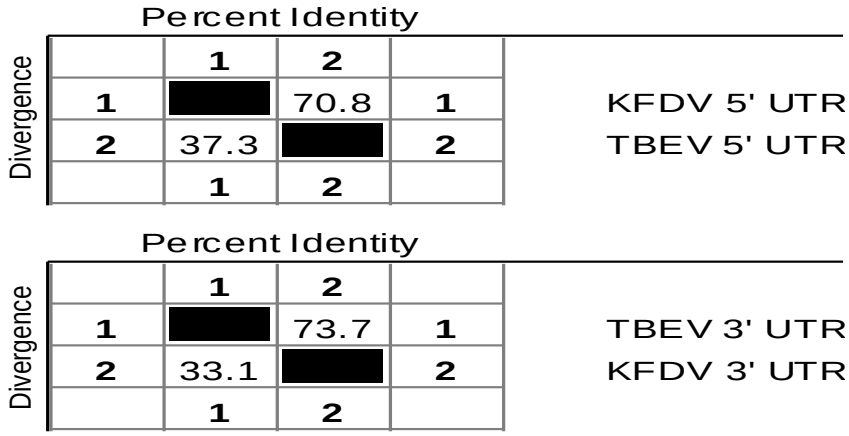


Figure A.3: Sequence comparisons of aligned KFDV and TBEV UTR regions.

Alignments were analyzed with ClustalW (DNAstar).

Literature Cited:

Adhikari Prabha, M. R., M. G. Prabhu, et al. (1993). "Clinical study of 100 cases of Kyasanur Forest disease with clinicopathological correlation." Indian J Med Sci **47**(5): 124-30.

Artsob, H. and R. Lindsay (2008). "Arboviruses." Elsevier Inc. International Encyclopedia of Public Health **One**(First Edition): 219-226.

Avirutnan, P., E. Mehlhop, et al. (2008). "Complement and its role in protection and pathogenesis of flavivirus infections." Vaccine **26 Suppl 8**: I100-7.

Bauman, R. W. (2009). Microbiology: With Diseases by Body System, Pearson Benjamin Cummings.

Baxevanis, A. D. and B. F. F. Ouelette (2001). Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins. New York, NY, USA, John Wiley & Sons, Inc.

Bhat, U. K. M. and M. K. Goverdhan (1973). "Transmission of Kyasanur Forest Disease Virus by the Soft Tick, Ornithodoros crossi." Acta Virologica.

Boyer, J. C. and A. L. Haenni (1994). "Infectious transcripts and cDNA clones of RNA viruses." Virology **198**(2): 415-26.

Bray, M., K. Davis, et al. (1999). "A mouse model for evaluation of prophylaxis and therapy of Ebola hemorrhagic fever." J Infect Dis **179 Suppl 1**: S248-58.

Calisher, C. H., N. Karabatsos, et al. (1989). "Antigenic relationships between flaviviruses as determined by cross-neutralization tests with polyclonal antisera." J Gen Virol **70 (Pt 1)**: 37-43.

Chang, H. H., H. F. Shyu, et al. (2002). "Facilitation of cell adhesion by immobilized dengue viral nonstructural protein 1 (NS1): arginine-glycine-aspartic acid structural mimicry within the dengue viral NS1 antigen." J Infect Dis **186**(6): 743-51.

Charrel, R. N., S. Fagbo, et al. (2007). "Alkhurma hemorrhagic fever virus in Ornithodoros savignyi ticks." Emerg Infect Dis **13**(1): 153-5.

Charrel, R. N., A. M. Zaki, et al. (2001). "Complete coding sequence of the Alkhurma virus, a tick-borne flavivirus causing severe hemorrhagic fever in humans in Saudi Arabia." Biochem Biophys Res Commun **287**(2): 455-61.

Charrel, R. N., A. M. Zaki, et al. (2005). "Low diversity of Alkhurma hemorrhagic fever virus, Saudi Arabia, 1994-1999." Emerg Infect Dis **11**(5): 683-8.

Ciesiolka, J., J. Wrzesinski, et al. (2001). "Ribozymes of the hepatitis delta virus: recent findings on their structure, mechanism of catalysis and possible applications." Acta Biochim Pol **48**(2): 409-18.

Clontech. Retrieved May 25, 2009 from http://www.clontech.com/images/pt/PT3266-5.pdf.

Cristea, I. M., J. W. Carroll, et al. (2006). "Tracking and elucidating alphavirus-host protein interactions." J Biol Chem **281**(40): 30269-78.

Dandawate, C. N., G. B. Desai, et al. (1994). "Field evaluation of formalin inactivated Kyasanur forest disease virus tissue culture vaccine in three districts of Karnataka state." Indian J Med Res **99**: 152-8.

Davies, M. V. and R. J. Kaufman (1992). "The sequence context of the initiation codon in the encephalomyocarditis virus leader modulates efficiency of internal translation initiation." J Virol **66**(4): 1924-32.

Denison, M. R. (2008). "Seeking membranes: positive-strand RNA virus replication complexes." PLoS Biol **6**(10): e270.
Dobler, G. "Zoonotic tick-borne flaviviruses." Vet Microbiol **140**(3-4): 221-8.
Dong, H., D. Ray, et al. (2007). "Distinct RNA elements confer specificity to flavivirus RNA cap methylation events." J Virol **81**(9): 4412-21.
Ebihara, H., A. Groseth, et al. (2005). "The role of reverse genetics systems in studying viral hemorrhagic fevers." Thromb Haemost **94**(2): 240-53.
Falconar, A. K. (1997). "The dengue virus nonstructural-1 protein (NS1) generates antibodies to common epitopes on human blood clotting, integrin/adhesin proteins and binds to human endothelial cells: potential implications in haemorrhagic fever pathogenesis." Arch Virol **142**(5): 897-916.
Fauquet, C. M., M. A. Mayo, et al. (2005). Virus Taxonomy. San Diego, USA, London, UK, Elsevier Inc.
Fedoruk-Wyszomirska, A., M. Giel-Pietraszuk, et al. (2009). "The mechanism of acidic hydrolysis of esters explains the HDV ribozyme activity." Mol Biol Rep **36**(7): 1647-50.
Filomatori, C. V., M. F. Lodeiro, et al. (2006). "A 5' RNA element promotes dengue virus RNA synthesis on a circular genome." Genes Dev **20**(16): 2238-49.
Flint, S. J., L. W. Enquist, et al. (2004). Principles of Virology: Molecular Biology, Pathogenesis and Control of Animal Viruses, ASM Press.
Flint, S. J., L. W. Enquist, et al. (2009). Principles of Virology. Washington, DC, American Society for Microbiology Press.
Fritz, R., K. Stiasny, et al. (2008). "Identification of specific histidines as pH sensors in flavivirus membrane fusion." J Cell Biol **183**(2): 353-61.
Gehrke, R., M. Ecker, et al. (2003). "Incorporation of tick-borne encephalitis virus replicons into virus-like particles by a packaging cell line." J Virol **77**(16): 8924-33.
Goodbourn, S., L. Didcock, et al. (2000). "Interferons: cell signalling, immune modulation, antiviral response and virus countermeasures." J Gen Virol **81**(Pt 10): 2341-64.
Goodman, J. L., D. T. Dennis, et al. (2005). Tick-Borne Diseases of Humans, ASM Press.
Grard, G., G. Moureau, et al. (2007). "Genetic characterization of tick-borne flaviviruses: new insights into evolution, pathogenetic determinants and taxonomy." Virology **361**(1): 80-92.
Gritsun, T. S., T. V. Frolova, et al. (2003). "Characterization of a siberian virus isolated from a patient with progressive chronic tick-borne encephalitis." J Virol **77**(1): 25-36.
Gruber, A., R. Lorenz, et al. (2008). "The Vienna RNA Websuite." Nucleic Acids Research.
Gualano, R. C., M. J. Pryor, et al. (1998). "Identification of a major determinant of mouse neurovirulence of dengue virus type 2 using stably cloned genomic-length cDNA." J Gen Virol **79 (Pt 3)**: 437-46.
Gubler, D. J. (2001). "Human arbovirus infections worldwide." Ann N Y Acad Sci **951**: 13-24.

Hanley, K. A. and S. C. Weaver (2010). Frontiers in Dengue Virus Research. Norfolk, UK, Caister Academic Press.
Harris, E., K. L. Holden, et al. (2006). "Molecular biology of flaviviruses." Novartis Found Symp **277**: 23-39; discussion 40, 71-3, 251-3.
Hayasaka, D., L. Ivanov, et al. (2001). "Distribution and characterization of tick-borne encephalitis viruses from Siberia and far-eastern Asia." J Gen Virol **82**(Pt 6): 1319-28.
Hinten, S. R., G. A. Beckett, et al. (2008). "Increased recognition of Powassan encephalitis in the United States, 1999-2005." Vector Borne Zoonotic Dis **8**(6): 733-40.
Holbrook, M. R., J. F. Aronson, et al. (2005). "An animal model for the tickborne flavivirus--Omsk hemorrhagic fever virus." J Infect Dis **191**(1): 100-8.
Howley, P. M. and D. M. Knipe (2007). Fields' Virology, Lippincott Williams & Wilkins.
Hurrelbrink, R. J., A. Nestorowicz, et al. (1999). "Characterization of infectious Murray Valley encephalitis virus derived from a stably cloned genome-length cDNA." J Gen Virol **80 (Pt 12)**: 3115-25.
Issur, M., B. J. Geiss, et al. (2009). "The flavivirus NS5 protein is a true RNA guanylyltransferase that catalyzes a two-step reaction to form the RNA cap structure." Rna **15**(12): 2340-50.
Jones, C. T., C. G. Patkar, et al. (2005). "Construction and applications of yellow fever virus replicons." Virology **331**(2): 247-59.
Kapoor, M., L. Zhang, et al. (1995). "Synthesis and characterization of an infectious dengue virus type-2 RNA genome (New Guinea C strain)." Gene **162**(2): 175-80.
Kapp, L. D. and J. R. Lorsch (2004). "The molecular mechanics of eukaryotic translation." Annu Rev Biochem **73**: 657-704.
Kawaoka, Y. (2004). Biology of Negative Strand RNA Viruses: The Power of Reverse Genetics, Springer-Verlag Berlin Heidelberg New York.
Khromykh, A. A., P. L. Sedlak, et al. (2000). "cis- and trans-acting elements in flavivirus RNA replication." J Virol **74**(7): 3253-63.
Khromykh, A. A. and E. G. Westaway (1994). "Completion of Kunjin virus RNA sequence and recovery of an infectious RNA transcribed from stably cloned full-length cDNA." J Virol **68**(7): 4580-8.
Kielian, M. (2006). "Class II virus membrane fusion proteins." Virology **344**(1): 38-47.
Kinney, R. M., S. Butrapet, et al. (1997). "Construction of infectious cDNA clones for dengue 2 virus: strain 16681 and its attenuated vaccine derivative, strain PDK-53." Virology **230**(2): 300-8.
Kobasa, D., A. Takada, et al. (2004). "Enhanced virulence of influenza A viruses with the haemagglutinin of the 1918 pandemic virus." Nature **431**(7009): 703-7.
Kofler, R. M., V. M. Hoenninger, et al. (2006). "Functional analysis of the tick-borne encephalitis virus cyclization elements indicates major differences between mosquito-borne and tick-borne flaviviruses." J Virol **80**(8): 4099-113.
Krishna, V. D., M. Rangappa, et al. (2009). "Virus-specific cytolytic antibodies to nonstructural protein 1 of Japanese encephalitis virus effect reduction of virus output from infected cells." J Virol **83**(10): 4766-77.

Kuhn, R. J., W. Zhang, et al. (2002). "Structure of dengue virus: implications for flavivirus organization, maturation, and fusion." Cell **108**(5): 717-25.
LaFemina, R. L. (2009). Antiviral Research: Strategies in Antiviral Drug Discovery, ASM Press.
Lai, C. J., B. T. Zhao, et al. (1991). "Infectious RNA transcribed from stably cloned full-length cDNA of dengue type 4 virus." Proc Natl Acad Sci U S A **88**(12): 5139-43.
Leung, J. Y., G. P. Pijlman, et al. (2008). "Role of nonstructural protein NS2A in flavivirus assembly." J Virol **82**(10): 4731-41.
Li, H., S. Clum, et al. (1999). "The serine protease and RNA-stimulated nucleoside triphosphatase and RNA helicase functional domains of dengue virus type 2 NS3 converge within a region of 20 amino acids." J Virol **73**(4): 3108-16.
Lin, D., L. Li, et al. (2003). "Analysis of the complete genome of the tick-borne flavivirus Omsk hemorrhagic fever virus." Virology **313**(1): 81-90.
Luo, D., T. Xu, et al. (2008). "Crystal structure of the NS3 protease-helicase from dengue virus." J Virol **82**(1): 173-83.
Madani, T. A. (2005). "Alkhumra virus infection, a new viral hemorrhagic fever in Saudi Arabia." J Infect **51**(2): 91-7.
Mahy, B. W. J. and M. H. Van Regenmortel (2010). Desk Encyclopedia of General Virology, Associated Press.
Mandl, C. W., M. Ecker, et al. (1997). "Infectious cDNA clones of tick-borne encephalitis virus European subtype prototypic strain Neudoerfl and high virulence strain Hypr." J Gen Virol **78 (Pt 5)**: 1049-57.
Mathews, D. H., J. Sabina, et al. (1999). "Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure." J Mol Biol **288**(5): 911-40.
Moss, B., O. Elroy-Stein, et al. (1990). "Product review. New mammalian expression vectors." Nature **348**(6296): 91-2.
Neumann, G., H. Feldmann, et al. (2002). "Reverse genetics demonstrates that proteolytic processing of the Ebola virus glycoprotein is not essential for replication in cell culture." J Virol **76**(1): 406-10.
Niwa, H., K. Yamamura, et al. (1991). "Efficient selection for high-expression transfectants with a novel eukaryotic vector." Gene **108**(2): 193-9.
Office of Laboratory Safety, P. H. A. o. C. M. S. D. S.-I. S. (2008) "Kyasanur Forest Disease Virus (KFDV)." **Volume**, DOI:
Pastorino, B. A., C. N. Peyrefitte, et al. (2006). "Mutagenesis analysis of the NS2B determinants of the Alkhurma virus NS2B-NS3 protease activation." J Gen Virol **87**(Pt 11): 3279-83.
Pattnaik, P. (2006). "Kyasanur forest disease: an epidemiological view in India." Rev Med Virol **16**(3): 151-65.
Pekosz, A., B. He, et al. (1999). "Reverse genetics of negative-strand RNA viruses: closing the circle." Proc Natl Acad Sci U S A **96**(16): 8804-6.
Perera, R. and R. J. Kuhn (2008). "Structural proteomics of dengue virus." Curr Opin Microbiol **11**(4): 369-77.
Peters, C. J. (2006). "Emerging infections: lessons from the viral hemorrhagic fevers." Trans Am Clin Climatol Assoc **117**: 189-96; discussion 196-7.

PlaneMad/Wikipedia. (2010). "Andaman and Nicobar Islands." Retrieved February 19, 2010, from http://en.wikipedia.org/wiki/Andaman_and_Nicobar_Islands.
Polo, S., G. Ketner, et al. (1997). "Infectious RNA transcripts from full-length dengue virus type 2 cDNA clones made in yeast." J Virol **71**(7): 5366-74.
Proutski, V., E. A. Gould, et al. (1997). "Secondary structure of the 3' untranslated region of flaviviruses: similarities and differences." Nucleic Acids Res **25**(6): 1194-202.
Rice, C. M., A. Grakoui, et al. (1989). "Transcription of infectious yellow fever RNA from full-length cDNA templates produced by in vitro ligation." New Biol **1**(3): 285-96.
Richman, D. D., R. J. Whitley, et al. (2009). Clinical Virology. Washington, DC, American Society for Microbiology (ASM) Press.
Sambrook, J., E. F. Fritsch, et al. (1989). Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press.
Sambrook, J. and D. W. Russell (2001). Molecular Cloning: A Laboratory Manual. New York, Cold Spring Harbour Laboratory Press.
Samsa, M. M., J. A. Mondotte, et al. (2009). "Dengue virus capsid protein usurps lipid droplets for viral particle formation." PLoS Pathog **5**(10): e1000632.
Schrauf, S., P. Schlick, et al. (2008). "Functional analysis of potential carboxy-terminal cleavage sites of tick-borne encephalitis virus capsid protein." J Virol **82**(5): 2218-29.
Stiasny, K., C. Koessl, et al. (2003). "Involvement of lipids in different steps of the flavivirus fusion mechanism." J Virol **77**(14): 7856-62.
Stiasny, K., C. Kossl, et al. (2007). "Characterization of a structural intermediate of flavivirus membrane fusion." PLoS Pathog **3**(2): e20.
Sumiyoshi, H., C. H. Hoke, et al. (1992). "Infectious Japanese encephalitis virus RNA can be synthesized from in vitro-ligated cDNA templates." J Virol **66**(9): 5425-31.
Teissier, E. and E. I. Pecheur (2007). "Lipids as modulators of membrane fusion mediated by viral fusion proteins." Eur Biophys J **36**(8): 887-99.
Theriault, S., A. Groseth, et al. (2004). "Rescue of Ebola virus from cDNA using heterologous support proteins." Virus Res **106**(1): 43-50.
Uchil, P. D., A. V. Kumar, et al. (2006). "Nuclear localization of flavivirus RNA synthesis in infected cells." J Virol **80**(11): 5451-64.
Venugopal, K., T. Gritsun, et al. (1994). "Analysis of the structural protein gene sequence shows Kyasanur Forest disease virus as a distinct member in the tick-borne encephalitis virus serocomplex." J Gen Virol **75 (Pt 1)**: 227-32.
Villordo, S. M. and A. V. Gamarnik (2009). "Genome cyclization as strategy for flavivirus RNA replication." Virus Res **139**(2): 230-9.
Volchkov, V. E., V. A. Volchkova, et al. (2001). "Recovery of infectious Ebola virus from complementary DNA: RNA editing of the GP gene and viral cytotoxicity." Science **291**(5510): 1965-9.
Wang, J., H. Zhang, et al. (2009). "Isolation of kyasanur forest disease virus from febrile patient, yunnan, china." Emerg Infect Dis **15**(2): 326-8.
White, D. O. and F. J. Fenner (1994). Medical Virology, Academic Press Inc.**:** 16-28, 74-80, 373-375, 434-449.

Xia, X. and Z. Xie (2001). "DAMBE: Data analysis in molecular biology and evolution." Journal of Heredity **92:371-373**.
Yon, C., T. Teramoto, et al. (2005). "Modulation of the nucleoside triphosphatase/RNA helicase and 5'-RNA triphosphatase activities of Dengue virus type 2 nonstructural protein 3 (NS3) by interaction with NS5, the RNA-dependent RNA polymerase." J Biol Chem **280**(29): 27412-9.
Yoshii, K. and M. R. Holbrook (2009). "Sub-genomic replicon and virus-like particles of Omsk hemorrhagic fever virus." Arch Virol **154**(4): 573-80.
Yun, S. I., S. Y. Kim, et al. (2003). "Development and application of a reverse genetics system for Japanese encephalitis virus." J Virol **77**(11): 6450-65.