Development of an Infectious Clone System to Study the Life Cycle of Hazara Virus

Jack Fuller

Submitted in accordance with the requirements for the degree of Doctor of Philosophy

The University of Leeds
Faculty of Biological Sciences

July 2020

The candidate confirms that the work submitted is his own and that appropriate credit has been given where reference has been made to the work of others.

Acknowledgements

I would like to thank my supervisors, John, Jamel and Roger for their endless support and fresh ideas throughout my PhD. A special thanks go to John for always providing enthusiasm and encouragement, especially during the tougher times of the project!

I would also like to thank everyone in 8.61, not only for making my time at Leeds enjoyable, but for helping me develop as a scientist through useful critique and discussion. A special thanks goes to Francis Hopkins, for welcoming me into the Barr group and providing a constant supply of humour and to Ellie Todd, for listening to my endless whines and gripes, and for providing welcome distractions in the form of her PowerPoint artwork.

Outside of the lab, my mum deserves a special mention for always being the voice of reason and support at the end of the phone, but also for pushing me to succeed throughout my entire education. I have no doubts I would not be in the position I am now without her! Finally, a huge thanks to my partner Hannah, who has been there to support me through all the highs and lows of my PhD, and has sacrificed many weekend trips to allow me to finish experiments and to tend to my viruses!

Abstract

Crimean-Congo hemorrhagic fever orthonairovirus (CCHFV) is a negative sense single stranded RNA virus, capable of causing fatal hemorrhagic fever in humans. Currently, an infectious clone system exists for CCHFV, however, owing to the high containment level required for experimentation with this virus, its potential is limited. Here, a highly efficient infectious clone system was developed for recovery of Hazara virus (HAZV), the first such orthonairovirus system able to be used in biosafety level 2 facilities, providing a valuable tool for increasing our understanding of these viruses.

Mechanisms of viral subversion of cellular trafficking pathways involved in viral entry, gene expression, assembly and egress are poorly understood for HAZV and CCHFV. The infectious clone system was adapted to express eGFP, enabling screening of an siRNA library targeting genes involved in cellular trafficking networks via live-cell fluorescent imaging, the first such screen for a nairovirus. Screening revealed an important role for subunits of the coat protein 1 vesicle coatomer (COPI), normally involved in regulation of Golgi / ER cargo trafficking. The effect was observed at multiple stages of the HAZV life cycle; an early stage prior to and including gene expression, and also a later stage during assembly and egress of infectious virus, with COPI-knockdown reducing titres by approximately 100-fold.

In addition to gain of function mutations, such as the reporter virus discussed above, the infectious clone system represents a powerful tool for exploring the role of individual nucleotides, amino acids and entire open reading frames within the viral genome. Both CCHFV and HAZV contain a well-documented, highly conserved caspase cleavage motif on the apex of the arm domain on the nucleoprotein (N). Whilst previous literature has demonstrated this motif to be cleaved as purified protein, limited data exists as to its role in the context of viral replication in a cellular system. To this end, the infectious clone system described above was used to create a panel of mutants targeting this conserved caspase cleavage motifs within HAZV N. HAZV bearing an uncleavable DQVE sequence rescued efficiently with growth rates equivalent to those of wild-type virus in both mammalian and tick cells, showing this site was dispensable for virus

multiplication. In contrast, substitution of the DQVD motif with the similarly uncleavable AQVA sequence could not be rescued despite repeated efforts. Together, these results highlight the importance of this caspase cleavage site in the HAZV life cycle but reveal the DQVD sequence performs a critical, as yet unknown, role aside from caspase cleavage.

Table of contents

Acknowledgements ..... *iii*
Abstract ..... *iv*
Table of figures ..... *ix*
Table of tables ..... *xi*
Abbreviations ..... *xii*
1 Introduction *1*
1.1 General introduction **1**
1.1.1 Discovery of CCHFV and HAZV 1
1.1.2 CCHFV and HAZV classification 2
1.1.3 Transmission, vectors and hosts 5
1.1.4 Epidemiology and global distribution of CCHFV 7
1.1.5 Diagnosis of CCHF 9
1.1.6 Clinical manifestations and treatment options 9
1.2 The virus **12**
1.2.1 Virus structure 12
1.2.2 Viral genome structure 13
1.3 The replication cycle of CCHFV and HAZV **14**
1.3.1 Binding and entry 14
1.3.2 Replication 16
1.3.3 Assembly and egress 18
1.4 Virally encoded proteins **19**
1.4.1 S segment 19
1.4.2 M segment 24
1.4.3 L Segment 26
1.5 Nairoviral reverse genetics systems **27**
1.5.1 Minigenome system 28
1.5.2 VLP system 31
1.5.3 Infectious clone 32
1.6 Apoptotic response to HAZV / CCHFV infection **35**
1.7 Project aims **36**
2 Materials and methods *38*
2.1 Materials **38**
2.1.1 Vectors 38
2.1.2 Bacterial cell lines 38
2.1.3 Mammalian cell lines 38

2.1.4 Tick cell lines 39
2.1.5 Hazara virus strain 39
2.2 General methods 39
2.2.1 Manipulation of recombinant DNA 39
2.2.2 Cell culture methods 42
2.2.3 Transfections 44
2.2.4 Analysis of protein expression 46
2.2.5 Virological techniques 50
2.2.6 Inhibitor studies 53
3 Development of a Hazara virus infectious clone 54
3.1 General introduction 54
3.2 Plasmid design 57
3.3 Optimising transfections 62
3.3.1 Transfection reagent 62
3.3.2 Ratio of transfection reagent to cDNA 63
3.3.3 Inclusion of exogenous T7 polymerase 65
3.4 Recovery of WT rHAZV 66
3.5 Comparison of recombinant HAZV to parental strain JC280 72
3.6 Confirmation of the cDNA source of rHAZV 73
3.7 Chapter summary and discussion 75
4 Generation of a fluorescent reporter virus 77
4.1 General introduction 77
4.2 Plasmid design 79
4.3 Recovery of rHAZV-eGFP 81
4.4 Comparison to WT rHAZV 85
4.5 Growth kinetics of eGFP HAZV in multiple cell lines 89
4.6 Identifying host membrane trafficking factors involved in HAZV multiplication 92
4.7 COPI components are important for HAZV infection 97
4.8 Identifying the role of COPI in HAZV infection 102
4.9 Dependence on COPI vesicles by HAZV is ARF1 and GBF1 independent 112

4.10 Chapter summary and discussion 116

5 Exploring Hazara nucleoprotein cleavage during infection 119

5.1 General introduction 119

5.2 Observation of multiple cleavage events in HAZV infection 121

5.2.1 Incorporation of cleavage products into virions 125

5.3 Induction of apoptosis by HAZV 125

5.4 Generation of cleavage site deficient rHAZV 127

5.4.1 DQVD and ENKD sites 127

5.4.2 DVMA 132

5.4.3 LFAA 135

5.5 Chapter summary and discussion 145

References 149

Appendix 168

1.1. Nucleotide sequences 168

1.1.1. Hazara S segment 168

1.1.2. Hazara M segment 169

1.1.3. Hazara L segment 171

1.1.4. EGFP-P2A gene insert 179

1.2. Primers 179

1.3. Membrane trafficking screen targets 181

1.4. Full siRNA screen results 205

Table of figures

Classification of Hazara orthonairovirus (HAZV) and Crimean-Congo hemorrhagic fever orthonairovirus (CCHFV) within the Bunyavirales order *4*
The natural life cycle of the CCHFV vector, Hyalomma ticks *6*
Geographic distribution of Crimean-Congo hemorrhagic fever *8*
Schematic diagram and electron micrographs of HAZV and CCHFV *13*
Schematic detailing negative sense gene expression *17*
Comparison of the crystal structures of HAZV N and CCHFV N *20*
Processing of the CCHFV glycoprotein precursor by cellular proteases *26*
CCHFV minigenome system schematic *30*
CCHFV virus-like particle system schematic *32*
CCHFV infectious clone system schematic *34*
Schematic of rHAZV rescue workflow *57*
Design of cDNAs for recovery of recombinant rHAZV *59*
Restriction digest confirmation of plasmid identity *60*
Confirmation of HAZV N expression from pMK-RQ-S *61*
Selection of transfection reagent *63*
Optimising ratio of transfection reagent to cDNA *64*
Optimisation of T7 polymerase expression *66*
Schematic detailing recovery of rHAZV *67*
Recovery of WT rHAZV *69*
qPCR detection of HAZV S segment RNA *71*
Comparison of rHAZV to parental strain *73*
Overview of insertion of a HindIII restriction site to S segment cDNA *74*
Recovery and confirmation of a silent mutant rHAZV *75*
eGFP S segment schematic *80*
Western blot confirmation of correct HAZV N expression from pMK-RQ-S-eGFP *81*
Western blot analysis of transfected and infected cell lysates confirms recovery of rHAZV(eGFP) *82*
Confirmation of the fluorescent nature of rHAZV(eGFP) via live cell imaging *84*
Plaque assay of rHAZV(eGFP) *85*
Multi-step growth kinetics of rHAZV vs rHAZV(eGFP) shows similar growth profiles *87*
Confirming co-expression of eGFP and HAZV N in infected SW13 cells *88*
Differential growth profiles of rHAZV(eGFP) infection in a range of cell lines *91*
Infection at multiple MOIs identifies an optimal MOI of 0.25 for infection of 96 well plates *93*
COPI and COPII vesicle schematic *98*
COPI associated genes are important for HAZV infection *99*
COPII associated genes are not required for HAZV infection *99*
qPCR confirmation of siRNA screen results using WT rHAZV *101*
Western blot analysis shows siRNA knockdown of COPI components reduces HAZV N translation *101*
COPI knockdown delays translation of HAZV N *104*
rHAZV(eGFP) infection does not cause redistribution of COPA *106*
COPI knockdown results in reduced infectious virus production *108*
Confocal imaging of COPA and HAZV N localisation in SW13 cells *111*
Role of Arf1 and GBF1 in COPI vesicle formation *113*
ARF1 and GBF1 are non-essential to rHAZV(eGFP) infection *114*
WT rHAZV is not sensitive to brefeldin A treatment. *115*
Differential cleavage of HAZV N in infection of SW13 and HAE/CTVM9 cell lines *123*
Caspases are responsible for HAZV N cleavage *124*
Caspase mediated HAZV N cleavage products are incorporated into infectious virions *125*
Differential activation of caspases following rHAZV infection of SW13 and HAE/CTVM9 cell lines *127*
Location of DQVD and ENKD cleavage sites on HAZV and CCHFV nucleoproteins *128*
Recovery of caspase motif deficient rHAZV *130*
Knockout of caspase motifs results in alternate cleavage of HAZV N *131*
Mutagenesis of the DQVD and ENKD sites does not affect ability of rHAZV to infect tick cells *132*
Location of the DVMD predicted cleavage site and conservation in alternate species *133*

Inability to recover rHAZV with a mutated DVMD motif shown via western blot and plaque assay 134
Location of the LFAD predicted cleavage site and conservation in alternate species..... 136
Recovery of rHAZV containing a mutated LFAD motif displaying altered plaque morphology 137
Recovery of rHAZV-eGFP containing a mutated LFAD motif demonstrates reduced fitness .. 139
Mutation of LFAD motif results in altered plaque morphology 141
rHAZV(eGFP-LFAA) displays delayed translation vs rHAZV(eGFP)..... 143
siRNA knockdown of SNX1 reduces rHAZV infectivity in SW13 cells 145

Table of tables

Table 2.1, List of primers and associated sequences used in qPCR 43
Table 2.2, List of primary antibodies used during the project 48
Table 3.1, Table of selected siRNA targets resulting in highest knockdown of total integrated intensity of eGFP signal 96

Abbreviations

ARF	ADP-ribosylation factor
BAR	Bin/amphiphysin/rvs
BFA	Brefeldin A
BSL	Biosafety containment level
BUNV	Bunyamwera virus
CAD	Caspase activate DNase
CHF	Crimean hemorrhagic fever
CCHF	Crimean-Congo hemorrhagic fever
CCHFV	Crimean-Congo hemorrhagic fever orthonairovirus
COPI	Coatamer protein complex I
CPE	Cytopathic effect
cRNA	Complementary RNA
DMEM	Dulbecco's modified Eagle's medium
DN	Dominant negative
EGFR	Epidermal growth factor receptor
Endo-H	Endoglycosidase H
ER	Endoplasmic reticulum
EREV	Erve virus
EV11	Echo virus 11
FBS	Foetal bovine serum
FL-PVDF	Fluorescence-compatible polyvinylidene difluoride
GAP	GTPase activating protein
GEF	Guanine exchange factor
GFP	Green fluorescent protein
GPC	Glycoprotein precursor
HAZV	Hazara orthonairovirus
HDR	Hepatitis delta virus ribozyme
IAV	Influenza A virus
iCLIP	Intramembrane cleaving proteases
IFA	Immunofluorescence assay
IFN	Interferon
IFNAR$^{-/-}$	Interferon α/β-receptor knockout
ISG	Interferon stimulated gene
JUNV	Junin virus
K^+	Potassium ion

KUPV	Kupe virus
L	Large
LACV	La Crosse virus
M	Medium
MLD	Mucin like domain
MOI	Multiplicity of infection
N	Nucleoprotein
NAAT	Nucleic acid amplification test
NOC	Nocodazole
NSDV	Nairobi sheep disease orthonairovirus
NSm	Non-structural M protein
NSs	Non-structural S protein
NTR	Non-translated region
ORF	Open reading frame
OTU	Ovarian tumour
P2A	Porcine Teshovirus-1 2A peptide linker
PAC	Paclitaxel
PARP	Poly(ADP-Ribose)
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PE	Promoter element
POC	Point of care
PS2	Pitstop-2
PtdIns3P	Phosphoinositide 3-monophosphate
PX	Phox homology
qPCR	Quantitative polymerase chain reaction
rCCHFV	Recombinant Crimean Congo hemorrhagic fever virus
RdRp	RNA dependent RNA polymerase
RDT	Rapid diagnostic test
rHAZV	Recombinant Hazara virus
RNP	Ribonucleoprotein
RSV	Respiratory syncytial virus
RT	Reverse transcriptase
RTP	Ribavirin-triphosphate
RVFV	Rift Valley fever virus
S	Small
SDM	Site directed mutagenesis

SFM	Serum free media
SKI-1/S1P	Subtilisin kexin isozyme-1/Site-1 proteases
SNX1	Sorting nexin 1
ssRNA	Single stranded RNA
TGN	*Trans*-Golgi network
TIIE	Total integrated intensity of eGFP
TSWV	Tomato spotted wilt orthospovirus
VAMP	Vesicle associated membrane protein
VLP	Virus-Like particle
vRNA	Viral RNA
VRP	Virus replicon particle
VSV	Vesicular stomatitis virus
WHO	World health organisation
ZsG	ZsGreen1

1 Introduction

1.1 General introduction

1.1.1 Discovery of CCHFV and HAZV

Crimean-Congo hemorrhagic fever (CCHF) was first described between 1944 and 1945 following an outbreak amongst approximately 200 Soviet military personnel in the Crimean Peninsula. The disease was initially named Crimean hemorrhagic fever (CHF). A viral aetiology was determined after psychiatric patients requiring pyrogenic therapy had CHF-like clinical symptoms induced following infection with a filterable agent isolated from CHF patients, termed Crimean hemorrhagic fever virus (Chumakov, 1945). The viral agent was first isolated following inoculation of new-born white mice with samples taken from a CHF patient in Astrakhan in 1967. The isolated strain (Drozdov) permitted development of antibodies and antigens to enable serological studies and comparisons to viruses isolated from outbreaks from various different geographical locations (Butenko *et al.*, 1968). Collaboration between M.P Chumakov and Jordi Casals identified the virus isolated from Crimea in 1945 to be antigenically indistinguishable from three stains of Congo virus taken from patients in the Congo and Uganda in 1956, resulting in the renaming of the virus to Crimean-Congo hemorrhagic fever virus which has since been adapted to Crimean-Congo hemorrhagic fever orthonairovirus (CCHFV) (Hoogstraal, 1979). The viral agent was also detected at a similar time in suspensions generated from *Hyalomma plubeum (Hyalomma marginatum)*, identifying ticks as the vector of the virus (Butenko *et al.*, 1968).

Despite the identification and characterisation of CCHFV in the last 50 years, similar CCHF-like symptoms have long been documented in areas of Tajikistan, dating back to the 12th century. Symptoms included blood in the urine, vomit, sputum, rectum and abdominal cavity, with a tough, small arthropod responsible for the spread of the disease (Hoogstraal, 1966). More recently, CCHF has been recognised under three different names by the indigenous population of southern Uzbekistan, which suggests that despite the relatively recent advances in knowledge surrounding CCHFV, it has been

the responsible agent of human hemorrhagic fever for centuries (Hoogstraal, 1979).

Hazara orthonairovirus (HAZV) was first isolated in 1970 from the vole *Alticola roylei* in subarctic terrain in the Kaghan Valley in the Hazara district of West Pakistan. Early serological tests showed it to be a distinct, but related virus to the Crimean and Congo strains that were renamed to CCHFV (Begum, Wisseman Jr and Traub, 1970). In line with CCHFV infection, HAZV infection of suckling mice progresses in a similar manner, with an often fatal disease outcome following intracerebral inoculation, with high viral loads detected in the liver and brain tissue (Smirnova *et al.*, 1977). However, there has been no evidence of HAZV infecting or causing disease in humans and limited information is available regarding its global distribution.

1.1.2 CCHFV and HAZV classification

CCHFV and HAZV are members of the order *Bunyavirales*, which was established in 2017 to accommodate related viruses exhibiting segmented, linear, single-stranded, negative or amibisense RNA genomes (Maes *et al.*, 2018). Initially 9 families were defined, however this was amended to 12 to permit efficient assignment of newly discovered species, most recently with the addition of the *Leishbuvidae* family and recreation of the *Tospoviridae* family in 2019 (Abudurexiti *et al.*, 2019). There are currently around 500 named isolates within this order, across 287 assigned species, making the *Bunyavirales* one of the largest taxonomic groupings. Prior to the formation of the order *Bunyavirales*, bunyaviruses were classified within the *Bunyaviridae* family, which comprised of 5 genera (*Nairovirus, Orthobunyavirus*, *Hantavirus*, *Phlebovirus* and *Tospovirus)*.

Both CCHFV and HAZV are members of the family *Nairoviridae* and are further classified into the *orthonairovirus* genus along with 13 other viruses (Figure 1.1). Notable examples from within this genus include Dugbe orthonairovirus, the type species of the genus and Nairobi Sheep Disease orthonairovirus

(NSDV), which causes 90 % mortality in sheep and goats, causing significant economic disruption (Lasecka and Baron, 2014).

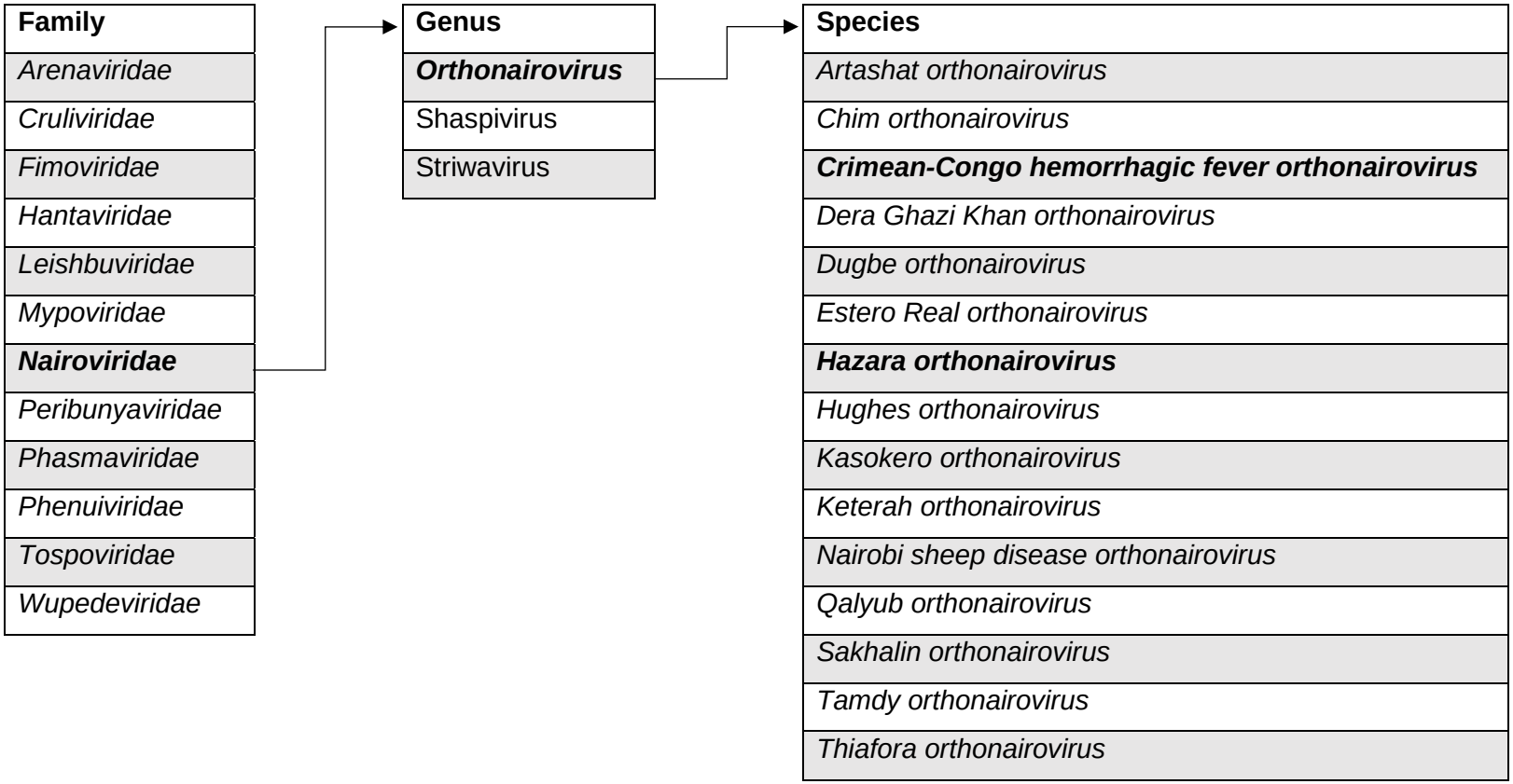


Figure 1.1, Classification of Hazara orthonairovirus (HAZV) and Crimean-Congo hemorrhagic fever orthonairovirus (CCHFV) within the Bunyavirales order

Figure shows the taxonomy of HAZV and CCHFV within the Bunyavirales order, Nairoviridae family, orthonairovirus genus. All species within the orthonairovirus genus are listed.

1.1.3 Transmission, vectors and hosts

CCHFV is a zoonotic infection that is maintained in several species of ixodid (hard) ticks through trans-stadial, transovarial and venereal transmission. Whilst CCHFV has been isolated from other arthropods, such as biting midges, this is believed to be a result of recent feeding on an infected mammal (Causey *et al.*, 1970). CCHFV is unable to be maintained in soft ticks, experimental infection of such ticks demonstrated the virus was unable to spread to tissues or be maintained trans-stadially (Shepherd *et al.*, 1989; Durden *et al.*, 1993). The virus is amplified during the spring and summer months during blood meals taken by ticks on small or large mammalian hosts. As ixodid ticks maintain levels of CCHFV throughout their entire life cycle, they are true hosts. Humans only transiently maintain levels of CCHFV and do not provide a source of infection for ticks and are therefore accidental “dead-end” hosts for the virus (Bente *et al.*, 2013).

Initial infection of the tick occurs when a bloodmeal is taken from an infected host, CCHFV then replicates in the lining of the tick’s midgut, before spreading to multiple tissues including the salivary glands and reproductive organs, where titres of CCHFV are highest (Nuttall *et al.*, 1994). Infected females may pass the virus to offspring by transovarial transmission. Upon hatching, larvae of the *Hyalomma spp.* tick feed on a small mammalian or avian host for their first blood meal. Following engorgement, the nymph drops off this small host and following molting, develops into the adult form. Adult ticks take a second blood meal on a larger animal host and mate whilst feeding. Female ticks subsequently drop off and find a suitable location for oviposition, where the cycle then repeats (Estrada-Peña and de la Fuente, 2014). A schematic of the life cycle of *Hyalomma spp.* can be seen in Figure 1.2.

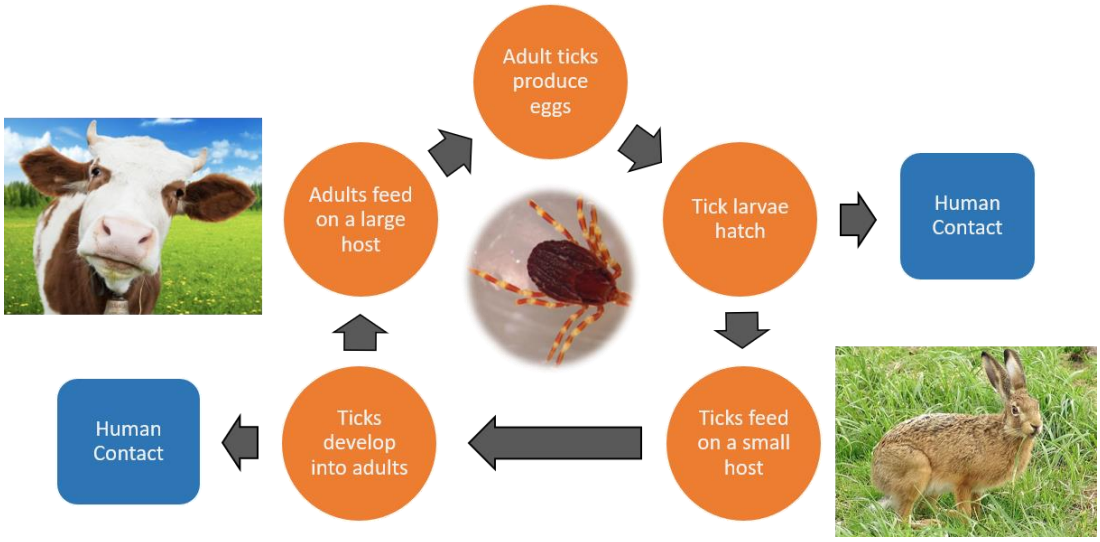


Figure 1.2, The natural life cycle of the CCHFV vector, Hyalomma ticks

The complete life cycle of the Hyalomma spp. tick requires a small and large host. A tick infected with CCHFV carries the virus for the full duration of its life cycle and can pass the virus vertically to offspring. Humans become infected following a bite with an infected tick at any stage of the life cycle, though questing adult Hyalomma ticks are thought to be responsible for the majority of human infections.

Infection of hosts can be assessed by serosurveys, which have been the traditional method of identifying regions of CCHFV circulation. Direct infections of multiple species of wild and domestic animals with CCHFV identified low levels of replication in horses, donkeys, sheep, cattle and ostriches (Shepherd *et al.*, 1987, 1989; Wilson *et al.*, 1990; Zeller, Cornet and Camicas, 1994). Whilst birds represent a common target for blood meals for the larvae and nymphs of the *Hyalomma spp.* ticks, there is limited evidence to suggest birds become viraemic following contact with an infected tick. Historical records suggest that whilst ticks isolated from birds were CCHFV positive, the birds themselves were not thought to be hosts for viral replication (Hoogstraal, 1979). Therefore, the likely role of birds in transmission of CCHFV is in increasing the global distribution of CCHFV positive ticks, rather than as a source for infecting naïve ticks.

Transmission to humans naturally occurs when people are bitten by virally infected ticks. The two-host life cycle of *Hyalomma spp.* ticks may have implications on the number of human cases of CCHF. Where there is an abundance of both small and large mammalian hosts for *Hyalomma spp.* ticks, infection rates observed in humans are low. However, where there is an imbalance in numbers of large hosts compared to small hosts, human cases

appear to increase. It is hypothesised that this is the reason for the outbreak among Soviet military personnel in the Crimean. Wild hare populations had recently proliferated but numbers of livestock had been sharply reduced, leaving a lack of non-human large mammalian hosts for *Hyalomma spp.* ticks to take blood meals from. The sudden increase in cases of CCHF seen in the military personnel was likely a result of questing *Hyalomma spp.* ticks feeding on humans as an alternative. In addition to tick bites, human transmission can occur during contact with virally infected blood or bodily fluids of an infected patient during the first seven to ten days of illness.

1.1.4 Epidemiology and global distribution of CCHFV

CCHFV is the second most widespread medically important arbovirus, after Dengue virus, responsible for over 140 outbreaks and more than 6000 cases worldwide since its discovery (Appannanavar and Mishra, 2011). Outbreaks have occurred across a vast geographical distribution, ranging from China to Africa, with recent cases also reported in India, Europe and the Middle East (Hoogstraal, 1979; Celikbas *et al.*, 2014; Makwana *et al.*, 2015; Spengler, Bergeron and Rollin, 2016) (Figure 1.3). However, the true geographic distribution of CCHFV is likely to be larger than current estimates suggest, due to an absence of surveillance systems. Previous first reports of virus, or human cases of CCHF, have coincided with the implementation of surveillance systems, supporting a larger distribution than is currently known (Mamuchishvili *et al.*, 2015). In addition, cases in endemic countries are rising annually, a study into cases of CCHF in Afghanistan between 2016 and 2018 reported increasing cases year on year, with 163 cases in 2016, 245 cases in 2017 and 483 cases in 2018. A potential explanation for this increase is the Eid-al-Adha festival, in which Muslims sacrifice cattle, sheep, goats or camels. For the next 10-15 years this festival is likely to be in the summer months when animals are likely to have peak viremia. Historically the festival has fallen in autumn or winter, during which fewer animals were infected, and thus posed less of an infection threat (Leblebicioglu *et al.*, 2015).

New cases are now also being reported in countries previously unaffected by the virus, which may be a result of the increased geographical distribution of the tick vector, with *Hyalomma* ticks detected for the first time in the UK and Norway by

the European Centre for Disease Prevention and Control (Jameson *et al.*, 2012; Hansford *et al.*, 2019). Additionally, countries have reported imported cases of CCHF from patients travelling to and from endemic regions. The UK confirmed two cases, one from a patient returning from Afghanistan in 2012 and another from Bulgaria in 2014 (Atkinson *et al.*, 2012; Lumley *et al.*, 2014). Given the ease of international travel in the modern era, it could be easily expected that imported cases could result in countries reporting first cases more frequently.

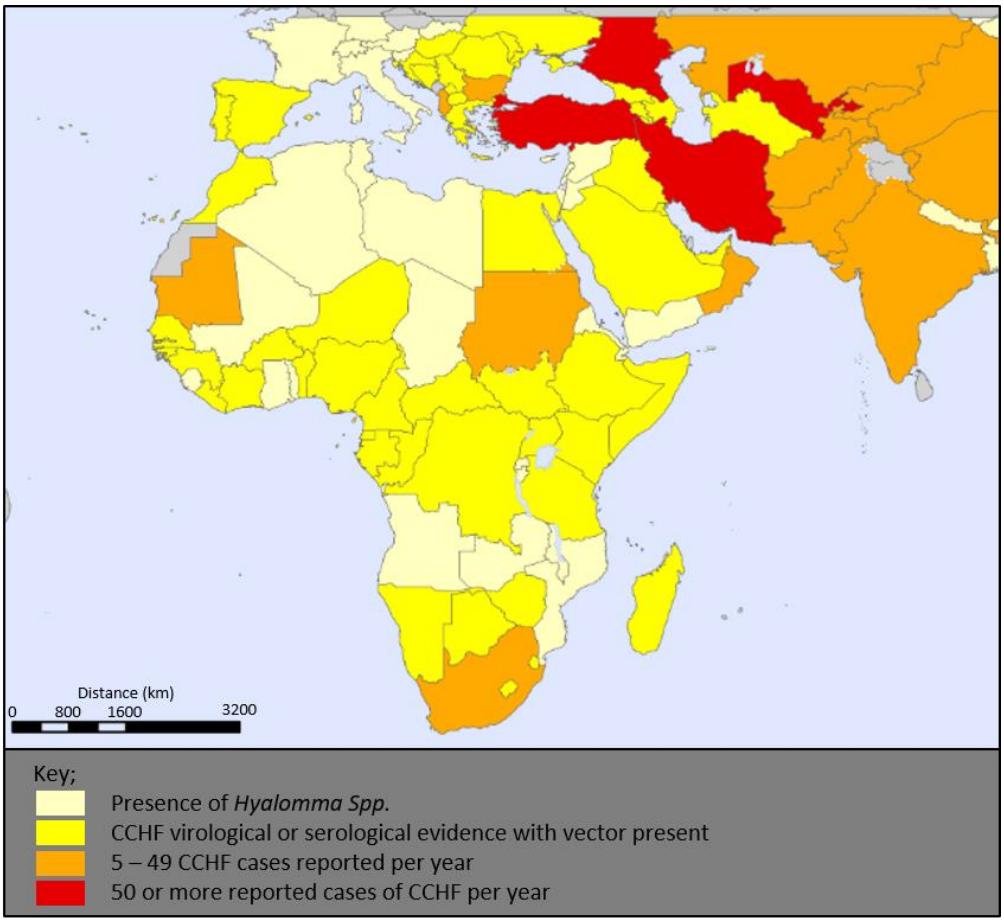


Figure 1.3, Geographic distribution of Crimean-Congo hemorrhagic fever

Countries in pale yellow report the presence of Hyalomma Spp. Countries in yellow have not reported cases of CCHF, but CCHFV has been isolated from ticks or animals in these countries, or serological evidence suggests its presence, in addition to presence of Hyalomma Spp. Orange countries report between 4 and 49 cases of CCHF per year. Red countries report more than 50 cases of CCHF per year. The horizontal line across the top of the map where countries are greyed out represents the northernmost limit of distribution of genus Hyalomma ticks. Image was adapted from an image created by the WHO in 2017.

(http://www.who.int/emergencies/diseases/crimean-congo-haemorrhagic-fever/en/).

1.1.5 Diagnosis of CCHF

Multiple methods of CCHF diagnosis exist, including reverse transcriptase (RT)-PCR, immunofluorescence assay (IFA), antibody and antigen capture ELISA and virus isolation. Due to the biosafety containment level 4 (BSL-4) requirements of CCHFV, virus isolation is rarely used as a diagnostic tool. RT-PCR is primarily used as it has the highest detection sensitivity at the earliest time point of infection (Mazzola and Kelly-Cirino, 2019). However, challenges to RT-PCR diagnosis of CCHF arise from the inherent variability in CCHFV genomic sequences, with the conserved genomic sequence required for effective RT-PCR not determined for all strains of CCHFV (Vanhomwegen *et al.*, 2012; Tezer and Polat, 2015). Therefore, the recommended diagnostic approach is a combination of a nucleic acid amplification test (NAAT), such as RT-PCR, with an immunological assay, such as an ELISA (Drosten *et al.*, 2003; Fernandez-Garcia *et al.*, 2014). Both NAATs and immunological approaches require a moderate to high level of biological containment infrastructure, which often does not exist at outbreak locations of CCHF. Rapid diagnostic tests (RDTs) that can diagnose at the point of care (POC) present an attractive option for initial screening in such locations. These POC devices are compact, self-contained and involve automated sample processing, meaning most healthcare workers can be trained in their use with minimal effort. RDTs utilise similar principles as NAATs and immunological based assays, combining them with a lateral flow format which enables much faster diagnosis, often within an hour of testing. The drawback of RDTs is their lower sensitivity of detection (Filippone *et al.*, 2013), especially with immunological based options, due to the low serological response observed in severe and fatal cases of CCHF.

1.1.6 Clinical manifestations and treatment options

Humans become infected with CCHFV following a tick bite or contact with contaminated livestock, often during the butchering process. Nosocomial infections are also widespread and well documented, with healthcare workers from Europe, Asia, and Africa all found to have contracted CCHFV (Burney *et al.*, 1980; Antoniadis and Casals, 1982; Watts *et al.*, 1988). Infection with CCHFV is often fatal, with 3-30 % of reported cases resulting in mortality, though previous

isolated outbreaks have seen significantly higher fatality rates of 70 % (Hoogstraal, 1979; Schwarz, Nsanze and Ameen, 1997; Mardani *et al.*, 2003).

Clinical progression of CCHF can be categorised into 4 distinct phases; incubation, pre-hemorrhagic, hemorrhagic and convalescence. The duration of these phases varies based on the source of infection, with delivery via tick bite resulting in the shortest duration of incubation (3.2 days) versus contact with infected blood of livestock (5 days) or humans (5.6 days) (Swanepoel *et al.*, 1987). The pre-hemorrhagic phase usually lasts an average of 3 days, though ranges from 1 to 7 days (Hoogstraal, 1979). During the pre-hemorrhagic phase symptoms of fever, headache, myalgia and dizziness are common, with some cases also causing diarrhoea, nausea and vomiting (Swanepoel *et al.*, 1987; Schwarz, Nsanze and Ameen, 1997; Karti *et al.*, 2004; Bakir *et al.*, 2005). The hemorrhagic phase develops rapidly, lasting between 2 and 3 days and has associated symptoms of petechiae and appearance of haematomas on mucus membranes and skin. Bleeding from the nose, gastrointestinal system, uterus and urinary tract are also common, though vaginal, gingival and cerebral hemorrhage have been documented (Swanepoel *et al.*, 1987; Ergönül, 2006). Convalescence occurs in survivors approximately 10 to 20 days after the first signs of illness; however, symptoms in this phase appear inconsistent across different outbreaks, including, but not limited to: tachycardia, labile pulse, loss of hearing and memory, polyneuritis and temporal complete loss of hair.

Almost 80 years following the first description of CCHF and the later association with CCHFV as the cause, supportive care still remains the mainstay of treatment for CCHF patients. Currently there are no licensed vaccines or CCHF specific therapeutics, however ribavirin has shown potential as an anti-viral, and although its efficacy is debated, it is recommended by the World Health Organisation (WHO) alongside general supportive care. Ribavirin is a guanosine analogue that undergoes phosphorylation within the target cell to ribavirin-triphosphate (RTP) (Witkowski *et al.*, 1972). RTP is then incorporated into the replicating RNA strand by the viral RNA dependent RNA polymerase (RdRp), as initially demonstrated with poliovirus, resulting in inhibition of chain elongation or chain termination and has since been used in treatment of Hepatits C virus (Crotty *et al.*, 2000; Maag

et al., 2001). A STAT-1 knockout mouse model has previously shown ribavirin to have a protective role against CCHFV challenge, with greater success observed when administered at early stages of infection and with lower viral titres (Bente *et al.*, 2010). However, another study assessing the efficacy of ribavirin and favipiravir, another nucleotide analogue also known as T-705, in an interferon (IFN) knockout mouse model found that despite a small number of mice surviving the CCHFV challenge, the majority succumbed to the disease and ribavirin had no significant effect on clinical outcome (Hawman *et al.*, 2018). This study found favipiravir to demonstrate statistically significant clinical benefit to subjects with CCHF, even when administered later following infection, though a small percentage of patients did experience onset of CCHF following conclusion of the treatment.

Despite the current lack of a globally approved vaccine, a potential CCHFV vaccine candidate has previously been trialled in Bulgaria. A virus preparation is grown in neonatal mouse brains then subsequently inactivated via chloroform treatment and heating to 58°C prior to absorbing it onto aluminium hydroxide (Vasilenko, 1976). Between 1953 and 1974, 1105 cases of CCHF were reported in Bulgaria with a fatality rate of 17 %. Following an immunisation programme in 1974, cases over the same period (1974 to 1996) dropped 4 fold to 279 with a fatality rate of 11.4 % and annual cases have remained relatively low (<20 cases per year) (Christova *et al.*, 2009; Maltezou *et al.*, 2010). However, this has only been authorised for use in Bulgaria where it was developed and has not met approval standards for other countries with at-risk populations due to safety concerns and the associated costs with scaling up production.

Several potential future vaccination options are currently being researched. Perhaps the most promising candidate is a virus replicon particle (VRP) incorporating the full length S and L genome segments from the IbAr10200 strain whilst omitting the M segment, instead directly transfecting in an expression plasmid to permit generation and amplification of VRPs. This was recently shown to offer heterologous protection versus multiple strains of CCHFV (IbAr10200, Oman-97 and Turkey), with 100 % of the interferon α/β-receptor knockout ($IFNAR^{-/-}$) mice surviving virus challenge with no signs of clinical disease

following a single dose of the VRP (Spengler *et al.*, 2019). Whilst other previous work has demonstrated complete protection vs CCHFV challenge in IFNAR$^{-/-}$ mice, all required multiple doses and not all offered protection against multiple strains (Buttigieg *et al.*, 2014; Hinkula *et al.*, 2017; Aligholipour Farzani *et al.*, 2019; Rodriguez *et al.*, 2019).

1.2 The virus

1.2.1 Virus structure

CCHFV and HAZV have a host-derived membrane and present as broadly spherical virions approximately 80 to 100 nm in diameter (Martin *et al.*, 1985). The structural glycoproteins Gn and Gc protrude through the membrane, extending 5 to 10 nm from the surface and are responsible for binding to cellular receptors during the entry phase (Bergeron, Vincent and Nichol, 2007). Within the virion, the viral genome segments are encapsidated within the viral nucleoprotein (N) and RdRp, with at least one copy of each genome segment within a single virion. A schematic of the HAZV virion and electron micrographs of CCHFV and HAZV are depicted in Figure 1.4, detailing the spherical nature and relative sizes of the viruses.

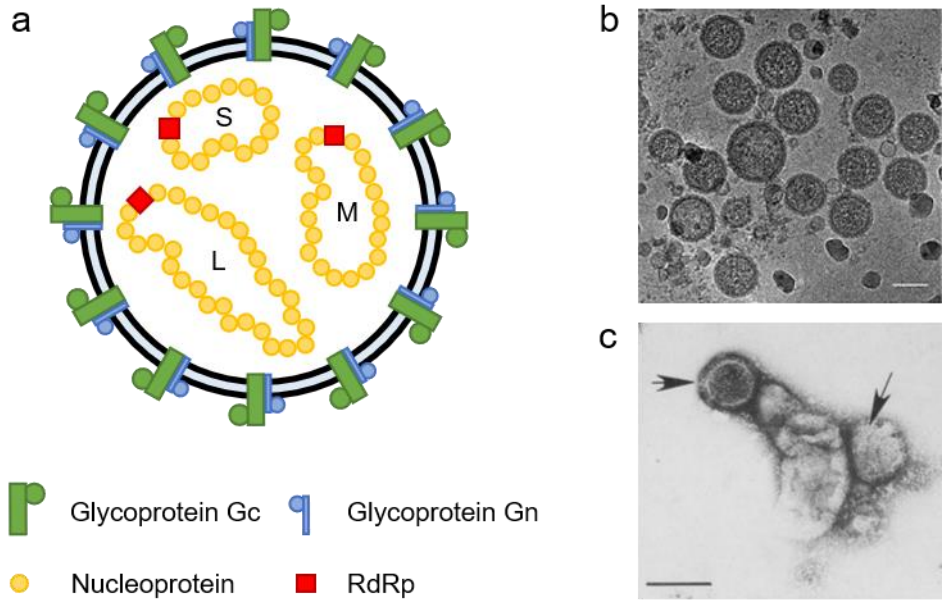


Figure 1.4, Schematic diagram and electron micrographs of HAZV and CCHFV

a) Schematic representation of a HAZV and CCHFV virion showing the glycoprotein studded membrane enclosing the three genome segments encapsidated with nucleoprotein and a copy of the viral RNA dependent RNA polymerase (RdRp) to form an RNP. b) Electron micrograph of HAZV virions, taken from (Punch *et al.*, 2018)*, scale bar 100 nm. c) Electron micrograph of CCHFV virions, taken from (Martin, Lindsey-Regnery et al. 1985), scale bar 100 nm.*

1.2.2 Viral genome structure

The genome of CCHFV and HAZV is comprised of three single stranded RNA (ssRNA) segments, termed small (S), medium (M) and large (L), that encode the viral N, glycoprotein precursor (GPC) and RdRp, respectively. The CCHFV S segment has been reported to encode an additional non-structural protein (NSs) using an ambisense coding strategy (Barnwal *et al.*, 2016). Sequence and amino acid conservation vary within CCHFV when geographically diverse strains are compared, with 20, 31 and 22 % nucleotide differences and 8, 27 and 10 % amino acid differences in the S, M and L segments, respectively (Deyde *et al.*, 2006).

The open reading frame (ORF) of each gene segment is flanked by a non-translated region (NTR) at the 5' and 3' termini. These NTRs contain *cis*-acting signals that acts as promoters for viral RNA synthesis activities (Matsumoto *et al.*, 2019; Mega *et al.*, 2020). NTR sequences are also postulated to have roles involved in several additional aspects of the viral lifecycle, including encapsidation of viral RNA (vRNA) by N into ribonucleoprotein (RNP) complexes

and packaging of these RNPs into virions (Kohl *et al.*, 2006). The length of the segment specific NTRs varies, however the 5' NTR in the vRNA is in all cases longer than its 3' partner. The terminal nine nucleotides (5'-UCUCAAAGA and 3'AGAGUUUCU) are entirely conserved between segments and across different nairoviruses and may form closed hairpins to enable binding to the viral RdRp as shown in related viruses such as La Crosse virus (LACV) (Gerlach *et al.*, 2015).

1.3 The replication cycle of CCHFV and HAZV

A complete and detailed description of the HAZV and CCHFV lifecycles remains unclear, due in part to the classification of CCHFV as a hazard level 4 pathogen, requiring BSL-4 facilities for its growth and study. In this section, current knowledge surrounding individual aspects of the life cycle of both HAZV and CCHFV is discussed.

1.3.1 Binding and entry

The initial phase of viral replication for enveloped viruses requires an interaction between viral and host cell components. For CCHFV and HAZV, this is thought to be mediated by the viral glycoproteins. It is hypothesised that Gc serves as the viral component in the interaction (Xiao *et al.*, 2011). Monoclonal antibodies against Gc, but not Gn, have been proven effective in preventing CCHFV infection during neutralisation assays *in vitro*. However, this is not consistent with *in vivo* results where monoclonal antibodies against Gn were found to offer more protective immunity against CCHFV in mice (Bertolotti-Ciarlet *et al.*, 2005). There is little information surrounding the cellular receptor for CCHFV. A functional interaction has been suggested between Gc and cell surface nucleolin, a protein normally found in the nucleus, via coimmunoprecipitation reactions between soluble Gc and CCHFV susceptible cell lines (Xiao *et al.*, 2011).

Following binding, CCHFV uptake into the host cell occurs in a clathrin, pH and cholesterol dependent manner, trafficking at least part way through the endosomal pathway (Simon, Johansson and Mirazimi, 2009; Garrison *et al.*, 2013). Dependence on clathrin was demonstrated by inhibition of the clathrin heavy chain using Pitstop-2 (PS2). Increasing the concentrations of PS2 administered to host cells led to reduced infectivity of CCHFV, implicating the

requirement of functional clathrin coated pits in CCHFV infection (Garrison *et al.*, 2013). Further evidence for the role of clathrin mediated endocytosis (CME) was provided by siRNA knockdown of the clathrin adaptor protein (AP) 2. AP2 is the second most abundant protein in clathrin coated pits and is not known to have additional roles outside CME (Blondeau *et al.*, 2004). siRNA-mediated knockdown of AP2 subunits inhibited CCHFV infection by ≈ 30 to 40 % (Garrison *et al.*, 2013). The role of pH in viral entry has been investigated for both CCHFV and HAZV. Treatment of cells with NH_4Cl, which prevents maturation of early endosomes to late endosomes via neutralisation of the luminal pH, reduced CCHFV infectivity in a dose-dependent manner, demonstrating a low pH requirement for infection (Garrison *et al.*, 2013). To assess how far through the endosomal pathway CCHFV trafficked, dominant negative (DN) forms of Rab5, a marker for early-endosomes, and Rab7, a marker for late endosomes were transfected into cells prior to infection. Rab5 and Rab7 associate with the membranes of their respective vesicles and shuttle cargo via interactions with effector proteins (Stenmark *et al.*, 1994; Mukhopadhyay *et al.*, 1997). DN forms of Rab5 and Rab7 contain a single amino acid change that renders the protein inactive, due to the inability to exchange GTP to GDP. Therefore, a virus requiring functional early-endosomes or late-endosomes would be negatively impacted by transfection of the respective DN form of Rab. CCHFV infectivity was reduced following transfection of DN Rab5 but not for DN Rab7, showing a requirement for early, but not late endosomes (Garrison *et al.*, 2013).

For HAZV, entry has been shown to be dependent on pH, potassium concentration (K^+) and cholesterol (Punch *et al.*, 2018; Charlton *et al.*, 2019). Punch *et al.* identified increased infectivity of HAZV in elevated K^+, with an optimal pH for infectivity at 7.3. Purification of HAZV particles and subsequent treatment at pH 7.3 with high or low K^+ concentrations revealed a conformational change in the surface glycoproteins via cryo-electron tomography. A hypothesised model suggested that under high K^+ conditions, either or both of the HAZV glycoproteins extend into a fusion ready state (Punch *et al.*, 2018). The HAZV dependence on cellular cholesterol was identified in a follow up study by Charlton *et al.* Here, the depletion of cellular cholesterol via treatment with U18666A or using methyl-beta cyclodextrin resulted in reduced infectivity of HAZV. Due to the previously

described requirement for HAZV fusion on K^+ and pH, the role of cholesterol in facilitating these conditions was also explored. Priming of HAZV prior to infection of cholesterol depleted cells recovered the infection, suggesting the HAZV dependence on cholesterol to be linked to establishing the pH and K^+ conditions within endosomes that facilitate fusion with host cell membranes (Charlton *et al.*, 2019).

1.3.2 Replication

There is little information surrounding the specific mechanism for CCHFV or HAZV RNA transcription and replication. However, in order to generate additional viral proteins and viral RNA, each viral genome segment encapsidated within an RNP must act as a template for two distinct RNA synthesis activities (Figure 1.5). For generation of viral proteins, the genomic segment must be transcribed into translatable mRNA. For the generation of additional viral genomes, the genomic segment must be replicated to synthesise full length complementary RNA (cRNA) copies for each viral genome segment, which may then act as templates for the synthesis of genomic vRNA. The two products, cRNA and mRNA, have distinct structures. Whilst the cRNA represents a full-length complementary version of the initial vRNA, the mRNA is extended at the 5’ end by the addition of a host-derived cap and truncated at the 3’ end in relation to the vRNA (Walter and Barr, 2011).

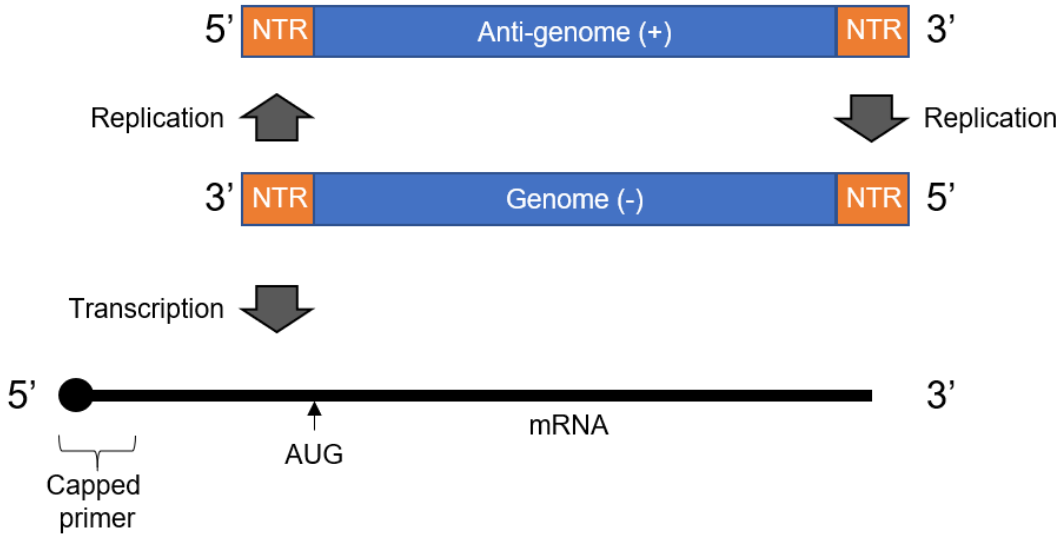


Figure 1.5, Schematic detailing negative sense gene expression

Schematic detailing how negative sense genomes act as a template for two discrete activities; RNA replication into anti-genomes and mRNA transcription.

Following cell entry and fusion, genomic RNPs are released into the cytosol, where the vRNA must serve as a template for the RdRp during primary transcription of mRNAs. Translation of mRNA requires a 5’ cap which acts to recruit initiation factors and in mediating binding to the 40S ribosomal subunit. Sequencing of bunyavirus mRNAs identified non-viral sequences at the 5’ termini, suggesting bunyaviruses utilise a cap-snatching mechanism to prime viral transcription (Patterson and Kolakofsky, 1984; Collett, 1986). Cap snatching occurs following the binding of a cap-binding domain of a viral polymerase to the cap structure of a cellular mRNA. Once bound, an endonuclease domain cleaves the RNA approximately 10-15 bases downstream of the cap-binding domain, generating a capped RNA fragment that can be used to prime viral transcription. Whilst unconfirmed for CCHFV, there is evidence that the RdRp contains a Mn^{2+} coordinating residue at position D693 that facilitates the cap snatching process, a process described in multiple negative sense viruses including influenza, arenaviruses and other bunyaviruses (Dias *et al.*, 2009; Morin *et al.*, 2010; Reguera, Weber and Cusack, 2010; Devignot *et al.*, 2015).

At a stage of the replication cycle that is currently undefined, the RdRp can begin replication of the input vRNA into cRNA. As the 5’ end of the cRNA is distinct to the 5’ end of mRNA, it is likely the pathways involved in their individual production

are distinct. Replication of vRNA to generate cRNA is known not to require capped primers as mutation of D693, whilst able to prevent RdRp transcription of mRNA, has no effect on replication (Devignot *et al.*, 2015). Replication of vRNA is therefore a process that initially requires the replication and encapsidation of uncapped vRNA, minimally involving N and the viral RdRp. This process results in the generation of cRNA that acts as a template for further amplification of vRNAs. During cRNA or vRNA synthesis, N subunits are added to the elongating strands to form antigenomic or genomic RNPs, respectively.

1.3.3 Assembly and egress

As with other stages of the CCHFV replication cycle, there is limited specific information available relating to the assembly and egress phases of infection. The subcellular localisation of mature glycoproteins can often provide insight into the location of viral assembly. A study into the localisation of CCHFV glycoproteins Gn and Gc identified Gn to localise to the Golgi, with Gc remaining in the ER following individual expression of both glycoproteins (Bertolotti-Ciarlet *et al.*, 2005). Upon co-expression of Gn and Gc, both glycoproteins co-localised with the Golgi, suggesting a Golgi localisation signal on Gn, and that Gc may localise to the Golgi via an interaction with Gn. Endoglycosidase H (Endo-H) treatment is commonly used to monitor post-translation modification of glycoproteins in the Golgi apparatus. CCHFV Gn and Gc were both found to be sensitive to Endo-H treatment, indicating they are targeted to early compartments of the Golgi (Bertolotti-Ciarlet *et al.*, 2005). This is also similar for members of the *orthohantavirus* genus, which suggests a conserved mechanism across multiple *Bunyaviridae* families (Ruusala *et al.*, 1992).

For members of the *Bunyaviridae* family, evidence supports the Golgi apparatus as the site of “virus factories”, in which virus particles assemble, bud and mature (Salanueva *et al.*, 2003). The model bunyavirus BUNV, is able to build factories surrounding the Golgi, consisting of virally induced tubular structures (Salanueva *et al.*, 2003). These structures have both viral and cellular origins and appear to be involved in multiple aspects of the viral lifecycle including genome replication, transfer of RNPs to assembly sites and viral morphogenesis (Fontana *et al.*, 2008). Detection of components indicative of viral replication, such as dsRNA,

BUNV N and viral polymerase within these structures supported the role of these structures as viral factories.

Egress of mature CCHFV virions is dependent on intact microtubules. Manipulation of microtubules using nocodazole (NOC) and paclitaxel (PAC) resulted in reduced egress of CCHFV, with less extracellular virus detected versus untreated controls (Simon *et al.*, 2009). The role of microtubules was suggested to be in both CCHFV assembly and egress. The manipulation of microtubules with NOC, which disrupts microtubule assembly/disassembly, reduced levels of intracellular virus, indicating a requirement of intact microtubules in virus assembly. Conversely, treatment with PAC, which acts to stabilise microtubules, had no significant impact on intracellular virus levels, suggesting that when intact microtubules were available, virus assembly could proceed normally (Simon *et al.*, 2009). CCHFV release has been shown to occur from the basolateral surface of polarised MDCK cells, with high titres detected released from the basolateral surface versus limited titres from the apical surface, indicating a preference for basolateral release (Connolly-Andersen, Magnusson and Mirazimi, 2007). However, the molecular mechanisms underpinning this preference are unknown, but may represent an interaction with a host cell protein, which are secreted in a polarised manner.

1.4 Virally encoded proteins

1.4.1 S segment

The S segment of HAZV and CCHFV minimally encodes the viral nucleoprotein, with CCHFV also recently reported to express an additional NSs using an ambisense coding strategy (Barnwal *et al.*, 2016).

1.4.1.1 Structure of HAZV N and CCHFV N

The structure of both CCHFV and HAZV N have been solved by multiple groups, including our own (Carter *et al.*, 2012; Guo *et al.*, 2012; Wang *et al.*, 2012, 2015; Surtees *et al.*, 2015). Both CCHFV and HAZV N display extensive similarities in structure, with two distinct domains, a globular domain and an arm domain (Figure 1.6, a-b). The globular domain of HAZV N is comprised of a central core of C-terminal helices α13 to α20, flanked by N-terminal helices α1 to α8. The arm

domain is comprised of two long helices, α11 and α12, with two supporting helices α9 and α10 (Surtees *et al.*, 2015). The globular domain of CCHFV is formed from the central core of C-terminal helices α18 to α29 flanked by N-terminal helices α1 to α11, with the arm domain formed again by two long helices α15 and α17 with a trio of supporting helices α12 to α14 (Carter *et al.*, 2012; Wang *et al.*, 2012, 2015).

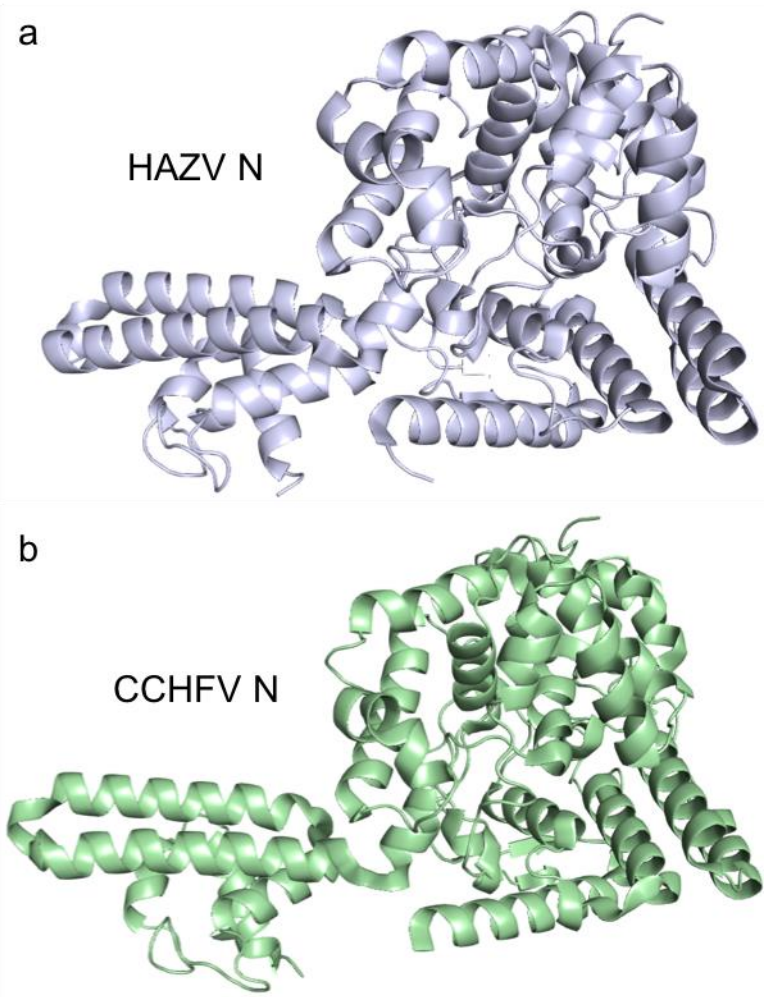


Figure 1.6, Comparison of the crystal structures of HAZV N and CCHFV N

Comparison of the crystal structure of a) HAZV N and b) CCHFV N shows the high similarity between the two nairoviral nucleoproteins in structure of both arm and globular domains.

The angle of the arm domain in relation to the globular domain is thought to be variable due to a flexible linker region between the two domains. In HAZV N the two domains are linked by an alpha helix, α13, and an additional disordered loop formed of residues 187 to 196 (Surtees *et al.*, 2015). The flexible nature of the

arm domain is perhaps displayed most evidently in the crystal structures of CCHFV N, all three solved structures, arising from the IbAr10200, Baghdad-12 and YL04057 strains, present the arm domain in a different orientation (Carter *et al.*, 2012; Guo *et al.*, 2012; Wang *et al.*, 2015). These variable positions may have functional implications, such as the ability of N to bind RNA.

During formation of the RNP, monomers of the HAZV N and CCHFV N interact with other monomers forming oligomers. For HAZV N, 6 residues have been identified in this interaction on the arm domain; Leu279, Trp 263, Lys 275, Val271, Glu270 and Phe216 and a single residue on the globular domain; Pro355 (Surtees *et al.*, 2015). The proline residue from the globular domain sits within a hydrophobic pocket formed by the 6 residues on the arm domain. This style of interaction between arm and globular domains is also observed in CCHFV N oligomers, suggesting this interaction may be responsible for the formation of long helical chain oligomers of N that may have functional relevance to their role in RNP formation (Wang *et al.*, 2012). It is unknown as to whether N binds RNA or forms oligomers first during the formation of the RNP. The RNA binding potential of both HAZV N and CCHFV N has been explored. Plotting of the surface electrostatic potential of HAZV N identified multiple regions of positive charge, with two significant areas; a platform located adjacent to the arm domain and a crevice on the opposite face of the protein, extending inwards and offering protection against degradation for any RNA sequestered within it (Surtees *et al.*, 2015). Comparisons with CCHFV show conservation of the platform and pocket regions (Carter *et al.*, 2012; Wang *et al.*, 2015). Given the flexible nature of the arm and the proximity of the platform to the arm domain, it is possible the flexibility of the arm may act as a “gate” mechanism, regulating the binding of RNA and formation of RNPs. Despite identifying the potential regions RNA could bind, an electrophoretic mobility shift assay, which assess DNA/RNA: protein interactions, showed weak affinity of CCHFV for a 24 nucleotide ssRNA fragment (Guo *et al.*, 2012). However, as this study utilised purified monomeric CCHFV N; it is possible either CCHFV N oligomers bind RNA more readily. Furthermore, the complementarity of the 5’ and 3’ NTRs suggests an interaction resulting in formation of a dsRNA panhandle structure that circularises the vRNA is possible. *In vitro* analysis of the ability of CCHFV to bind to this dsRNA structure

demonstrated a high affinity interaction specifically between the stalk domain of CCHFV N and vRNA panhandle, with CCHFV N able to bind to both HAZV and CCHFV vRNA, but not vRNA from the related Andes virus of the *Hantavirus* genus (Jeeva *et al.*, 2019).

1.4.1.2 Additional roles of HAZV N and CCHFV N

In addition to the formation of RNPs, as described in section 1.4.1, previous research has identified both HAZV N and CCHFV to express caspase-3 cleavage motifs, ^{269}DQVD^{272} and ^{266}DEVD^{269} respectively, located on the tip of the arm domain on a flexible loop that links the two helices. The exact biological role of this motif is unclear, however *in-vitro* analysis of purified active caspase-3 with purified HAZV N or CCHFV N demonstrated the ability of N to be cleaved into two products of approximate mass 20 and 32 kDa (Carter *et al.*, 2012; Wang *et al.*, 2012; Surtees *et al.*, 2015). It is therefore expected that N acts a substrate decoy in a similar manner to the Junin nucleoprotein, which via expression of multiple caspase motifs is able to reduce levels of caspase 3 cleavage and thereby delay progression of apoptosis (Wolff, Becker and Groseth, 2013). Interference with immune system pathways is a tactic employed by many viral families, multiple viruses have evolved mechanisms to both promote or down-regulate the induction of apoptosis. For viruses that delay apoptosis, examples include inhibition of the pro-apoptotic tumour-suppressor p53 by the DNA virus SV40 and the direct inhibition of effector caspases by baculovirus p35 protein, which following cleavage remains irreversibly bound to the active site of the caspase (Zhou *et al.*, 1998; Ali and DeCaprio, 2001).

In addition to interacting with host cell proteins, HAZV N and CCHFV N, alongside other Bunyavirales members, have been shown to interact with cellular nucleic acids. CCHFV N has previously been shown to display intrinsic nuclease activity in the presence of Mn^{2+}, Co^{2+} or Mg^{2+} cations on both single and double stranded DNA, in a sequence nonspecific manner. This nuclease activity was specific to DNA, as attempts to digest single or double stranded RNA were unsuccessful. Responsible amino acids for this activity were identified as Y374, R384, E387, K411, H453 and Q457 via site-directed mutagenesis, each with variable significance of impact on the endonuclease activity (Guo *et al.*, 2012). These

residues sit within a positively charged pocket in the globular domain of CCHFV. Further structural analyses on HAZV, Kupe virus (KUPV) and Erve virus (ERVEV) all identified similar positively charged pockets, however in ERVEV a loop region folds over the entrance to the pocket, blocking solvent access. This structural blockade correlated with endonuclease activity, both KUPV and HAZV displayed similar levels of endonuclease activity to CCHFV, whilst ERVEV demonstrated significantly lower enzymatic activity (Wang *et al.*, 2012).

In addition to N, the S-segment of CCHFV has also recently been reported to express a second protein, NSs (Barnwal *et al.*, 2016). The ORF of CCHFV NSs expresses a 150 amino acid product coded in a positive sense, overlapping the ORF of CCHFV N. Ectopic expression of CCHFV NSs resulted in activation of apoptotic pathways, with the detection of the characteristic apoptotic makers active caspase 3/7 and cleaved poly(ADP-ribose) (PARP). Mutagenesis of NSs determined the residues responsible for apoptosis activation to be Leu127 and Leu13, which triggered apoptosis via association with the mitochondrial membrane and disruption of the membrane potential, resulting in the release of cytochrome C into the cytoplasm. No such evidence exists for expression of NSs in HAZV, however an ORF in a similar position does exist and could present an interesting study to identify if NSs expression is a shared feature between these viruses.

Expression of NSs is not unique to CCHFV within the Bunyavirales. The type species of the family, bunyamwera virus (BUNV) NSs is required for efficient BUNV replication in mosquito cells (Szemiel, Failloux and Elliott, 2012). BUNV NSs also counteracts the immune response, via delaying interferon regulatory factor 3 mediated apoptosis, preventing cell death in the early stages of BUNV infection (Kohl *et al.*, 2003). Rift Valley Fever Virus (RVFV) NSs also interacts with cellular factors, with the ability to shut off host cell transcription via interaction with the transcription factor IIH (Le May *et al.*, 2004). For non-human viruses, members of the Tospoviridae also express an NSs. Tomato spotted wilt orthospovirus (TSWV) encodes NSs using an ambisense strategy and has been identified as a virulence factor both in host plant cells and the insect vector. TSWV NSs is injected into plants in the saliva of infected thrips, where it interferes with

the RNA silencing or antiviral RNAi response (Kormelink *et al.*, 1991; Nagata *et al.*, 1999).

1.4.2 M segment

The M segment encodes the GPC, a large, 1689 amino acid polyprotein that is processed via complex post-translational modifications to yield the structural glycoproteins Gn and Gc and in the case of CCHFV, the non-structural glycoproteins GP160, GP85, GP38 and NS_m, GPmuc and proGc (Figure 1.7, a) (Sanchez, Vincent and Nichol, 2002; Sanchez *et al.*, 2006; Altamura *et al.*, 2007). It was reported that HAZV encodes three structural glycoproteins within the GPC of apparent molecular weights 84, 45 and 30 kDa, however this has not been confirmed since (Foulke, Rosato and French, 1981).

1.4.2.1 GPC processing

The GPC is heavily glycosylated to facilitate cleavage by host cell proteases to yield the structural and non-structural products. This enables members of the proprotein convertase family of serine proteases to process the GPC (Bergeron, Vincent and Nichol, 2007). The N-terminal region of the GPC contains a signal peptide that directs synthesis to the secretory pathway. The exact termini of this signal peptide are strain specific, with signalase cleavage predicted to occur in the region of amino acids 21-27, dependant on strain (Sanchez, Vincent and Nichol, 2002). The signal peptide is located within a highly variable mucin-like domain (MLD). The MLD of CCHFV, strain Matin, is encoded by amino acids 1-248, the five C-terminal residues of this domain (243-248) are highly variable in relation to the remainder of the M ORF, with 23.8 to 56.4% amino acid difference among strains, resulting from a nucleotide variability of 14.3 to 37.2% (Papa *et al.*, 2002; Sanchez, Vincent and Nichol, 2002). The signal peptide is predicted to be aa 1 to 24 and is removed in the secretory pathway (Sanchez, Vincent and Nichol, 2002). Complete processing of the CCHFV GPC requires multiple cellular proteases and extensive co- and post-translational modifications. Further sequence analysis of the GPC revealed 5 transmembrane domains and 10 *N*-linked glycosylation sites, with 3 present in Gc, 1 present in Gn, 2 in GP38 and 4 in the above mentioned MLD (Sanchez, Vincent and Nichol, 2002).

A series of primary cleavage events to the GPC generate preGn, NS_m and preGc. These occur close to the 2nd and 4th transmembrane domain and are reported to be due to the action of intramembrane cleaving proteases (iCLIPs) (Sanchez, Vincent and Nichol, 2002; Altamura *et al.*, 2007) (Figure 1.7, a). These products, preGn, NS_m and preGc, are then trafficked to the Golgi where the MLD of preGn is *O*-glycosylated and subsequently cleaved early in the secretory pathway, either in the ER or on route to the *cis*-Golgi, by the subtilisin kexin isozyme-1/site-1 proteases (SKI-1/S1P) at a specific RRLL motif generating GP160/85 and Gn (Vincent *et al.*, 2003) (Figure 1.7, b). The protease responsible for the cleavage event resulting in generation of mature Gc from preGc is unknown, however the presence of the motif RKPL, which is the responsible cleavage site of the Guanarito arenavirus by SKI-1/S1P, suggests a similar protease may act on CCHFV preGc (Rojek and Kunz, 2008). GP160/85 is processed further in the *trans*-Golgi network (TGN) by furin at a RSKR motif located between the GP38 and MLD (Figure 1.7, b) (Sanchez, Vincent and Nichol, 2002; Bergeron *et al.*, 2015).

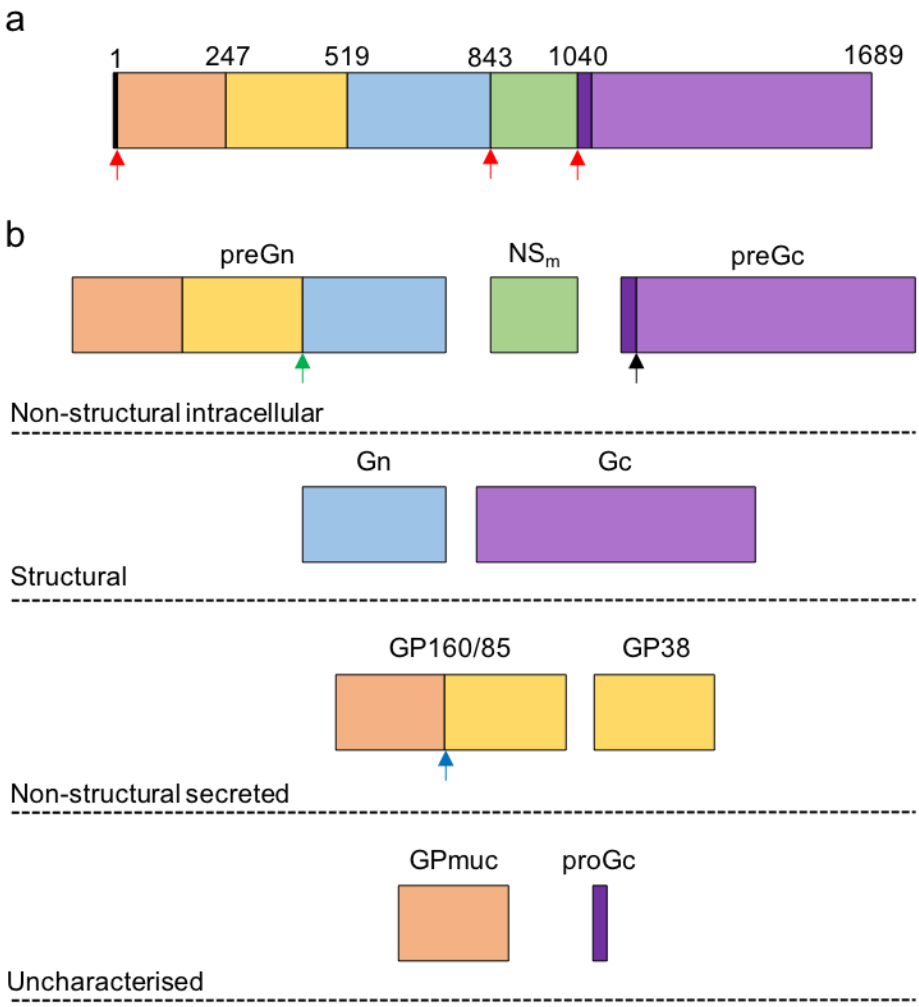


Figure 1.7, Processing of the CCHFV glycoprotein precursor by cellular proteases

a) CCHFV GPC is expressed as a 1689 amino acid polyprotein from which all structural and non-structural glycoproteins are derived. Primary cleavage of GPC is a result of the action of iCLIPs and signal peptidases (red arrows). b) Primary cleavage generates the non-structural intracellular preGn, NS_m and preGc proteins. Cleavage by SKI-1/S1P (green arrow), SKI-1/S1P-like (black arrow) and furin (blue arrow) yields the structural proteins Gc and Gn, non-structural secreted proteins GP160/85 and GP38 and the uncharacterised GPmuc and proGc proteins.

1.4.3 L Segment

The L segments of HAZV and CCHFV are both approximately 12 kb in length, containing a 11835 nucleotide ORF, a 76 nucleotide 3' NTR and a 253 nucleotide 5' NTR (Kinsella *et al.*, 2004). The ORF encodes a polyprotein approximately 448 kDa, making it significantly larger than related bunyavirus L proteins (Honig, Osborne and Nichol, 2004).

The N terminal region of the RdRp encodes a functional ovarian tumor (OTU)-like protease domain, with additional roles in deubiquitination and immune evasion (Frias-Staheli *et al.*, 2007). This OTU domain was shown to be dispensable for RNA replication in a minigenome system, but was essential for recovery of infectious virus, explained in more detail in section 1.5 (Bergeron *et al.*, 2010; Scholte *et al.*, 2017). This domain was further identified to suppress the cellular immune response by acting as an IFN antagonist, following studies identifying the RdRp bound to interferon stimulated gene (ISG) 15 and ubiquitin (James *et al.*, 2011).

At the time of writing, no full length orthonairovirus L protein structure has successfully been solved by x-ray crystallography. The polymerase of the related LACV was solved by electron microscopy and x-ray crystallography, complexed with the 5’ and 3’ NTRs (Gerlach *et al.*, 2015). Interestingly, despite the complementary nature of the terminal 9 bases of the 5’ and 3’ NTRs, the solved structure showed the 5’ NTR forming a stem-loop structure and binding to the RdRp in a distinct location to the 3’ NTR, in a similar manner previously identified for influenza virus (Pflug *et al.*, 2014). Binding of the NTRs to the polymerase in this way, rather than using complementary Watson-Crick base pairing to form a “pan-handle” structure between 5’ and 3’ NTRs suggests an alternative mechanism for circularisation of RNPs in the *Bunyavirales* order. In addition, the structures of the Lassa and Machupo arenavirus polymerases have been solved and shown to display similar architecture to LACV and influenza virus polymerases (Peng *et al.*, 2020).

1.5 Nairoviral reverse genetics systems

In classical, “forward”, genetics the starting point is a mutant phenotype. Researchers work backwards from the phenotype and identify the mutation responsible for it, thereby associating the phenotype to a gene. With reverse genetics, advances in recombinant DNA technologies and DNA sequencing enabled these forward genetics approaches to be reversed, allowing a researcher to begin with a gene of interest, insert mutations within it and analyse gene function from the resulting phenotype. Due to the reversion of classical genetics, this approach has been termed “reverse genetics”. This approach can

be applied to virology to understand the role of viral genes and non-coding regions during replication, immune evasion and additional host-pathogen interactions. Reverse genetics systems are especially useful in uncovering the basics of viral genome transcription, replication, assembly and egress, whilst also permitting mutational analysis and the development of reporter viruses. Minigenome systems have been developed for multiple bunyaviruses, namely, BUNV, Uukuniemi virus, Hantaan virus, RVFV and LACV (Dunn *et al.*, 1995; Flick and Pettersson, 2001; K. Flick *et al.*, 2003; Klemm *et al.*, 2013).

For HAZV and CCHFV, a number of different reverse genetics systems exist. For CCHFV, a minigenome system, virus-like particle (VLP) system and infectious clone system have all been reported and used to uncover detailed information surrounding CCHFV replication that will be discussed further in the following sections (Bergeron *et al.*, 2010, 2015; Devignot *et al.*, 2015; Zivcec *et al.*, 2015). For HAZV, at the outset of this project, no minigenome, VLP or infectious clone systems had yet been described. However, in 2019 a minigenome system was successfully developed and reported for HAZV, allowing detailed study into viral transcription (Matsumoto *et al.*, 2019).

1.5.1 Minigenome system

Minigenome systems for negative sense viruses use a genome-like reporter RNA flanked by viral NTRs encapsidated by N, allowing recognition and transcription by the viral RdRp. The minigenome system therefore enables study into gene expression and genome replication. As the viral coding regions are replaced with reporters, the system is non-infectious and as such can be used in low containment facilities.

The first report of a CCHFV minigenome system was published in 2003, where the CCHFV S segment was transcribed by pol I and rescued via coinfection with a helper virus. This reporter encoding S segment was also packaged into CCHFV virions, representing the first recombinant CCHFV system (R. Flick *et al.*, 2003). A more comprehensive CCHFV minigenome system was developed in 2010, where CCHFV coding regions for S, M and L segments were replaced with eGFP and *Gaussia* luciferase coding regions. Initially, the viral NTRs flanking the

reporters were found to contain sufficient signals for transcription, replication and packaging, following superinfection of BSR-T7/5 cells with CCHFV following transfection of the minigenome reporter. This initial study using a helper-virus also identified a gradient in efficiencies between the CCHFV S, M and L segment NTRs, with L having the highest and S the lowest signal (Bergeron *et al.*, 2010). Further developments of the system in the same study removed the requirement of the helper virus. Transfection of the CCHFV RdRp and CCHFV N, as opposed to a superinfection, alongside the S, M or L minigenome RNAs showed the segment specific NTRs supported similar levels of replication, suggesting the gradient in signal observed with the earlier superinfection was linked to variable packaging efficiencies across the segment specific NTRs (Bergeron *et al.*, 2010). The minigenome system also demonstrated its use in screening anti-viral compounds that affect the genome replication phase of infection. Ribavirin, previously described in section 1.1.6, was shown to have a dose dependent inhibitory effect on CCHFV in cell culture models (Bergeron *et al.*, 2010).

The HAZV minigenome system, in which the HAZV coding regions for the S, M and L segments were replaced by *Renilla* luciferase, was first described in a study exploring the role of complementarity and sequence in NTRs during viral transcription (Matsumoto *et al.*, 2019). The study identified that in addition to the terminal 9-11 nucleotides, termed promoter element (PE) 1, a second region directly adjacent to PE1 was also required for efficient transcription. This region, termed PE2, contains an initial region of non-complementary bases at positions 12 to 15/16 from the genome end, followed by a region of at least 9 complementary bases. Mutational analysis identified the non-complementary region to be critical, with the 3’ and 5’ ends independently changed to give complementary base pairing with the opposite end resulting in severely reduced reporter signal. Furthermore, the region of complementarity following this initial region was shown to be both sequence and complementarity specific (Matsumoto *et al.*, 2019). This revealed a significant difference between NTRs of the *Nairoviridae* and other Bunyaviruses, in which no region of ssRNA exists following PE1. It is hypothesised that this ssRNA region enables flexibility between PE1 and the dsRNA sequence of PE2, potentially enabling it to bind to an additional RNA binding site on the RdRp.

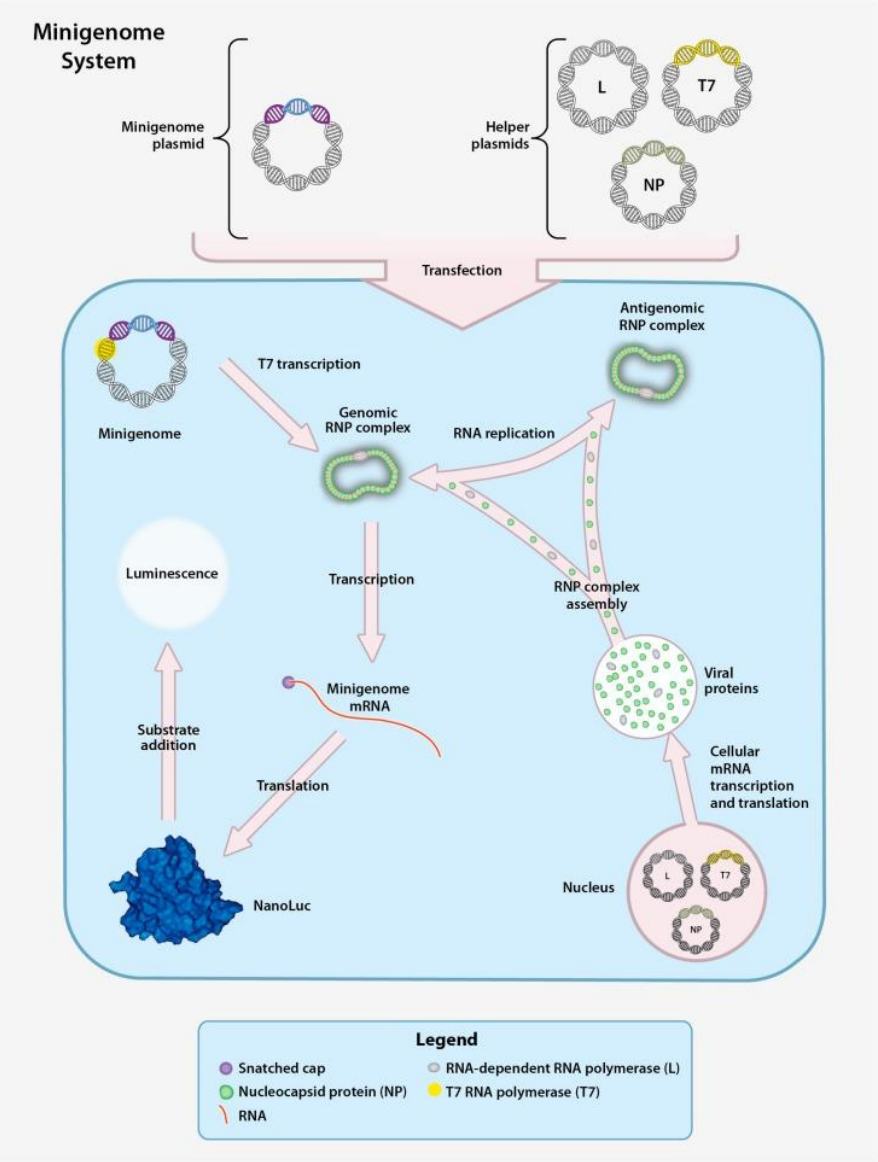


Figure 1.8, CCHFV minigenome system schematic

The CCHFV minigenome system comprises of a reporter encoded on a minigenome plasmid and three support ("helper") plasmids, L, NP and T7 encoding the viral polymerase, nucleoprotein and T7 RNA polymerase downstream of an RNA pol II promoter. The minigenome plasmid encodes a reporter, such as NanoLuc, Renilla or Gaussia luciferase flanked by the 5' and 3' viral non-translated regions (NTRs) from either the S, M or L genome segment. Co-transfection of support plasmids provides the corresponding viral proteins for transcription of vRNA from the minigenome plasmid. This vRNA is encapsidated by viral proteins expressed from the support plasmids into genomic RNPs that permit transcription of mRNA and its translation into the reporter protein capable of generating a luminescent signal that can be detected.

Figure adapted from (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4848600/)

1.5.2 VLP system

The VLP system expands upon the principles of the minigenome system by co-expression of viral glycoproteins alongside the viral NTR flanked reporter RNA sequences, viral N and polymerase. This leads to self-assembly of transcriptionally competent VLPs that can be internalised into fresh cells, allowing study of assembly, attachment and entry processes.

Whilst no such system currently exists for HAZV, a CCHFV VLP system was described in 2015 (Devignot *et al.*, 2015). Using the previously established minigenome system as a basis, the secreted *Gaussia* luciferase previously used was exchanged for *Renillia* luciferase due to its suitability for a VLP system. Additionally, passage of the VLP system through cells expressing N, the RdRp and GPs allows propagation of the VLP stocks alongside examination of genome replication, secondary transcription, particle assembly and release. The VLP system was employed in this study to identify the endonuclease domain of the CCHFV RdRp involved in cap snatching for mRNA during mRNA transcription. Substitution of D693 to an inactive alanine prevented the synthesis of mRNA, whilst still permitting genome replication and assembly of RNPs (Devignot *et al.*, 2015).

A follow up study describing an improved VLP system in 2015 built upon the original CCHFV VLP system, but incorporated NanoLuc as the reporter and utilised codon-optimised support plasmids for the RdRp and GPs (Zivcec *et al.*, 2015). Rather than try to gain knowledge of CCHFV replication, this study utilised the VLP system to screen antiviral drugs and neutralising antibodies in a high-throughput manner. Neutralisation of CCHFV VLPs by mAbs could differentiate strain specific and broad acting mAbs with varying effectiveness, whilst small compound inhibitors such as ribavirin and chloroquine demonstrated more consistent inhibition between the two compounds (Zivcec *et al.*, 2015). The ability to screen CCHFV inhibitors for multiple stages of virus replication, without the use of high containment facilities, is an obvious advantage of the VLP system.

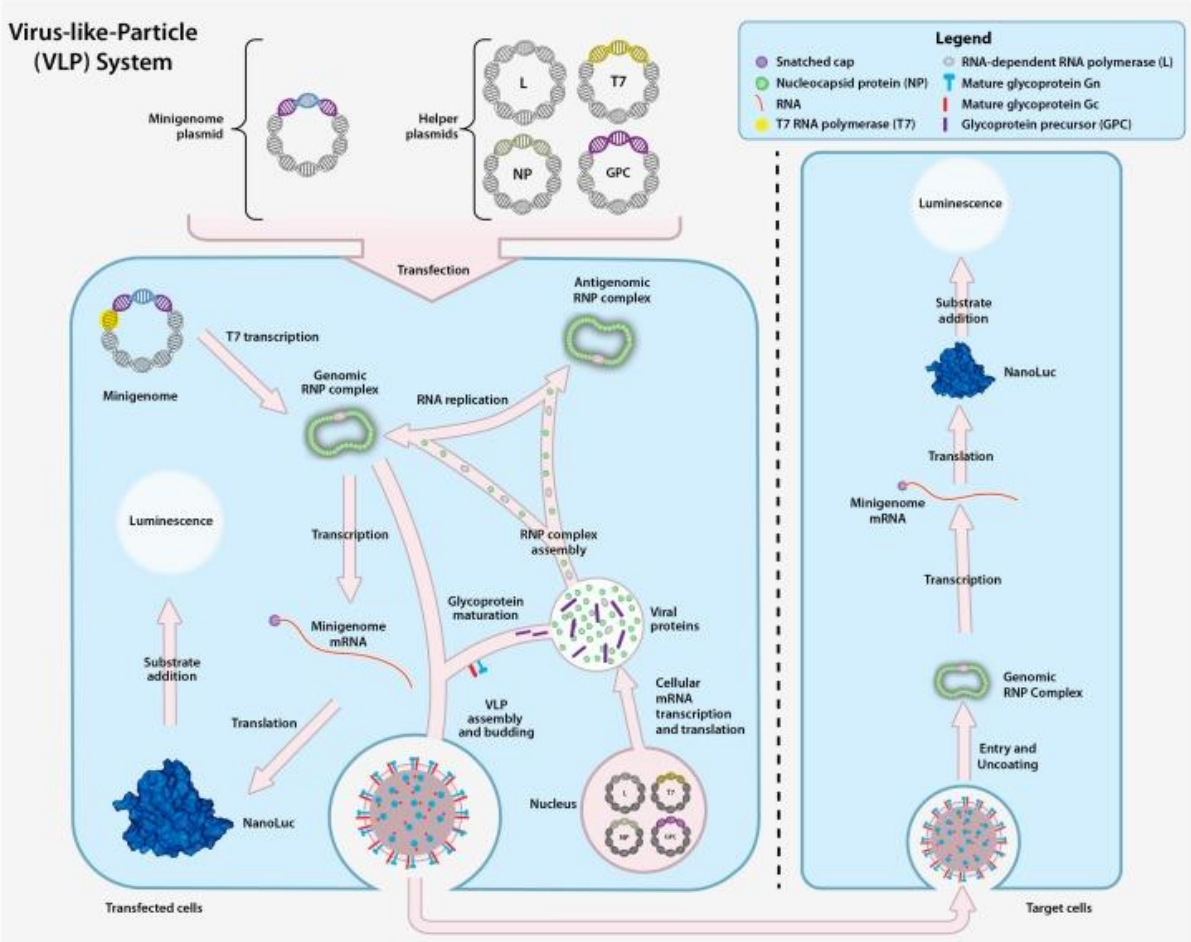


Figure 1.9, CCHFV virus-like particle system schematic

The CCHFV virus-like particle system is composed of the minigenome plasmid, encoding a reporter flanked by non-translated regions (NTRs) from either S, M or L segments, and four support ("helper") plasmids. The support plasmids encode the T7 RNA polymerase and viral nucleoprotein, RNA dependent RNA polymerase (RdRp) and glycoprotein precursor (GPC) downstream of an RNA pol II promoter. These viral proteins and T7 RNA polymerase drive transcription of vRNA from the minigenome plasmid, which is subsequently encapsidated in N and a copy of the RdRp to form genomic RNPs. Translation of mRNAs expressed by support plasmids yields the required proteins for packaging genomic RNPs into VLPs capable of infecting recipient cells. Following infection, the genomic RNPs are released into the cytoplasm and express the minigenome, resulting in generation of a luminescent signal from the encoded reporter.

Figure adapted from (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4848600/)

1.5.3 Infectious clone

The CCHFV infectious clone system permits the recovery of infectious recombinant CCHFV (rCCHFV), meaning all precautions and high containment facilities associated with working with CCHFV are required. The infectious clone system represents the most comprehensive model available for study of the viral replication cycle, as the produced recombinant virus must minimally complete all

the basic stages of CCHFV replication for successful recovery of virus. However, introduction of a mutation that prevents completion of any critical stage will prevent recovery of the virus, with no indication as to which stage the mutation interrupted. In these cases, the systems described in sections 1.5.1 and 1.5.2 can be used to elucidate the effects on individual stages.

rCCHFV was recovered in 2015, via transfection of plasmids expressing positive sense RNAs associated with the full length S, M and L segments in combination with codon optimised expression plasmids for N and the RdRp (Bergeron *et al.*, 2015). The system was used to explore non-structural glycoprotein processing during CCHFV infection by substitution mutagenesis of the furin-like proprotein convertase RSKR cleavage site to a cleavage resistant ASKA motif. Increased WT rCCHFV production was observed in furin expressing cells over the cleavage deficient ASKA mutant. This suggested that whilst furin cleavage at the RSKR was not essential for rCCHFV recovery, it enhanced virion production (Bergeron *et al.*, 2015). The ability to recover viruses from cloned complementary DNA allows such precisely determined genomic sequences to be incorporated into infectious virions with relative ease. Whilst this technology exists for CCHFV, there is no similar system for any member of the *Nairoviridae* that can be used at BSL-2 conditions, including HAZV.

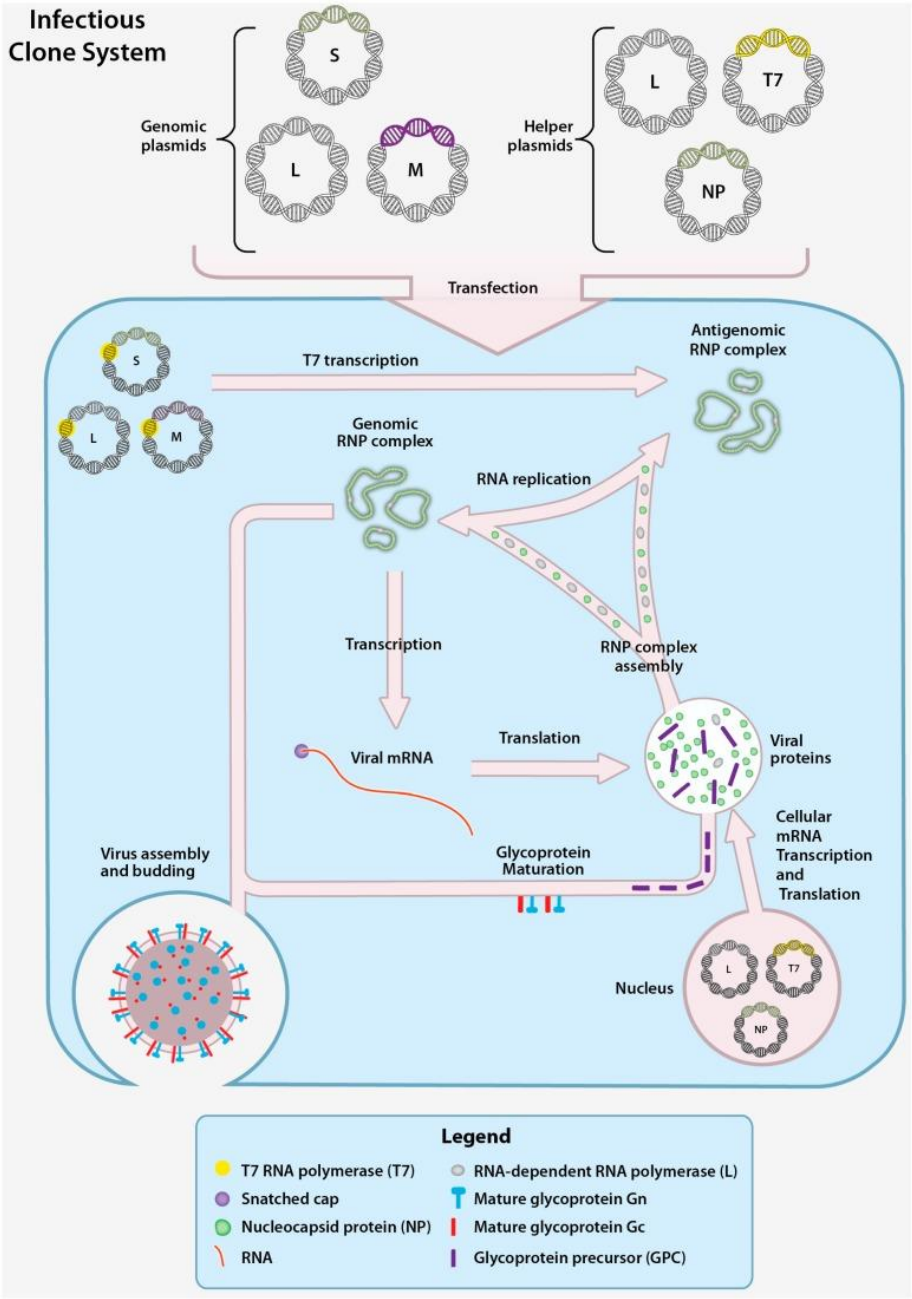


Figure 1.10, CCHFV infectious clone system schematic

The CCHFV infectious clone system comprises of 3 plasmids, each containing the entire sequence from one of the S, M and L genome segments and 3 support ("helper") plasmids encoding the T7 RNA polymerase, viral nucleoprotein and RNA dependent RNA polymerase (RdRp) downstream of a T7 RNA pol II promoter. Following transfection of all plasmids, T7 polymerase drives transcription of the genomic plasmid generating cRNA which is subsequently encapsidated by viral proteins, encoded on the support plasmids, into antigenomic RNPs. Antigenomic RNPs form a template for replication of genomic RNPs which assemble with viral proteins to form infectious virions.

Figure adapted from (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4848600/)

1.6 Apoptotic response to HAZV / CCHFV infection

Apoptosis forms a key part of the immune response against a variety of physiological and pathophysiological stimuli, including viral infection. It permits regulation of the immune response by deletion of self-killing immune cells through cytotoxic killing (Ekert and Vaux, 1997). Two distinct apoptotic pathways have been described, the intrinsic pathway, which is activated by intracellular stimuli and the extrinsic pathway, which is activated by extracellular stimuli. Both intrinsic and extrinsic pathways ultimately result in disruption of the mitochondrial membrane, permitting the release of pro-apoptotic proteins such as cytochrome c (Liu *et al.*, 1996). Release of cytochrome c, alongside others, leads to activation of caspases, a family of cysteine proteases. Two groups of caspases exist initially as inactive pro-caspases, the initiator pro-caspases 8, 9 and 10, which function to activate the second family of executioner pro-caspases 3, 6 and 7 through cleavage. Activation of executioner caspases triggers the execution phase of apoptosis, which results in cell death as a result of cleavage of cellular substrates such as PARP by caspase 3. PARP plays a critical role in a number of essential cellular processes such as DNA repair, transcription and replication (Morales *et al.*, 2014). Therefore, cleavage of PARP by caspase 3 can act as a useful biomarker for apoptotic activity through generation of unique PARP fragments.

Viruses must employ mechanisms to subvert or delay the apoptotic response to complete their replication cycles. The baculovirus p35 protein, following a cleavage event, is able to bind irreversibly to the active site of caspases, preventing further progress of the apoptotic pathway (Zhou *et al.*, 1998). Within the Bunyavirales, LACV and BUNV also regulate apoptosis at an early stage through expression of NSs (Kohl *et al.*, 2003; Blakqori *et al.*, 2007). CCHFV infection has previously been reported to induce apoptosis in mammalian cell lines, with detectable activation of the executioner caspase 3 (Karlberg, Tan and Mirazimi, 2011). Furthermore, caspase mediated cleavage of CCHFV N has been shown to occur at a conserved DEVD motif on the apex of the arm domain by multiple groups (Karlberg, Tan and Mirazimi, 2011; Carter *et al.*, 2012; Wang *et al.*, 2012). Expression of active caspase 3 and detection of DNA fragmentation can be detected 48 hours post-infection with CCHFV, with levels rising to 72 hours at which point visible cytopathic effect (CPE) was observed. Whilst the

exact mechanism in which CCHFV triggers apoptosis is unclear, infection with UV-inactivated CCHFV failed to illicit an apoptotic response, suggesting induction was replication dependent. CCHFV induced ER stress, as indicated by activation of the unfolded protein response, may also be responsible for activation of the intrinsic apoptotic response (Rodrigues *et al.*, 2012). It is likely that CCHFV infection induces apoptosis through multiple pathways and therefore must be able to modulate these induced pathways in order to complete the replication cycle prior to cell death.

1.7 Project aims

Reverse genetics systems, such as those involving the generation of an infectious clone, facilitate and accelerate research of the target organism. Previously, a rescue system for the highly pathogenic CCHFV was established, though due to the high containment facilities required, the usefulness of such a system is limited. Such a system in HAZV would prove invaluable to nairovirus research, with no current infectious clone systems available for a BSL-2 member of this family. Therefore, the initial aim of the project was to develop an infectious clone system for HAZV, to act as a model system for CCHFV.

Following development of an infectious clone system for HAZV, the aim was to utilise the system to increase knowledge and understanding surrounding nairoviral replication. Previous infectious clone systems have incorporated reporter genes to facilitate live-cell imaging, which could prove a useful means of screening potential therapeutics. The ability of the infectious clone system to identify critical residues involved in virus replication, due to their inability to be recovered following knockout mutagenesis, would also provide valuable insight into the roles of the viral proteins in virus replication. In this sense, a further aim of the project was to utilise the infectious clone system to further our understanding of the HAZV N interaction with caspases. By creation of a panel of mutants in which the caspase site was altered, the requirement of the DQVD motif, and additional caspase motifs present on HAZV N, during viral replication can be assessed.

The final aim of the project was to design and subsequently validate the development of a reporter strain of HAZV expressing eGFP. As with the WT infectious clone system, no such reporter system exists for a BSL-2 nairovirus member and would prove invaluable in high-throughput screening, such as antivirals. In regards to this project, the system was designed in order to screen a large panel of host factors involved in entry, in order to identify those utilised by HAZV.

In summary, the aims of this project were to develop an infectious clone system for HAZV, then modify it to increase our understanding of HAZV replication, via mutagenesis of conserved regions of interest, or gain of function mutations such as incorporation of reporter genes.

2 Materials and methods

2.1 Materials

2.1.1 Vectors

Full-length cDNAs representing the S, M, and L segments were synthesized by R. A. Surtees (Genewiz) using the HAZV strain JC280 (GenBank accession numbers M86624.1, DQ813514.1, and DQ076419.1, respectively) as reference and were incorporated into the pMK-RQ plasmid, resulting in the generation of pMK-RQ-S, pMK-RQ-M, and pMK-RQ-L, respectively, able to express S, M, and L segment-specific RNAs. Modifications were performed upon receipt as per section 2.2.1 to remove an additional unrequired promoter. The T7 polymerase expressing pCAG-T7pol (Addgene plasmid number 59926) was a gift from Ian Wickersham. The eGFP-P2A containing pUC57-Kan-eGFP-P2A was synthesized and purchased (Genewiz), prior to insertion into pMK-RQ-S. Sequences for all plasmids can be found in appendix 1.1.

2.1.2 Bacterial cell lines

Plasmid DNA constructs were amplified through competent *Escherichia coli* (*E. coli*) cells. Strains used were either DH5alpha (New England Biolabs) for work involving pMK-RQ-S or JM109 cells (Promega) for work involving pMK-RQ-M and pMK-RQ-L.

2.1.3 Mammalian cell lines

SW13 cells, derived from a human adrenal cortex carcinoma line, were used in experimental procedures and for propagation of viral stocks. BSR-T7 cells, derived from a baby hamster kidney cell line that constitutively express T7 polymerase, were used in transfections. A549 cells, derived from the human alveolar basal epithelium, were used in experimental procedures. Huh7 cells, derived from a human hepatocellular carcinoma line, we used in experimental procedures.

2.1.4 Tick cell lines

HAE/CTVM9 cells, derived from the tick *Hyalomma anatolicum*, were used in experimental procedures and were a kind gift from Dr Lesley Bell-Sakyi, The University of Liverpool.

2.1.5 Hazara virus strain

The HAZV strain JC280 used in this project was supplied as an infectious cell culture supernatant from Dr Emma Punch. Sequences of the cRNAs associated with this strain for the S, M and L segments are available from GeneBank with the respective accession numbers; M86624.1, DQ813514.1, DQ076419.1

2.2 General methods

2.2.1 Manipulation of recombinant DNA

2.2.1.1 Bacterial transformations

For DH5alpha cells, both plasmids and cells were thawed on ice immediately prior to transformation. 5 ng of plasmid was added to 50 µl of DH5alpha cells and incubated on ice for 30 minutes. The transformation mixture was then heat shocked at 42 °C for 30 seconds. Eppendorf tubes were then cooled on ice for 2 minutes before 450 µl of 37 °C SOC media was added to each reaction. The transformation mixture was then incubated at 37 °C for 1 hour with shaking. The mixture was then split into 20 µl and 100 µl aliquots to spread on antibiotic selective kanamycin / ampicillin LB media plates and incubated overnight at 37 °C.

For JM109 cells, both plasmid and cells were thawed on ice immediately prior to transformation. 100 ng of plasmid DNA was added to 65 µl of JM109 cells and incubated on ice for 30 minutes. The transformation mixture was then heat shocked at 42 ºC for 30 seconds. Eppendorf's were then cooled on ice for 2 minutes before 300 µl of room temperature LB broth was added to each reaction. The transformation mixture was then incubated at 30 ºC for 1 hour with shaking. The mixture was then split into 300 µl and 50 µl aliquots to spread on antibiotic selective kanamycin / ampicillin LB media plates. Plates were then incubated overnight at 30 ºC.

2.2.1.2 Starter cultures

Single colonies were picked from antibiotic selective LB plates and used to inoculate 5 mL LB broth containing the appropriate antibiotic (kanamycin 50 µg/mL or ampicillin 50 µg/mL). Cultures were incubated at 30°C for 8 hours with shaking. These were then scaled up to either 10 mL (mini-prep), 50 mL (midi-prep) or 100 mL (maxi-prep) cultures and incubated overnight at 30°C or 37°C for JM109 and DH5alpha cell lines respectively.

2.2.1.3 Plasmid DNA amplification

Overnight cultures of 100 mL LB with kanamycin at a concentration of 50 µg/mL were pelleted by centrifugation at 5000 rpm for 15 minutes. A Plasmid Miniprep, Midiprep or Maxiprep kit (Qiagen) was used to isolate the plasmid DNA from the pelleted bacteria according to the manufacturer's protocol, where isolation of plasmid DNA is based on the alkaline lysis method followed by isopropanol precipitation. Concentration of plasmid DNA was identified following elution in ddH_2O using a NanoDrop 1000 (Thermo Scientific) by spectrophotometry.

2.2.1.4 The polymerase chain reaction (PCR)

Following first strand synthesis of extracted viral RNA, sequences of interest were amplified by PCR to permit sequencing confirmation of mutations in recovered virus. PCR reactions were performed in a 25 µL reaction volume using the Q5 High-Fidelity 2X Master Mix (New England Biolabs), containing 25 ng template DNA, and 0.5 µM forward and reverse primers. Reaction cycles were carried out as follows; initial denaturation at 98°C for 5 minutes, 35 cycles of denaturation at 98°C for 10 seconds, primer annealing at 50-72°C (primer specific) for 30 seconds, extension at 72°C for 30 s/kb, then a final extension at 72°C for 2 minutes. PCR products were then purified using the Monarch® PCR & DNA Cleanup Kit (New England Biolabs) according to manufacturers' instructions.

2.2.1.5 Reverse transcription quantitative PCR (RT-qPCR)

RNA was harvested from cells of interest via Trizol extraction. Briefly, cells were washed in nuclease free phosphate buffered saline (PBS) then resuspended in

TRI reagent™ (Invitrogen), chloroform was then added, mixed vigorously and left at room temperature for 2 minutes. Samples were centrifuged at 12,000 g and the resulting aqueous phase was collected and added to isopropanol to precipitate RNA for 5 minutes at room temperature. Samples were spun at 12,000 g and the resulting pellet was washed in ice cold 75 % EtOH prior to resuspension in nuclease free H_2O. qPCR was carried out using the One step MESA GREEN qRT-PCR MasterMix for SYBR® Assay (Eurogentec) according to the manufacturer's instructions, with samples normalised to GAPDH expression. Primer sequences used can be found in Table 2.1.

2.2.1.6 Agarose gel electrophoresis

DNA was examined for quality and purity, or purified from 1 % agarose gels composed of 0.5 g agarose, 50 mL of 1 x TAE buffer (40 mM Tris-acetate, 1 mM ethylene-diamine-tetraacetic acid (EDTA)) and SYBR Safe DNA stain (Life Technologies) diluted 1:10,000. Samples were mixed with 6 X DNA loading dye (New England Biolabs) and ran alongside the Quick-load 2-Log DNA Ladder (New England Biolabs) at 110 V for 1 hour in 1 X TAE buffer

2.2.1.7 Site directed mutagenesis

Generation of mutant plasmids incorporating deletions, small insertions and/or substitutions was achieved using the Q5 site-directed mutagenesis kit (New England Biolabs), according to the manufacturer's instructions, with all mutant plasmid sequences confirmed via sequencing (Genewiz). Primers used in generation of mutant sequences can be found in appendix 1.2.

2.2.1.8 Restriction digests

Restriction digests were carried out using the indicated restriction enzymes (New England Biolabs) using the manufacturer's instructions. Briefly, 1 µg of DNA was digested using 1 µL of restriction enzyme in a 50 µL total volume at 37°C for 1 hour. Restriction digestion products were then purified via agarose gel electrophoresis followed by gel extraction using the NEB Gel extraction kit (New England Biolabs) as per the manufacturer's instructions.

2.2.1.9 Ligations

Ligations were carried out in 20 µL total reaction volumes containing: 3:1 molar ratio of insert: vector DNA, 1 x Ligase Reaction Buffer (30 mM Tris-HCl (pH 7.8), 10 mM MgCl2, 10 mM DTT and 1 mM ATP; Promega) and 1 unit of T4 DNA Ligase (New England Biolabs). Ligase reactions were incubated at room temperature for 1 hour or 4 °C overnight, then transformed into DH5alpha cells.

2.2.2 Cell culture methods

2.2.2.1 Mammalian cell culture

SW13, A549 and Huh7 cells were all maintained at 37°C with 5 % CO_2 in humidified incubators with Dulbecco's modified Eagle's medium (DMEM, Sigma-Aldrich) supplemented with 10 % foetal bovine serum (FBS, Invitrogen), 100 IU/mL penicillin and 100 µg/mL streptomycin. Cell lines were passaged every 2 to 3 days using trypsin-EDTA (Sigma) based on the level of confluency.

2.2.2.2 Arthropod cell culture

HAE/CTVM9 cells were maintained in L15 medium containing 20 % FBS, 10 % tryptose phosphate broth, 2 mM L-glutamine, 100 IU/mL penicillin and 100 µg/mL streptomycin at 30 °C and passaged as required every fortnight.

Primer	Direction	Sequence (5’ to 3’)	Supplier
HAZV S segment	Forward	CAA GGC AAG CAT TGC CAC AC	IDT[1]
HAZV S segment	Reverse	GCT TTC TCT CAC CCC TTT TAG GA	IDT
GAPDH	Forward	TGT GGT CAT GAG TCC TTC CAC GAT	Sigma Aldrich
GAPDH	Reverse	AGG GTC ATC ATC TCT GCC CCC TC	Sigma Aldrich
COPA	Forward	CCA CTA TCA GAA TGC CCT ATA CC	IDT
COPA	Reverse	CCA CAA ACC CAT CTT CAT CC	IDT
COPB1	Forward	ACA GAGA GAA AGA GGC AGC AGA	IDT
COPB1	Reverse	GCA AGG TATA CAC TGG TTT GGT TC	IDT
COPB2	Forward	GTG GGG ACA AGC CAT ACC TC	IDT
COPB2	Reverse	GTG CTC TCA AGC CGG TAG G	IDT

Table 2.1, List of primers and associated sequences used in qPCR

Table displays qPCR primers used during the project, detailing their sequences and supplier. [1]Integrated DNA technologies.

2.2.3 Transfections

2.2.3.1 TransIT-LT1 mediated transfection of BSR-T7 cells with a single plasmid

Twelve-well plates were seeded with $1x10^5$ BSR-T7 cells/well 1 day prior to transfection in 1 mL DMEM supplemented with 2.5 % FBS. 24 hours later, cells were transfected with 1 μg of the indicated cDNA and 2.5 μl Mirus TransIT-LT1 (Mirus Bio) transfection reagent per μg DNA in 200 μL Opti-MEM (Life Technologies) following a 20 minute incubation period at room temperature. Cells were then incubated at 37 °C for 24 hours prior to collection of total cell lysate for analysis.

2.2.3.2 TransIT-LT1 mediated transfection of BSR-T7 cells with WT rHAZV cDNAs

Six-well plates were seeded with $2x10^5$ BSR-T7 cells/well 1 day prior to transfection,

in 2 mL DMEM supplemented with 2.5 % FBS, 100 U/mL penicillin and 100 μg/mL streptomycin. 16 to 24 hours later, cells were transfected with a transfection mixture containing 1.2 μg pMK-RQ-S, pMK-RQ-M, and pMK-RQ-L and 0.6 μg pCAG-T7pol, 2.5 μl Mirus TransIT-LT1 transfection reagent per microgram of DNA and 200 μl Opti-MEM. This transfection mixture was applied to cells following a 20 minute incubation of the mixture at room temperature. At 24 hours post transfection media containing the transfection mix was removed and replaced with fresh DMEM supplemented with 2.5 % FBS, 100 U/mL penicillin and 100 μg/mL streptomycin. A control sample, in which transfection of pMK-RQ-L was omitted, was set up alongside each experiment as a negative control. Supernatant was harvested at 72, 96 and 120 hours post transfection and examined for presence of virus.

2.2.3.3 TransIT-LT1 mediated transfection of BSR-T7 cells with mutant rHAZV cDNAs

For mutant recovery, the same transfection protocol was followed as described in section 2.2.3.2, with the mutated cDNA used in the place of the

corresponding WT cDNA. In addition to the negative control omitting pMK-RQ-L, a full WT rHAZV recovery was carried out alongside each mutant recovery to provide a positive control.

2.2.3.4 Reverse transfection of Silencer™ Human Membrane Trafficking siRNA library

The Silencer™ Human Membrane Trafficking siRNA Library (Thermo Fisher Scientific, 0.25 nmol) was initially prepared prior to use. 96 well plates containing stock siRNAs were briefly centrifuged, reconstituted in 125 μL nuclease free H_2O and mixed via thorough pipetting to generate a stock plate of 2 μM. Working stocks were prepared via combining 50 μL of the stock plate with 50 μL nuclease free H_2O, diluting the stock 2 fold, in a 96 well plate, to achieve a working stock concentration of 1 μM (1 pmol / μL). Plates were stored at -80 °C for long term storage and number of freeze thaw cycles was recorded.

Reverse transfection of the siRNAs was carried out as follows; trypsinised SW13 cells in DMEM containing 2.5 % FBS and 100 U/mL penicillin and 100 μg/mL streptomycin were counted using a hemocytometer and used to make a cell suspension containing 1 x 10^5 cells/mL. A master mix containing 0.3 μL Lipofectamine® RNAiMAX reagent (Invitrogen) and 16.7 μL Opti-MEM per well as required was created and 17 μL of this master mix was pipetted into each well of a 96-well plate. 3 μL of working stock siRNA (1 μM) was pipetted into the transfection master mix and gently pipetted up and down to ensure good distribution and final concentration of 3 pmol of siRNA per well. 100 μL of the 1 x 10^5 cell suspension was then applied per well, to give a final count of 1 x 10^4 cells per well. Cells were incubated with the transfection mix for 24 hours at 37°C, then 60 μL of the media was removed and replaced with 140 μL of fresh DMEM containing 2.5 % FBS, 100 U/mL penicillin and 100 μg/mL streptomycin to dilute out any potential toxic effects of the siRNAs or transfection reagent. 6 hours post-dilution, media was removed and cells were washed in PBS prior to infection with rHAZV(eGFP) at a MOI of 0.25 in 100

μL of DMEM containing 2.5 % FBS, 100 U/mL penicillin and 100 μg/mL streptomycin.

2.2.3.5 Reverse transfection of COPI specific siRNAs

To validate the knockdown observed in the siRNA screen, siRNAs specific to COPA, COPB1 and COPB2 were purchased (Ambion) and reverse transfected into SW13 cells. Briefly, 75 pmol of siRNA was mixed with 4.5 μL Lipofectamine® RNAiMAX in 200 μL Opti-MEM and incubated at room temperature for 20 minutes. Following incubation, the transfection mix was added dropwise to a 6 well plate and overlaid with 2 mL of cell suspension containing 1 x 10^5 cells/mL in 2.5 % DMEM. 24 hours post transfection, media was changed to fresh 2.5 % DMEM for 6 hours then infected with rHAZV at a MOI of 0.1. 24 hours post infection lysates were collected for analysis via western blot or qPCR.

2.2.4 Analysis of protein expression

2.2.4.1 Preparation of whole cell lysate

For preparation of cell lysates, monolayers were washed twice in ice-cold PBS, followed by incubation in ice-cold RIPA buffer (150 mM sodium chloride, 1.0 % NP-40 alternative, 0.1 % SDS, 50 mM Tris [pH 8.0]) and agitated for 120 s. Cells were then harvested via cell scraping and were transferred to prechilled Eppendorf tubes, after which lysates were centrifuged at 20,000 X g for 15 min to pellet insoluble material. Supernatant was then collected and stored at -20 °C prior to analysis

2.2.4.2 SDS polyacrylamide gel electrophoresis (SDS PAGE)

Proteins were separated on a 5 mL 12 % SDS-polyacrylamide resolving gel comprised of 4 mL 30 % bis-acrylamide, 2.5 mL 1.5 M Tris-HCl pH 8.8, 3.3 mL ddH_2O, 100 μL 10 % SDS, 100 μL 10 % ammonium persulphate (APS), 10 μL TEMED and a 5 % stacking gel containing 0.83 mL 30 % bis-acrylamide, 0.63 mL 1M Tris-HCl pH 6.8, 50 μL 10 % SDS, 50 μL 10 % APS and 5 μL TEMED. Protein samples were mixed 1:1 with 2 X denaturing sample buffer (60 mM Tris pH 6.8, 25 % (v/v) glycerol, 2 % (w/v) SDS, 5 % β-mercaptoethanol, 0.01

% (w/v) bromophenol blue) supplemented with 50 mM DTT, then denatured via heating to 95°C for 10 minutes immediately prior to loading. Samples were then loaded into gels alongside the Color Prestained Protein Standard, Broad Range (11-245 kDa) (New England Biolabs) ladder and electrophoresis was performed in SDS running buffer (25 mM Tris, 192 mM glycine, 0.1 % (w/v) SDS) at 180 V for 1 hour.

2.2.4.3 Western blot analysis

Proteins were transferred from SDS PAGE gels (section 2.2.4.2) to fluorescence-compatible polyvinylidene difluoride (FL-PVDF) membranes (Millipore) using a Trans-Blot semi-dry cell (Bio-Rad) in Towbin buffer (25 mM Tris, 192 mM glycine, 20 % (v/v) methanol) for 1 hour at 15 V. FL-PVDF membranes were blocked in 1:1 Odyssey® Blocking Buffer in TBS, TBS (50 mM Tris-HCl pH 7.5, 150 mM NaCl) for 1 hour at room temperature with gentle rocking. Blocking buffer was then replaced with primary antibody diluted to the appropriate concentration in 1:1 Odyssey® Blocking Buffer in TBS, TBS-T (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.1 % Tween 20) for 1 hour at room temperature with gentle rocking. A full list of primary antibodies and associated concentrations used to detect immobilised proteins is provided in Table 2.2.

FL-PVDF membranes were extensively washed in TBS, then incubated with the appropriate fluorescently labelled secondary antibody at a 1:10,000 dilution under the same conditions as for the primary antibody incubations. FL-PVDF membranes were washed three times in TBS-T, once in water, then fluorescently labelled secondary antibodies were directly visualised using a LiCor Odyssey Sa Infrared imaging system (LiCor).

Target	Species	Supplier	Dilution
Actin	Mouse	Sigma Aldrich	1:10000
Caspase 3	Rabbit	Cell Signaling Technologies	1:1000
Caspase 7	Rabbit	Cell Signaling Technologies	1:1000
Caspase 9	Mouse	Cell Signaling Technologies	1:1000
Cleaved caspase 3	Rabbit	Cell Signaling Technologies	1:1000
Cleaved caspase 7	Rabbit	Cell Signaling Technologies	1:1000
Cleaved caspase 9	Rabbit	Cell Signaling Technologies	1:1000
Cleaved Poly ADP-ribose polymerase	Rabbit	Cell Signaling Technologies	1:1000
GAPDH	Mouse	GeneTex	1:10000
Enhanced Green fluorescent protein (eGFP)	Rabbit	BioVision	1:1000
Hazara nucleoprotein (HAZV N)	Sheep	In-house[1]	1:3000
Poly ADP-ribose polymerase	Rabbit	Cell Signaling Technologies	1:1000

Table 2.2, List of primary antibodies used during the project

Table displays antibodies used during the project, detailing their targets, host species, supplier and dilution used at during experimentation. [1]HAZV N was previously generated in-house by R. A. Surtees (Surtees *et al.*, 2015).

2.2.4.4 Immunofluorescence microscopy

Trypsinised SW13 cells were seeded onto 16 mm round glass coverslips (VWR) in a 12-well plate at 5 x 10^4 cells/well and incubated at 37°C. 16-24 hours later, cells were infected at a MOI of 0.1 in DMEM containing 100 U/mL penicillin and 100 μg/mL streptomycin for 1 hour at 37°C. Following infection, media was removed and cells were washed in PBS, DMEM containing 2.5 % FBS, 100 U/mL penicillin and 100 μg/mL streptomycin was then applied and returned to incubate at 37°C for a further 24 hours. At this point media was removed and cells were washed twice in PBS prior to fixation in 4 % paraformaldehyde in PBS (v/v) for 10 minutes at room temperature. Following fixation, cells were washed twice in PBS and then incubated in permeabilization buffer (0.1 % Triton-X100 [v/v], 1 % BSA [w/v] in 1X PBS) for 15 minutes at room temperature. Cells were then blocked in blocking buffer (1 % BSA [w/v] in 1X PBS) for 45 minutes then incubated with the required primary antibody at a specified dilution for 1 hour at room temperature. Cells were washed three times in PBS then incubated with the corresponding secondary AlexaFluor® antibody (Life Technologies, 1:500 in blocking buffer) for 1 hour at room temperature in a light protected vessel then washed three times with PBS. Coverslips were then mounted onto glass slides with the addition of ProLong Gold Antifade Reagent with DAPI (Thermo Fisher Scientific), sealed and stored at 4°C. Images were then taken on the LSM 700 or LSM 880 confocal microscope and processed using Zen (Blue Edition) software.

2.2.4.5 Detection of fluorescent rHAZV

The IncuCyte® Live Cell Imaging System (Essen BioSciences) was used to detect infection by rHAZV(eGFP) or rHAZV(eGFP-LFAA). SW13 cells were infected as described in section 2.2.5.1 and imaged at indicated time points to permit analysis of total green cell count or total integrated intensity of eGFP (TIIE) over time. Captured images were analysed using the associated software for fluorescent protein expression. The fluorescent thresholds for analysis were determined using control wells and were used to identify fluorescent positive objects from background fluorescence

2.2.5 Virological techniques

2.2.5.1 Virus infections

SW13, A549, BHK or Huh7 monolayers were infected with HAZV at the specified multiplicity of infection (MOI) in serum-free DMEM (SFM) at 37 °C. After 1 hour, the inoculum was removed and cells washed in PBS, fresh DMEM containing 2.5 % FBS, 100 U/mL penicillin and 100 μg/mL streptomycin was then applied for the duration of the infection.

Infections of HAE/CTVM9 cells of tick origin were carried out at a MOI of 0.01 in a similar manner to mammalian cell infections, but with an incubation temperature of 30 °C and using L15 media containing 20 % FBS, 10 % tryptose phosphate broth, 2 mM L-glutamine, 100 IU/mL penicillin and 100 μg/mL streptomycin.

During recovery of recombinant virus, reinfections were carried out as follows; Six-well plates were seeded with $2x10^5$ SW13 cells/well, 1 day prior to infection, in 2mL DMEM supplemented with 10 % FBS, 100 U/mL penicillin and 100 μg/mL streptomycin. Cell supernatants from transfected BSR-T7 cells (section 2.2.3) were collected 72, 96, and 120 hours post transfection, and 300 μL supernatant was passaged for 48 hours in a 6 well plate of SW13 cells grown in DMEM supplemented with 10 % FBS, 100 U/mL penicillin and 100 μg/mL streptomycin.

2.2.5.2 Virus growth curves

To assess viral fitness over time, $2x10^6$ SW13 cells were seeded into 75 cm^2 flasks one day prior to infection in 8 mL DMEM supplemented with 10 % FBS. 24 hours later, virus was used to infect flasks at the specified MOI in 5 mL SFM for 1 hour with shaking, an aliquot of the infection media was collected to permit back-titration, ensuring equal MOI comparisons were made. Following infection, media containing virus was removed, cells were washed with PBS and 7 mL DMEM supplemented with 2.5 % FBS was reapplied and returned to incubated at 37 °C for the remainder of the experiment. At 24 hour intervals, 200 μL supernatant was removed and stored at -80 °C until all samples had been collected and stored in a similar manner. Following collection of all timepoints, samples were analysed for titer of infectious virus as described in section 2.2.5.4.

2.2.5.3 Virus purification

For propagation of virus stock, 4 x 175 cm^2 flasks were seeded with $3x10^6$ BHK cells, one day prior to infection in 25 mL DMEM supplemented with 10 % FBS. 24 hours later, WT rHAZV was used to infect flasks at a MOI of 0.001 in 5 mL SFM for 1 hour with shaking. Following infection, media containing virus was removed, cells were washed in PBS and 15 mL DMEM supplemented with 2.5 % FBS was reapplied and flasks were incubated at 37 °C for 72 hours with occasional shaking. Following incubation, supernatant containing virus was removed and clarified at 4500 X g for 30 minutes, passed through a 0.45 μM filter then re-clarified as above. Clarified and filtered supernatant was then loaded above a 20 % sucrose cushion and spun at 150,000 X g for 3 hours. Following centrifugation, supernatant was removed and the tube was inverted to permit air-drying of the pellet prior to overnight resuspension in 1 mM $MgCl_2$ and 1mM $CaCl_2$ at 4 °C with shaking.

2.2.5.4 Viral titration

Determination of virus titers was achieved through plaque assays. Supernatant was collected at the time titration was required and serially diluted by a factor of 10 from neat to create dilutions of 10^{-1} to10^{-6} to infect fresh monolayers of SW13 cells in a 6-well plate. Following infection, medium containing virus was removed, cells were washed in ice cold PBS and a 1:1 ratio of 2.5 % FBS DMEM to 1.6 % methylcellulose (Sigma) was reapplied, and cells were returned to incubate for a further 6 days prior to fixing and staining with crystal violet (13.5 % formaldehyde, 2.5 % crystal violet, 20 % ethanol). Plaques formed where HAZV had replicated and caused CPE and were counted, and virus titers were determined using the equation;

$$pfu\ /\ mL = \frac{Number\ of\ plaques}{Dilution\ factor\ x\ Volume\ (mL)}$$

2.2.5.5 Generation of cDNA from infectious virus

Viral RNA was extracted from virus containing cell-free supernatant using the QIamp viral RNA kit (Qiagen) and subsequently treated with DNase to remove any contaminating DNA. A cDNA copy was generated using the ProtoScript II reverse transcriptase kit (New England Biolabs), according to the manufacturer's instructions. Amplification of the cDNA was carried out using PCR, as described in section 2.2.1.4. To rule out plasmid carryover from transfection media, a no reverse-transcriptase control was also carried out in the same manner as described, with the omission of the reverse transcriptase.

2.2.5.6 rHAZV(eGFP) MOI optimisation

SW13 cells were seeded at $1x10^4$ cells/well in a 96-well plate and incubated at 37°C for 24 hours to replicate conditions generated according to the siRNA reverse transfection protocol. Media was then removed and cells were infected with rHAZV(eGFP) at range of MOIs (0.1, 0.25 and 1.0) in SFM containing 100 U/mL penicillin and 100 μg/mL streptomycin. Cells were incubated for a further 48 hours at 37°C with hourly imaging using the IncuCyte® Live Cell Imaging System to detect the number of eGFP positive cells.

2.2.5.7 Quantifying viral expression levels using eGFP signal

Following a 24 hour infection, GFP fluorescence intensity was determined using the IncuCyte® Live Cell Imaging System as a measure of virus infectivity. Total green integrated intensity (GCU x μm^2/image) was first normalised to confluency per well, then analysed as a percentage of the total green integrated intensity in positive control wells containing virus but omitting siRNA and transfection reagent. Normalised values within each technical repeat were averaged and then averaged across two biological repeats. Statistical analyses were performed using an unpaired, two-tailed Student's *t*-test assuming variance in standard deviation between the means of control and experimental sample groups.

2.2.6 Inhibitor studies

2.2.6.1 Inhibition of retrograde transport

Trypsinised A549 cells were seeded into 12 well plates at $1x10^5$ cells/well and incubated at 37°C. 16 to 24 hours later, cells were pretreated with brefeldin A (BFA) at the indicated concentrations for 45 minutes in SFM prior to infection with rHAZV or rIAV at a MOI of 0.1 for 1 hour at 37°C. Following the infection period, media containing virus was removed and cell monolayers washed three times with PBS. Fresh 2.5 % DMEM was then reapplied containing the indicated concentration of BFA for a further 24 hours. At this point lysates were collected and analysed via western blot (2.2.4.3).

2.2.6.2 Inhibition of caspases

SW13 monolayers were pretreated for 45 min with 20 µM Z-FA-FMK or Z-VAD-FMK in SFM at 37°C prior to infection with HAZV at a MOI of 0.1. Following infection, virus and inhibitor were removed, cells were washed three times in PBS, and 2.5% FBS-DMEM containing 20 µM Z-FA-FMK was reapplied. At 24, 32, and 48 hour time points, total cell lysates were harvested and prepared for western blotting.

2.2.6.3 Viability assays

Trypsinised A549 cells were seeded into 96 well plates at $1x10^4$ cells/well and incubated at 37°C. 16 to 24 hours later, cells were treated with inhibitor at the indicated concentrations for 1 hour 45 minutes in SFM, to replicate the pre-treatment and infection timings. After this period 100 µL of fresh 2.5 % DMEM was then reapplied containing the indicated concentration of inhibitor for a further 24 hours. At this point the CellTitre 96® AQ_{ueous} One Solution Cell Proliferation Assay (Promega) was used to measure cell viability according to the manufacturer's instructions. Briefly, 20 µL was added to each well and returned to incubate for a further 4 hours, at which point absorbance was measured at 490 nm using a plate reader (Tecan) and used to determine viability via comparison to untreated and media only controls.

3 Development of a Hazara virus infectious clone

3.1 General introduction

The ability to perform detailed genetic analysis of any virus is dramatically increased by the provision of a reverse genetics system, which allows any nucleotide or amino acid residue to be specifically altered, allowing any resulting change in phenotype to be identified. Such a system has been developed for CCHFV (Bergeron *et al.*, 2015), which has been exploited to expand the current understanding of nairovirus molecular and cellular biology (Scholte *et al.*, 2017; Welch *et al.*, 2017). However, the classification of CCHFV as a hazard group 4 pathogen presents additional and significant experimental hurdles, such as the requirement of BSL-4 facilities, which significantly increase experimental cost and study times. Therefore, there is an urgent need to generate similar systems in related, lower containment level nairoviruses such as HAZV to provide a platform to study the function of the viral genes and genetic elements. A HAZV infectious clone system would also enable the recovery of so-called 'gain of function' viruses, such as incorporation of reporter genes to facilitate efficient analysis of virus growth kinetics. Any findings of interest discovered in the HAZV infectious clone system could then be replicated in CCHFV, streamlining the discovery process and therefore reducing the length of time required to make significant advances into our knowledge of CCHFV replication.

The first recovery of an infectious RNA virus from cDNA occurred in 1981, with the successful recovery of poliovirus (Racaniello and Baltimore, 1981). Since then, multiple RNA viruses from a broad range of families have had reverse genetics systems developed enabling rapid advances in understanding. Due to the inability of negative sense RNA viruses to have their genomes directly translated into proteins and the requirement of the genome to be associated with viral N to be replication competent, advancement of reverse genetics systems for this group of viruses has been slower, with the first such development published in 1994, for the recovery of infectious rabies virus (Conzelmann and Schnell, 1994; Schnell, Mebatsion and Conzelmann, 1994). In the years since, multiple infectious clone systems have been set up for multiple members of the *Bunyavirales* order across four of the five families; The *Hantaviridae* family are

the only group to have no such system, despite repeated attempts to recover recombinant virus. BUNV, CCHFV and RVFV all have well established infectious clone systems (Bridgen and Elliott, 1996; Bridgen *et al.*, 2001; Bird, Albariño and Nichol, 2007; Bergeron *et al.*, 2015). These systems have enabled discovery of many features surrounding the viral lifecycles, including the previously described processing mechanisms of the GPC of CCHFV in section 1.4.2.1 and BUNV requirement of cellular proteases such as furin and signal peptide peptidase (Bergeron *et al.*, 2015; Shi *et al.*, 2016). However, despite multiple members of the *Bunyavirales* order having infectious clone systems established, no hazard group 2 members of the *Nairoviridae* currently have such systems. Reverse genetics systems provide an invaluable means to study the role of individual nucleotides, amino acids, entire genes and NTRs by enabling targeted mutagenesis, via manipulation of the DNA source. These reverse genetics approaches allow the user to decide on the region for mutation and then identify any resulting phenotypic effect, rather than relying on the forward genetics method of sequencing naturally occurring mutants and working backwards from phenotype to identify the responsible mutations. This dependence of forward genetics on mutations resulting in a viable phenotype limits the study into critical elements of the viral genome.

Recovery of infectious negative sense viruses can be achieved via the use of DNA plasmids expressing positive or negative sense primary RNA transcripts corresponding to viral genome segments. The previously described rescue system for recovery of rCCHFV utilised plasmids expressing positive sense primary RNA transcripts. Upon transfection of a positive sense plasmid into a T7 polymerase expressing cell line, such as BSR-T7 cells, T7 RNA polymerase recognises the T7 promoter positioned upstream of the gene of interest and begins transcription, generating either translatable mRNAs or antigenomic vRNAs from a single transcription reaction. Being directly translatable, the mRNAs are able to drive translation of the viral products. In the case of HAZV, they permit translation of HAZV N and L, which are critical for packaging the antigenomic vRNAs forming antigenomic (coding) RNPs which can act as replication templates to generate genomic (non-coding) RNPs that can subsequently be packaged into maturing virions (Figure 3.1). When mutations

are introduced that remain viable for virus replication, the infectious clone system has the advantage of being able to study the impact of the viable mutation across all aspects of the viral lifecycle. The fundamental drawback occurs when an alteration, such as large insertions, deletions or mutations to critical amino acids, prevent the virus from carrying out any essential part of the viral lifecycle. In cases where this occurs, recovery of infectious virus is not possible and the only conclusion that can be drawn is that the alteration affected a critical domain of the viral genome or prevented the virus from carrying out an essential phase of the lifecycle. In these situations, alternatives, such as the mini-genome and VLP systems, are well suited to the necessary further genetic analysis (Bergeron *et al.*, 2015; Devignot *et al.*, 2015; Matsumoto *et al.*, 2019).

The aims of this chapter were as follows;

1. Generate an infectious clone for rHAZV
2. Compare rHAZV growth to that of the parental virus
3. Demonstrate the ability of the infectious clone to incorporate mutations

Successful completion of these aims would yield a valuable tool for studying HAZV and by extension other members of the *Nairoviridae* family, in a more efficient manner without the need for high level containment facilities. More specifically, it would provide an excellent model for increasing understanding the replication cycle and pathogenicity of CCHFV, in which initial characterization of orthonairovirus properties would be performed using HAZV with selected properties of interest later characterized using CCHFV.

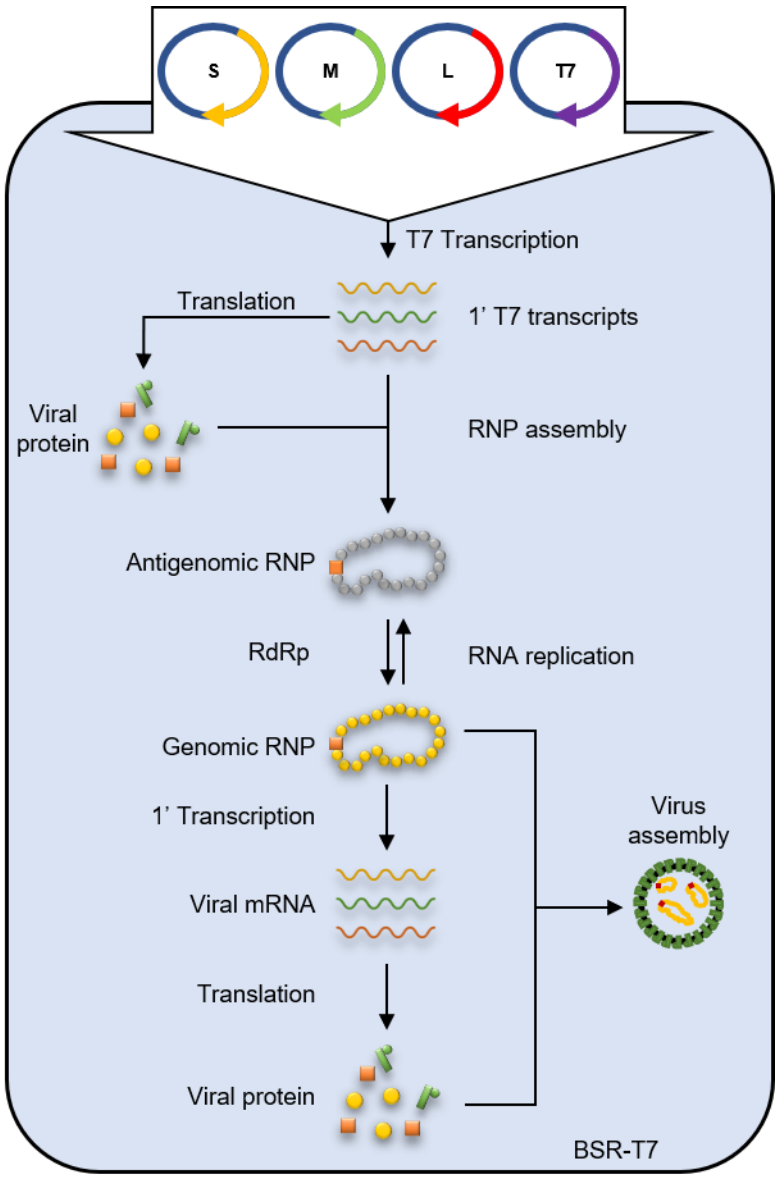


Figure 3.1, Schematic of rHAZV rescue workflow

Plasmid cDNAs representing HAZV segments are transfected into actively replicating BSR-T7 cells. The action of both endogenous and exogenous T7 RNA polymerase drives primary transcription of RNAs that represent S, M and L segment antigenomes, which also act as templates for translation of viral structural and non-structural proteins. Translated viral proteins then associate with primary T7 RNA polymerase transcripts to form positive sense antigenomic RNPs, which are in turn replicated by the HAZV RdRp to form negative sense genome segments. These genomes then act as templates for either mRNA transcription or RNA replication, permitting accumulation of negative sense RNPs and viral proteins, subsequently resulting in assembly of infectious virions which then exit the cell and can be collected in the supernatant.

3.2 Plasmid design

The infectious clone system requires transfection of a copy of the viral genome, carried either on a bacterial plasmid, or on an *in-vitro* transcript derived from it into cultured cells. The use of plasmid DNA in recovery of CCHFV is described

in sections 1.5.3 and 3.1. The use of plasmid DNA in infectious clone systems has multiple advantages over transfection of RNA. Firstly, as the plasmid is continuously transcribed within the cell, more copies are generated from a cDNA plasmid than a finite amount of RNA. Also, the direct use of a cDNA plasmid removes the requirement for *in-vitro* transcription and subsequent purification of RNA, reducing the time and costs involved with infectious clone recovery. Finally, DNA has increased stability over RNA and therefore handling, storage and transfection of plasmid DNA over RNA is made more reliable.

Plasmids for the recovery of HAZV had previously been designed within our group to express positive sense primary viral RNA transcripts corresponding to S (Figure 3.2, a), M (Figure 3.2, b) and L (Figure 3.2, c) segments, with each segment cDNA encoding both viral 5' and 3' NTRs flanking the ORF encoding the associated viral gene. Upstream of the 5' NTR, a bacteriophage T7 RNA polymerase promoter sequence was included to permit transcription within a T7 polymerase expressing cell line, in our case BSR-T7 cells. Directly downstream of the 3' NTR a hepatitis delta virus ribozyme (HDR) sequence and T7 terminator sequence were included to terminate transcription and permit self-cleavage of the resulting RNA product. This HDR cleavage event is a transesterification reaction, which generates products with either a 2', 3'-cyclic phosphate or 5'-hydroxyl termini (Ferré-D'Amaré and Scott, 2010). Therefore, via the use of the HDR in our system, we are able to exactly define the 3' end of our transcribed RNA sequence. A kanamycin resistance gene and origin of replication were both incorporated to permit selection of bacterial colonies following transformations and provide an initiation sequence for replication of dsDNA. The designed inserts for S, M and L segments were then incorporated into the pMK-RQ backbone to generate pMK-RQ-S, pMK-RQ-M and pMK-RQ-L respectively.

Plasmids expressing positive sense primary RNA transcripts, including directly translatable mRNAs, were utilised over those expressing negative sense RNA transcripts, despite the negative sense polarity of the HAZV genome. This is as the uncapped positive sense T7 transcripts generated by T7 polymerase have previously been shown in other bunyavirus recovery systems to be sufficient in initiating viral replication (Lowen *et al.*, 2004; Bergeron *et al.*, 2015). In addition,

should the use of support plasmids be required, utilising plasmids expressing a positive sense RNA corresponding to the viral genome segments eliminates issues arising from dsRNA formation between complementary negative sense RNA genome transcripts and the positive sense RNA support transcripts.

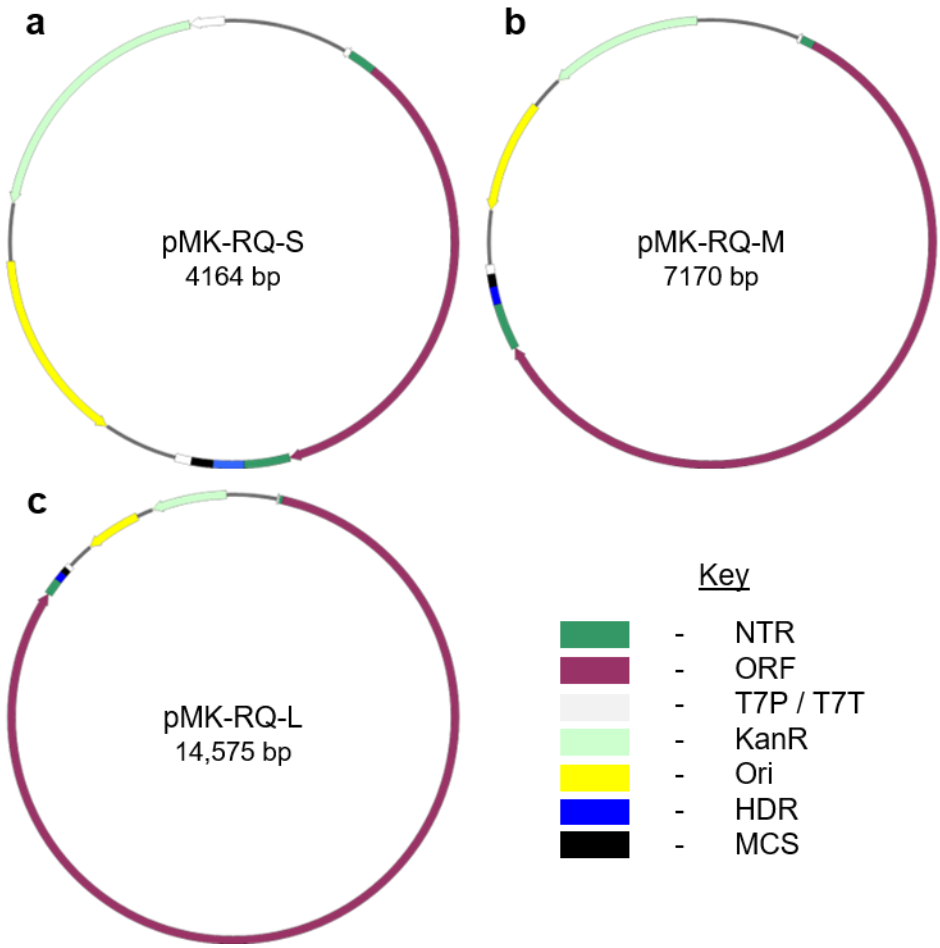


Figure 3.2, Design of cDNAs for recovery of recombinant rHAZV

Schematic representation of cDNA plasmids designed for expression of HAZV S (a), M (b) and L (c) segments. Virus-encoded proteins are flanked by their respective non-translated regions (NTRs). T7 promoter (T7P, T7T) and terminator sequences drive primary transcription in transfected cells, with the hepatitis delta virus ribozyme (HDR) responsible for generating a defined 3' terminus. A kanamycin resistance gene (KanR), origin of replication (Ori) and multiple cloning site (MCS) were included to permit plasmid amplification, cloning and antibiotic selection.

Modification following receipt of the plasmids was carried out to remove an additional unnecessary T7 promoter that was identified in the backbone of each plasmid following sequencing. Presence of this secondary T7 promoter would result in extension of the viral 3' NTR sequence by 51 bases, changing its terminal nucleotide sequence, which would likely have a significantly negative effect on

rescue efficiency. Therefore, primers were designed back to back, flanking the additional promoter region and it was removed via site-directed mutagenesis (SDM). As an initial test to check whether the mutagenesis had the desired outcome, restriction digests were performed on each plasmid. pMK-RQ-S was digested with *Kpn*I and *Pst*I, pMK-RQ-M was digested with *Kpn*I and *Xho*I and pMK-RQ-L was digested with *Xho*I, *Nco*I and *Eco*RI. These digests generate a predicted set of cleavage products, listed in Figure 3.3, that allowed an initial check of SDM success. Following this initial check, the plasmids were fully sequenced, including the backbone, to ensure complete confidence in the source material for the infectious clone system.

a

Plasmid	Enzymes used	Expected Products
pMK-RQ-S	*Kpn*I, *Pst*I	2713, 1451
pMK-RQ-M	*Kpn*I, *Xho*I	4023, 3147
pMK-RQ-L	*Xho*I, *Nco*I, *Eco*RI	9173, 3748, 1654

b

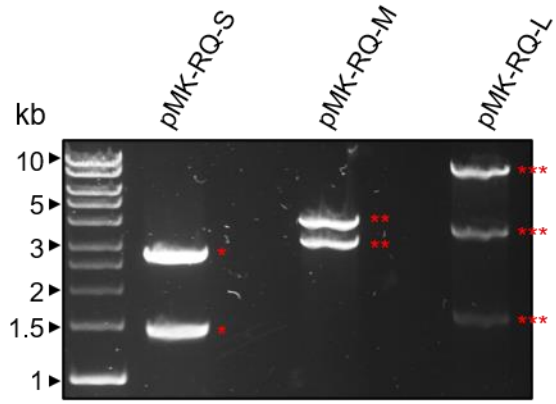


Figure 3.3, Restriction digest confirmation of plasmid identity

a) Expected sizes of products generated following restriction digest of pMK-RQ-S, pMK-RQ-M and pMK-RQ-L. pMK-RQ-S was digested with KpnI and PstI, pMK-RQ-M was digested with KpnI and XhoI and pMK-RQ-L was digested with XhoI, NcoI and EcoRI. b) A 2 % agarose gel was used to separate the products of the restriction digests outlined in part (a) for pMK-RQ-S (*), pMK-RQ-M (**) and pMK-RQ-L (***).

Due to the availability of HAZV N antisera, confirmation that pMK-RQ-S could express the correct virally encoded products was achieved by transfection of pMK-RQ-S into BSR-T7 cells and analysing expression of HAZV N compared to a WT HAZV infection of SW13 cells via western blot. BSR-T7 cells were

transfected at 60-70 % confluency with 1.5 μg pMK-RQ-S and 3.75 μL TransIT-LT1 and returned to a 37°C incubator for 24 hours. For the WT HAZV infection, 60-70 % confluent SW13 cells were infected at a MOI of 0.1 and incubated at 37°C for 24 hours. Following the transfection and infection, total cell lysate was collected, lysed in RIPA buffer, clarified via centrifugation then analysed via western blot (Figure 3.4). Utilising anti-HAZV N serum, visualisation of a band at approximately 54 kDa, representing full length HAZV N was achieved following transfection, with a similar molecular weight band also detected in lysates collected from the WT HAZV infection, as expected. In addition to the 54 kDa product, a slightly smaller HAZV N specific product was detected in both transfected and infected cell lysates, potentially representing a cleavage or degradation product, suggesting that the HAZV N expressed by pMK-RQ-S was a similar molecular weight and able to be processed in the same way as HAZV N.

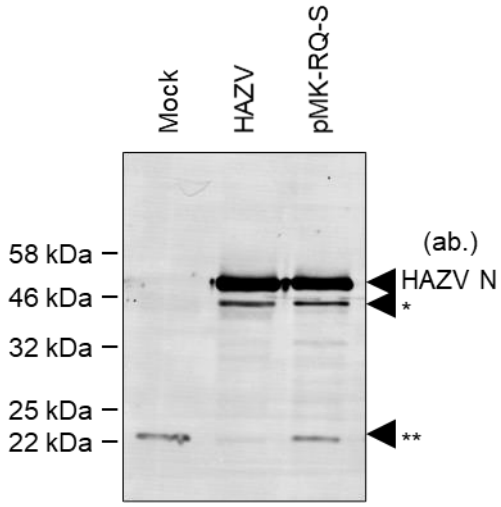


Figure 3.4, Confirmation of HAZV N expression from pMK-RQ-S

Comparison of HAZV N expressed via transfection of the pMK-RQ-S plasmid into BSR-T7 cells versus HAZV infection of SW13 cells. BSR-T7 cells were transfected at 60-70 % confluency using the TransIT-LT1 transfection reagent. SW13 cells were infected at a MOI of 0.1 in SFM for 1 hour prior to washing and reapplication of 2.5 % DMEM. Following a 24 hour incubation period, total cell lysate for both SW13 cells and BSR-T7 cells was collected and analysed for HAZV N expression via western blot using HAZV N antisera. A negative control of BSR-T7 cells with no virus added was also included to permit identification of non-specific HAZV N antisera reactive bands.

3.3 Optimising transfections

In order to generate recombinant virus, the efficiency of delivering cDNAs into the BSR-T7 cell line was optimised to increase chances of successful recovery of infectious virus. Despite protocols being provided by manufacturers for efficient use of their transfection reagents, inherent variability in the conditions required for efficient transfection into both closely related and different cell lines, requires optimisation prior to commencing experiments (Rose, 2003). Factors such as transfection reagent, ratios of DNA mass to reagent quantity and inclusion of additional plasmids to aid transcription were all explored.

3.3.1 Transfection reagent

In the recovery of infectious CCHFV, the Mirus Bio transfection reagent, TransIT-LT1, was used to transfect cDNA plasmids into BSR-T7 cells (Bergeron *et al.*, 2015). To ensure this reagent was optimal for cDNA plasmids encoding HAZV genome segments, we compared it to Lipofectamine 2000, the most heavily cited transfection reagent available at the time of writing.

Equal amounts of pMK-RQ-S were transfected into 60-70 % confluent BSR-T7 cells using either the TransIT-LT1 or Lipofectamine 2000 reagents according to the manufacturer's instructions, 24 hours post transfection lysates were harvested and analysed via western blot (Figure 3.5, a). Detection of HAZV N was used as a marker of transfection efficiency to permit effective comparison between the two. A non-transfected control was included and levels of GAPDH were also measured to confirm equal loading of samples. Densitometry performed on independent triplicate repeats showed expression of pMK-RQ-S via Lipofectamine 2000 to have a mean expression rate 30 % lower than that using TransIT-LT1 and was therefore chosen as the reagent of choice in all future DNA transfections (Figure 3.5, b).

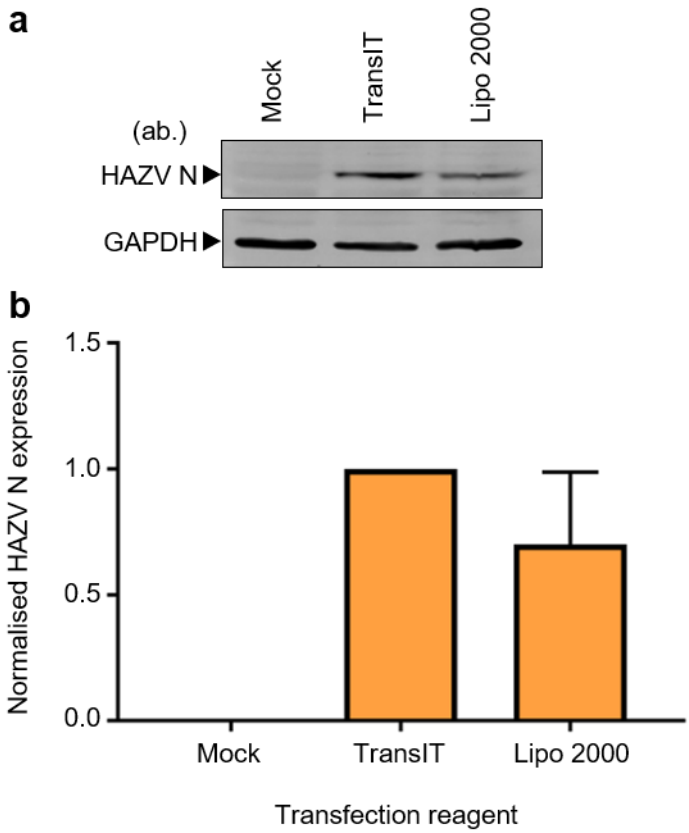


Figure 3.5, Selection of transfection reagent

a) Comparison of HAZV N expression levels with either TransIT-LT1 or Lipofectamine 2000 (Lipo 2000) transfection reagent. 1.5 µg of pMK-RQ-S was transfected into BSR-T7 cells at ≈ 60-70 % confluency using each transfection reagent at the manufacturers indicated ratio. At 24 hours post transfection cell lysates were collected for analysis by western blot analysis using HAZV N antisera. GAPDH expression levels were also examined, using antisera, as a loading control. b) Densitometric analysis of HAZV N expression in part (a) using ImageJ plotted using GraphPad Prism software.

3.3.2 Ratio of transfection reagent to cDNA

Following selection of a reagent, optimisation of the ratio of transfection reagent to plasmid DNA was explored to minimise the volume required to achieve effective plasmid expression. This minimises not only cost, but also toxic effects to the cells that are an inherent side-effect of transfection reagents. This allows cells to remain healthy and thereby increase expression.

A range of ratios of DNA to TransIT-LT1, including a negative control of 0:1, from 1:1 to 1:3.5 were used to transfect 60-70 % confluent BSR-T7 cells. Following a 24 hour incubation period, lysates were harvested and HAZV N expression analysed via western blot (Figure 3.6, a). As with the transfection reagent

optimisation, HAZV N expression was used as a marker of transfection efficiency to permit comparison between ratios. Levels of GAPDH were also analysed to confirm equal loading of samples. Densitometry showed HAZV N expression to plateau at ratios of 2.5 and above (Figure 3.6, b), with no significant benefit arising from further addition of reagent. Expression of HAZV N at a DNA/transfection reagent ratio of 1:1 was used as baseline expression and the other ratios compared to this value. The mean relative expression values were determined via independent triplicate repeats for ratios 1:2 (1.88), 1:2.5 (2.43), 1:3 (2.45) and 1:3.5 (2.23). For this reason, a ratio of 1:2.5 of plasmid DNA to TransIT-LT1 was used for further transfections.

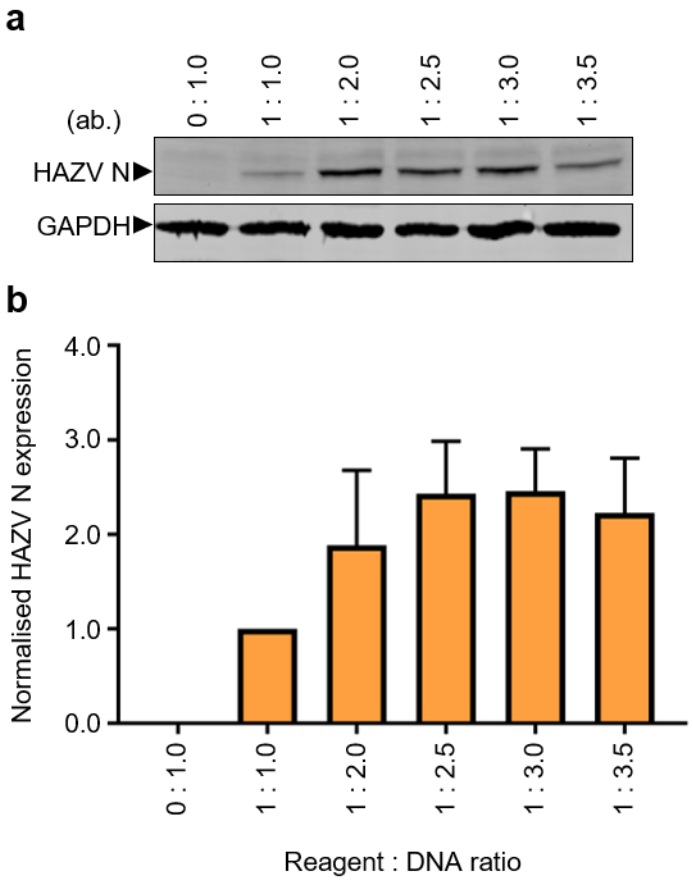


Figure 3.6, Optimising ratio of transfection reagent to cDNA

a) Determination of optimal cDNA to TransIT-LT1 ratio for transfection of pMK-RQ-S into BSR-T7 cells, measured via expression of HAZV N, using the TransIT-LT1 transfection reagent at 60-70 % confluency according to the product protocol. At 24 hours post transfection cell lysates were collected for analysis by western blot analysis with HAZV N antisera, GAPDH expression levels were also examined, using antisera, as a loading control. b) Densitometric analysis of HAZV N expression in part (a) using ImageJ plotted using GraphPad Prism software.

3.3.3 Inclusion of exogenous T7 polymerase

Previous reports have indicated the fluctuation and unreliable nature of T7 polymerase expression within BSR-T7 cells (Albariño *et al.*, 2009). Previous recovery of infectious CCHFV was also aided by inclusion of a T7 polymerase expression plasmid, with no associated negative side effects observed (Bergeron *et al.*, 2015). Therefore, to increase the efficiency of recovery of HAZV, we aimed to ensure maximal expression of our HAZV genome cDNAs via increased expression of T7 polymerase using a T7 RNA polymerase expression plasmid.

For these reasons, we trialled the inclusion or omission of pCAGGS-T7pol, a T7 polymerase expression plasmid under the control of both a chicken actin promoter and T7 promoter, alongside transfection of pMK-RQ-S. To ensure fair comparison the same amount and ratio of DNA to transfection reagent was maintained across the experiment. After 24 hours following transfection of either pMK-RQ-S or pMK-RQ-S segment plus pCAGGS-T7pol into BSR-T7 cells, lysates were harvested and analysed via western blot (Figure 3.7, a). A non-transfected control was included and levels of GAPDH were also detected to confirm equal loading of samples. As with previous optimisations, HAZV N expression was used as a marker of transfection efficiency to permit comparison between samples, with GAPDH also included to ensure equal loading of samples. Densitometric analysis using ImageJ was used to compare expression of HAZV N, with omission of pCAGGS-T7pol used as baseline expression (Figure 3.7, b). Independent triplicate repeats revealed inclusion of the T7 expression plasmid to substantially increase expression of HAZV N, with a mean relative fold increase of 3.82.

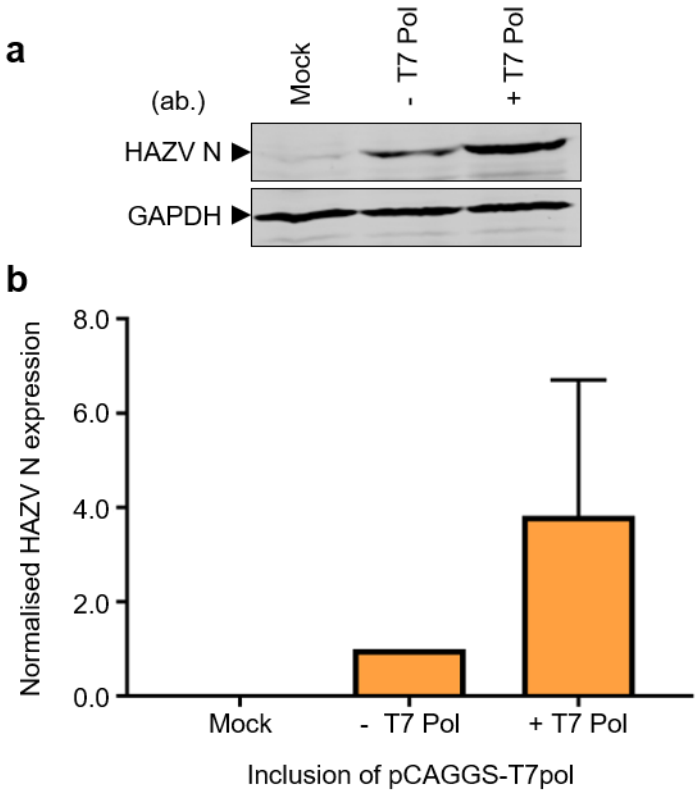


Figure 3.7, Optimisation of T7 polymerase expression

a) Comparison of expression of HAZV N with or without pCAG-T7pol, 1.5 µg pMK-RQ-S and +/- 1 µg pCAG-T7pol were transfected into BSR-T7 cells at 60-70 % confluency. At 24 hours post transfection cell lysates were collected for analysis by western blot analysis with HAZV N antisera, GAPDH expression levels were also examined, using antisera, as a loading control b) Densitometric analysis of HAZV N expression in part (a) using ImageJ plotted using GraphPad Prism software.

3.4 Recovery of WT rHAZV

An attempt was then made to recover infectious recombinant HAZV (rHAZV) using the optimisations identified in sections 3.3.1, 3.3.2 and 3.3.3. BSR-T7 cells were transfected with pMK-RQ-S, pMK-RQ-M and pMK-RQ-L alongside pCAGGS-T7pol at 50-60 % confluency, this is referred to as a “complete” transfection from herein. Additional transfections in which pMK-RQ-L was omitted were also included, to act as recovery controls incapable of generating infectious virus, this is referred to as a “control” transfection from herein. Despite the previously described rCCHFV system supplementing the individual genome segment specific cDNAs with support plasmids expressing the CCHFV N and CCHFV RdRp, recovery of rHAZV was attempted without the HAZV equivalents in an attempt to reduce the number of plasmids required to a minimum, thus reducing the complexity of transfection (Bergeron *et al.*, 2015). Further reasoning

for omission of support plasmids was provided by the rescue of BUNV from cDNA, which had previously incorporated the use of support plasmids alongside plasmids expressing full length antigenomic vRNAs; In a subsequent follow-up paper, attempts to recover infectious virus without addition of support plasmids proved successful, demonstrating the possibility of rescuing members of the order *Bunyavirales* without additional transfection of support plasmids (Bridgen and Elliott, 1996; Lowen *et al.*, 2004). A lower confluency of cells was used versus the transfection optimisations previously described in section 3.3 to allow for amplification of cell numbers within the time frame of the rescue experiments. Attempts made with higher confluency cells were met with visible CPE, potentially due to a combination of over confluency, toxicity arising from transfection reagent and virus-induced CPE. At regular time intervals post transfection, namely 72, 96 and 120 hours, supernatant was harvested from BSR-T7 cells and used to infect fresh monolayers of SW13 cells for a further 48 hour infection, following this infection, supernatant was harvested and stored for downstream applications (Figure 3.8). In addition, supernatant samples were also harvested from the transfected BSR-T7 cells at the 72, 96 and 120 hour time points and used to titrate virus levels to assess the optimal time to harvest recovered virus following transfection of the plasmid DNA into BSR-T7 cells.

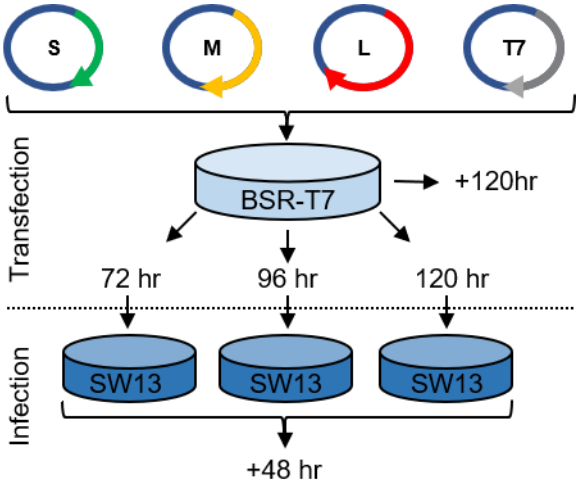


Figure 3.8, Schematic detailing recovery of rHAZV

Rescue of rHAZV was achieved by BSR-T7 cell transfection of plasmids pMK-RQ-S, pMK-RQ-M and pMK-RQ-L designed to express small (S), medium (M), large (L) genome segments, respectively, alongside a T7 RNA polymerase expression plasmid pCAG-T7pol (T7). At the indicated timepoints, supernatant was used to infect fresh SW13 cell monolayers for 48 hours, when supernatants and cell lysates were harvested for downstream application and analysis.

Following the 48 hour SW13 cell infections, lysates were collected and analysed for expression of HAZV N as a marker of viral transcription and translation by SDS PAGE and western blotting. Significant amounts of HAZV N were detected after the 48 hour infections set up using supernatant collected 96 and 120 hours post transfection arising from the complete rescue attempt (Figure 3.9, a, lanes 1-4). Supernatant at each of these time points was collected and stored at −80°C for future downstream applications. As expected, no HAZV N was detected in any of the SW13 cells infected with supernatant collected from control transfected BSR-T7 cells (Figure 3.9, a, lanes 5-8). Titre of infectious rHAZV was determined via plaque assay, with increasing titre of rHAZV detected in the supernatant taken from transfected BSR-T7 cells at 72, 96 and 120 hours post transfection (Figure 3.9, b), extensive CPE of the BSR-T7 cells created difficulties with collection of supernatant past this point. It is important to note that the titre following transfection should not be used as a measure of viral fitness; the requirement of multiple plasmids to be transfected into a single cell may lead to some cells acquiring the required cDNAs for virus production much sooner than others. However, in all recoveries of rHAZV, the titre of infectious virus generated peaked at day 5 post transfection, in a similar manner to recovery of rCCHFV (Bergeron *et al.*, 2015). Interestingly, rCCHFV was only recoverable in the presence of support plasmid expressing the RdRp, whereas our data shows HAZV not to require the inclusion of any supports, potentially due to differences in length of the L segment, though this was not explored further. However as previously mentioned, the ability of HAZV to be recovered without the use of support plasmids is not unique to the *Bunyavirales* order (Lowen *et al.*, 2004). Plaque assays, which rely on viral lysis of cells in localised areas, also confirmed the infectious nature of the virus produced from cDNAs (Figure 3.9, c), allowing downstream comparisons to parental virus to be carried out at equal MOI to ensure accurate conclusions to be drawn on their phenotypes. Taken together, the expression of viral products coupled with presence of CPE in the plaque assays shows successful recovery of a rHAZV infectious clone that, to our knowledge, represents the first such recovery of this virus.

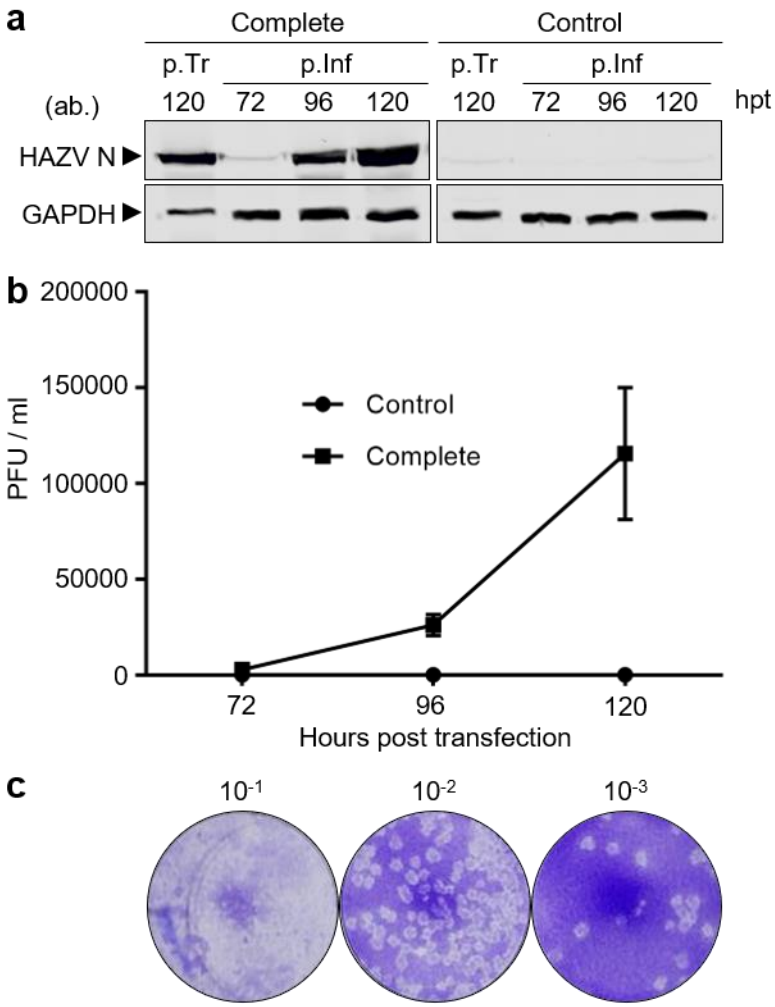


Figure 3.9, Recovery of WT rHAZV

a) Detection of HAZV N post-transfection (p.Tr) of BSR-T7 cells and subsequent 48 hour post-infections (p.Inf) by western blot using HAZV N antisera. Supernatant samples collected from transfected BSR-T7 cells at 96 hours post transfection were used to infect monolayers of SW13 cells. At 48 hours post infection, lysates were collected and analysed by SDS-PAGE. Recovery of rHAZV using S, M and L rescue expressing plasmids ('complete') was carried out alongside control transfections in which L segment plasmid was omitted. Detection of GAPDH abundance by western blotting was included as a loading control. b) Titre of infectious rHAZV released into supernatant at 72, 96 and 120 hours post transfection of 'complete' and control plasmids was assessed via plaque assay. c) Representative plaque assays at 10^{-1} to 10^{-3} dilutions from rHAZV titration in part (b). Serial dilutions were generated from virus harvested 120 hours post infection and used to infect SW13 cell monolayers for 1 hour prior to addition of a methyl cellulose overlay. Cells were then fixed and stained 6 days post infection and counted.

Following successful titration of virus via the plaque assay method, quantitative polymerase chain reaction (qPCR) was carried out to confirm that the virus causing the plaque assay CPE was rHAZV. Primers were designed to be specific to the S segment of HAZV (5' CAA GGC AAG CAT TGC ACA AC 3' and 5' GCT TTC TCT CAC CCC TTT TAG GA 3') and a set of standard RNA was transcribed

in vitro to permit quantification of S segment copies. Following an infection of SW13 cells with rHAZV at a range of MOIs including mock, 0.01 and 0.1 total RNA was isolated via Trizol extraction and analysed via qPCR. Detection of the *in-vitro* transcribed RNA was as expected, with relative threshold cycle values (ct) increasing as the copy number of RNA decreased (Figure 3.10, a). Consistent with expected results, the RNA isolated from the highest MOI infection of 0.1 had the lowest ct values, followed by the 0.01 MOI infection (Figure 3.10, b). Unfortunately, amplification of RNA did occur in RNA isolated from the mock infected SW13 cells. However, the ct values corresponding to the mock infected cells were outside the range of the RNA standards and is therefore possibly a result of non-specific binding of primers. As a control to ensure total cellular RNA was consistent between samples, primers against GAPDH were designed and incorporated into the qPCR assay. Primer sequences can be found in Table 2.1. As expected, ct values for all samples were similar, demonstrating a similar concentration of starting material was present in the samples (Figure 3.10, c). Inclusion of the GAPDH also permitted normalisation of the data relative to the 0.10 MOI infection. Consistent with the 10 fold reduction in MOI, RNA isolated from the 0.01 MOI infection contained approximately 7 % of the S segment specific RNA detected in the 0.10 MOI infection, consistent with expectations (Figure 3.10, d).

Having an established qPCR system may prove beneficial in future experiments where mutations incorporated into the recovered virus affect its ability to form plaques, providing an alternative and faster means of titration. In addition, when used in combination with cellular lysates rather than supernatants it could provide useful information on viral transcription levels and is therefore a valuable tool in furthering understanding into HAZV infection.

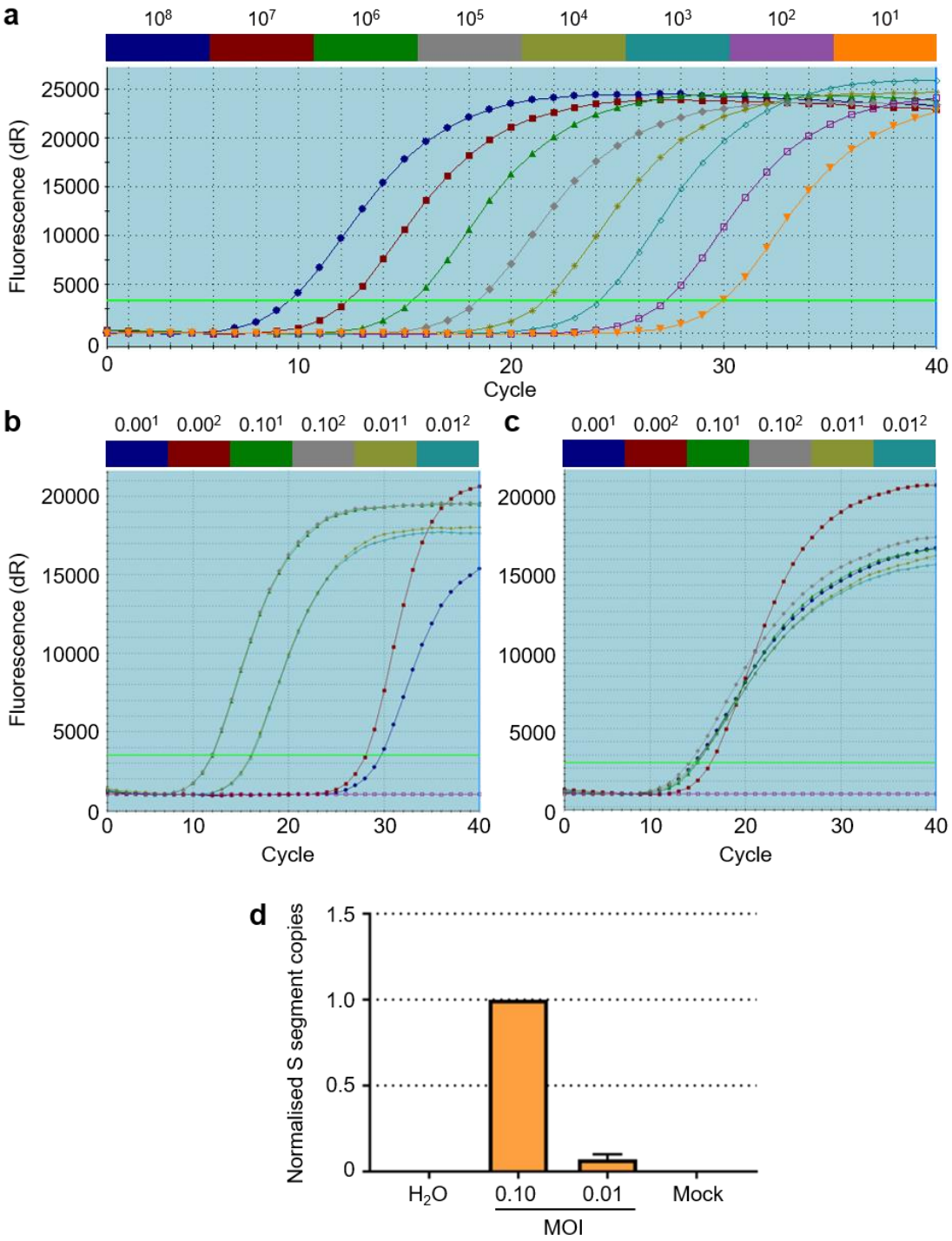


Figure 3.10, qPCR detection of HAZV S segment RNA

qPCR readout for (a) in vitro transcribed S RNA using HAZV S primers. RNA extracted from SW13 cells infected at 0.1, 0.01 and mock infected MOIs using (b) HAZV S segment primers and (c) using GAPDH primers. Axis display fluorescence versus cycle number, threshold was automatically determined by qPCR software. d) Normalised S segment copies for each of the experimental conditions plotted using GraphPad Prism software.

3.5 Comparison of recombinant HAZV to parental strain JC280

To permit accurate conclusions to be drawn about the parental strain from the infectious clone, the recombinant virus must be confirmed as a suitable model for the parental strain. To this end, rHAZV was compared directly against in-house parental HAZV stocks, examining both fitness and plaque morphology.

Briefly, to examine the growth profiles of HAZV and rHAZV, SW13 cell monolayers were infected at an equal MOI of 0.001 in SFM for 1 hour. At this time infection media was removed, cells were washed in PBS and fresh media was reapplied. Supernatant samples were harvested at 24, 48, 72 and 96 hours post infection, at which point and viral titre was determined via plaque assay. Plotting of infectious titre versus time displayed both HAZV and rHAZV to share similar growth profiles, with both titres increasing from the point of infection until 72 hours post infection, with titres observed to be approximately 1-2 x 10^6 pfu / mL, after this point, virus titre began to drop off in a similar manner with both HAZV and rHAZV (Figure 3.11, a).

As virus titre was determined via plaque assay, it also permitted analysis of plaque morphologies for both HAZV and rHAZV. Plaque morphology is a useful measure of viral fitness as variation in dimensions, clarity and intensity of staining can all indicate altered fitness. As with the growth profiles of the two viruses, plaque morphologies were similar (Figure 3.11, b), presenting as circular, with well-defined edges and a "bullseye" centre. This bullseye forms when the initially infected cells are not killed to the same degree to subsequently infected cells surrounding them, leaving a portion of cells still able to take up the crystal violet stain. As to why the initially infected cells are not killed by rHAZV or HAZV remains unknown, but may be linked to the reduced viral load initially present in the infection. These data both supported the rHAZV infectious clone as a suitable model system to study HAZV and its suitability to be utilised in further studies.

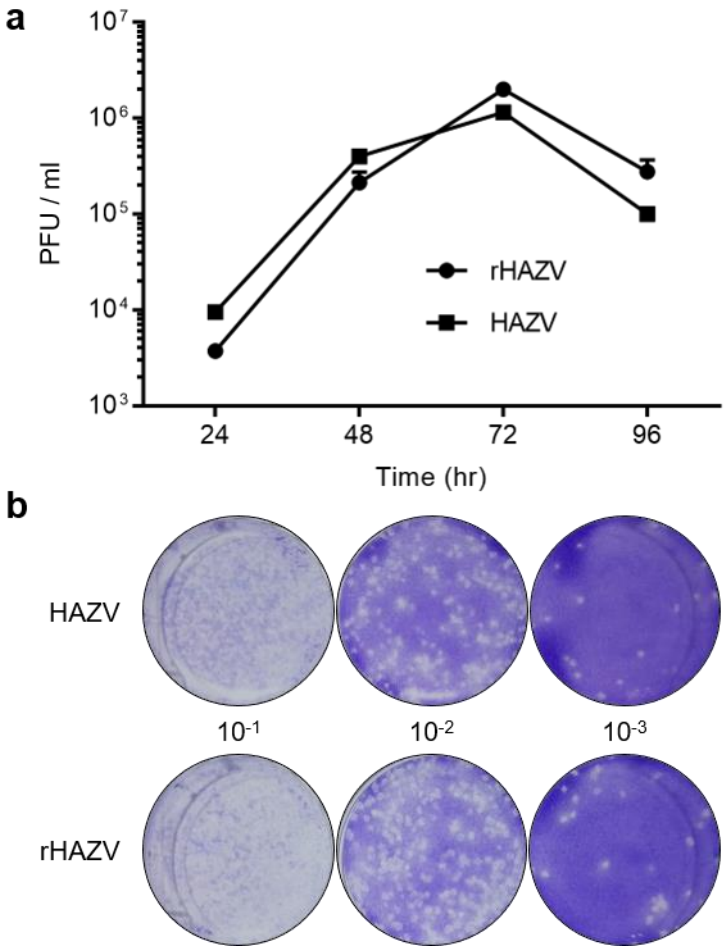


Figure 3.11, Comparison of rHAZV to parental strain

a) Titre of infectious rHAZV (circle data points) and its parental strain of HAZV (square data points) harvested at 24 hour intervals up to 96 hours post infection of SW13 cells at an equal MOI of 0.001. b) Representative plaque assays from the 10^{-1} to 10^{-3} dilutions used to determine titre of infectious HAZV and rHAZV in (a) showing similar plaque morphology in both instances. Serial dilutions were generated from virus harvested 24 hours post infection and used to infect SW13 cell monolayers for 1 hour prior to addition of a methyl cellulose overlay. Cells were then fixed and stained with crystal violet 6 days post infection and plaques counted.

3.6 Confirmation of the cDNA source of rHAZV

In order to confirm the origin of the rescued HAZV as having a cDNA source, rather than as a laboratory contaminant, site directed mutagenesis (SDM) was conducted on pMK-RQ-S resulting in the insertion of a “silent” mutation as a genetic marker, which was a *Hind* III restriction site at nucleotide position 723 via switching of guanine to thymine, giving the new plasmid termed pMK-RQ-S-nG723T. This change was silent at the amino acid sequence level, as evidenced by the resulting codon changing from CTG to CTT, both encoding leucine residues (Figure 3.12, a). Following recovery of the mutant rHAZV, primers were

designed to flank this region at position 415 and 1217 in the S genome to permit PCR amplification of an 800 bp sequence of DNA, following RNA extraction and first strand synthesis, that could then be examined for the restriction site via sanger-sequencing or restriction digest (Figure 3.12, b).

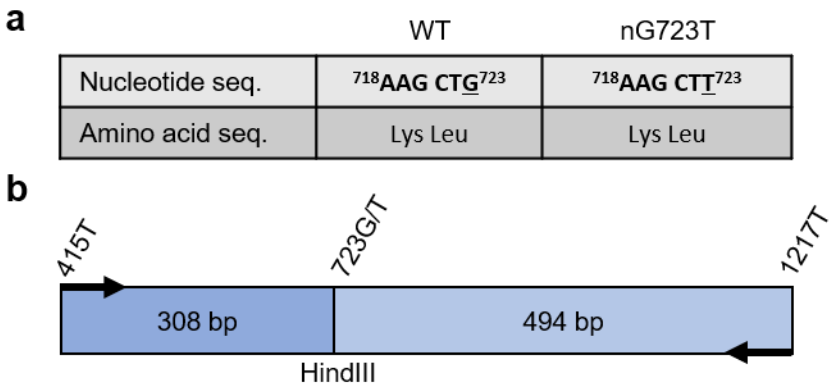


Figure 3.12, Overview of insertion of a HindIII restriction site to S segment cDNA

a) Table showing the change (underlined) to both DNA sequence of the HAZV N ORF and the resulting amino acid sequence. b) Schematic showing the location of the inserted Hind III restriction site into the S-segment ORF.

The plasmid containing the nG723T mutation, pMK-RQ-S-nG723T, was then utilised in mutant recovery in place of pMK-RQ-S alongside control and complete recoveries as previously described in section 3.4. As expected, the silent nature of the mutagenesis meant no phenotypic effect was observed and the nG723T mutant rescued alongside the WT recovery, with strong HAZV N expression detected in both transfected BSR-T7 cells and the subsequent infections of SW13 cells using supernatant harvested at 96 and 120 hour post transfection (Figure 3.13, a). Following recovery and subsequent amplification to generate high titre virus, an RNA extraction, first strand synthesis and PCR amplification using the primers described above permitted isolation of an 800 bp fragment of DNA. Restriction digest using *Hind* III cleaved the fragment isolated from rHAZV(nG723T) into 300 and 500 bp fragments, whilst the fragment isolated from WT rHAZV remained intact (Figure 3.13, b). A control omitting the reverse transcription stage was also included to rule out possible plasmid contamination as the source of the PCR amplified DNA. This proved beyond reasonable doubt that rHAZV was cDNA-derived.

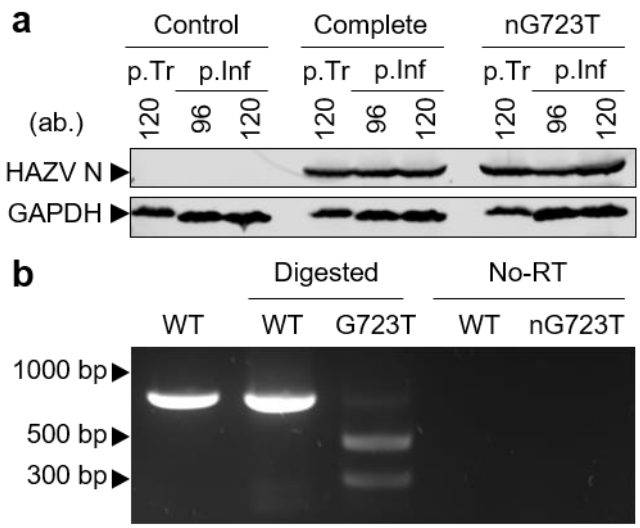


Figure 3.13, Recovery and confirmation of a silent mutant rHAZV

a) Detection of HAZV N by western blotting following transfection (p.Tr) of BSR-T7 cells and subsequent 48 hour infections (p.Inf). Supernatant samples collected from transfected BSR-T7 cells at 96 hours post transfection were used to infect monolayers of SW13 cells. Following a 48 hour infection, lysates were collected and ran on SDS-PAGE. Recovery of rHAZV containing a HindIII restriction site (rHAZV(G723T)) was carried out alongside a complete and control recovery of rHAZV. Detection of GAPDH abundance was included as a loading control. b) A 2 % agarose gel was used to separate the products of the Hind III restriction digest of the dsDNA generated following RNA extraction of rHAZV and rHAZV(nG723T) containing supernatants, first strand synthesis and PCR amplification of viral genetic material. A control reaction in which no reverse transcriptase was added to the reaction mixture ("No-RT") was also included to rule out plasmid carryover as the source of the amplification.

3.7 Chapter summary and discussion

In this chapter, an infectious clone system was designed, optimised and utilised to generate infectious rHAZV from positive sense cDNAs in a robust manner. Previous work has developed a similar system for CCHFV, however due to the highly pathogenic nature of this virus, the CCHFV system requires restrictive containment level 4 protocols (Bergeron *et al.*, 2015). The classification of HAZV as a hazard level 2 pathogen permits its use in much more easily accessible BSL-2 facilities and, given it is the first orthonairovirus of this hazard level to be recovered from cDNAs, represents a major advance in the study of orthonairovirus molecular and cellular biology. The potential uses of an infectious clone system for HAZV are broad; the above described system for CCHFV has already been utilised to increase understanding on glycoprotein processing, the requirement of the OTU domain and more recently in aiding anti-viral discovery (Bergeron *et al.*, 2015; Scholte *et al.*, 2017; Welch *et al.*, 2017). Recovery

systems are not just limited to small scale changes to the viral genomes; the RVFV infectious clone system was utilised to generate a NSm deficient virus, demonstrating NSm to be dispensable for both virulence and lethality (Bird, Albariño and Nichol, 2007). Another important aspect of an infectious clone is its ability to aid vaccine development. Attenuation of domains critical for viral pathogenesis, or alteration of the optimal replication temperature or additional host requirements, has already been explored in both RVFV and respiratory syncytial virus (RSV) in pursuit of vaccine candidates (Collins *et al.*, 1995; Smith *et al.*, 2018). With extensive similarities between HAZV and CCHFV, aside from the shared human pathology, incorporation of the CCHFV antigen into a HAZV virion could present an interesting vaccine candidate, assuming the pathogenic domains of CCHFV are not a result of an overlapping region with the antigenic section.

The system developed in this chapter has been demonstrated to be highly efficient, generating titres of rHAZV over $1.2x10^5$ in primary transfected supernatants, which can be used to further create higher titre stocks through a second round of passaging once the titre of the initial stock has been determined. Furthermore, the recovery of WT rHAZV was successful in 100 % of rescues using the sequence confirmed pMK-RQ-S/M/L and pCAGGS-T7pol plasmids, with a minimum of 5 recoveries attempted, demonstrating the robust and repeatable nature of the system. To test the ability of the infectious clone system to tolerate mutations, and also to confirm the source of the infectious virus as cDNA and not contaminating cDNA, a silent mutation was inserted, resulting in the generation of a *Hind* III restriction site. Following recovery of this mutant virus, the altered cDNA source was confirmed, and contamination ruled out, via RT-PCR and restriction digest. This ability to tolerate mutations, with high efficiency and repeatability of recovery will greatly aid recovery of mutant viruses with attenuated fitness and permit identification of mutations resulting in fatal phenotype.

This system will provide the basis of the following 2 chapters, which will utilise the WT system and adapt it to incorporate additional features and to explore the effects of mutations to the WT sequence on viral pathogenesis.

4 Generation of a fluorescent reporter virus

4.1 General introduction

Following the successful recovery of a WT recombinant clone for HAZV, detailed in chapter 3, the ability to modify the viral genome sequence was explored further. In addition to studying the importance of individual nucleotides, amino acids, NTRs and ORFs of the virus, the recombinant system can also be manipulated to add additional features to the virus. To this end the HAZV S segment expression plasmid pMK-RQ-S was modified to incorporate an eGFP ORF linked to HAZV N via a porcine teshovirus-1 2A peptide linker (P2A) self-cleaving sequence, creating pMK-RQ-S-eGFP (Figure 4.1). This construct permits expression of both eGFP and HAZV N as independently translated proteins from a single mRNA transcribed from the same viral genome segment.

It was decided to co-express eGFP and HAZV N from the same gene segment as it prevented the virus simply excluding an additional eGFP segment during packaging, due to the critical requirement of HAZV N in HAZV infection. The independent expression is a result of a ribosomal skipping event which occurs between the terminal glycine and proline residues in the P2A peptide, with no peptide bond forming during the continuation of translation, ensuring expression of the downstream gene (Donnelly *et al.*, 2001). Independent expression of both eGFP and HAZV N has multiple advantages. Firstly, as the final viral product is expressed without any additional residues, it should function as a WT viral protein, resulting in minimal loss of fitness to HAZV overall. Secondly, the untagged form of eGFP is not limited to regions where HAZV N localises to, allowing its expression to be ubiquitous within the cell, aiding detection of eGFP positive cells. Previous studies have successfully incorporated P2A linkers into recombinant virus systems, perhaps the most relevant example being in CCHFV, where a ZsGreen1 (ZsG) fluorescent marker was introduced in a similar manner, which permitted screening of a number of potential anti-viral compounds in an efficient and quantitative manner (Welch *et al.*, 2017). The CCHFV infectious clone expressing ZsG led to the identification of 2’-deoxy-2’flurocytidine as a potent inhibitor of CCHFV replication, with efficacy higher than ribavirin or favipiravir (Welch *et al.*, 2017). Given the similarities between HAZV and CCHFV,

and the lower containment level required for HAZV propagation, rHAZV(eGFP) presents an attractive means of assessing candidate compounds for CCHFV, without the high-level containment laboratory requirements.

It is worth noting that an inherent disadvantage to the P2A linker compared with post-translational cleavage systems arises from the imperfect nature of the ribosomal skipping event. During the ribosomal skipping event three potential outcomes are observed. The desired outcome is a successful skipping event resulting in two independently translated products, however previous studies have also identified cases where the skipping does not occur, giving a single polyprotein. Additionally, cases have been reported where the ribosome falls off the mRNA during translation at the glycine to proline site, leading to expression of only the first ORF, in our case eGFP (Liu *et al.*, 2017). Taking these factors into account, with many translation events occurring, a skew in gene expression occurs, with ORFs upstream of the P2A linker regions sometimes having higher expression rates than the downstream versions, a study examining expression of fluorescent reporters in bi-, tri- and tetra-cistronic vectors found that expression of the fluorophores decreased with progression towards the end of the construct (Liu *et al.*, 2017).

The aims of this chapter were as follows;

1. Develop an rHAZV infectious clone expressing eGFP
2. Validate rHAZV(eGFP) as a suitable reporter for rHAZV infection
3. Utilise the reporter virus to screen a panel of membrane trafficking factors
4. Follow up any significant results to uncover original findings on HAZV infection

Successful generation of an eGFP expressing HAZV infectious clone would enable multiple avenues of investigation into many aspects of HAZV infections in addition to those mentioned above and would therefore be a valuable tool in furthering knowledge into HAZV.

4.2 Plasmid design

To generate a cDNA plasmid encoding the eGFP-P2A ORF upstream of HAZV N, a cDNA was commercially synthesised and purchased (Genewiz) encoding the eGFP-P2A nucleotide sequence in a pUC57-Kan backbone, hereby known as pUC57-eGFP-P2A. The full sequence for this insert can be found in appendix 1.1.4. The purchased synthesised sequence for eGFP-P2A included a 5' *Eco*RI and 3' *Not*I restriction site to permit ligation into pMK-RQ-S. The S segment was chosen to express both HAZV N and eGFP due to its small size and prior availability of a marker, antiserum against HAZV N. As HAZV is able to package the much larger M and L segments, it was hypothesised that there would be no issues extending the length of the S segment RNP, the only factor that may be an issue would be fitting the enlarged RNP into a virion alongside M and L RNPs. However, as incorporation of a larger S segment RNP in the CCHFV fluorescent clone had no significant impact on viral fitness, it was assumed similar results would be observed in HAZV (Welch *et al.*, 2017).

To facilitate ligation of the eGFP-P2A insert from pUC57-eGFP-P2A into pMK-RQ-S, SDM was carried out on pMK-RQ-S to insert an *Eco*RI and *Not*I restriction site at the 5' end of the HAZV N ORF, immediately after the 3' end of the NTR, resulting in generation of pMK-RQ-S(*Eco*/*Not*). Following transformation of the plasmids after SDM, colonies were isolated, grown up and the DNA purified to permit sequence validation of the insertion of both restriction sites in the correct orientation. Upon confirmation, pUC57-eGFP-P2A and pMK-RQ-S(*Eco*/*Not*) were digested with *Eco*RI and *Not*I and relevant bands excised and purified following agarose gel electrophoresis. These digestion products were then used in a T4 ligation reaction to generate the plasmid displayed in Figure 4.1. A final set of SDM reactions were then carried out to remove the *Eco*RI and *Not*I restriction sites flanking the insert, to ensure HAZV N was kept in frame with the eGFP ORF, resulting in generation of pMK-RQ-S-eGFP. Confirmation of SDM success and lack of off target mutations was achieved via sequencing of the entire cDNA.

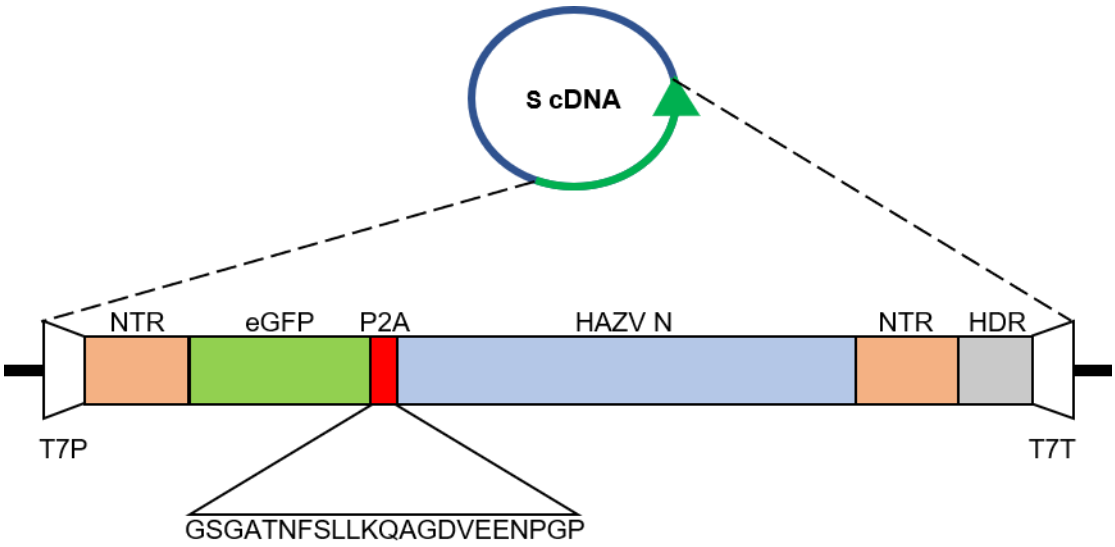


Figure 4.1, eGFP S segment schematic

Schematic of S segment plasmid design incorporating the eGFP ORF (green) immediately downstream of the 3' NTR (orange) and upstream of the HAZV N ORF (blue). Transcription of this plasmid by T7 RNA polymerase generates a positive sense antigenome. The P2A linker sequence (red) was inserted between the eGFP and HAZV N ORFs in frame to permit translation of both gene products. The amino acid sequence of the P2A linker is indicated below the schematic. A hepatitis delta virus ribozyme (HDR) ensured correct 3' end formation of transcribed RNAs. The entire transcriptional unit is under control of T7 promoter (T7P) and T7 terminator (T7T) sequences (white).

To confirm the correct expression of HAZV N following generation of pMK-RQ-S-eGFP, lysates collected from a WT rHAZV infection of SW13 cells were run alongside BSR-T7 cell lysates collected 24 hours following transfection with pMK-RQ-S-eGFP. Comparisons were made to ensure the P2A linker was correctly generating an untagged form of HAZV N, therefore the transfected lysates should contain a similar sized HAZV N to those of the rHAZV infected lysates. As seen in (Figure 4.2), HAZV N observed in transfected cell lysates was indistinguishable from WT rHAZV infected lysates, demonstrating that pMK-RQ-S-eGFP expression of HAZV N results in a product that is the same observed molecular weight as WT HAZV N and is recognisable by HAZV N antiserum.

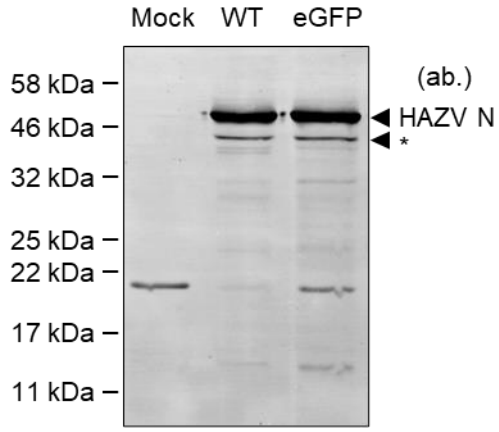


Figure 4.2, Western blot confirmation of correct HAZV N expression from pMK-RQ-S-eGFP

Comparison of HAZV N expressed by WT rHAZV during infection of SW13 cells (WT) versus expression of HAZV N following transfection of the pMK-RQ-S-eGFP plasmid into BSR-T7 cells (eGFP). BSR-T7 cells were transfected at 60 - 70 % confluency using the TransIT-LT1 transfection reagent. SW13 cells were infected at a MOI of 0.1 in SFM for 1 hour prior to washing and reapplication of 2.5 % DMEM. Following a 24 hour incubation period, total cell lysate for both SW13 cells and BSR-T7 cells was collected and analysed for HAZV N expression via western blot using anti HAZV N serum. A negative control in BSR-T7 cells with no virus added was also included to identify specific bands (*) from non-specific bands (**). Relevant molecular sizes are indicated in kDa.

4.3 Recovery of rHAZV-eGFP

The pMK-RQ-S-eGFP construct described in section 4.2 was synthesised and transfected alongside pMK-RQ-M and pMK-RQ-L into BSR-T7 cells at 50-60 % confluency. As with the WT rHAZV system described in section 3.4, a control recovery was set up with the omission of the pMK-RQ-L. As with recovery of WT rHAZV, supernatants were collected at 72, 96 and 120 hour time points following transfection and used to infect fresh monolayers of SW13 cells. Following a 48 hour infection, cell lysates were collected and analysed via western blot for the presence of HAZV N that would indicate successful recovery of rHAZV (Figure 4.3). Presence of a HAZV N specific band in both re-infections of SW13 cells using supernatant harvested at 96 and 120 hours post transfection indicated successful recovery of rHAZV(eGFP) and WT rHAZV utilising the engineered pMK-RQ-S-eGFP plasmid and WT pMK-RQ-S plasmid respectively. To ensure the virus was able to generate eGFP, western blot analysis using an eGFP antibody was also carried out. Detection of a band specific to eGFP was observed in lanes corresponding to the initially transfected BSR-T7 cells and in the SW13 cells re-infected with supernatant harvested at 120 hours post transfection

(Figure 4.3). Critically, detection of both HAZV N and eGFP was observed at their expected molecular weights, confirming their independent expression and ruling out an eGFP-HAZV N polyprotein expression system.

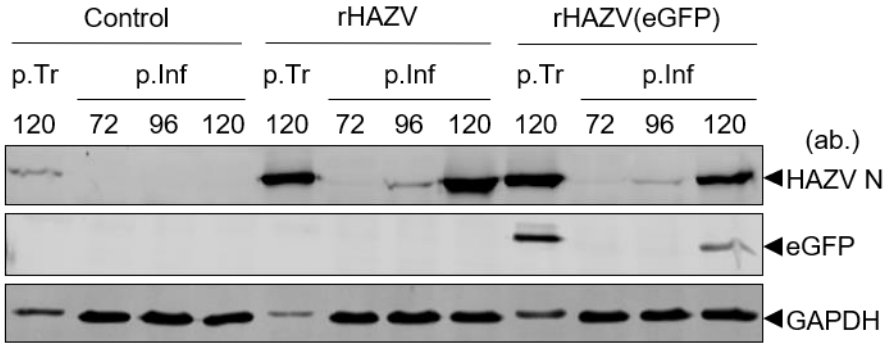


Figure 4.3, Western blot analysis of transfected and infected cell lysates confirms recovery of rHAZV(eGFP)

Detection of HAZV N and eGFP following transfection (p.Tr) of BSR-T7 cells and subsequent 48 hour infections of SW13 cells (p.Inf) using supernatant samples collected from transfected BSR-T7 cells at 72, 96 and 120 hours post transfection. Following infection, lysates were collected and analysed via western blot. Recovery of rHAZV containing the eGFP-P2A ORF was carried out alongside a WT rHAZV and control recovery of rHAZV omitting the L segment. Detection of GAPDH abundance was included as a loading control.

To confirm the fluorescent nature of the recovered rHAZV(eGFP), the 48 hour re-infection of SW13 cells, using supernatant harvested from BSR-T7 cells at 120 hours post-transfection, was carried out in the IncuCyte live cell imaging system. This enabled images to be taken and analysed for green cell count every hour throughout the infection (Figure 4.4, a). Representative images from the time course at 0 and 48 hours post infection for complete and rHAZV(eGFP) are provided showing the highly fluorescent nature of the virus (Figure 4.4, b). No visible fluorescence was observed in wells infected with WT rHAZV (Complete), indicating fluorescence was an attribute unique to the eGFP expressing virus. Analysis of total green cell count per well showed a two-step increase of eGFP positive cells, the first occurring between 15-20 hours post infection and the second at 40-45 hours post infection (Figure 4.4, a). The observation of a two step increase allows the estimation of the length of the infectious cycle in SW13 cells to be approximately 20 to 25 hours. This can be further supported by the detection of infectious virus in the supernatant at 24 hours post infection, as previously shown in Figure 3.11 As expected, in both the control and WT rHAZV infections, no eGFP positive cells were detected, with minor fluorescence visible

from 40 hours onwards, likely as a consequence of auto-fluorescence from dying cells as the cell culture becomes over confluent. The ability to indirectly measure viral activity in real time, by observing eGFP levels, will provide an excellent tool for analysing the effect of siRNAs on viral infection, streamlining the sample processing stages required prior to analysis.

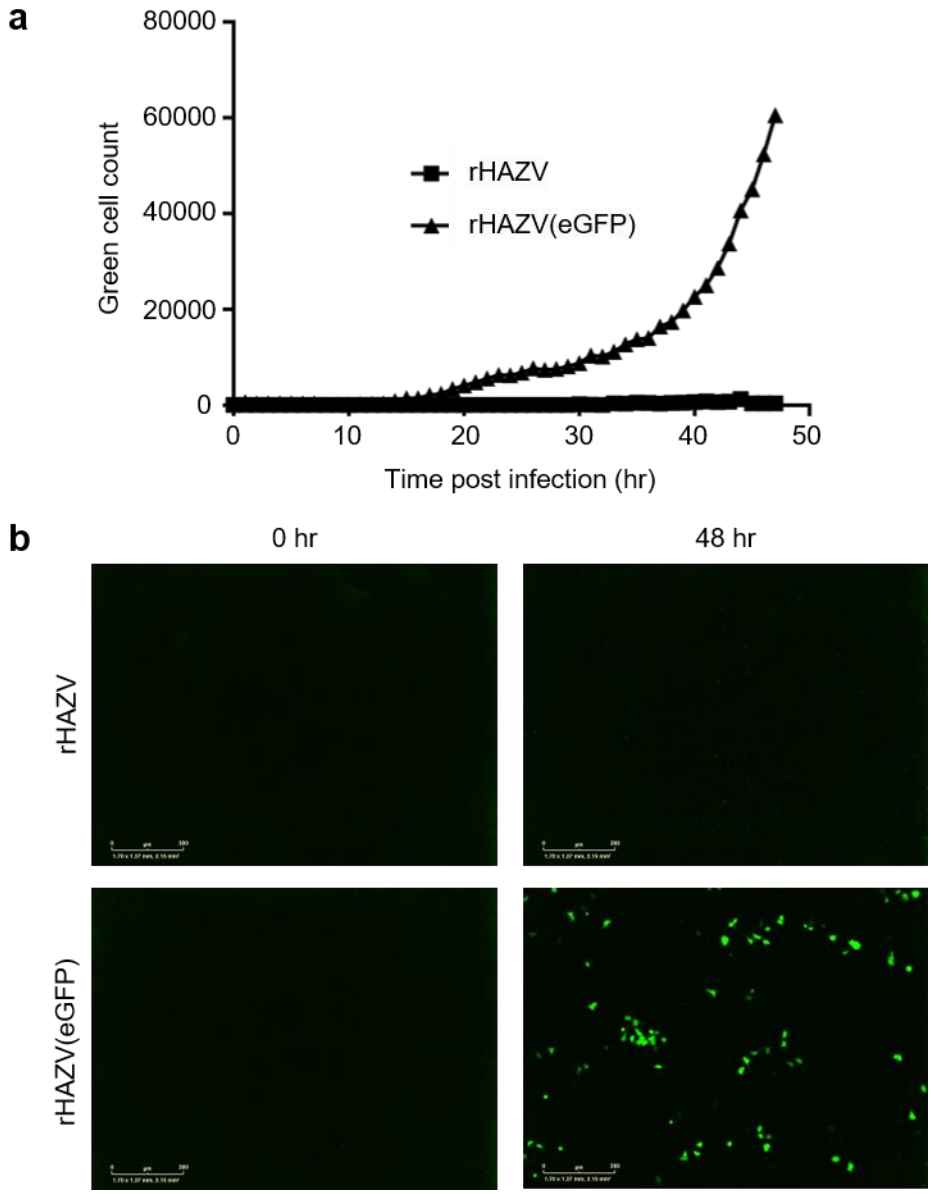


Figure 4.4, Confirmation of the fluorescent nature of rHAZV(eGFP) via live cell imaging

a) Graphical representation of green cell count over time during the 48 hour re-infection of SW13 cells using supernatant harvested from BSR-T7 cells 120 hours post transfection. Analysis was carried out using the IncuCyte® live cell imaging system and plotted using GraphPad Prism. b) Demonstration of the fluorescent nature of rHAZV(eGFP) shown in images taken at 0 and 48 hours post infection from the experiment carried out in (a) from rHAZV and rHAZV(eGFP) infections. Scale bar represents 300 µM.

Titre of supernatant harvested from BSR-T7 cells 120 hours post-transfection was measured via plaque assay in SW13 cells to permit further propagation of the virus stock at a known MOI to generate a larger titre in future experimentation (Figure 4.5). Visualisation of plaques 6 days post infection also confirmed the lytic nature of rHAZV(eGFP), as was previously observed with WT rHAZV in section

3.4. Taken together, this demonstrated successful recovery of a rHAZV mutant, capable of expressing eGFP independently of viral proteins that can be used for multiple applications, including streamlining screening of siRNAs without the need for western blotting, immunofluorescence or qPCR-based techniques.

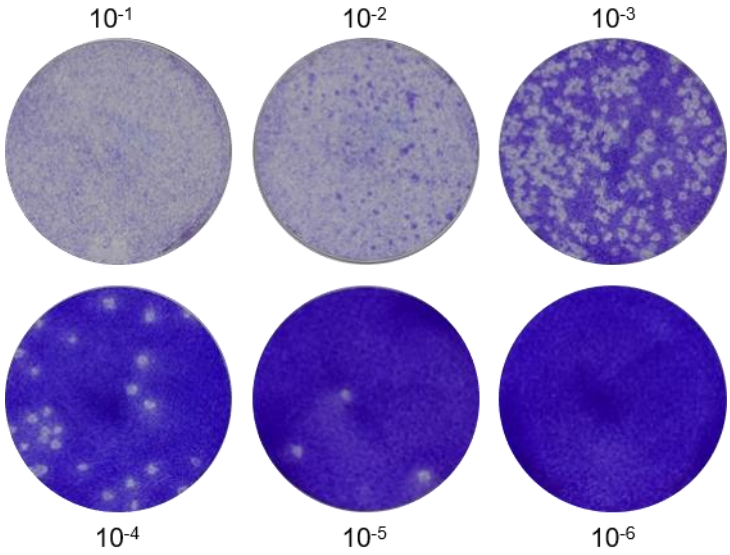


Figure 4.5, Plaque assay of rHAZV(eGFP)

Representative plaque assays of rHAZV(eGFP) used to titre virus also show lytic nature of the recovered recombinant reporter virus. Images show dilutions from 10^{-1} to 10^{-6} in SW13 cells stained 6 days post infection.

4.4 Comparison to WT rHAZV

The nature of the P2A linkage of eGFP to HAZV N in rHAZV(eGFP) should result in independent expression of both proteins and therefore have little or no effect on the viral lifecycle, however previous evidence has suggested this linker is not 100 % efficient in generating two independent products of translation (Liu *et al.*, 2017). To ascertain whether the inserted eGFP ORF was having an effect on virus growth kinetics, comparison of rHAZV(eGFP) and WT rHAZV was required in order to validate rHAZV(eGFP) as a model system. To achieve this, the multistep growth kinetics for WT rHAZV and rHAZV(eGFP) was compared in SW13 cells infected at an equal MOI of 0.01. At regular 24 hour intervals post infection, supernatant samples from each infection were harvested and used to titre infectious virus in supernatants via plaque assay. Comparison of viral titres displayed a marginal increase in viral titre for WT rHAZV versus rHAZV(eGFP) (Figure 4.6, a). However, the growth profile remained consistent between both strains, peaking at 72 hours post infection, then decreasing at subsequent

timepoints. Mean values for plaque forming units per mL (PFU / mL) were as follows; 24 hours: 5400 (rHAZV) and 3750 (rHAZV(eGFP)), 48 hours: 310,000 (rHAZV) and 80,500 (rHAZV(eGFP)), 72 hours: 440,000 (rHAZV) and 190,000 (rHAZV(eGFP)), 96 hours: 90,000 (rHAZV) and 44,500 (rHAZV(eGFP)), 120 hours: 30,000 (rHAZV) and 7850 (rHAZV(eGFP)). This observed difference between the WT and reporter viruses may be a result of allocation of cellular resources away from expression of viral products during synthesis of eGFP during rHAZV(eGFP) infection, resulting in generation of fewer infectious virions. The increased length of the S segment, due to the inclusion of the eGFP and P2A sequences, will result in slower transcription and translation times. Previous reports have also identified lower expression of proteins downstream of P2A linker sequences, potentially due to the disadvantages associated with P2A linkers identified in section 4.1, which could result in translation of less HAZV N (Liu *et al.*, 2017). These explanations would point to the ability of rHAZV to produce a higher number of infectious particles than rHAZV(eGFP), rather than a difference in fitness between the fully assembled particles produced by either virus. Plaque morphology remained largely similar for rHAZV and rHAZV(eGFP), with only a slight loss of clarity (Figure 4.6, b). However overall, rHAZV and rHAZV(eGFP) displayed similar growth properties and phenotype, demonstrating the suitability of rHAZV(eGFP) as a reporter strain for HAZV.

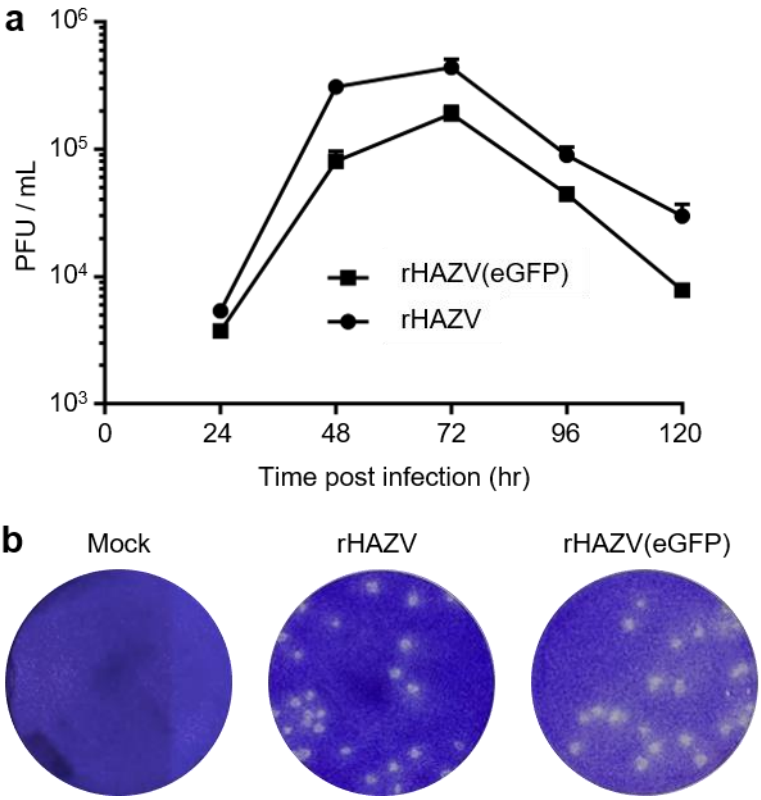


Figure 4.6, Multi-step growth kinetics of rHAZV vs rHAZV(eGFP) shows similar growth profiles

a) Graphical representation of titres of infectious supernatant harvested at 24, 48, 72, 96 and 120 hours from SW13 cells infected with either rHAZV or rHAZV(eGFP) at a MOI of 0.01. Error bars represent a minimum of three technical repeats. Plotted using GraphPad prism. b) Representative plaque assays taken from samples used to calculate data in (a), Images show SW13 cells stained 6 days post infection with either rHAZV or rHAZV(eGFP). A mock infected was included as a negative control.

The location of the eGFP ORF upstream of HAZV N within the S segment encoding cDNA should rule out exclusion of eGFP were it to be expressed on a unique viral genome segment due to the critical requirement of HAZV N to the viral lifecycle. However, due to the error prone nature of the viral RdRp (Steinhauer, Domingo and Holland, 1992), it is possible that mutations could go unchecked within the eGFP ORF, as expression of eGFP is not essential for virus replication. In this case, rHAZV(eGFP) may become capable of infecting cells without expression of functional eGFP, which may result in false negative results following analysis of eGFP derived signals. To determine whether this required addressing, a comparison of eGFP positive cells to HAZV N positive cells was performed. This was achieved by infection of SW13 cells with rHAZV(eGFP) and imaging via immunofluorescence microscopy, targeting HAZV N with antisera

and eGFP using its natural fluorescence. Visualisation of images taken showed all eGFP positive cells contained a HAZV N signal. This suggests that no mutations affecting critical regions of eGFP had occurred (Figure 4.7). However, this experiment was carried out with passage 5 rHAZV(eGFP), which is relatively low and therefore this potential for loss of eGFP signal should be closely monitored following extensive passage of the virus. Further evidence for stability of the insert is shown by the recombinant CCHFV infectious clone expressing ZsG, in which analysis of the ZsG-P2A insertion showed no mutations at passage 10 (Welch *et al.*, 2017). Therefore, as long as the virus is used at low passages, this reporter strain will no doubt prove invaluable to future studies, providing a means of assessing viral infection in an efficient and easy manner.

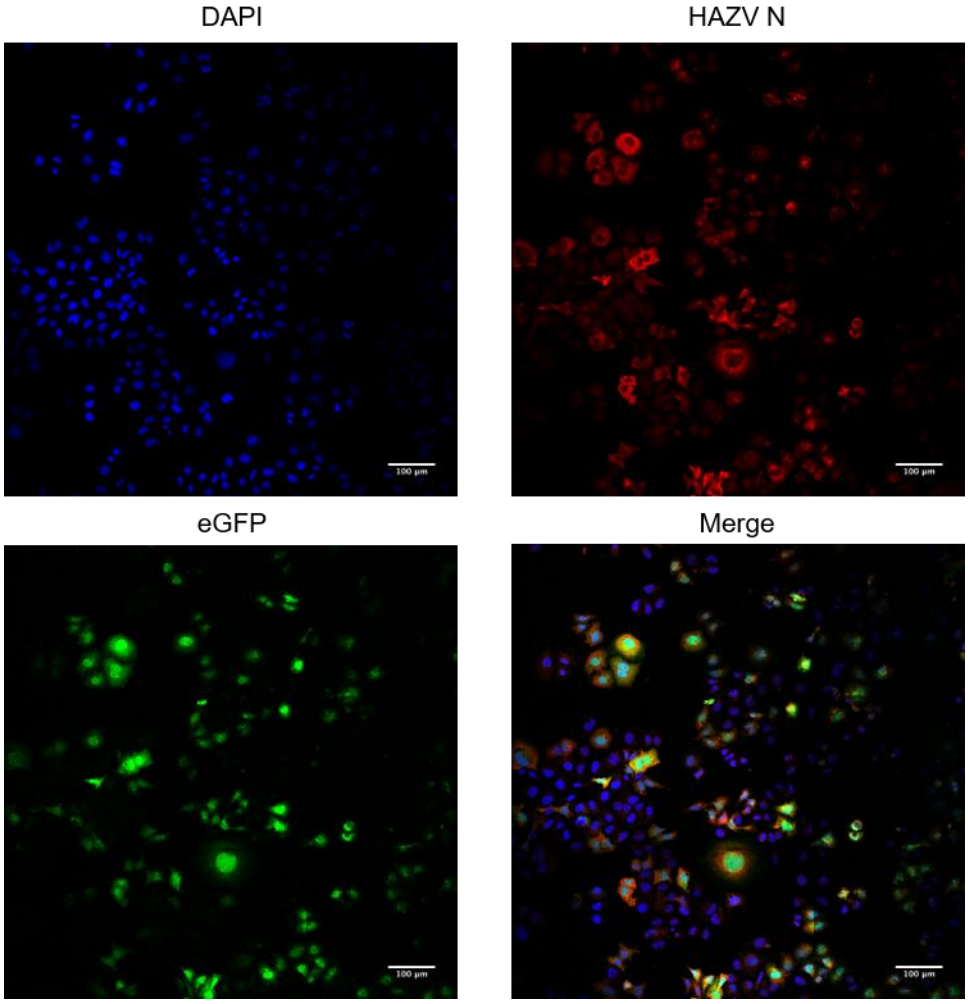


Figure 4.7, Confirming co-expression of eGFP and HAZV N in infected SW13 cells

Tile scan of 5 x 5 immunofluorescent images of rHAZV(eGFP) infected SW13 cells stained using HAZV N antisera, taken on a Zeiss LSM880 confocal microscope using x40 magnification. Nuclei are shown in blue (DAPI), HAZV N is shown in red, eGFP signal is shown in green and a combined merged image is displayed. Scale bar represents 100 µM.

4.5 Growth kinetics of eGFP HAZV in multiple cell lines

At the time of writing, no previous studies have extensively analysed the growth profile of HAZV in multiple cell lines. Given the ability of rHAZV(eGFP) to be used with live cell imaging on the IncuCyte® live cell imaging system, the mammalian cell lines SW13, A549, Huh7 and BHK were infected with rHAZV(eGFP) at a MOI of 0.01 and relative growth was examined via hourly detection of eGFP positive cells for 48 hours. SW13 cells, originating from a human adrenal cortex carcinoma cell line, are commonly described in the literature for HAZV research, with multiple groups reporting data using this cell line (Surtees *et al.*, 2015; Hover *et al.*, 2016; Matsumoto *et al.*, 2019). A549 cells, originating from the adenocarcinomic human alveolar basal epithelium, have been extensively used in study of the BUNV, the model species of the *Bunyavirales order*, having been used to uncover the dependence of this family of viruses on ion channels and cholesterol during the entry stages of infection (Hover *et al.*, 2016; Charlton *et al.*, 2019). Huh7 cells, derived from a human hepatocellular carcinoma, are more commonly used for CCHFV research and were used during recovery of the WT and fluorescent clones for CCHFV (E. Bergeron 2015, S. R. Welch 2017). The BHK cell line, originating from baby hamster kidney cells, was used as it originates from a small mammal, which represent natural hosts of CCHFV during its lifecycle (Spengler, Bergeron and Rollin, 2016). In addition, previous members of our group have achieved the highest titres of virus from the BHK cell line during purification of HAZV, though why this line specifically is able to generate such high titres of virus is unknown (Punch *et al.*, 2018).

Following infection of the indicated cell lines, images were taken hourly and total green cell count per well was analysed on the IncuCyte® live cell imaging system and plotted versus time (Figure 4.8, a, b). As expected, rHAZV(eGFP) was able to replicate efficiently in all four cell lines, as evidenced by detection of eGFP positive cells in all four infections. Despite all cell lines supporting rHAZV(eGFP) growth, the growth profiles varied significantly. In SW13 cells, eGFP was first detected approximately 13 hpi and increased until 48 hpi, with a mean green cell count of 125,621. Interestingly, at 9 hpi, A549 cells provided the shortest time before eGFP detection, in line with previous data in which HAZV translation

products were detected prior to 10 hpi (Punch *et al.*, 2018). However, following a rapid period of growth between 9 and 27 hpi, the total green cell count then gradually began to decrease, from a mean green cell count of 37,906 at 27 hpi to 30,589 at 48 hpi. Further research into the immune response within this cell line may yield answers as to this apparent halt in virus production. The BHK cell line displayed a similar growth profile to the SW13 cells, however despite the initial detection of fluorescence occurring earlier at the 10 hpi timepoint, total growth at the 48 hpi timepoint was lower, with a final mean green cell count of 53,449. This slowed growth in BHK cells versus that observed in the SW13 cell line may provide a longer window in which the cell is able to release virus before CPE is observed, which might provide an explanation to the high titres observed whilst using this cell line. Huh7 cells also permitted earlier detection of eGFP positive cells than the SW13 cells, with signal also detected at 10 hpi. However, despite only varying from the SW13 by 3 hours, the speed of subsequent infections was much faster with a steady increase in detectable green cells occurring up to 48 hpi, resulting in an infected cell count of 127,260.

The utility of rHAZV(eGFP) into facilitating examination of virus growth within multiple cell lines simultaneously with minimal sample processing highlights the potential of rHAZV(eGFP) to be used as a tool to further knowledge into HAZV infection. In a follow up publication to the development of ZsG expressing CCHFV, the CCHFV reporter was used to identify tissue tropism and the early cell targets during an *in-vivo* experimental system (Welch *et al.*, 2019). Similar experiments using rHAZV(eGFP) are feasible and may provide an insight into the differences in pathogenesis between HAZV and CCHFV despite their large similarities in other areas.

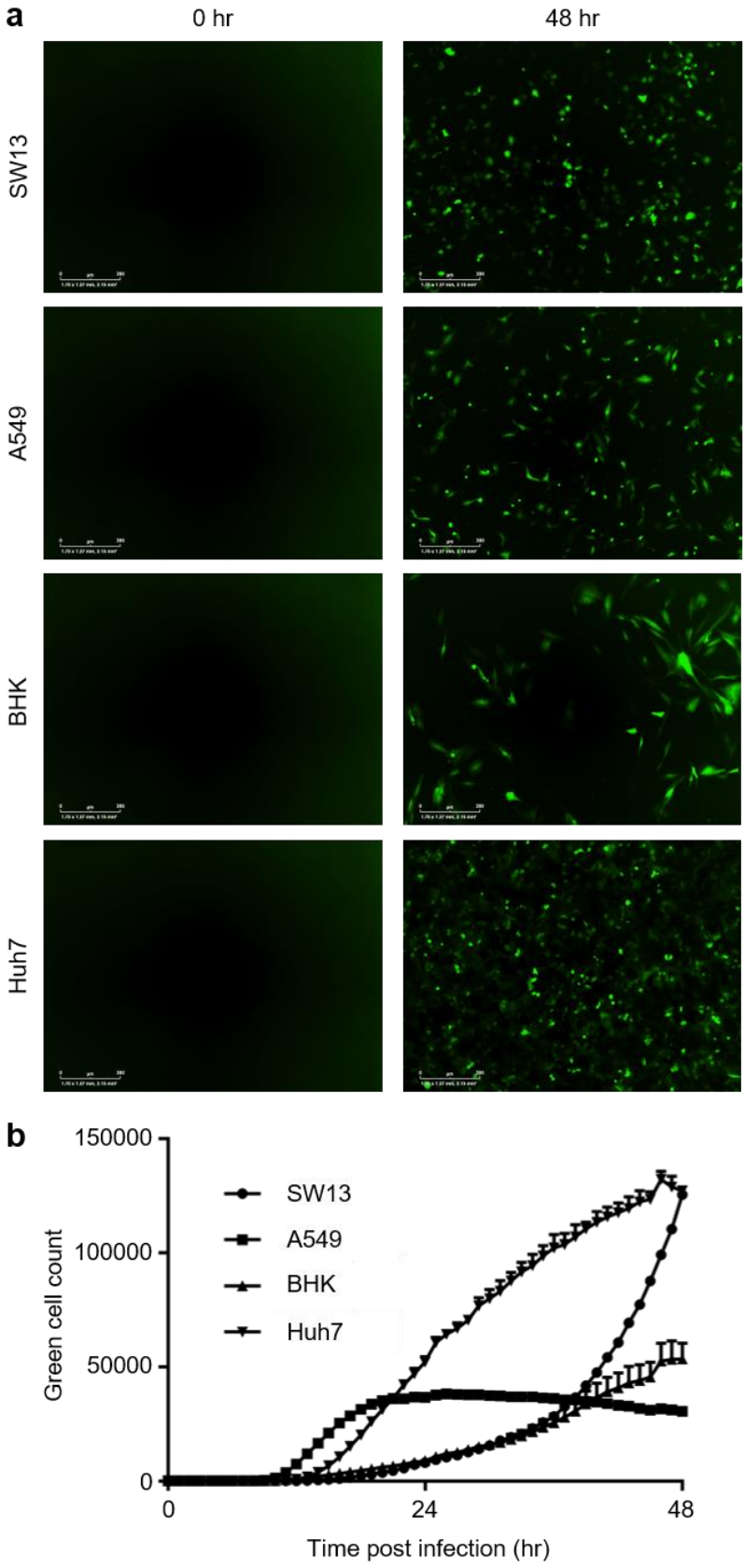


Figure 4.8, Differential growth profiles of rHAZV(eGFP) infection in a range of cell lines

a) Representative images of infected SW13, A549, BHK and Huh7 cell lines by rHAZV(eGFP) at 0.01 MOI 0 and 48 hours post infection taken on the IncuCyte® live cell imaging system. Scale bar represents 300 µM. b) Graphical representation of green cell

count over time during the 48 hour infection of SW13, A549, BHK and Huh7 cell lines with rHAZV(eGFP) at a MOI of 0.01. Analysis was carried out using the IncuCyte® live cell imaging system and plotted using GraphPad Prism.

4.6 Identifying host membrane trafficking factors involved in HAZV multiplication

Following establishment of the reporter system in sections 4.1 to 4.4, rHAZV(eGFP) was used to aid understanding of HAZV infection. In order to gain a greater understanding of the host membrane trafficking factors with importance to the HAZV lifecycle, rHAZV(eGFP) was utilised in a siRNA screen targeting 142 independent genes associated with a range of cellular membrane trafficking factors. Each gene was targeted by three independent siRNA sequences to improve detection reliability. A full list of gene names and the associated siRNA sequences can be found in appendix 1.3. The use of rHAZV(eGFP) over WT rHAZV enabled a more-streamlined workflow, removing the need for fixation, permeabilization and antibody staining of the reverse transfected cells and, should it be required, permitted live cell imaging during the course of the infection.

In order to generate high quality data, the optimal MOI for infection of 96 well plates, in which the screening would take place, was established. SW13 cells were infected at a range of MOIs (0, 0.1, 0.25 and 1) with rHAZV(eGFP) and infection was allowed to progress for 48 hours. Consistent with the previous infections using rHAZV(eGFP), fluorescence was observed in all infected wells, with no detectable fluorescence in the mock infected wells (Figure 4.9, a). As expected, increasing the MOI correlated with increased green cell count, with the highest green cell count in the highest MOI (1) and the lowest green cell count in the lowest MOI (0.10) (Figure 4.9, b). An MOI of 0.25 was chosen as a suitable for the siRNA screen due to the presence of approximately 1500 infected cells at 24 hours post infection. The 24 hour timepoint was used over the later timepoint of 48 hours to minimise the risk of the siRNAs degrading during the infection and also to reduce the stress on the cells during the experiment, preventing them from becoming over confluent. Whilst having more infected cells would give a more reliable output for analysis during the screen, the lower 0.25 MOI was chosen versus the MOI of 1 to conserve low passage virus stocks.

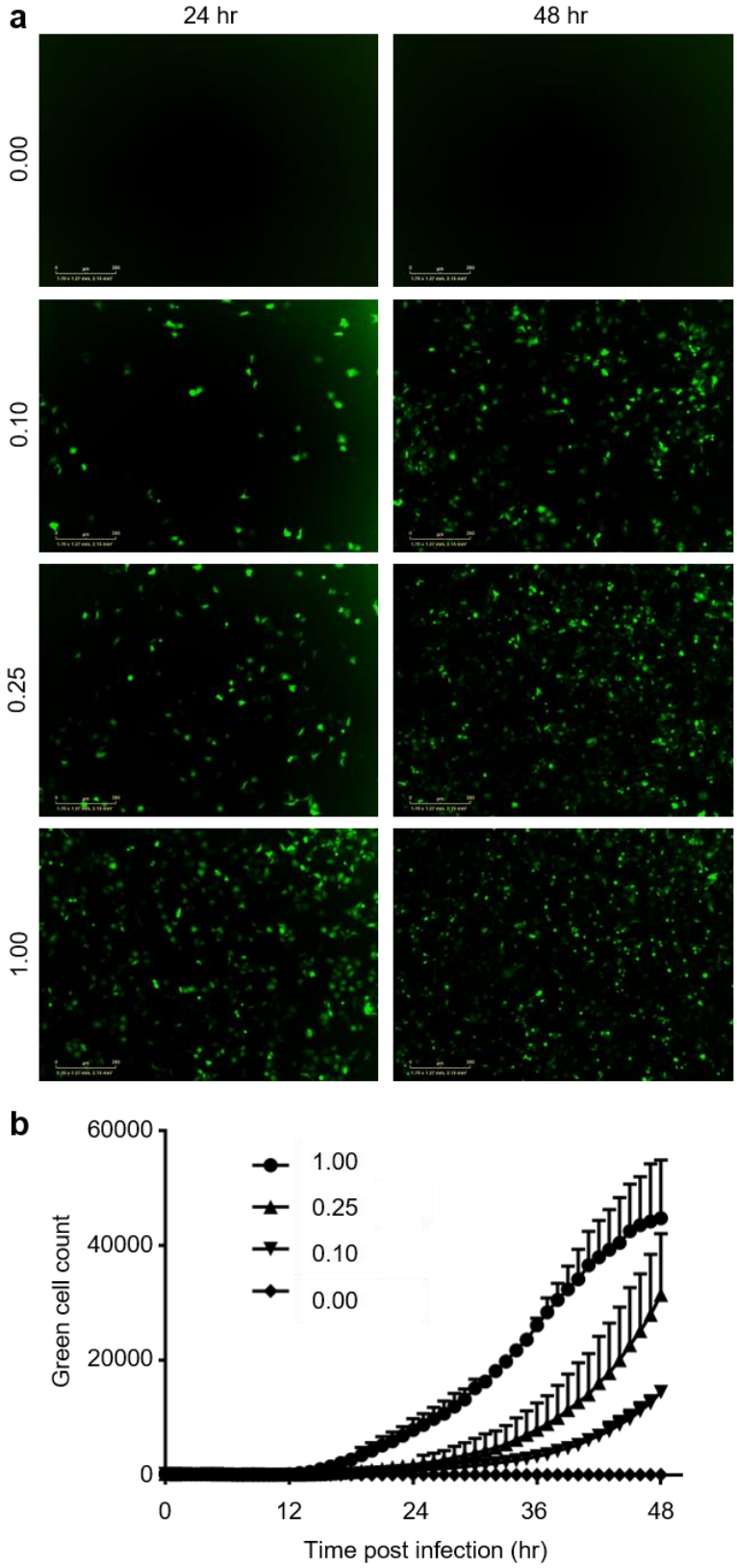


Figure 4.9, Infection at multiple MOIs identifies an optimal MOI of 0.25 for infection of 96 well plates

a) Detection of eGFP signal over a 48-hour infection with rHAZV-eGFP at MOIs of 1, 0.25 and 0.1. A control MOI of 0 was included to rule out non-specific fluorescence. Following infection cells were incubated in the IncuCyte® live cell imaging system at 37℃ and images collected every hour for 48 hours. Scale bar represents 300 µM. b) Graphical

representation of green cell count over time during the 48 hour infection for each MOI. Analysis carried out using the IncuCyte® live cell imaging system software and plotted against time using GraphPad Prism software with error bars representing three independent repeats.

The siRNA screen utilised a reverse transfection protocol, outlined in the materials and methods section 2.2.3.4, in which cells were seeded over the top of the pre-mixed transfection mixture containing OptiMEM, Lipofectamine RNAiMax and a unique siRNA. Suitable controls were determined, including cells infected (Virus) and mock infected (Mock) with rHAZV(eGFP) and no reagents or siRNA, cells infected with rHAZV(eGFP) in the presence of Lipofectamine RNAiMax (Control) and cells infected with rHAZV(eGFP) in the presence of Lipofectamine RNAiMax and a non-specific scrambled siRNA primer (Scrambled). eGFP expression in the control and scrambled tests was found to be insignificant compared to the virus control, and therefore any effect observed with the gene specific siRNAs can reliably be deemed a target specific effect.

Due to the large number of siRNA probes included in the screen, reliable comparisons between the plates were essential. To this end, each plate containing siRNA probes against specific genes also contained a full set of control wells, this enabled a relative value for TIIE to be calculated, which was then averaged across a minimum of three independent experiments. Results were then plotted against the virus control, which was set at a baseline of 1.0, gene specific siRNAs that significantly affected rHAZV(eGFP) were then identified and ranked based on their ability to knockdown the TIIE when averaged across the three unique siRNAs against each gene, with the top 25 hits listed in Table 3.

It is worth noting that the design of rHAZV(eGFP) only permits deductions on stages of rHAZV infection up to and including translation as after this point eGFP is not incorporated into the viral products due to its independent expression. Whilst it is unable to identify the specific stage at which infection is perturbed, it encompasses virus binding, entry into the endosomal pathway, uncoating, fusion, replication and translation phases. Also, the 24 hour duration of the siRNA screen infection phase is limited to detection of a single round of infection, therefore siRNAs targeting genes linked to virus assembly, maturation and exit would not

be identified by this system in any case. Extending the duration of the screen would permit examination of further infection rounds, however information surrounding the stability of the siRNAs in media for longer than 24 hours was unavailable therefore, in combination with the data surrounding the optimal MOI, the decision was made to keep to a 24 hour infection period.

The results of the siRNA screen can be found in full in appendix 1.4. Consistency was largely similar between unique siRNAs targeting the same gene, though there were cases in which one of the three siRNAs displayed a substantial difference in TIIE. In these cases, it highlights the importance of triplicate siRNAs against each gene, with a minimum of 4 experimental repeats.

Initial analysis of the results identified a number of families of genes that either affected rHAZV(eGFP) or played no role in its infection. Families of genes such as syntaxins, coatamer protein complex I (COPI) components, vesicle associated membrane proteins (VAMPs) and synpatogamins all demonstrated significant knockdown of TIIE following infection with rHAZV(eGFP). Conversely, other gene families displayed little to no effect on the TIIE caused by infection. Interestingly, despite the previously described entry mechanism of CCHFV as clathrin mediated endocytosis (Garrison *et al.*, 2013), no effect on TIIE was observed following siRNA knockdown of clathrin light or heavy chain components. Similarly, no effect was observed following knockdown of the clathrin adaptor complex 1 and 3.

The top five genes that displayed the largest effect on rHAZV(eGFP) infection, measured via TIIE, were COPA (26.04 %), COPB2 (56.69 %), COPB1 (70.67 %), AP2B1 (70.72 %), VAMP1 (72.21 %) and FAM62B (72.23 %) (Table 3). With three of the top five hits linked to COPI components, it was decided to pursue this family of genes further. Whilst the other gene families may well represent interesting avenues of investigation, due to time constraints they were not pursued further than the initial screen.

Gene ID	siRNA 1	siRNA 2	siRNA 3	Avg. all
COPA	31.18%	24.13%	22.82%	26.04%
COPB2	43.03%	23.72%	103.31%	56.69%
COPB1	68.28%	80.33%	63.40%	70.67%
AP2B1	65.15%	63.45%	83.57%	70.72%
VAMP1	77.16%	66.05%	73.43%	72.21%
FAM62B	62.91%	82.48%	71.30%	72.23%
VPS13B	89.40%	51.05%	77.82%	72.76%
STX18	68.75%	83.03%	69.71%	73.83%
VPS33B	83.04%	81.98%	57.82%	74.28%
EHD2	59.58%	87.51%	76.92%	74.67%
ARHGAP10	74.70%	68.96%	81.25%	74.97%
AP2A2	80.49%	75.21%	69.78%	75.16%
STX17	67.27%	87.59%	78.28%	77.71%
SYT17	89.74%	73.37%	74.82%	79.31%
COPZ2	64.69%	80.32%	93.43%	79.48%
STXBP1	83.38%	59.11%	97.98%	80.16%
COPG1	45.03%	93.74%	102.70%	80.49%
STX1B	83.26%	88.50%	70.07%	80.61%
HUWE1	71.20%	71.90%	99.66%	80.92%
VAMP7	84.95%	69.97%	88.39%	81.10%
COPE	70.86%	90.84%	82.79%	81.50%
COPG2	80.33%	81.40%	87.50%	83.08%
SYT3	70.47%	85.51%	93.55%	83.18%
AP1S3	107.01%	93.64%	49.21%	83.28%
FAM62C	62.01%	86.58%	103.29%	83.96%

Table 3.1, Table of selected siRNA targets resulting in highest knockdown of total integrated intensity of eGFP signal

Table format of the top 25 hits from the siRNA membrane trafficking screen, ranked via percentage of total integrated intensity of eGFP versus the control. Each percentage for the unique siRNAs (1-3) represents an average of four independent experimental repeats. The average column represents the average across the three unique siRNAs.

4.7 COPI components are important for HAZV infection

COPI complexes are primarily involved in retrograde transport from the Golgi apparatus to the ER, however they have also been shown to have roles in anterograde transport, endosomal transport and expression of cell surface receptors and lipid composition (Aniento *et al.*, 1996; Misselwitz *et al.*, 2011; Wang *et al.*, 2015). They are formed from seven subunits, α, β, β′, γ, δ, ε, and ζ. In addition to vesicular transport between the Golgi and ER, during retrograde transport, COPI complexes can also form tubular structures, which have been linked to anterograde transport between Golgi cisternae (Yang *et al.*, 2011). The role of COPI components in the endosomal pathway has also been proven. The β and ε subunits were found to be present on endosomes in a pH dependent manner and acted to regulate endosomal membrane transport (Aniento *et al.*, 1996). The distribution of cholesterol, a lipid required by HAZV and BUNV during endosomal escape, and the Rho GTPases Cdc42 and Rac1, can be influenced by COPI components, showing the broad range of roles COPI components is not limited simply to retrograde transport within the cell (Misselwitz *et al.*, 2011; Charlton *et al.*, 2019).

COPI vesicles are distinct from COPII vesicles, which are responsible for anterograde transport from the ER to the Golgi apparatus (Figure 4.10). COPI vesicles form upon detection of a KKXX or KDEL motif, for membrane bound and soluble proteins respectively. The GPC of HAZV encodes five such KKXX motifs, four of which are located in the coding region for Gc (KKIS, KKIC, KKFS and KKDA) with two each side of the predicted transmembrane domain, placing them in a potentially accessible position for COPI vesicle formation. It is unknown whether COPI binds directly or indirectly to the KKXX motif, however the resultant high local abundance of COPI causes curvature of the Golgi membrane, which eventually pinches off forming a vesicle.

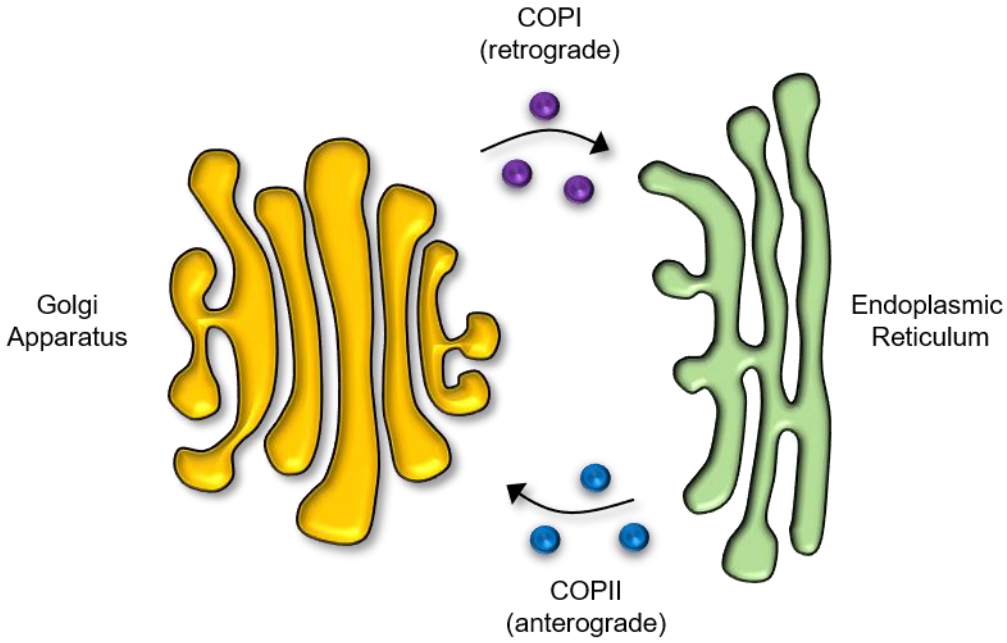


Figure 4.10, COPI and COPII vesicle schematic

COPI vesicles function to transport cargo from the Golgi apparatus to the ER, known as retrograde transport, whilst COPII vesicles transport cargo from the ER to the Golgi apparatus, known as anterograde transport

Analysis of TIIE signal identified multiple genes encoding COPI subunits as important to viral infection. COPA, COPB1, COPB2 and COPE all resulted in a significant knockdown of eGFP signal in at least two of the three siRNAs against each gene (Figure 4.11). siRNA knockdown of COPA had the largest effect on eGFP signal with all three unique siRNAs against COPA displaying ~70 to 80 % reduction versus the untreated control (siRNA 1 = 0.311 (p=0.0003), siRNA 2 = 0.241 (p=<0.0001), siRNA 3 = 0.228 (p=0.0001)). COPB1 knockdown also caused significant reduction of eGFP signal with all three siRNAs, but to a lesser degree, displaying ~20 to 40 % reduction versus the untreated control (siRNA 1 = 0.682 (p=0.0127), siRNA 2 = 0.803 (p=0.0078), siRNA 3 = 0.634 (p=0.0008)). COPB2 knockdown reduced eGFP signal significantly in 2 of 3 siRNAs (siRNA 1 = 0.430 (p=0.0289), siRNA 2 = 0.237 (p=0.0009), siRNA 3 = 1.033 (p=0.6497)), as did COPE knockdown ((siRNA 1 = 0.709 (p=0.0043), siRNA 2 = 0.908 (p=0.2701), siRNA 3 = 0.828 (p=0.0146)).

Conversely those genes associated with COPII vesicles were not as critical to rHAZV(eGFP) infection, with only SAR1A knockdown having a statistically significant effect on TIIE signal in two of three siRNAs (siRNA 1 = 0.893

(p=0.1153), siRNA 2 = 0.817 (p=0.0256), siRNA 3 = 0.873 (p=0.0123), though even with this significance, the knockdown effect was much less at ~10 to 20 % reduction versus the untreated control (Figure 4.12).

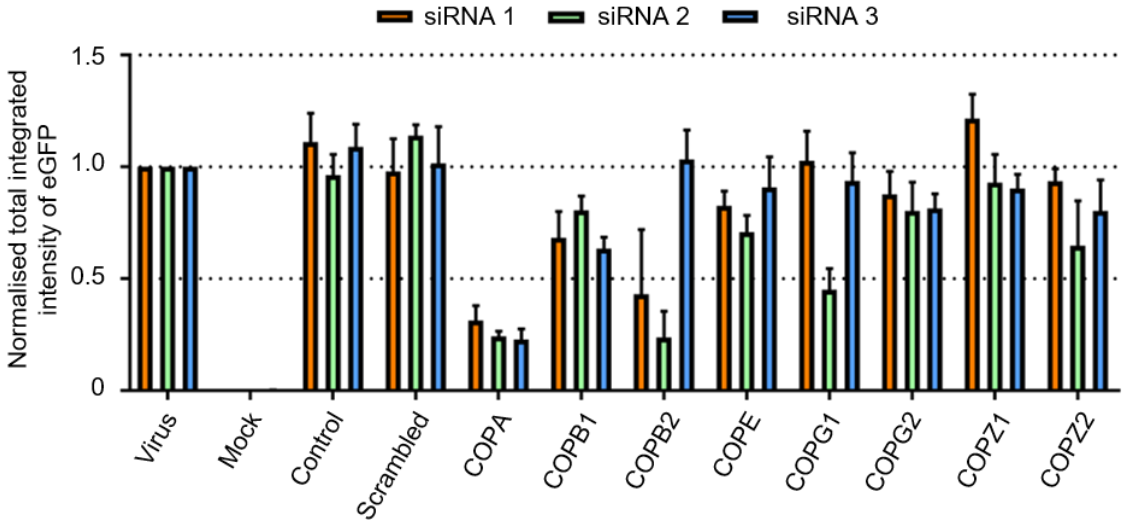


Figure 4.11, COPI associated genes are important for HAZV infection

Histogram of the relative fold change in total integrated intensity of eGFP signal in SW13 cells infected with rHAZV(eGFP) at a MOI of 0.25 following independent reverse transfection of triplicate siRNAs (siRNA 1 – orange bars, siRNA 2 - green bars, siRNA 3 - blue bars) against the subunits of the coatomer protein complex 1 (COPI). Transfection reagent only and scrambled siRNA controls were included for reference. Error bars represent data from 4 experimental repeats. Plotted using GraphPad Prism.

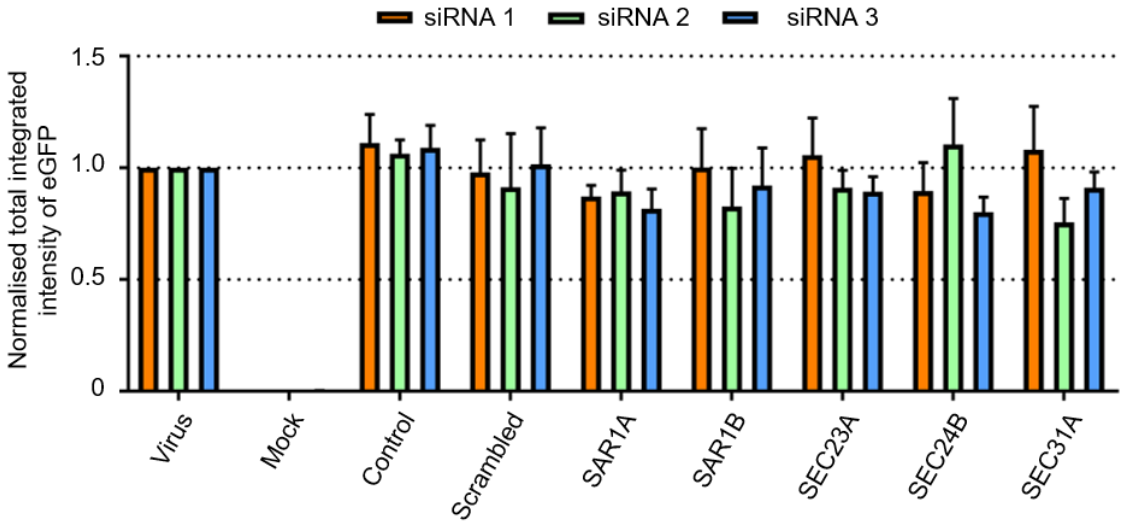


Figure 4.12, COPII associated genes are not required for HAZV infection

Histogram of the relative fold change in total integrated intensity of eGFP signal in SW13 cells infected with rHAZV(eGFP) at a MOI of 0.25 following independent reverse transfection of triplicate siRNAs (siRNA 1 – orange bars, siRNA 2 - green bars, siRNA 3 - blue bars) against the subunits of the coatomer protein complex 2 (COPII). Transfection reagent only and scrambled siRNA controls were included for reference. Error bars represent data from 4 experimental repeats. Plotted using GraphPad Prism.

The finding that knockdown of many components of the COPI vesicle negatively impacted TIIE was a strong indicator as to the importance of these factors in the viral lifecycle. Before conclusions could be made in respect of this, validation of the screen was required. Confirming knockdown of virus gene expression alongside knockdown of the genes targeted by the siRNAs was carried out using qPCR and western blotting. The siRNAs against COPA, COPB1 and COPB2 were selected for validation as they represented the three highest hits on the screen. In addition, the α-COP in combination with β′-COP and ε-COP form a subcomplex that forms the backbone of the COPI cage, therefore by validating the role of the core structure of the vesicle, the role of the additional COPI subunits also gain credibility. To enable validation, a larger volume of one of the triplicate siRNAs for each of COPA, COPB1 and COPB2 was purchased from the same supplier as the siRNA membrane trafficking library, with the same sequence as used in the screen. This enabled knockdown of each target in a larger format that would facilitate the validation experiments. As rHAZV(eGFP) had identified a smaller pool of target siRNAs, the validation was carried out with WT rHAZV to rule out eGFP as a contributing factor to COPI linked inhibition of HAZV. To ensure the siRNAs were successfully knocking down their target genes, qPCR was also used to target COPA, COPB1 and COPB2 in treated and untreated cells following knockdown of the COPI specific targets in SW13 cells. As expected, all three siRNAs displayed knockdown of the target genes versus the untreated controls (Figure 4.13, a).

Validation of the effect of the siRNA knockdown on rHAZV was then assessed. SW13 cells were infected with WT rHAZV according to the same timings and concentrations used in the initial screen. After 24 hpi the supernatant was removed and the cells lysed in Trizol reagent to maintain the integrity of RNA in the lysate whilst solubilising cellular components. RNA was subsequently extracted via a phenol chloroform extraction and analysed via qPCR using the One step MESA GREEN qRT-PCR kit. Primers used in the development of the qPCR system for rHAZV, detailed in Chapter 3, were used to detect the S segment of HAZV as an indicator of virus infection with primers against GAPDH also included to permit normalization between replicates. Analysis of the normalized data confirmed the findings of the screen, with COPA, COPB1 and

COPB2 knockdown all resulting in significant decreased copy number of HAZV S segment RNA versus an untreated control (Figure 4.13, b). Further validation of the siRNA screen and qPCR results was achieved via western blotting of cell lysates collected following the same experimental procedures as the siRNA screen (Figure 4.14, a). As expected, knockdown of COPA, COPB1 and COPB2 resulted in reduced expression of HAZV N, however to a much lesser extent than seen at the transcriptional level (Figure 4.14, b).

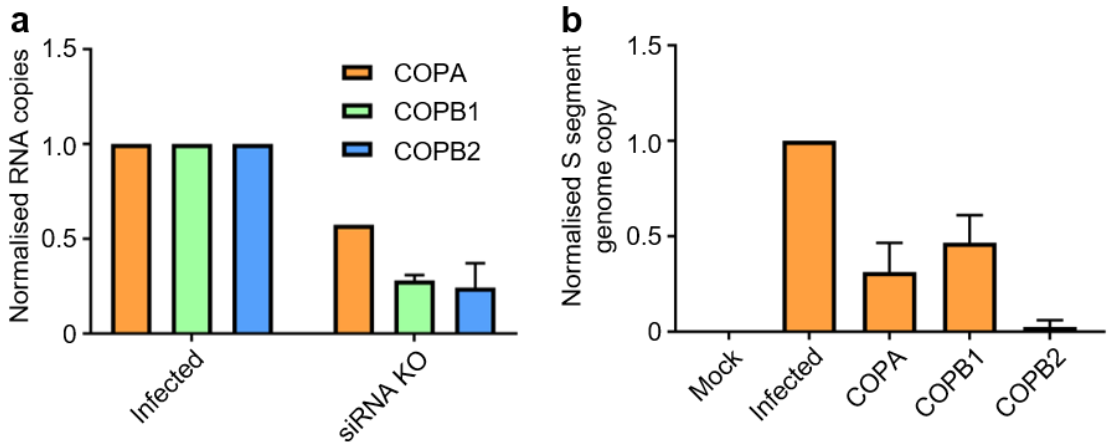


Figure 4.13, qPCR confirmation of siRNA screen results using WT rHAZV

a) Normalised HAZV S segment copy number isolated from SW13 cells transfected with siRNAs targeting COPA, COPB1 and COPB2 then infected with HAZV at 0.1 MOI. Error bars represent data from 2 experimental repeats. Plotted using GraphPad Prism. b) Normalised COPI subunit RNA copy numbers isolated from SW13 cells previously transfected with siRNAs targeting COPA, COPB1 and COPB2. Error bars represent data from 2 experimental repeats. Plotted using GraphPad Prism.

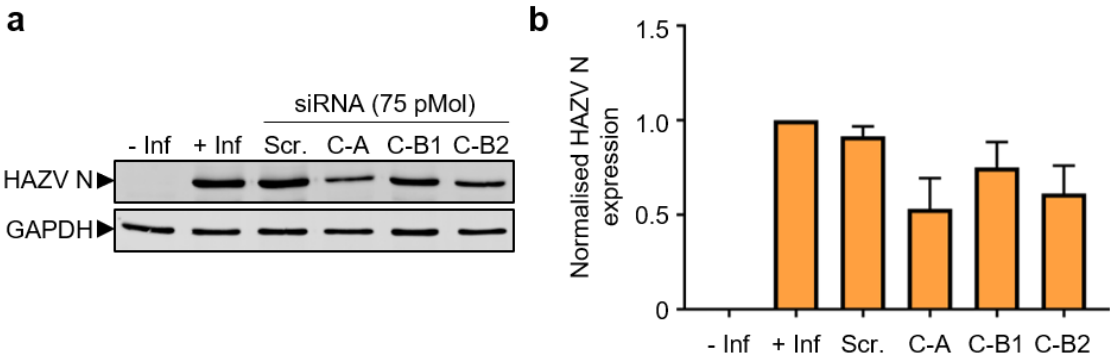


Figure 4.14, Western blot analysis shows siRNA knockdown of COPI components reduces HAZV N translation

a) Western blot analysis of lysates collected from SW13 cells treated with siRNAs targeting COPA (C-A), COPB1 (C-B1), COPB2 (C-B2), a scrambled control (Scr.) was also included alongside a mock infected (- Inf) and infected (+ Inf) sample in which transfection reagent but no siRNA was added to both. Samples were probed for HAZV N and GAPDH as a loading control. b) Histogram displays normalised densitometric

analysis of western blot data from (a), calculated using the ImageJ software package, error bars represent data from 3 experimental repeats. Plotted using GraphPad Prism.

4.8 Identifying the role of COPI in HAZV infection

Due to the previously described dependence on membrane cholesterol by HAZV and the evidence that COPI components can influence the distribution of cholesterol on the plasma membrane, the role of COPI components in HAZV entry was first explored (Misselwitz *et al.*, 2011; Charlton *et al.*, 2019). A role for COPI complexes at this early stage of the lifecycle was supported by the data in Figure 4.13 and Figure 4.14, which revealed both mRNA transcription and translation were both affected by COPI knockdown, indicating a stage prior to these events, thus implicating virus entry. To confirm this, a 24 hour infection was set up using rHAZV(eGFP) to infect SW13 cells in which different COPI components had been knocked down, with eGFP expression measured at hourly intervals using the IncuCyte® live cell imaging system. SW13 cells were treated with either a scrambled siRNA or COPI specific siRNA according to the same protocol used in the initial siRNA screen and subsequent validations. Analysis of green cell count over time revealed initial detection of eGFP signal in the controls and COPB1 knockdown cells to occur at approximately 12 to 13 hpi (Figure 4.15, a), in line with previous observation for SW13 cells in section 4.5. In addition to the total green cell count in COPA, COPB1 and COPB2 knockdown cells being significantly lower at 24 hpi, detection of eGFP signal in COPA and COPB2 samples occurred later, at 15 to 16 hours post infection, suggesting the knockdown of these genes to have delayed a phase in the viral lifecycle at a pre-translation stage, or reduced the effectiveness of viral translation, leading to an increased duration until detection of eGFP. This is visible by eye when looking at images taken at 14 hours into the infection, with green cells visible in wells containing the controls and COPB1 knockdowns, but not visible in those wells containing COPA or COPB2 knockdowns (Figure 4.15. b). When visualising the same location in those wells at 24 hours, green cells become visible in all wells, with reduced numbers in the COPA and COPB2 wells versus the controls. This apparent disruption to viral infection by COPI has previously been documented for different pathogens. In vesicular stomatitis virus (VSV), an enveloped, negative sense, ssRNA virus, depletion of the COPI component γ-COP resulted in reduced VSV invasion (Misselwitz *et al.*, 2011). This same study also identified

the mechanism of VSV dependence on COPI to be linked to the role of COPI in controlling the lipid composition of the membrane, specifically of sphingolipid GM1 and cholesterol. Cholesterol abundance within the plasma membrane has also been documented to affect HAZV among other viruses such as Semliki Forest virus, Andes virus and influenza virus (Phalen and Kielian, 1991; Kleinfelter *et al.*, 2015; Chlanda *et al.*, 2016). The role of cholesterol in HAZV entry was shown to be linked to formation of a K^+ rich environment within endosomes that facilitated escape into the cytoplasm, demonstrated via depletion of cholesterol and K^+ priming of HAZV virions in cholesterol depleted cells (Charlton *et al.*, 2019). Therefore, one possible explanation of the dependence on COPI by HAZV in early stages of the life cycle is as a result of the native role of COPI in regulating cholesterol distribution.

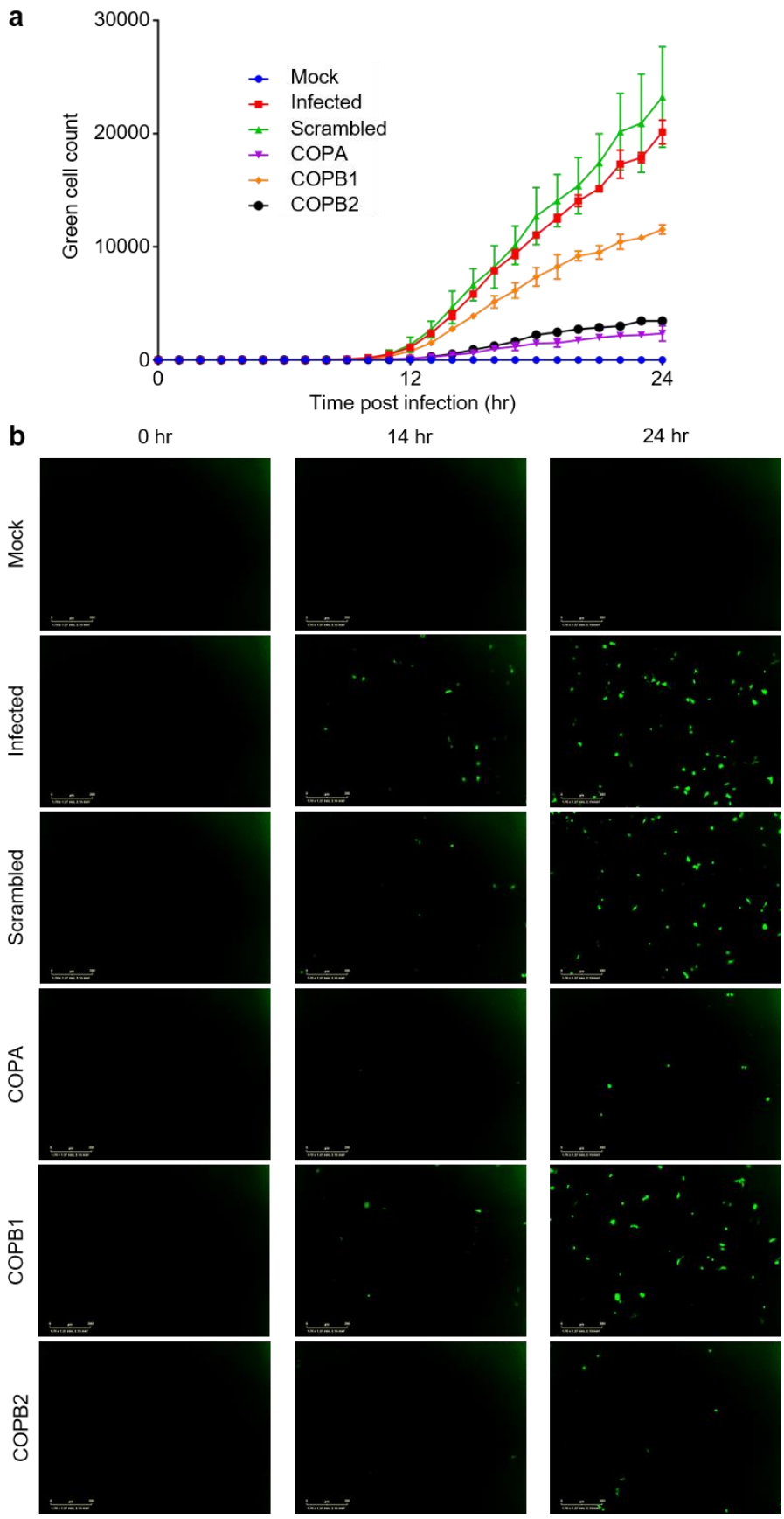


Figure 4.15, COPI knockdown delays translation of HAZV N

a) Total green cell count for each siRNA knockdown condition was determined using the IncuCyte® live cell imaging system software and plotted against time using GraphPad

Prism software with error bars representing two independent repeats. B) Representative images of siRNA treated infected SW13 cells by rHAZV(eGFP) at 0, 14 and 24 hours post infection taken on the IncuCyte® live cell imaging system. Scale bar represents 300 µM

If the role of COPI in HAZV entry was linked to its native role in cholesterol distribution, then infection with rHAZV should not require COPI to become redistributed compared to uninfected cells. To confirm this was the case, SW13 cells were either mock infected or infected with rHAZV(eGFP) at an MOI of 0.1 and imaged at 24 hpi using a confocal microscope. Analysis of eGFP, as an indicator of infected cells, and COPA demonstrated no significant difference in distribution of COPA between mock and infected cells, indicating HAZV was not hijacking COPI vesicles from their native cellular location (Figure 4.16).

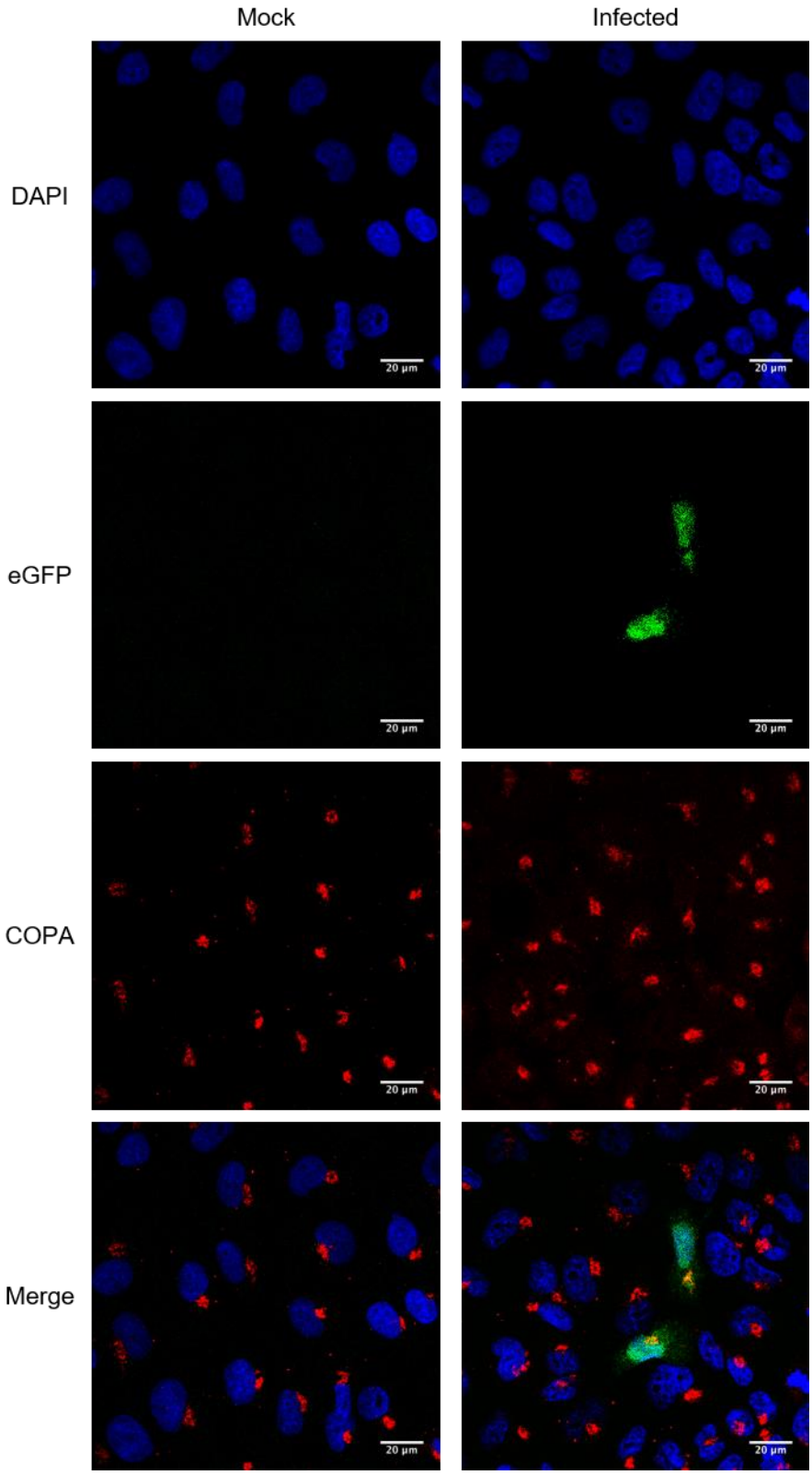


Figure 4.16, rHAZV(eGFP) infection does not cause redistribution of COPA

Immunofluorescent images of rHAZV(eGFP) infected SW13 cells taken on a Zeiss LSM880 confocal microscope using x40 magnification. Nuclei are shown in blue (DAPI), eGFP signal is shown in green, HAZV N is shown in red and a combined merged image is displayed. Scale bar represents 20 µM.

To assess if the delay observed in translation of HAZV N had an impact on generation of infectious virions, titration of rHAZV collected from supernatant of infected SW13 cells, following knockdown of COPA, COPB1 and COPB2, was carried out. It was expected that due to the observed decreases in viral RNA (Figure 4.13) and translation products (Figure 4.14) all occurring within the range of a 10 fold reduction following a 24 hour infection, that a similar result would be observed with virus titre. As before, the experimental procedure followed the same concentrations and timings as previous experiments. Analysis of the number of infectious viruses present in the supernatant revealed a much higher decrease than initially expected. In the infected and scrambled controls, virus titre was estimated in the range of 10^4 copies per mL, versus titre in the range of 10^2 for COPA, COPB1 and COPB2, representing a decrease of almost 100 fold following COPI knockdown (Figure 4.17, a, b).

Comparison of the observed 2 fold change in N protein expression with the 100 fold change in released infectious virus following COPA knock down strongly points towards COPA involvement in virus assembly. It is plausible that COPI vesicles also affect an additional part of the HAZV lifecycle. As mentioned in section 4.7, the HAZV glycoprotein Gc contains multiple KKXX motifs that can be recognised by COPI vesicles and therefore have the potential to interact with COPI at a post translational stage. Further evidence for this comes from CCHFV. CCHFV glycoproteins are extensively processed between the ER and Golgi apparatus (Vincent *et al.*, 2003; Haferkamp *et al.*, 2005; Sanchez *et al.*, 2006), a detailed description can be found in section 1.4.2.1,. In a separate study, CCHFV Gc was identified to colocalise with β-COP, at the time being used as a marker for the Golgi and was therefore not pursued further at the time (Erickson *et al.*, 2007). It is possible that COPI components perhaps play a more significant role in HAZV glycoprotein processing or transport during infection, however due to the unavailability of a Gc specific antibody, this hypothesis could not be confirmed experimentally.

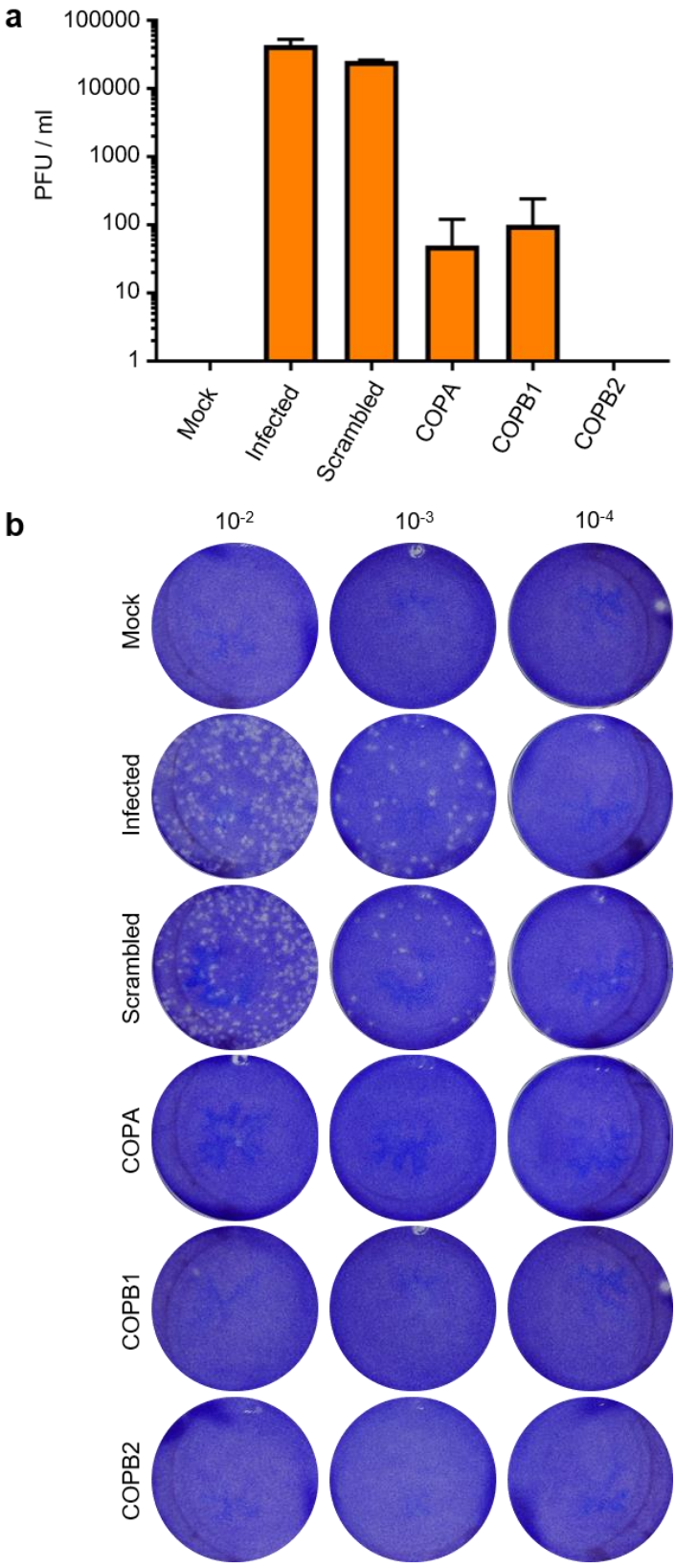


Figure 4.17, COPI knockdown results in reduced infectious virus production

a) Histogram displays plaque forming units per mL (PFU / mL) of rHAZV released into supernatant following siRNA knockdown of COPI components COPA, COPB1 and COPB2 in SW13 cells. Both mock infected (Mock) and infected controls were included, alongside a scrambled siRNA control. Plotted using GraphPad Prism software, error bars

represent 2 independent repeats. b) Representative plaque assays used to determine the titre in part (a). SW13 cells were infected and fixed then stained 6 days post infection.

As multiple roles of COPI in the HAZV lifecycle now seemed more likely, these alternate mechanisms were further explored. Multiple viruses have been shown to utilise COPI vesicles in formation of virus replication factories. Echo virus 11 (EV11) infection of cells treated with BFA, a compound that blocks formation of COPI vesicles via inactivation of ADP-ribosylation factor (ARF) 1, led to a block in virus replication and RNA synthesis. Further studies also revealed colocalisation of sites of EV11 RNA synthesis with β-COP early in infection (Gazina *et al.*, 2002). In other viruses, a viral replication factory dependence on COPI but not COPII components has been documented. Depletion of COPI components inhibited formation of virus replication factories for both Drosophila C virus and polio virus, an effect not observed with COPII knockdown (Belov, Fogg and Ehrenfeld, 2005; Cherry *et al.*, 2006). The prototypic bunyavirus, BUNV, is able to hijack the Golgi apparatus to form viral factories around the Golgi complex, incorporating Golgi stacks, mitochondria and rough ER elements (Fontana *et al.*, 2008). Therefore, given the ability of COPI components to form tubular and vesicular structures in proximity of the Golgi, it is plausible HAZV and CCHFV may utilise COPI vesicles in virus replication factory formation.

As it was observed that the natural distribution of COPA, and by extension COPI vesicles, was unchanged by rHAZV(eGFP) infection (Figure 4.16), whether HAZV forms replication factories around the Golgi was explored by using the COPA subunit as a marker. Viral components are normally concentrated in virus factories, due to the increased efficiency of replication afforded by having essential cellular components in a localised area (Novoa *et al.*, 2005). Therefore, SW13 cells were infected with rHAZV at a 0.1 MOI and following a 24 hour infection, were stained for COPA and HAZV N then examined by immunofluorescence microscopy (Figure 4.18, a). It was hypothesised that a replication factory would have a high local concentration of HAZV N due to its association with RNP segments. However, confocal imaging of COPA and HAZV N localisation within SW13 cells was inconclusive. Line scans of COPA and HAZV N staining were analysed (Figure 4.18, b, i-iii) and suggest colocalisation in some infected cells, as indicated by coinciding peaks in COPA and HAZV N

signal (Figure 4.18, b, i). However, line scans in other cells suggested that no such colocalisation occurred (Figure 4.18, b, ii) thus if an interaction between HAZV N and COPA is required for virus growth, it may be transient, and only occurring at a specific stage of the HAZV replication cycle. Based on the evidence provided by the confocal microscopy it is unlikely COPA is involved in formation of HAZV replication factories.

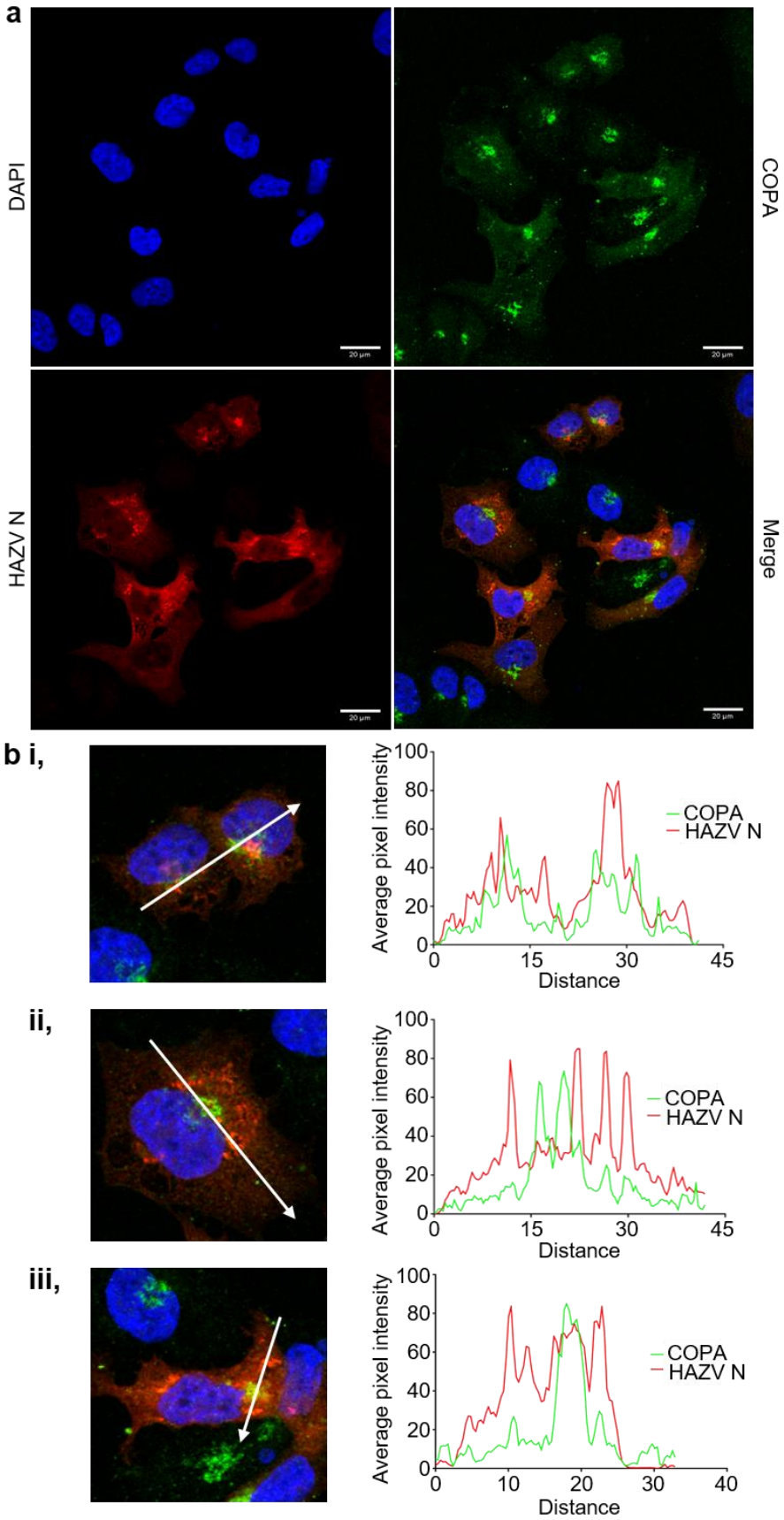


Figure 4.18, Confocal imaging of COPA and HAZV N localisation in SW13 cells

a) Immunofluorescent images of WT rHAZV infected SW13 cells taken on a Zeiss LSM880 confocal microscope using x40 magnification. Nuclei are shown in blue (DAPI),

COPA signal is shown in green, HAZV N is shown in red and a combined merged image is displayed. Scale bar represents 20 µM. b, i-iii) intensity profiles of COPA and HAZV N signal along the white arrows indicated on the respective confocal panel. Plotted using GraphPad Prism software.

4.9 Dependence on COPI vesicles by HAZV is ARF1 and GBF1 independent

COPI vesicle formation is a highly regulated and complex process. Initially COPI subunits are recruited to the Golgi membrane via the action of an Arf, a GTPase. Arf can be present in a GDP or GTP bound state, regulated via the action of guanine exchange factors (GEFs), which convert it to the GTP form, and GTPase activating proteins (GAPs), that return it to its GDP bound state (Donaldson and Jackson, 2000; Casanova, 2007; Inoue and Randazzo, 2007). Arf-GEFS therefore play a key role in the temporal regulation and distribution of Arf on the Golgi membrane, facilitated by a key Sec7 catalytic domain that enables the nucleotide transfer (Jackson and Casanova, 2000). GBF1 is a high molecular weight Arf GEF that acts on class I and II Arf proteins and is found localised in early Golgi compartments (Claude *et al.*, 1999; Kawamoto *et al.*, 2002). Once in its GTP bound state, ARF1 is able to recruit the coatomer to the membrane, where the KKXX motif is then able to associate with either the α-COP or β'-COP. Therefore, GBF1 plays a key role in the recruitment of ARF1 and subsequent formation of COPI vesicles (Figure 4.19).

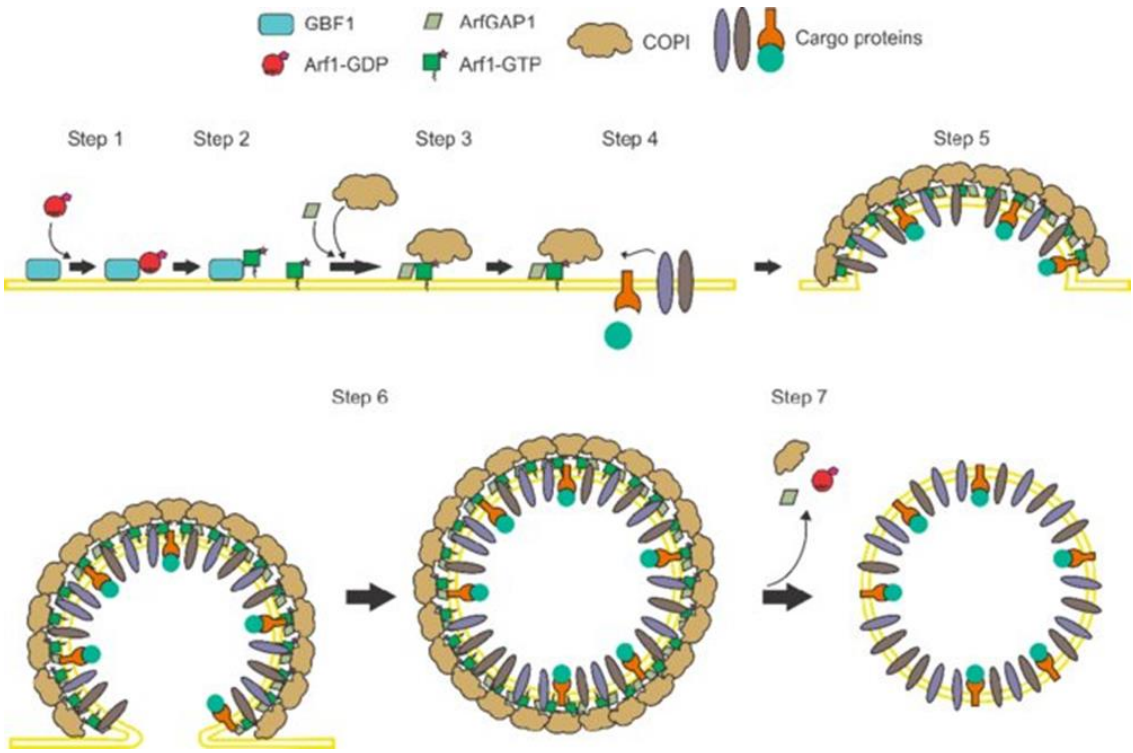


Figure 4.19, Role of Arf1 and GBF1 in COPI vesicle formation

Formation of COPI vesicles is initiated by the interaction of Arf1-GDP with GBF1 (step 1). This interaction results in the activation of Arf1 by catalysing the exchange of GDP to GTP, exposing a myristoyl group that enables association of the protein with the membrane (step 2). Membrane bound Arf1-GTP recruits the pre-formed COPI coat (step 3) which stimulates the recruitment and concentration of cargo (step 4), inducing curvature of the membrane into a vesicle (step 5). The formation of the vesicle completes with budding from the Golgi membrane (step 6). Vesicles are disassembled following hydrolysis of Arf1-GTP into Arf1-GDP by ArfGAP1 (step 7)

Figure adapted from (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7354614/)

Both ARF1 and GBF1 were targeted by siRNAs during the membrane trafficking screen carried out in section 4.6. Due to the demonstrated dependence of HAZV on multiple subunits of the COPI complex, it was expected that ARF1 and GBF1 would also demonstrate a similar effect. However, analysis of the TIIE following siRNA knockdown of ARF1 and GBF1 demonstrated no significant effect (Figure 4.20) indicating that the use of COPI vesicles during the early phases of HAZV infection is in an ARF1 / GBF1 independent manner. However, it is important to remember that rHAZV(eGFP) in conjunction with the siRNA screen is only able to provide information up to and including viral translation. For CCHFV, translation of the GPC occurs on ER resident ribosomes, in contrast to the cytoplasmic ribosomes thought to translate N and the RdRp (Sanchez, Vincent

and Nichol, 2002). Therefore, it is likely that virus translation of the GPC and eGFP, expressed concurrently with HAZV N, occur in discrete locations to each other. The GPC, via a complex series of processing stages described in detail in section 1.4.2.1, is processed into the structural and non-structural glycoproteins. This occurs in the Golgi, where it is proposed CCHFV also assembles (Bertolotti-Ciarlet *et al.*, 2005; Sanchez *et al.*, 2006). As COPI vesicle assembly occurs in the Golgi apparatus, it is likely that any effect on the virus by ARF1 and GBF1 would not be observed until the viral lifecycle required the Golgi, i.e. during glycoprotein processing.

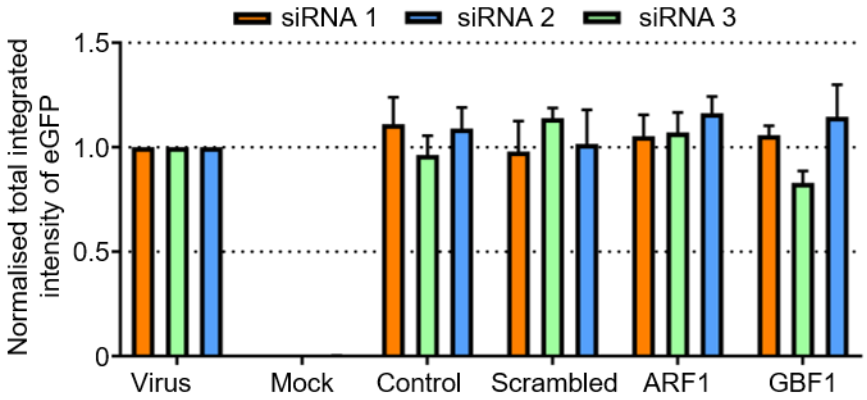


Figure 4.20, ARF1 and GBF1 are non-essential to rHAZV(eGFP) infection

Histogram of the relative fold change in total integrated intensity of eGFP signal in SW13 cells infected with rHAZV(eGFP) at a MOI of 0.25 following independent reverse transfection of triplicate siRNAs (siRNA 1 – orange bars, siRNA 2 - green bars, siRNA 3 - blue bars) against ARF1 and GBF1. Transfection reagent only and scrambled siRNA controls were included for reference. Error bars represent data from 4 experimental repeats. Plotted using GraphPad Prism.

Further validation of the ARF1 and GBF1 independent nature of HAZV COPI dependence at a pre-translational stage was achieved via the use of BFA. BFA is a non-competitive inhibitor of GBF1 that functions via formation of an abortive ARF1-GBF1-BFA complex, preventing formation of COPI vesicles by stabilising GBF1 on Golgi membranes (Niu *et al.*, 2005). Therefore, to validate the siRNA data in Figure 4.20, BFA was used to treat A549 cells prior to, and during, infection with WT rHAZV. Due to previous data supporting a role of COPI vesicles in influenza A virus (IAV) infection, IAV was used as a positive control to ensure BFA was functional. Western blot analysis of lysates collected following infections of the BFA treated A549 cells with WT rHAZV and rIAV displayed no significant reduction in HAZV N, whilst IAV N, as expected, was significantly reduced in a

dose dependent manner (Figure 4.21, a, b). A549 cells were chosen due to their ability to support replication of both rHAZV and rIAV, removing the need for multiple cell lines in the comparison. To rule out toxicity arising from BFA treatment of the A549 cells, an MTS viability assay was carried out using concentrations of 0-1000 ng / mL. Up until this point no significant effect was seen on cell viability by BFA, therefore it is assumed the decrease in IAV N is a BFA specific effect and that HAZV is unaffected by BFA at a pre-translational stage.

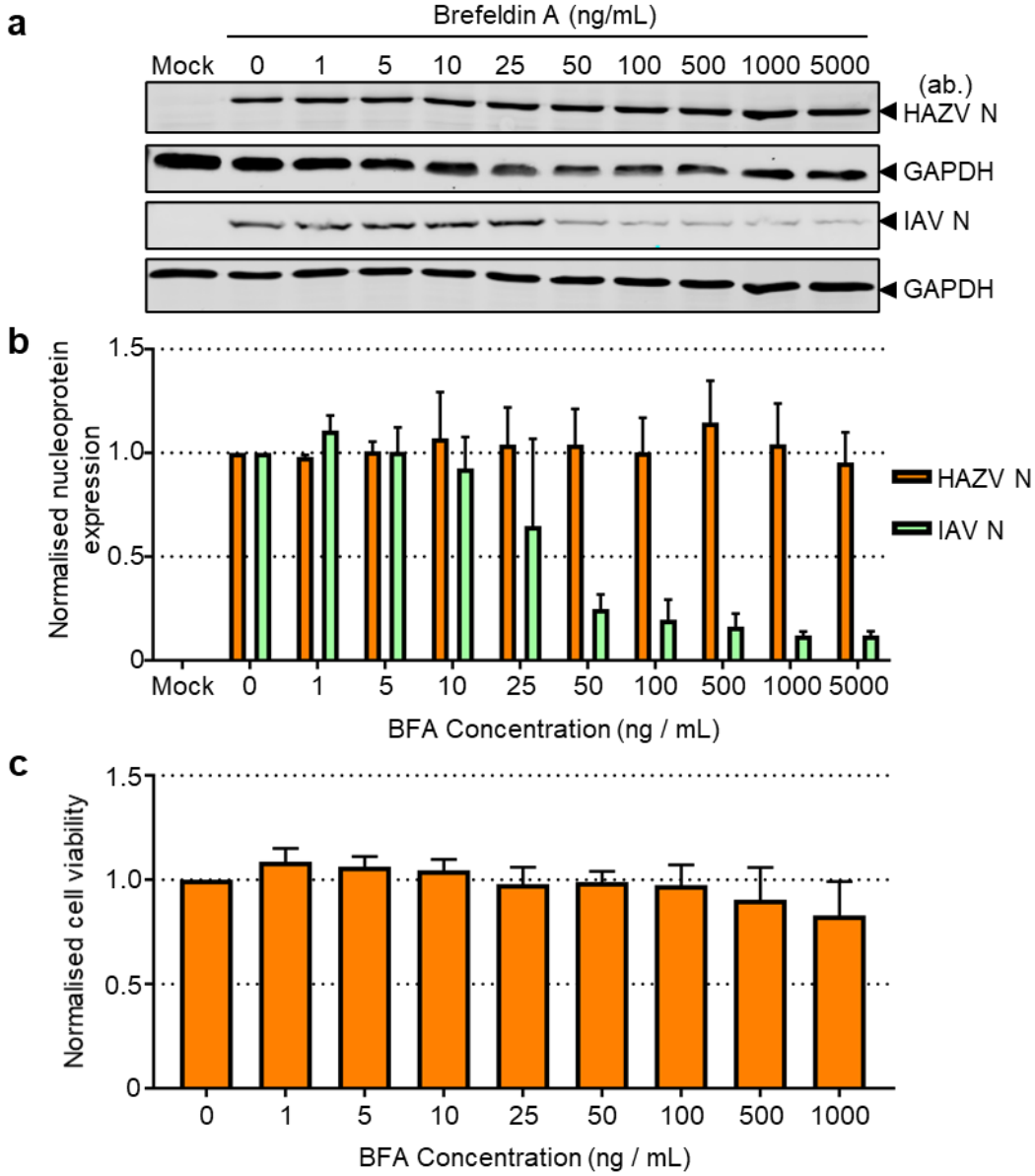


Figure 4.21, WT rHAZV is not sensitive to brefeldin A treatment.

a) Western blot analysis of lysates collected from A549 cells pre-treated with the indicated concentration of brefeldin A (BFA) for 45 minutes prior to infection with rHAZV or rIAV at a MOI of 0.1. Samples were probed for their respective nucleoproteins (HAZV N or IAV N) and GAPDH as a loading control. b) Normalised densitometric analysis of western blot data from (a), calculated using the ImageJ software package, error bars represent data from 3 experimental repeats, plotted using GraphPad Prism. c) Histogram

of the relative fold change in cell viability following treatment with BFA at a range of concentrations from 0 to 1000 ng / mL. Plotted using GraphPad Prism.

4.10 Chapter summary and discussion

In this chapter, the infectious clone system described in chapter 3 was modified to incorporate an eGFP ORF, upstream of the HAZV N protein, into the S segment expressing cDNA. Incorporation of this plasmid into the rescue protocol enabled the recovery of rHAZV(eGFP), which to the best of our knowledge represents the first recovery of a HAZV variant able to express a reporter protein. Expression of the eGFP reporter used a P2A linker, which promotes ribosome skipping and thus permits independent expression of HAZV N and eGFP from the same mRNA. Subsequently rescued rHAZV(eGFP) was shown to be a suitable model for studying HAZV, exhibiting a barely significant drop in multi-step virus growth versus the parental rHAZV WT strain and no detectable loss of eGFP expression following repeated virus passage.

The potential applications of such a reporter virus are many; one such use is in determining virus growth parameters. Use of rHAZV(eGFP) in conjunction with the IncuCyte® live cell imaging system reduced the workload and allowed real-time comparison of virus growth in multiple cell lines, identifying different patterns of growth without the requirement for cell lysate sample processing techniques such as western blotting. A recent study using a similar CCHFV reporter virus allowed examination of CCHFV mediated pathology and disease progression *in-vivo* using mouse models, successfully identifying the initial cells and organs targeted during CCHFV infection (Welch *et al.*, 2019).

In terms of examining the molecular and cellular virology of nairoviruses, the rHAZV(eGFP) system described here offers an important benefit compared to the CCHFV system. The rHAZV(eGFP) system is amenable to low CL-2 containment facilities in comparison to the CL-4 facilities required for CCHFV experimentation, which are costly and represent a major barrier to experimental progress. This rHAZV(eGFP) system will allow a streamlined work flow in which results can initially be collected using HAZV, then based on these findings, selected experiments with CCHFV in high containment can be performed, thus reducing overall costs.

The major application of rHAZV(eGFP) described in this chapter was screening a library of membrane trafficking components via siRNA knockdown. A total of 142 genes were targeted by 426 unique siRNA primers and following infection with rHAZV(eGFP), were able to be analysed for beneficial or negative impact on virus replication at stages prior to and including translation. Targets were graded based on % reduction of TIIE and sorted in ascending order, quickly identifying a number of genes associated with COPI vesicles having a significant effect on viral replication when knocked down. Validation of these targets was achieved via qPCR and western blot analysis, confirming a role of COPI vesicles in HAZV infection. Further analysis revealed this role was initially likely to be linked with an early stage of the HAZV lifecycle, possibly in endosomal escape. This was hypothesised due to the recently described dependence of HAZV on membranous cholesterol and the ability of COPI components to regulate this (Misselwitz *et al.*, 2011; Charlton *et al.*, 2019).

A further role of COPI components in HAZV infection was identified following titration of released virus following COPI knockdown in SW13 cells. A substantial increase in the fold reduction of infectious virus produced in COPI knockdown cells versus control cells could not be explained when combined with the fold reduction seen in the transcriptional and translational assays previously mentioned. This pointed to an additional role of COPI vesicles at a post-translational stage of the HAZV life cycle. As previous evidence has supported an interaction between CCHFV Gc and the COPI component β-COP (Erickson *et al.*, 2007) and the observation of multiple KKXX motifs, required for transport of membrane bound proteins between the Golgi and ER, it is likely the additional role of COPI components occurs at a glycoprotein processing or transportation stage.

Finally, the identification of an ARF1 / GBF1 independent mechanism of COPI dependence provided the most interesting finding. ARF1 is a well characterised component required for the formation of COPI vesicles, siRNA knockdown of ARF1 and GBF1, an activator of ARF1, resulted in no significant knockdown of TIIE following infection with rHAZV(eGFP). This was further supported via

treatment of cells with BFA, an inhibitor acting on ARF1, followed by infection with rHAZV. BFA displayed no inhibitory effects on translation of HAZV N. However, due to the limitations of rHAZV(eGFP) it was unable to rule out a post-translational role for ARF1 or GBF1, which given the evidence around COPI vesicle formation, is most likely when ARF1 and GBF1 would have effect.

Unfortunately, due to time constraints, further investigation into the role of the COPI components in virus infection was not pursued as extensively as hoped, with many potential validations and follow up experiments still to complete. Developing a method of exploring glycoprotein processing and trafficking, either via the generation of a tagged form of Gc / Gn or via specific antibodies against Gc, would enable more in depth understanding of the post-translational role of COPI vesicles in HAZV infection. Furthermore, examining the role of BFA on generation of infectious virus in a similar manner to the siRNA knockdown of COPI genes in Figure 4.17 would confirm whether the dependence of HAZV on COPI components is truly an ARF1 / GBF1 independent one.

The reporter strain developed and utilised in this chapter successfully achieved the aims set out at the beginning of the project. rHAZV(eGFP) serves as an extremely valuable tool for furthering understanding into the lifecycle of HAZV, as evidenced by its identification of the importance of COPI vesicles and will no doubt be useful in uncovering more knowledge into HAZV infection. The uses of the system described in this chapter, in combination with the WT system described in chapter 3, are further explored in chapter 5, where mutations to the WT sequence are introduced.

5 Exploring Hazara nucleoprotein cleavage during infection

5.1 General introduction

The development of the HAZV infectious clone system as described in chapter 3 was used in subsequent experiments described in chapter 4 for the generation and use of a recombinant HAZV able to express eGFP. However, another key use of an infectious clone system is the ability to discover the role of nucleotides, amino acids or ORFs within the viral genome using the principles of reverse genetics.

In this chapter, the highly-conserved caspase 3 cleavage motif, DQVD, in HAZV was selected as the initial target of mutagenesis. The DQVD motif is located at position 269 to 272, at the tip of the arm domain on HAZV N. (Surtees *et al.*, 2015; Wang *et al.*, 2015). CCHFV also expresses a similar site, DEVD, at position 266 to 269 (Carter *et al.*, 2012; Wang *et al.*, 2015). The importance of the cleavage site in CCHFV has been previously demonstrated, however neither study confirmed the exact role of the cleavage event. A study in 2011 identified cleavage of CCHFV N by caspase 3 occurred during infection of mammalian cell lines, following activation of apoptotic pathways. Inhibition of caspases was shown to have a beneficial effect on viral infection, marked by an increase in titre of released virus, suggesting a protective role of the motif against the cellular immune response (Karlberg, Tan and Mirazimi, 2011). A recent study, published during the course of this project, identified the DEVD motif in CCHFV to be dispensable for infection of mammalian cells, but was required for efficient replication in cells of tick origin (Salata *et al.*, 2018). Caspase mediated cleavage of viral proteins is not uncommon, the Junin virus (JUNV) encodes caspase cleavage motifs on its nucleoprotein as decoy substrates to subvert and delay the immune response (Wolff, Becker and Groseth, 2013). All human, but not avian, strains of influenza A and B also contain conserved caspase cleavage motifs within the nucleoprotein, with associated cleavage accompanied by increased caspase activity and development of apoptosis (Zhirnov *et al.*, 1999; Lipatov *et al.*, 2008). However, due to the previous lack of reverse genetics

systems for HAZV, the importance of the DQVD site in HAZV has never been fully explored.

Caspases can be activated via three main pathways; the intrinsic pathway, the extrinsic pathway and the granzyme B pathway (Taylor, Cullen and Martin, 2008). However, regardless of the initial activation mechanism all three initiation routes lead to the activation of caspase 3 and result in DNA fragmentation, degradation of cellular proteins and ultimately cell death in a controlled manner (Elmore, 2007). Caspases can be split into two groups, initiator and executioner. They are largely expressed as an inactive proenzyme which following activation can activate additional proenzymes, causing an activation cascade that amplifies and accelerates cell death. All caspases recognise and cleave proteins at aspartic acid residues, though the specificities vary between caspases as to the composition of neighbouring amino acids, most significantly with the 4 amino acids occupying the positions 1 to 4 (P_1-P_4) directly upstream of the cleaved aspartic acid. The preferential substrate site for caspase 3 has been well explored; the optimal cleavage sequence has been reported as DEVDG, though the caspase will cleave a DxxD sequence efficiently (Stennicke *et al.*, 2000). Whilst highly inefficient, it has also been demonstrated that caspase 3 can tolerate alternate amino acids at P_4 (Thornberry, 1997). Caspase 3 is widely regarded as the most important of the executioner caspases, it is able to activate caspase activated DNase (CAD) via cleavage of its inhibitor, ICAD (Sakahira, Enari and Nagata, 1998). CAD is then freely able to degrade chromosomal DNA and induce chromatin condensation.

The immune response to viral infection therefore must be either prolonged or prevented in order for viral infection to be successful, premature cell death prior to completion of the viral life cycle would inhibit the ability of the virus to produce progeny virions. Therefore, viruses have evolved highly complex mechanisms to evade the immune system at multiple stages. A number of viruses target the Bcl2 family of pro-apoptotic regulators, with examples including adenoviruses and Epstein-Barr virus (Henderson *et al.*, 1993; Han *et al.*, 1996). Caspases can also be directly targeted by direct interaction with the caspase active site by a viral protein or via competitive inhibition of signalling proteins (Skaletskaya *et al.*,

2001; Filippova *et al.*, 2007). As it has previously been shown that CCHFV is able to induce an immune response that results in the cleavage of CCHFV N, this chapter will focus on characterising the immune response and subsequent cleavage in relation to HAZV infection.

The aims of this chapter were as follows;

1. Determine the caspase cleavage characteristics of HAZV N in mammalian and tick cell lines
2. Utilise the infectious clone to perform a reverse genetics analysis of caspase cleavage sites necessary for virus infection
3. Identify the role of caspase cleavage sites in HAZV infection

Identification of an essential conserved motif within HAZV N would provide an excellent target for antiviral therapy, due to the inability of the virus to evolve resistance. This would also underline the importance of the development of the recombinant system for HAZV in furthering research into the HAZV and CCHFV replication.

5.2 Observation of multiple cleavage events in HAZV infection

Previous studies have identified both HAZV N and CCHFV N to be cleaved by activated caspase 3 (Carter *et al.*, 2012; Surtees *et al.*, 2015; Wang *et al.*, 2015). However, the importance of this cleavage event in relation to the HAZV life cycle has been poorly characterised to date. Therefore, to identify at what stage of virus replication this cleavage event was occurring, SW13 cells were infected at a MOI of 0.001 with rHAZV and total cell lysate was harvested at regular intervals between 0 and 96 hpi. Lysates were examined for HAZV N expression via western blotting and detection of GAPDH, using corresponding antisera, was used as a loading control. Interestingly, in addition to the expected cleavage products relating to caspase 3 cleavage at the DQVD site of approximately 30 and 22 kDa, multiple additional HAZV N specific cleavage products were observed at variable timepoints throughout the infection (Figure 5.1, a). An initial cleavage event was observed resulting in generation of a product with a mass of approximately 45 kDa in the first 32 hpi following infection. This 45 kDa band then

decreased in abundance through the next 16 hpi, before becoming almost undetectable at 72 hpi. In contrast, the expected cleavage products resulting from caspase 3 cleavage at 30 and 22 kDa appeared following the decrease in abundance of the initial 45 kDa cleavage product. Levels of the 30 and 22 kDa products increased throughout the infection, with the most abundant levels seen at 48 to 72 hpi. This apparent switch in processing of the viral protein may correlate with activation or progression of a host cell defence mechanism, such as the progression from activation of initiator to executioner caspases, which target different motifs, during the apoptotic response to viral infection. This observation suggests HAZV N is cleaved in a much more complex manner than originally predicted, with the function and origin of these additional cleavage products unknown.

As HAZV can replicate in both mammalian and tick vectors, whether HAZV was also cleaved in a similar manner in the tick cell line, HAE/CTVM9, was also explored. HAE/CTVM9 cells originate from the tick *Hyalomma anatolicum* and therefore provide an excellent physiologically relevant cell line for HAZV studies (Bell-Sakyi, 1991). Strikingly, despite strong detection of full length HAZV N, very limited N cleavage in these cells was detected, as judged from the lack of HAZV N specific bands of less than the full-length mass of 54 kDa (Figure 5.1, b). Detection of HAZV N in HAE/CTVM9 cells was also slightly delayed versus infection of the SW13 cell line, with first obvious detection of HAZV N occurring at 14 hpi compared to 10 hpi in SW13 cells. This demonstrates a striking difference in viral protein processing and duration of the attachment to translation phases between two physiologically relevant cell types for the same virus.

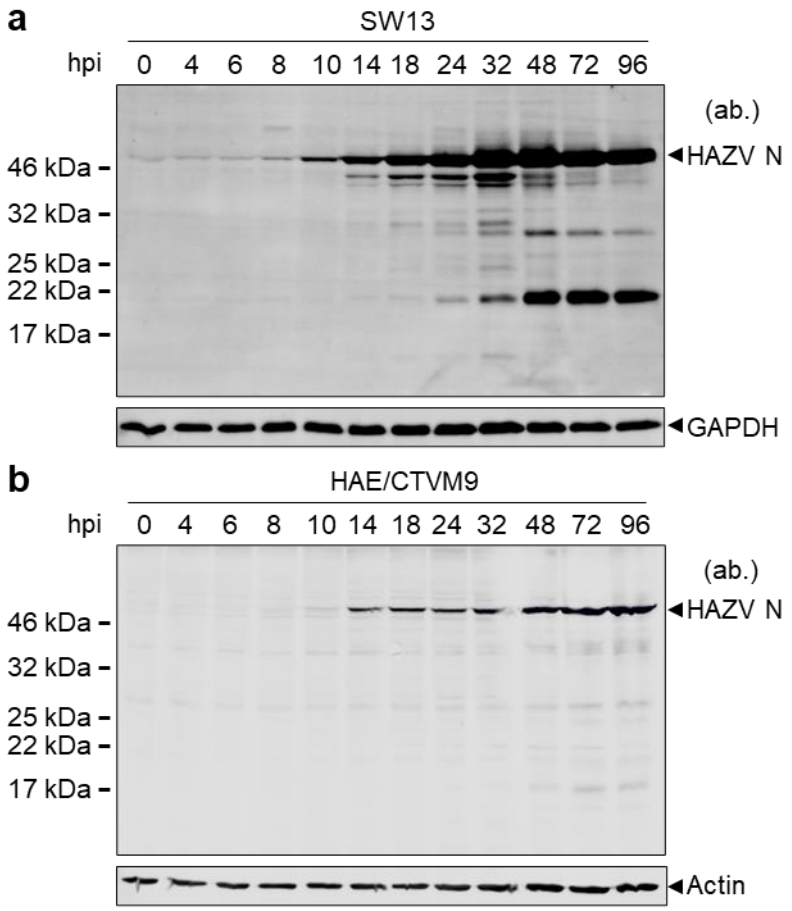


Figure 5.1 Differential cleavage of HAZV N in infection of SW13 and HAE/CTVM9 cell lines

Monolayers of SW13 (a) or HAE/CTVM9 (b) cells were infected with an HAZV at a MOI of 0.001. At the indicated time points post-infection, total cell lysates were collected and analyzed for HAZV N expression by western blotting with HAZV N antisera. Detection of GAPDH in SW13 cells and actin in HAE/CTVM9 cells, using corresponding antisera, were used as loading controls.

To identify the origin of these cleavage products, SW13 cells were pre-treated with the caspase inhibitors, Z-FA-FMK and Z-VAD-FMK then infected with rHAZV to identify if the additionally detected cleavage products were generated in a similar manner as the previously described 30 and 22 kDa products. Z-FA-FMK and Z-VAD-FMK are peptides that bind irreversibly to the catalytic site of caspases, inhibiting caspase-mediated apoptosis with minimal cytotoxic effects and are therefore useful tools for studying the role of caspases in virus infection (Van Noorden, 2001; Lawrence *et al.*, 2006). Analysis of infected cell lysates via western blot demonstrated neither Z-FA-FMK (Figure 5.2, a) or Z-VAD-FMK (Figure 5.2, b) to have any major effect on full length HAZV N expression. However, Z-FA-FMK was able to block cleavage of the 45 kDa (*), 30 kDa (**)

and 20 kDa (***) products, as evidenced by failure to detect the corresponded bands in treated lysates versus untreated controls. Interestingly, Z-VAD-FMK was less able to block the cleavage of HAZV N, with similar levels of the 30 kDa (**) and 20 kDa (***) cleavage products detected in treated and untreated control lysates. Despite the inability of Z-VAD-FMK to prevent production of the caspase 3 specific cleavage products, it was able reduce production of the 45 kDa (*) product, with lower abundance detected in treated lysates versus untreated controls. The ability of both caspase inhibitors to block detection of the 45 kDa product suggests the responsible host cell protease for the cleavage event resulting in generation of the 45 kDa product during infection to be a caspase.

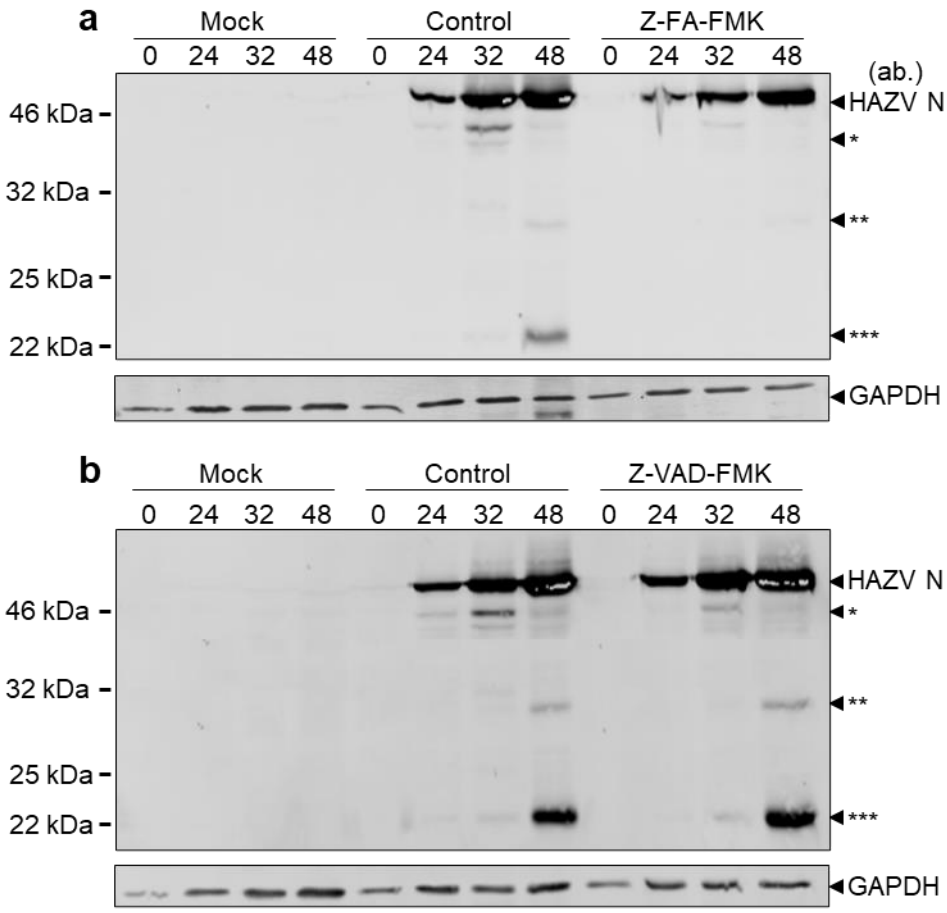


Figure 5.2, Caspases are responsible for HAZV N cleavage

Monolayers of SW13 cells were pre-treated with either (a) Z-FA-FMK or (b) Z-VAD-FMK prior to infection with rHAZV at 0.01 MOI. At the indicated time points post-infection total cell lysates were collected and analysed for HAZV N expression and its specific cleavage products at 45 kDa (*), 30 kDa (**) and 20 kDa (***) by western blotting with HAZV N antisera. detection of GAPDH, using antisera, was used as a loading control.

5.2.1 Incorporation of cleavage products into virions

To assess whether these cleavage products were incorporated into infectious virions, rHAZV was purified across a 20 % sucrose cushion via ultracentrifugation at 150,000 RPM. An aliquot of the resuspended pellet containing virus was analysed via western blot for HAZV N specific products. In addition to detection of full length HAZV N, a range of smaller molecular weight products were also detected in the purified virions

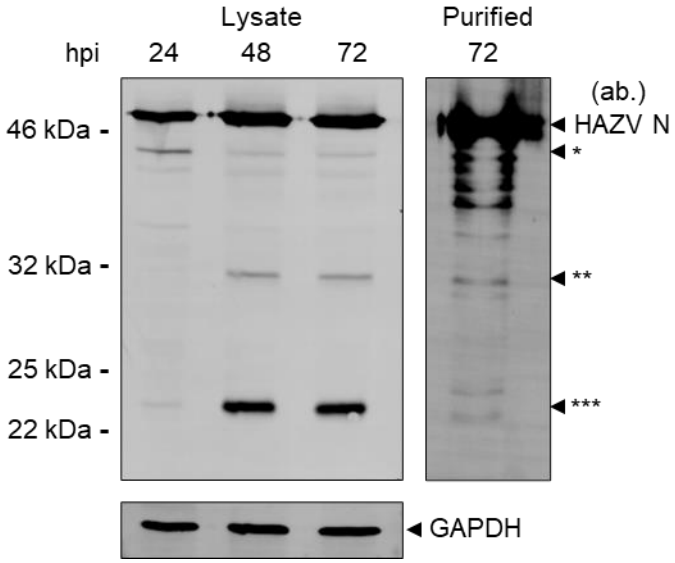


Figure 5.3, Caspase mediated HAZV N cleavage products are incorporated into infectious virions

Left panel lysates collected from monolayers of SW13 cells infected with an rHAZV at a 0.1 MOI. At the indicated time points post-infection, total cell lysates were collected and analyzed for HAZV N expression by western blotting with HAZV N antisera, detection of GAPDH, using antisera, was used as a loading control. Cleavage products were detected at approximately 45 kDa (*), 30 kDa (**) and 22 kDa (***). Purified rHAZV collected 72 hours post infection was denatured at 95℃ and analysed for HAZV N expression by western blotting with HAZV N antisera.

5.3 Induction of apoptosis by HAZV

Previous work has identified caspase 3 as the responsible protease for cleavage of HAZV N (Wang *et al.*, 2015), however little work has been done to characterise this interaction in a live-cell model. To this end, SW13 cells were infected at 0.001 MOI and total cell lysate was harvested daily at 24, 48, 72 and 96 hpi. As a negative control, a sample was also taken at 0 hpi. Lysates were first checked for HAZV infection and equal loading via detection of HAZV N and GAPDH (SW13 cells) or actin (HAE/CTVM9 cells). As expected, no HAZV N was detected at the

0 hpi time point, but was detectable in all lysates collected 24 hr post infection and above. To determine whether the caspase pathway was activated in mammalian cells in response to HAZV infection, lysates were again analysed via western blot and probed for caspase 3, 7, 9 and PARP, along with their associated activated cleaved forms (Figure 5.4). Activation of caspases was detected in mammalian cell lysates at a minimum of 48 hpi, with all caspases having cleaved forms present in the lysates at 96 hpi (Figure 5.4, left panel). Strikingly, the same was not observed in the HAE/CTVM9 tick cell line, despite observation of caspases in the lysate, no activated forms were detected at any of the time points sampled (Figure 5.4, right panel).

The inability to detect activated caspases in tick cells may provide an explanation for the lack of associated HAZV N cleavage observed previously during infection of tick cells in Figure 5.1. However, it is also possible that an unknown mechanism may delay apoptosis in tick cells, allowing the virus to persist longer in the vector. Delaying induction of apoptosis has been demonstrated by BUNV NS_S, via detection of interferon activated apoptotic pathways earlier in NS_S knockout BUNV vs WT BUNV infection (Kohl *et al.*, 2003). However, following transfection of the CCHFV NS_S into SW13 cells it was shown that this viral protein induced apoptosis via both intrinsic and extrinsic pathways. The ability of these proteins to induce or delay apoptosis may be dependent on the host cell type. It would make sense that in the cell line derived from the vector, HAE/CTVM9, that the virus would be unable to induce apoptosis, allowing persistence of the infection. However, in a cell line derived from the dead-end host, SW13, the virus induces apoptosis and resulting pathogenesis.

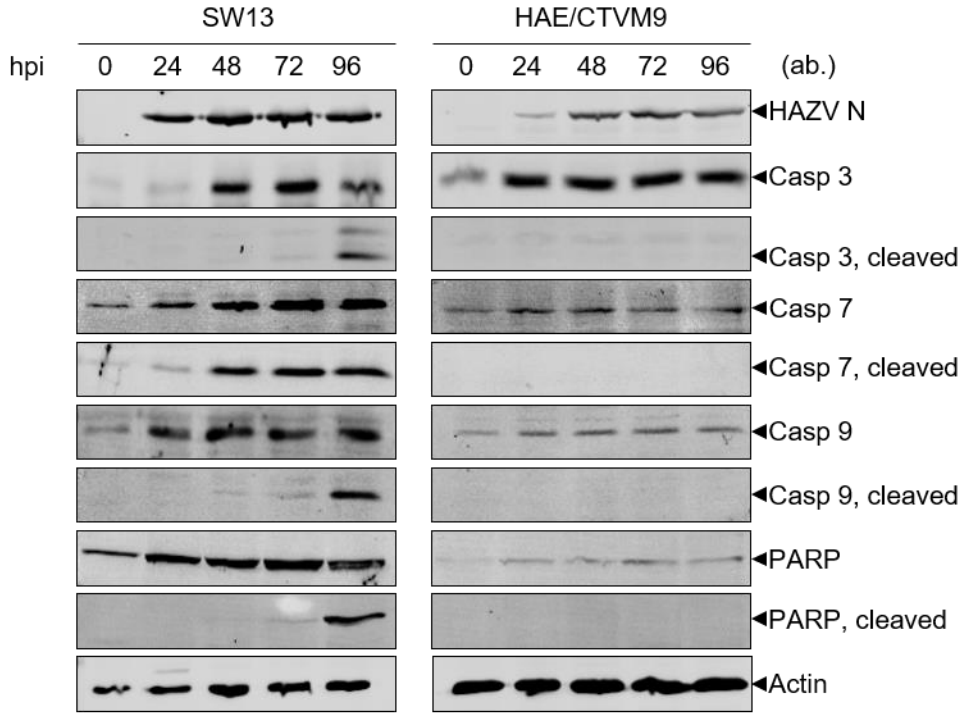


Figure 5.4, Differential activation of caspases following rHAZV infection of SW13 and HAE/CTVM9 cell lines

Monolayers of SW13 cells (left column) or HAE/CTVM9 cells (right column) were infected with HAZV at 0.001 MOI. At the indicated time points post infection, total cell lysates were collected and analyzed for expression of HAZV N, Caspases 3, 7, 9, PARP and their associated cleavage products by western blotting with HAZV N antisera or purchased polyclonal antibodies. Detection of Actin, using antisera, was used as a loading control in both SW13 and HAE/CTVM9 cells.

5.4 Generation of cleavage site deficient rHAZV

In order to assess the importance of the individual cleavage motifs, a panel of mutations were introduced into the S segment rescue plasmid cDNA to knockout the ability of each of the motifs to be cleaved. This was achieved via SDM of the position 1 aspartic acid residue to either alanine or glutamic acid. Alanine was selected to completely knockout the cleavage ability, whereas glutamic acid was selected to knockout cleavage ability whilst maintaining a degree of structural integrity to the region.

5.4.1 DQVD and ENKD sites

It is well established that HAZV N contains a conserved caspase 3 cleavage motif at the tip of the arm domain, with the amino acid sequence ^{269}DQVD^{272} (Surtees *et al.*, 2015; Wang *et al.*, 2015). However, it is often overlooked that the position 4 aspartic acid residue also forms the position 1 residue in an overlapping

caspase motif, ^{266}ENKD^{269}. Whilst not the stereotypical DxxD motif that is preferred by active caspase 3, previous research has shown amino acid flexibility at the 2, 3 and 4 positions of caspases can be tolerated and indeed are described in the literature (Thornberry, 1997). Conservation of the motif exists between HAZV (DQVD) and CCHFV (DEVD), with the two sites occupying the same spatial position on the arm domain of their respective proteins, despite being located three amino acids apart (Figure 5.5).

The DEVD motif is conserved in CCHFV across all fully sequenced isolates of CCHFV, both in its location on the apex of the protein and in sequence, suggesting a critical role of the motif in the viral lifecycle (Deyde *et al.*, 2006).

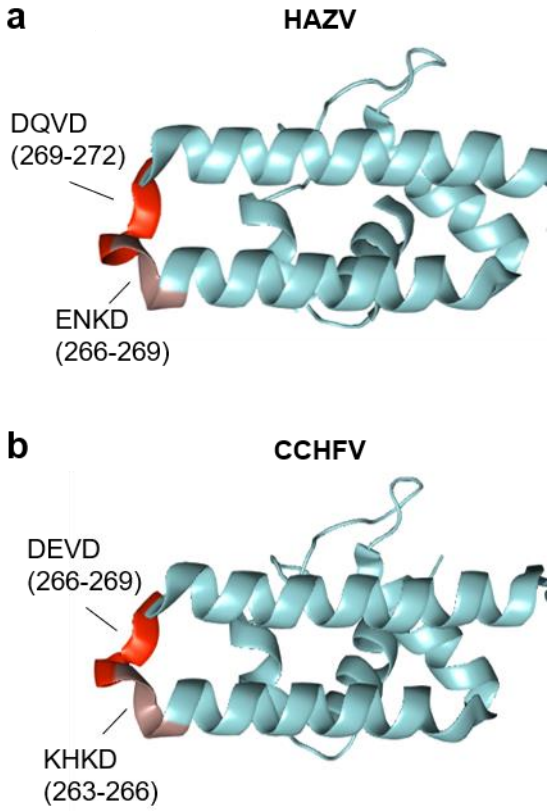


Figure 5.5, Location of DQVD and ENKD cleavage sites on HAZV and CCHFV nucleoproteins

Schematic showing locations of caspase cleavage motifs DEVD/DQVD and ENKD/KHKD on the apex of the arm domains of HAZV (a) and CCHFV (b) respectively, with amino acid positions indicated numerically.

In order to assess the importance of the overlapping ENKD and DQVD motifs, site directed mutagenesis was carried out, substituting the respective position 1 aspartic acid residue for glutamic acid, rendering the motif uncleavable by

caspases. An additional mutant in which both position 1 residues were substituted for alanine, giving an AQVA motif, was also included to assess the effect of knocking out both caspase cleavage sites concurrently and also to permit a direct comparison to a similar mutation described using the recombinant CCHFV system (Salata *et al.*, 2018). Recovery of rHAZV containing the ENKE, DQVE or AQVA motif was then carried out as described in section 3.4, substituting out the WT S segment cDNAs for the associated caspase cleavage site deficient S segment cDNAs. As previously described, following transfection of BSR-T7 cells, supernatants were harvested at 72, 96 and 120 hours post transfection and used to re-infect fresh monolayers of SW13 cells. Following a 48 hour reinfection, total cell lysates were collected and analysed for expression of HAZV N as a marker of viral replication. Presence of HAZV N in the reinfection lanes of both rHAZV(ENKE) and rHAZV(DQVE) mutants suggested a non-essential role for the ENKD and DQVD motifs independently; however, the inability to detect HAZV N in reinfections using the AQVA mutant S segment (Figure 5.6, a) suggests a requirement for at least one of the aspartic acid residues to be present on the arm domain.

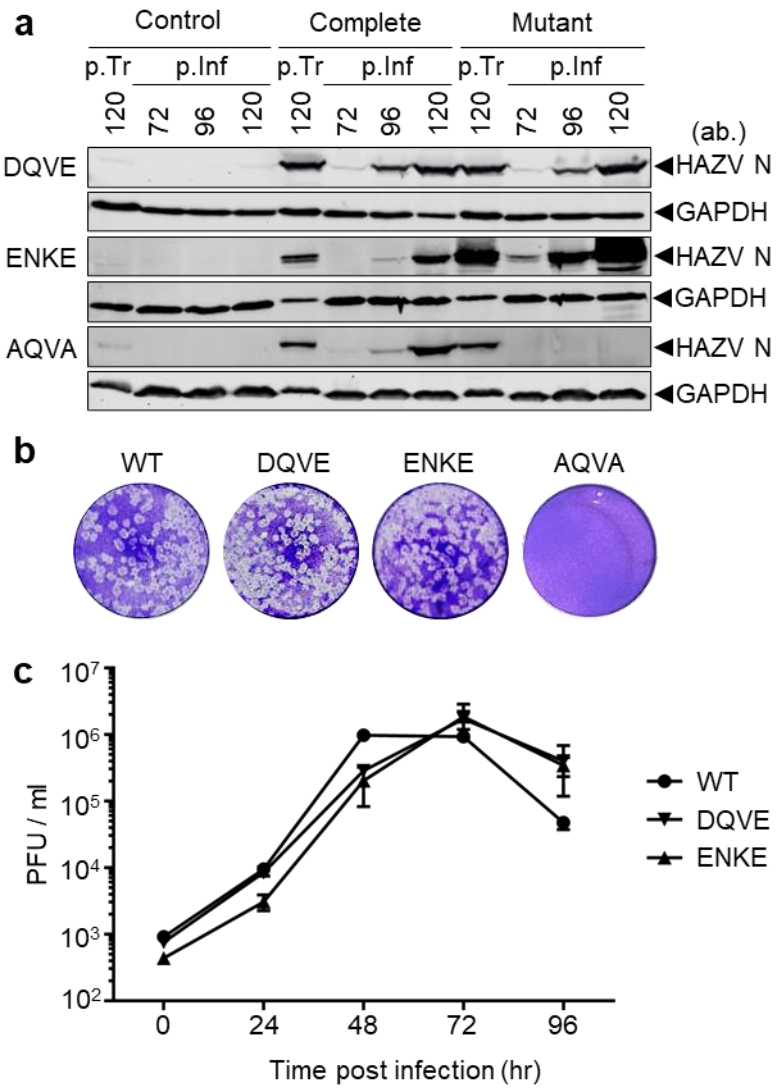


Figure 5.6, Recovery of caspase motif deficient rHAZV

a) Detection of HAZV N protein for rHAZV(DQVE), rHAZV(ENKE), and rHAZV(AQVA) mutants in lysates collected post transfection (p.Tr) of BSR-T7 cells and 48 hours post infection (p.Inf) of SW13 cells, infected using supernatants harvested 72, 96, and 120 hour post transfection. Recovery of all mutants was carried out alongside independent complete control recovery of WT rHAZV and a control in which transfection of L segment plasmid was omitted. detection of GAPDH, using antisera, was used as a loading control. b) Representative plaque assays of supernatants taken 120 h post transfection, displaying plaque morphology for recovered viruses. c) Comparison of titers of infectious WT rHAZV, rHAZV(DQVE) and rHAZV(ENKE) at 24 hour intervals following infection of SW13 cells at a MOI of 0.001 over a 4 day period, plotted using GraphPad Prism.

Of note, recovery of the two mutants containing DQVE and ENKE motifs displayed similar plaque morphology to WT rHAZV, displaying the same clarity and bullseye characteristics (Figure 5.6, b). Comparison of fitness of mutants was achieved via a multistep growth curve over 4 days, the time at which titer of HAZV begins to decrease in supernatants from cell culture. Analysis of plaque assays

showed similarities in the growth profiles of WT rHAZV, rHAZV(DQVE) and rHAZV(ENKE) mutants, with peaks observed at approximately 10^6 PFU / mL at 72 hpi (Figure 5.6, c). To confirm the cleavage site had been rendered redundant, western blot analysis of lysate collected from SW13 cells infected with WT, rHAZV(DQVE) and rHAZV(ENKE) was analysed. Surprisingly, despite the disappearance of the larger ≈ 32 kDa molecular weight product specific to HAZV N in the infection with rHAZV(DQVE), an alternate cleavage product was detected at a marginally larger molecular weight (Figure 5.7). In addition, the lower molecular weight product at ≈ 20 kDa remained unexplainably unchanged. No change to the cleavage products generated was observed with the rHAZV(ENKE) infection, suggesting the ENKD motif to be unassociated with caspase cleavage. Taken together, it suggests the entire DQVD motif may have an additional role to a caspase cleavage motif within HAZV infection.

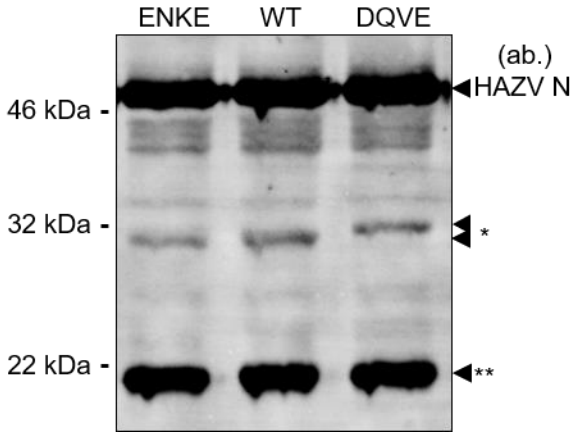


Figure 5.7, Knockout of caspase motifs results in alternate cleavage of HAZV N

Western blot detection of full length HAZV N and its associated cleavage products at 30 and 32 kDa (*) and 20 kDa (**) following a 48 hour infection of SW13 cells at a MOI of 0.01 with WT rHAZV, rHAZV(ENKE) and rHAZV(DQVE). Relevant molecular sizes are indicated in kDa.

To determine whether the caspase cleavage site was a requirement in cells originating from the natural tick vector, rather than mammalian cells, tick-origin HAE/CTVM9 cells were infected with WT rHAZV alongside rHAZV(ENKE) and rHAZV(DQVE) mutants. The time course displayed in Figure 5.1 identified a strong HAZV N band at 48 hpi and so this time point was used to determine whether the mutants were capable of expression of viral products. HAE/CTVM9 cells were infected at a MOI of 0.001 and incubated for 48 hours. At this point total cell lysate was collected and analysed for HAZV N expression via western

blot. Actin was also detected as a loading control and mock infected HAE/CTVM9 cells were also set up as a negative control. Visualisation of strong HAZV N specific bands in all infected wells of equal density demonstrates the ability of the rHAZV mutants rHAZV(ENKE) and rHAZV(DQVE) to bind, enter, uncoat, transcribe and translate viral products efficiently within HAE/CTVM9 cells suggesting the role of the HAZV caspase cleavage motif is not required for replication in these cells.

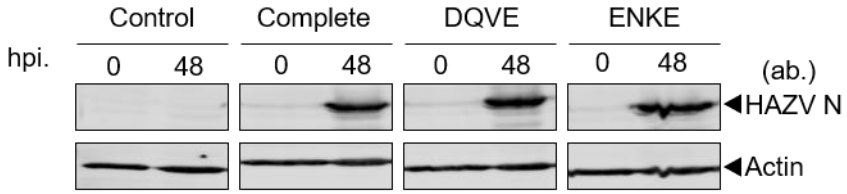


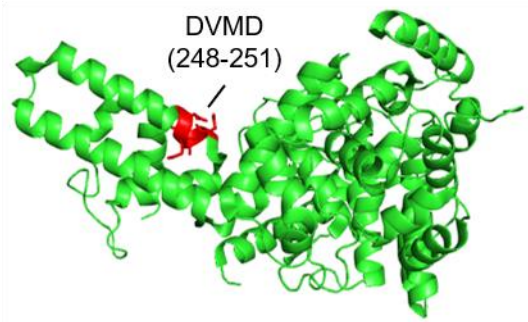
Figure 5.8, Mutagenesis of the DQVD and ENKD sites does not affect ability of rHAZV to infect tick cells

Western blot analysis of lysates collected 0 and 48 hours post infection (hpi.) of HAE/CTVM9 cells with WT rHAZV, rHAZV(DQVE) and rHAZV(ENKE) at 0.01 MOI via detection of HAZV N using anti HAZV N serum. Detection of Actin, using antisera, was used as a loading control.

5.4.2 DVMA

Caspase cleavage prediction software, GraBCas, identified an alternative motif to DQVD that may act as a substrate for caspase cleavage in HAZV. The DVMD motif at position 248-251 met the optimal consensus cleavage motif for caspase 3, DxxD and was suitably located in an assumedly accessible region to permit interaction with the caspase and subsequent cleavage (Figure 5.9, a). Examination of related viruses revealed the site was very loosely conserved, with minimal conservation observed other than the P1 aspartate residue, which was located between positions 248-251 in CCHFV, NSDV, ERVE and TOFV (Figure 5.9, b).

a



b

Virus	P1' Position	Amino Acid Seq.
HAZV	251	DVMD ↓
CCHFV	248	GVFD ↓
NSDV	251	VAED ↓
ERVEV	251	QAAD ↓
TOFV	248	EGPD ↓

Figure 5.9, Location of the DVMD predicted cleavage site and conservation in alternate species

a) 3D representation of HAZV N, with the location of the DVMD motif indicated in red at position 248 to 251. Image generated using PyMOL. b) Conservation of the DVMD motif across related viruses within the Bunyavirales.

In order to assess the importance of this domain in the HAZV life-cycle, a plasmid encoding a mutated DVMD motif, in which the P1 aspartate was replaced with alanine to generate an uncleavable DVMA motif (pMK-RQ-S-DVMA). Recovery of infectious rHAZV was attempted as previously described utilising a complete set of WT plasmids as a positive control. In the recovery attempt of the DVMA mutant the WT S segment pMK-RQ-S plasmid was replaced with the mutated version, pMK-RQ-S-DVMA, and transfected into BSR-T7 cells. Subsequent attempts to re-infect SW13 cells at 72, 96 and 120 hours post transfection resulted in no infection with rHAZV(DVMA), indicated by the lack of HAZV N signal following western blot analysis (Figure 5.10, a). Failure to recover rHAZV(DVMA) was shown not to be due to inefficiencies in the recombinant system as triplicate repeats ended with similar results, with WT rHAZV recovered in all cases, indicating the failure of rHAZV(DVMA) rescue to be a result of the mutation. Further validation that no mutant virus was recovered was provided by plaque assay, demonstrating recovery of WT rHAZV via observation of characteristic plaque morphology, with no detectable plaques in the rHAZV(DVMA) plaque assay.

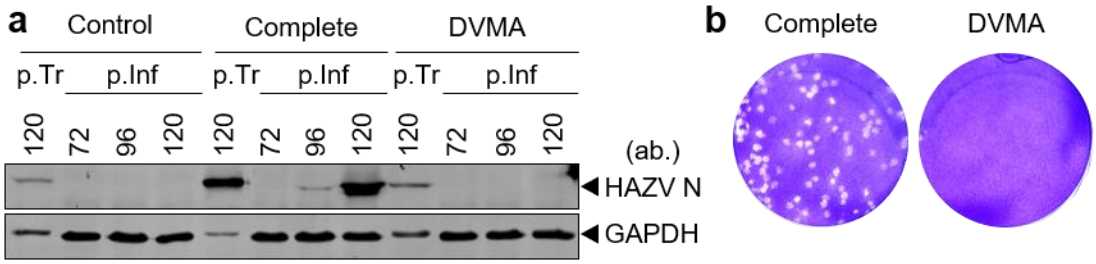


Figure 5.10, Inability to recover rHAZV with a mutated DVMD motif shown via western blot and plaque assay

a) Detection of HAZV N for the rHAZV(DVMA) mutant in lysates collected post transfection (p.Tr) of BSR-T7 cells and 48 hours post infection (p.Inf) of SW13 cells, infected using supernatants harvested 72, 96, and 120 hour post transfection. Recovery of the rHAZV(DVMA) mutant was carried out alongside independent complete control recovery of WT rHAZV and a control in which transfection of L segment plasmid was omitted. Detection of GAPDH, using antisera, was used as a loading control. b) Representative plaque assays of supernatants taken 120 h post transfection

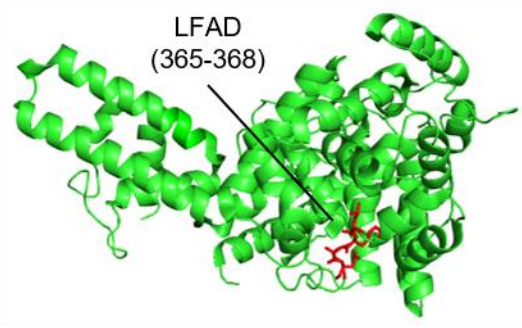
It is unknown why mutagenesis of the DMVD motif resulted in such a surprising, fatal phenotype. The minimal conservation of the motif, other than the P1 aspartic acid, suggests that caspase cleavage may not be the reason behind the failure to recover infectious virus. The limitation of the infectious clone system prevents the identification of any loss of cleavage product that would support the role of caspases associated with DVMD cleavage. Recovery of a defective virus, in which one of the cleavage products identified in Figure 5.1 was no longer present, would have provided strong evidence for the role of HAZV N cleavage during viral infection. However due to the poor conservation of the P2-P4 residues within the motif across multiple strains, it is unlikely that caspase cleavage at the DVMD site in HAZV would have such a critical role. The availability of a minigenome system could provide further understanding as to the role of DVMD during HAZV RNA synthesis, however such a system was not available during this project and attempts to generate one were unsuccessful. An alternative reason for the importance of the P1 aspartic acid residue may be structural. The location of the motif lies at the interface between the arm and globular domains of HAZV N, the conversion of D to A may result in an alteration in structure, such as the angle of the arm domain in relation to the globular domain that inhibits a critical function of HAZV N. Nevertheless, regardless of the way in which the DMVD motif affects HAZV infection, the infectious clone system was successfully used to demonstrate the critical nature of it to viral replication.

5.4.3 LFAA

5.4.3.1 Recovery of LFAA rHAZV

Alignment of the amino acid sequences for the N protein of HAZV, CCHFV, NSDV, ERVEV and TOFV revealed the total conservation of a motif, LFAD, located at position 365 to 368 in HAZV that displayed 100 % conservation through the P1 to P4 positions. The motif is located on the globular domain of HAZV N, in an easily accessible position (Figure 5.11, a). Location of the motif was also similar, with the P1 aspartate residue located at position 360, 365 or 368, dependent on species (Figure 5.11, b). At the time of writing, no previous research describing the importance of such a well-conserved motif in HAZV or CCHFV has been reported and therefore presented an attractive area to investigate. Occurrence of a LFAD motif in a non-viral context is observed in the sorting nexin 1 (SNX1) protein. SNX1 is part of a larger family of sorting nexins, and contains a conserved phox homology (PX) domain that permits binding to phosphoinositide 3-monophosphate (PtdIns3P) (Carlton *et al.*, 2005). In addition to a PX domain, SNX1 also contains a Bin/Amphiphysin/Rvs (BAR) domain that is able to sense and induce membrane curvature within endosomal compartments in the formation of transport vesicles (Peter *et al.*, 2004). SNX1 has previously been shown to have a role in endosomal sorting of epidermal growth factor receptor (EGFR) between endosomes, the Golgi and lysosomes as part of the retromer complex comprising vacuolar-protein-sorting proteins 5, 17, 26, 29 and 35 (Vps5p, Vps17p, Vps26p, Vps29p, Vps35p) (Seaman *et al.*, 1997). Critically, during the apoptotic response SNX1 is cleaved at the LFAD motif, preventing its ability to sort endosomal receptors, a process that might be required by HAZV and related viruses during infection (Duclos, Lavoie and Denault, 2017). To this end, it is possible that conservation of the LFAD motif in orthonairoviruses may be as a decoy substrate to prevent cleavage of SNX1 during virally-induced apoptosis as described in section 5.3.

a **b**



Virus	P1' Position	Amino Acid Seq.
HAZV	368	LFAD ↓
CCHFV	365	LFAD ↓
NSDV	365	LFAD ↓
ERVEV	360	LFAD ↓
TOFV	365	LFAD ↓

Figure 5.11, Location of the LFAD predicted cleavage site and conservation in alternate species

a) Model displays a 3D representation of HAZV N, with the location of the LFAD motif indicated in red at position 365 to 368. Image generated using PyMOL. b) Conservation of the LFAD motif across related viruses within the Bunyavirales.

To assess the role of the LFAD motif in HAZV infection, the infectious clone system detailed in section 3.4 was modified to generate a rHAZV clone in which the LFAD motif was mutated to an uncleavable motif, LFAA. The mutated LFAA cDNA was generated via SDM of the pMK-RQ-S plasmid, converting the P1 aspartate to alanine (pMK-RQ-S-LFAA). Recovery of mutant rHAZV was achieved as described in section 3.4, with pMK-RQ-S substituted for pMK-RQ-S-LFAA. Recovery of rHAZV(LFAA) was carried out alongside recovery of WT rHAZV, in which WT encoding plasmids were used as a positive control for the recovery. Confirmation of successful recovery of rHAZV(LFAA) was achieved via the detection of HAZV N signal following western blot analysis of lysates taken from SW13 cells infected with supernatant harvested from BSR-T7 cells 120 hours post transfection (Figure 5.12, a), as expected HAZV N signal was also detected in the rHAZV reaction.

Despite recovery of both viruses, the intensity of HAZV N expressed following the 48 hour reinfection of SW13 cells was markedly lower in the infection with rHAZV(LFAA) compared to rHAZV, with no recovery observed at all in infections established using supernatant collected 96 hours post transfection from rHAZV(LFAA), despite detection of rHAZV signal under the same experimental conditions (Figure 5.12, a). Further differences were observed between WT rHAZV and rHAZV(LFAA) upon analysis of plaque assays set up using

supernatant collected 120 hr post transfection. As in previous recoveries of WT rHAZV, plaques presented as clearly defined circles, with a ‘bulls-eye’ centre, however plaque morphology of rHAZV(LFAA) was substantially altered, with poorly defined plaques barely visible (Figure 5.12, b). This phenotype suggests severely reduced virus fitness, with this mutant virus unable to cause similar levels of CPE in SW13 cells as the WT equivalent. However, the extent of the decrease in fitness observed could not be experimentally determined effectively due to the inability to accurately determine the titre of the stock due to the issues surrounding detection of plaques, therefore a solution that enabled accurate titration of the viruses was required.

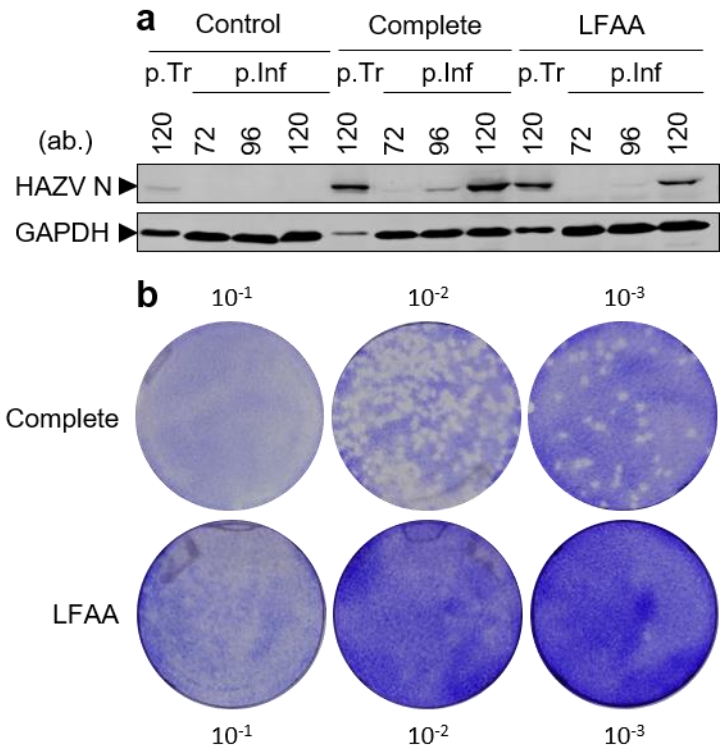


Figure 5.12, Recovery of rHAZV containing a mutated LFAD motif displaying altered plaque morphology

a) Detection of HAZV N for the rHAZV(LFAA) mutant in lysates collected post transfection (p.Tr) of BSR-T7 cells and 48 hours post infection (p.Inf) of SW13 cells, infected using supernatants harvested 72, 96, and 120 hour post transfection. Recovery of the rHAZV(LFAA) mutant was carried out alongside independent complete control recovery of WT rHAZV and a control in which transfection of L segment plasmid was omitted. Detection of GAPDH, using antisera, was used as a loading control. b) Representative plaque assays of supernatants taken 120 h post transfection

5.4.3.2 Recovery of fluorescent LFAA rHAZV

To resolve the issues surrounding titration of rHAZV(LFAA), the eGFP infectious clone detailed in section 4.3 was used. The same SDM reaction used to generate pMK-RQ-S-LFAA was repeated, using pMK-RQ-S-eGFP as the template cDNA, generating pMK-RQ-S-LFAA-eGFP which was used to recover the mutant rHAZV(eGFP-LFAA) in the same way as described previously. As rHAZV(eGFP) had previously been shown to be comparable to WT rHAZV, it was expected that rHAZV(eGFP-LFAA) would be recoverable and demonstrate similar properties to rHAZV(LFAA). This was confirmed via western blot analysis of lysates taken from SW13 cells infected with supernatant harvested from BSR-T7 cells 96 and 120 hours post transfection, with HAZV N and GFP signal detected in both complete rHAZV(eGFP) and rHAZV(eGFP-LFAA) reactions (Figure 5.13, a)

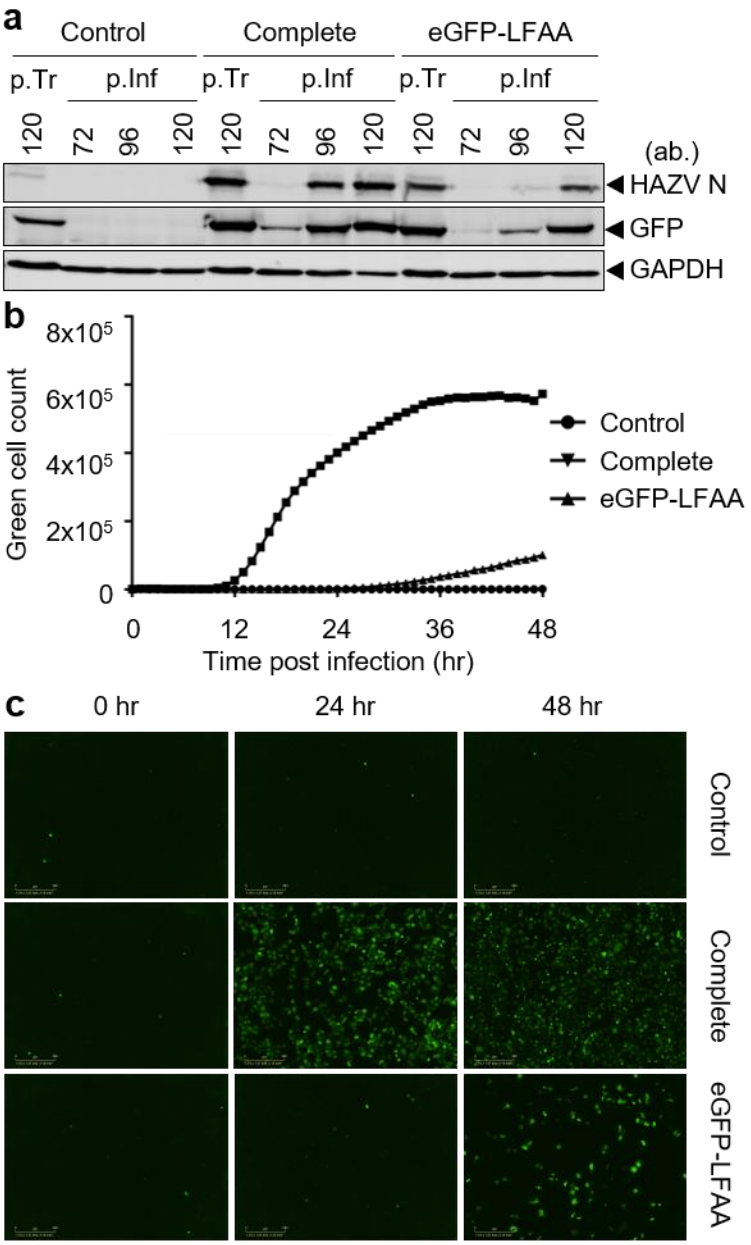


Figure 5.13, Recovery of rHAZV-eGFP containing a mutated LFAD motif demonstrates reduced fitness

a) Detection of HAZV N and GFP for the rHAZV(eGFP-LFAA) mutant in lysates collected post transfection (p.Tr) of BSR-T7 cells and 48 hours post infection (p.Inf) of SW13 cells, infected using supernatants harvested 72, 96, and 120 hour post transfection. Recovery of the rHAZV(eGFP-LFAA) mutant was carried out alongside independent complete control recovery of WT rHAZV and a control in which transfection of L segment plasmid was omitted. Detection of GAPDH, using antisera, was used as a loading control. b) Graphical representation of green cell count over time during the 48 hour re-infection of SW13s using supernatant harvested from BSR-T7 cells transfected for control, complete and rHAZV(eGFP-LFAA) 120 hours post transfection. Analysis was carried out using the IncuCyte® live cell imaging system and plotted using GraphPad Prism. c) Representative images of infected SW13 cells by rHAZV(eGFP-LFAA) at 0, 24 and 48 hours post infection taken on the IncuCyte® live cell imaging system. Scale bar represents 300 µM.

As with the non-fluorescent mutant rHAZV(LFAA), the resulting rHAZV(eGFP-LFAA) also appeared to demonstrate obvious reduced fitness issues following its recovery from cDNA. Abundance of both HAZV N and GFP signal via western blot were reduced versus the WT rHAZV(eGFP) controls in both 96 and 120 hour post transfection reinfections. This was supported by data collected from the incucyte scans of the 120 hour post transfection reinfection (Figure 5.13, b, c), with the time taken for GFP signal to be detected occurring between the 24 and 48 hpi timepoints, versus a pre 24 hpi value for rHAZV(eGFP).

5.4.3.3 Alterations to LFAD motif reduce viral fitness

To determine whether the addition of eGFP expression to rHAZV(LFAA) would enable titration of the mutant form, a plaque assay was set up for both rHAZV(eGFP) and rHAZV(eGFP-LFAA) and imaged at 72 hpi. Observation of the patterns of infected cells, marked by eGFP expression, revealed a substantial difference in plaque morphology between rHAZV(eGFP) and rHAZV(eGFP-LFAA). The most obvious difference was observed to be the diameter of the plaques, with the average diameter of a WT plaque being 534 μm (± 27.85, n=13) versus an average diameter of 348.6 μm (±16.67, n=12) for rHAZV(eGFP-LFAA), giving a highly significant difference between WT and mutant ($P = <0.0001$) (Figure 5.14, a). In addition to the decreased diameter of plaques generated by rHAZV-LFAA(eGFP), the plaques appeared less “full” than those observed with WT rHAZV(eGFP) (Figure 5.14, b), which may explain why the plaques observed previously in rHAZV-LFAA (section 5.4.3.1) were much more poorly defined than the WT equivalents.

Importantly, the visualisation of clear-cut foci of infection enabled titration of the rHAZV(eGFP-LFAA) virus stock. Using the EVOS microscope, which permits whole well fluorescence imaging, virus stocks of rHAZV(eGFP-LFAA) were able to be accurately titrated. A whole well image is generated via sequential imaging of the entire well, which are then stitched together by the EVOS software to create a final image (Figure 5.14, c). Plaques can then be counted as with a crystal violet plaque assay and the titre determined. The ability to accurately titre rHAZV(eGFP-LFAA) enables comparative infections at the same MOI as WT

rHAZV, allowing any difference in fitness to be attributed to the mutation rather than a different starting concentration of virus.

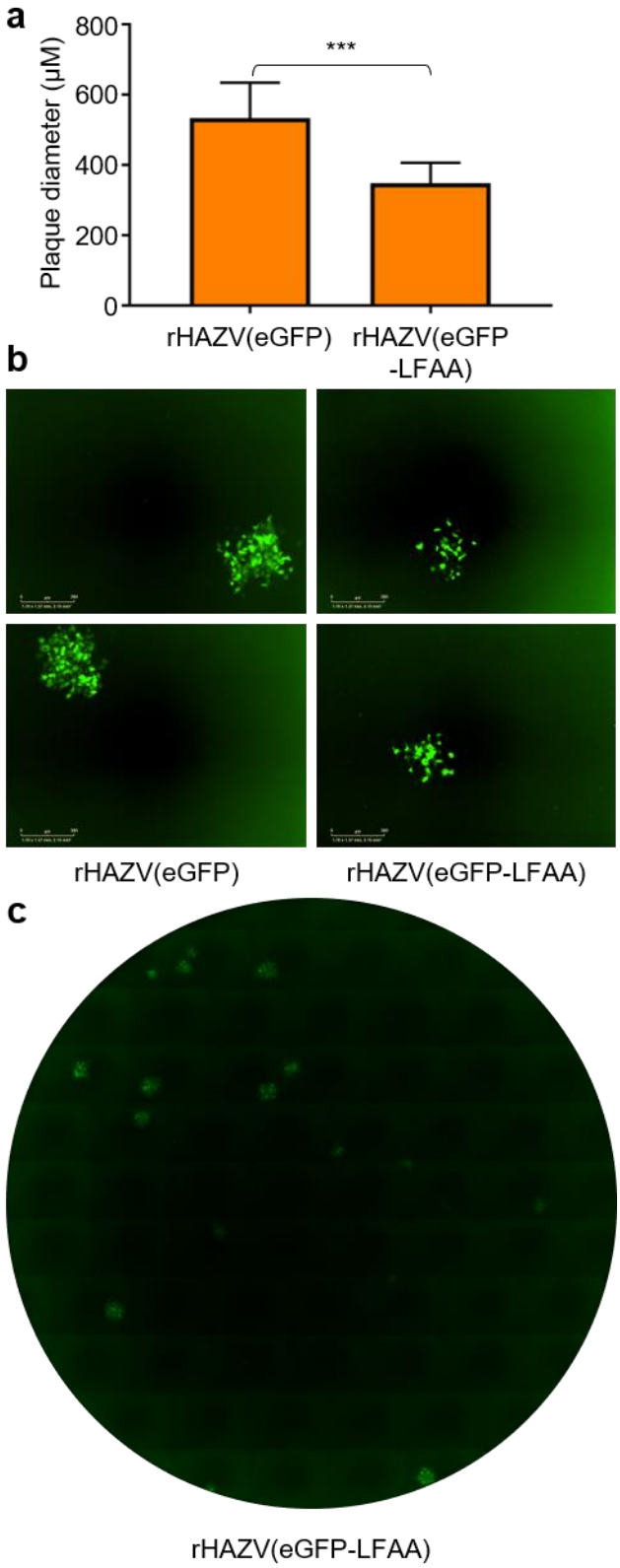


Figure 5.14, Mutation of LFAD motif results in altered plaque morphology

a) Histogram displays the diameter of the fluorescent plaques generated 72 hours post infection of SW13 cells by rHAZV(eGFP) or rHAZV(eGFP-LFAA). Minimum of 10

plaques were measured for each virus, significance determined by T-Test (P = <0.0001). Plotted using GraphPad Prism. b) Representative images of fluorescent plaques used to generate data in (a), imaged on the IncuCyte® live cell imaging system. c) Representative image of rHAZV(eGFP-LFAA) titration on EVOS microscope, taken 72 hours into infection of SW13 cells.

Following the establishment of a means to calculate titres of rHAZV(eGFP) and rHAZV(eGFP-LFAA), an infection was set up in SW13 cells and imaged hourly on the IncuCyte® live cell imaging system. This enabled insight into the time taken between applying virus to cells and translation of viral transcripts. Plotting of green cell count versus time post infection identified a delay in detection of eGFP signal in rHAZV(eGFP-LFAA) infections compared to rHAZV(eGFP) infections (Figure 5.15, a). Whilst detection of eGFP in rHAZV(eGFP) occurred at 13 to 14 hpi, in line with that observed previously in section 4.5, eGFP signal was detected much later in infections with rHAZV(eGFP-LFAA) at 25 to 26 hpi. This was evident from examining the images collected during the infections, with green cells barely visible in images taken 24 hpi for rHAZV(eGFP-LFAA), but multiple strong intensity green cells present for rHAZV(eGFP) (Figure 5.15, b). Whilst the initial translation event was delayed, the rate at which the green cell count increased following the initial detection of eGFP was similar for both viruses and reached a similar peak, indicating once rHAZV(eGFP-LFAA) had overcome the delay caused by the mutation, it did not hinder further stages of the viral life cycle. This would suggest a role of the LFAD motif in a stage prior to translation.

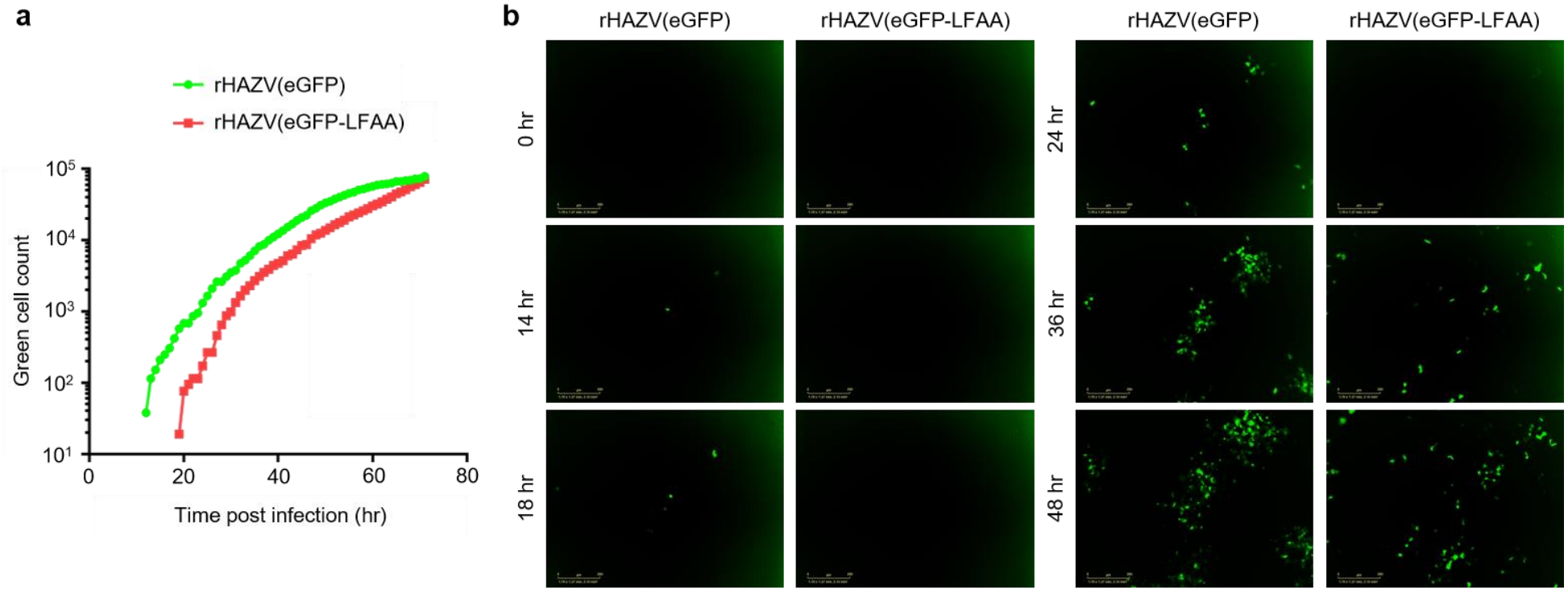


Figure 5.15, rHAZV(eGFP-LFAA) displays delayed translation vs rHAZV(eGFP)

a) Graphical representation of green cell count versus time post infection of SW13 cells infected with rHAZV(eGFP) or rHAZV(eGFP-LFAA) for 72 hours. Analysis was carried out using the IncuCyte® live cell imaging system and plotted using GraphPad Prism. b) Representative images of infections described in (a) taken at 0, 14, 18, 24, 36 and 48 hours post infection on the IncuCyte® live cell imaging system.

This would support the relationship between the LFAD motif and virus entry and could be explained by the native role of SNX1. The inability of rHAZV(eGFP-LFAA) to prevent cleavage of SNX1 during an apoptotic response to viral infection may impact the ability of rHAZV(eGFP-LFAA) to enter the cell. However, confusion arises regarding the timing of the delay in relation to SNX1. If the LFAD role is to protect SNX1 from apoptosis mediated cleavage, then it would follow that rHAZV had already entered and triggered apoptosis and therefore would not require the membrane receptors recycled by SNX1.

To confirm whether SNX1 was a factor in rHAZV infections, siRNA knockdown was carried out. SW13 cells were pretreated with an siRNA targeting SNX1, using the same protocol as described in section 4.6. Following a 24 hour infection, SW13 cell lysates were collected and analysed for SNX1 and HAZV N expression via western blot (Figure 5.16, a). SNX1 levels were visibly reduced from the western blot analysis in lysates collected from SW13 cells that had been treated with either 15 or 30 nM siRNA. Analysis of normalised densitometric data of HAZV N revealed a trend in reduction of HAZV N expression with increasing concentration of the siRNA (Figure 5.16, b). The reporter rHAZV(eGFP) was also used to provide an alternative method of quantifying the effect of SNX1 on rHAZV infections (Figure 5.16, c). In this case analysis of the TIIE in SNX1 knockdown SW13 cells infected with rHAZV(eGFP) resulted in a similar trend observed in Figure 5.16, a. However, with the analysis this time carried out by the IncuCyte® live cell imaging system software the results were significant, indicating a role of SNX1 in HAZV infection at a stage including or prior to translation (Figure 5.16, d). Whether the role of SNX1 in HAZV replication is linked to the LFAD motif present in HAZV N was not confirmed experimentally in this project, though further investigation into this area could uncover important aspects of HAZV entry.

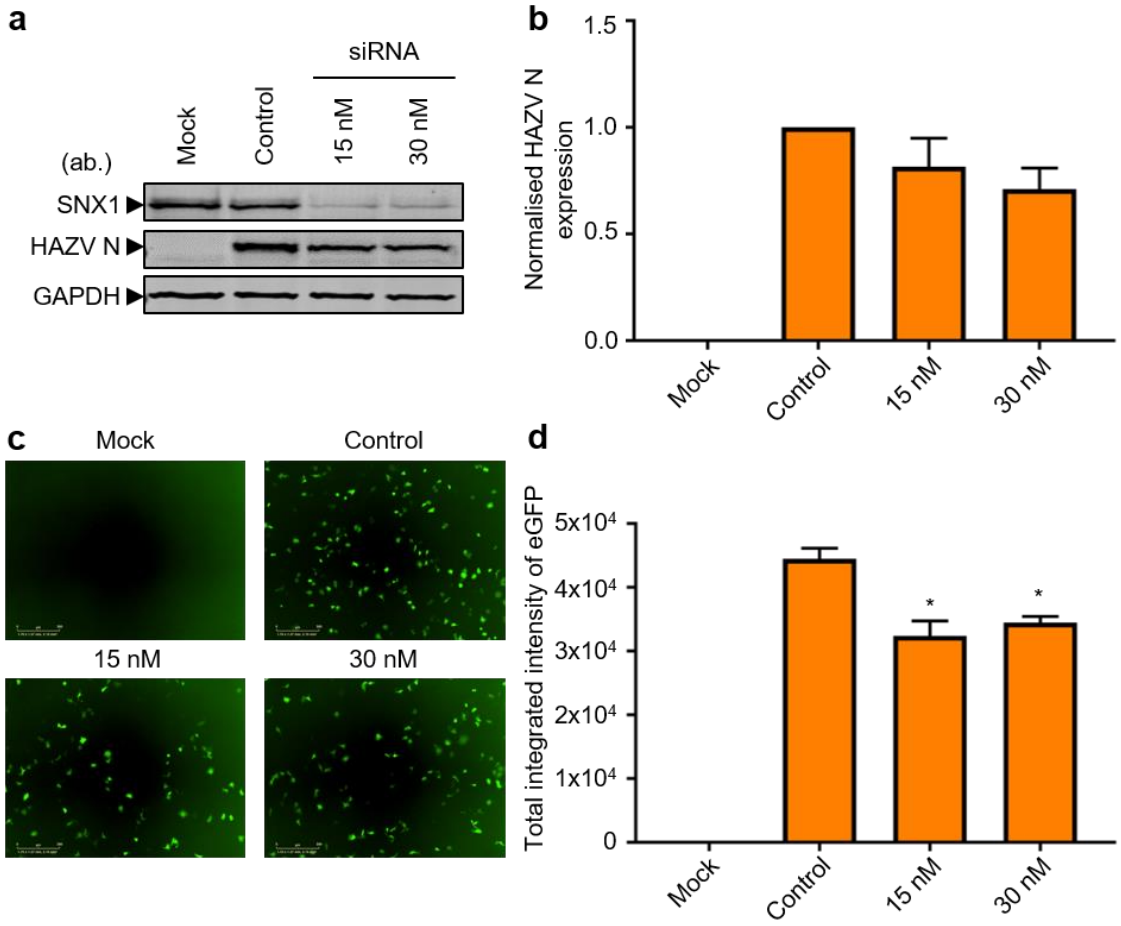


Figure 5.16, siRNA knockdown of SNX1 reduces rHAZV infectivity in SW13 cells

a) Western blot analysis of SNX1 knockdown SW13 cells infected with WT rHAZV. Cells were pretreated with 15 or 30 nM of siRNA then infected with rHAZV at an MOI of 0.1 for 24 hours. Total cell lysates were analysed for SNX1 and HAZV N, detection of GAPDH, using antisera, was also used as a loading control. A control in which SNX1 was not targeted by siRNA and infected with rHAZV was also included (Control). b) Histogram displays normalised densitometric analysis of western blot data from (a), calculated using the ImageJ software package, error bars represent data from 2 experimental repeats. Plotted using GraphPad Prism. c) Representative images of rHAZV(eGFP) infections taken 24 hours post infection, imaged on the IncuCyte® live cell imaging system. d) Histogram displays total integrated intensity of eGFP from images taken and analysed by the IncuCyte® live cell imaging system, error bars represent data from 2 experimental repeats. Plotted using GraphPad Prism.

5.5 Chapter summary and discussion

In this chapter, the infectious clone system developed as described in chapter 3 then modified as described in chapter 4 was further utilised to explore the role of conserved cleavage sites within HAZV N. Initially, extensive cleavage of HAZV N was identified in SW13 cells, with multiple previously undocumented cleavage products identified. Critically cleavage of HAZV N was not observed in

HAE/CTVM9 cells, uncovering a stark difference in processing of the viral protein in two cell types with physiological relevance to the HAZV life cycle. Further characterisation of the cleavage products via inhibition of activated caspases identified the origin of the highest abundance products, specifically those at ≈ 45, 30 and 20 kDa, to be caused by caspase cleavage. Interestingly, despite extensive conservation of the caspase cleavage motif on the tip of the arm domain of HAZV N and CCHFV N, inhibition of its cleavage had no effect on virus infection up to the point of translation. This would support the role of cleavage resulting in these specific cleavage products is as a decoy substrate, as the lack of the cleaved form of the protein did not appear to affect viral replication.

The infectious clone system was then utilised to modify multiple proposed cleavage sites on HAZV N. Initially, the caspase 3 cleavage motif DQVD was mutated to change the P1 and/or P4 aspartic acids (which also formed the P1 residue of an adjacent cleavage motif) to either alanine or glutamic acid, resulting in an ENKE motif, a DQVE motif and an AQVA motif. This was chosen to knock out the ability of the motif to be cleaved, whilst maintaining structural integrity. A previous study, published during the course of this project, attempted a similar task using the infectious clone model for CCHFV, in which the DEVD motif was converted to AEVA. They found that despite the motif being dispensable for replication in mammalian cell lines, the mutation of both P1 and P4 aspartic acid residues rendered the virus incapable of infecting tick cells (Salata *et al.*, 2018). Therefore, it was a surprising and substantial discovery that the similar mutant generated in this project, AQVA, was unable to be recovered, despite multiple attempts. In both of the mutants where a single change to the DQVD motif was made, successful recovery of virus was achieved. The fact that both of the resulting viruses, rHAZV(ENKE) and rHAZV(DQVE), were viable and demonstrated multi step growth kinetics that were indistinguishable from those of WT rHAZV is a critical observation. As the DQVE mutant is cleaved at neither the DQVE site nor the adjacent ENKD site, as shown by the shifted band following western blot analysis, these findings show cleavage within the arm apex is not a requirement for virus viability. The alternative cleavage shown in infection with rHAZV(DQVE) shows that alternative cleavage sites exist that are amenable to cleavage once the dominant sites are removed. Deciphering the caspase

cleavage profiles of nairoviral N proteins is a complex task, as there are a total of 28 aspartate residues that could potentially be cleaved in HAZV N alone.

The failure to rescue the rHAZV-AQVA mutant despite repeated efforts suggests it is deficient in a critical function. The lack of rescue of rHAZV(AQVD), when contexed with the successful rescue of the cleavage deficient rHAZV(ENKE) and rHAZV(DQVE), shows it cannot be due to loss of cleavability of the arm apex. The results within this chapter show the DQVD site is important for rescue but not for caspase cleavage, which provides a starting point for studies to elucidate the critical role of this DQVD motif and provide insight into how CCHFV is able to remain viable following a similar mutation. A possibility is that the DQVD site is involved in an interaction with a host factor, and the availability of such a component may also explain the differential outcomes of CCHFV infection in mammalian and tick cells.

In addition to the DQVD motif, the DVMD motif located between the arm and globular domains of HAZV was also investigated for its importance in HAZV replication. Again, mutagenesis of the motif to an uncleavable form rendered rescue of rHAZV(DVMA) impossible. Though as the location of this motif lies at an interaction site between two domains of the protein it may be due to the loss of structural integrity rather than any active role within infection.

Finally, the role of a completely conserved motif, LFAD, was explored. Despite initial recovery of rHAZV(LFAA) containing a caspase cleavage deficient motif, the virus had severe fitness issues. Plaques generated by rHAZV(LFAA) displayed poor clarity and resulted in failure to accurately titrate viral stocks to permit comparisons. Incorporation of the LFAA mutation into rHAZV(eGFP) resolved this issue, permitting equal MOI infections of rHAZV(eGFP) and rHAZV(eGFP-LFAA) to be made. Further investigation into the role of the LFAD motif revealed a substantial delay in translation when the motif was knocked out, recovery of expression levels comparable to the WT motif following this delay suggested a pre-translational effect. SNX1, located within endosomes, is responsible for sorting of membrane receptors and inducing curvature of endosomal membranes (Peter *et al.*, 2004; Carlton *et al.*, 2005). Interestingly, it

is cleaved during the apoptotic response at an LFAD motif (Duclos, Lavoie and Denault, 2017), therefore it is plausible that HAZV N is also acting as a decoy for this protein as well. siRNA knockdown of SNX1 prior to infection with rHAZV(eGFP) demonstrated a slight effect on rHAZV, supporting this statement, however the scale of the reduction of rHAZV infection was not close to the loss of fitness displayed by rHAZV(LFAA), despite almost total knockdown of detectable SNX1. Further research into understanding the role of the LFAD motif in HAZV N may reveal a critical role of the lifecycle that could be targeted by anti-viral drugs, due to the easily accessible location of the motif on the surface of HAZV N.

The WT infectious clone system, in combination with the eGFP reporter virus have again proven to be extremely powerful tools in this chapter. The use of the system demonstrated significant differences between conserved motifs on HAZV N and CCHFV N and uncovered novel motifs of importance in relation to HAZV infection of mammalian cells. The ability to efficiently mutate multiple motifs within HAZV N demonstrates the utility of the recombinant system, allowing detailed experimentation into highly specific regions of the virus, which will no doubt prove invaluable in studies into any of the alternate proteins expressed by the virus.

References

Abudurexiti, A. *et al.* (2019) ‘Taxonomy of the order Bunyavirales: update 2019’, *Archives of Virology*, 164(7), pp. 1949–1965. doi: 10.1007/s00705-019-04253-6.

Albariño, C. G. *et al.* (2009) ‘Efficient reverse genetics generation of infectious Junin viruses differing in glycoprotein processing’, *Journal of virology*. Am Soc Microbiol, 83(11), pp. 5606–5614.

Ali, S. H. and DeCaprio, J. A. (2001) ‘Cellular transformation by SV40 large T antigen: interaction with host proteins.’, *Seminars in cancer biology*. England, 11(1), pp. 15–23. doi: 10.1006/scbi.2000.0342.

Aligholipour Farzani, T. *et al.* (2019) ‘Immunological Analysis of a CCHFV mRNA Vaccine Candidate in Mouse Models.’, *Vaccines*, 7(3). doi: 10.3390/vaccines7030115.

Altamura, L. A. *et al.* (2007) ‘Identification of a Novel C-Terminal Cleavage of Crimean-Congo Hemorrhagic Fever Virus PreG_N That Leads to Generation of an NS_M Protein’, *Journal of Virology*, 81(12), pp. 6632 LP – 6642. doi: 10.1128/JVI.02730-06.

Aniento, F. *et al.* (1996) ‘An endosomal beta COP is involved in the pH-dependent formation of transport vesicles destined for late endosomes.’, *Journal of Cell Biology*, 133(1), pp. 29–41. doi: 10.1083/jcb.133.1.29.

Antoniadis, A. and Casals, J. (1982) ‘Serological evidence of human infection with Congo-Crimean hemorrhagic fever virus in Greece.’, *The American journal of tropical medicine and hygiene*. United States, 31(5), pp. 1066–1067. doi: 10.4269/ajtmh.1982.31.1066.

Appannanavar, S. B. and Mishra, B. (2011) ‘An update on crimean congo hemorrhagic Fever’, *Journal of global infectious diseases*. Medknow Publications, 3(3), pp. 285–292. doi: 10.4103/0974-777X.83537.

Atkinson, B. *et al.* (2012) ‘Sequencing and phylogenetic characterisation of a fatal Crimean - Congo haemorrhagic fever case imported into the United Kingdom, October 2012.’, *Euro surveillance : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin*. Sweden, 17(48).

Bakir, M. *et al.* (2005) ‘Crimean-Congo haemorrhagic fever outbreak in Middle Anatolia: a multicentre study of clinical features and outcome measures.’,

Journal of medical microbiology. England, 54(Pt 4), pp. 385–389. doi: 10.1099/jmm.0.45865-0.

Barnwal, B. *et al.* (2016) 'The Non-structural Protein of Crimean-Congo Hemorrhagic Fever Virus Disrupts the Mitochondrial Membrane Potential and Induces Apoptosis', *Journal of Biological Chemistry*, 291(2), pp. 582–592. doi: 10.1074/jbc.M115.667436.

Begum, F., Wisseman Jr, C. L. and Traub, R. (1970) 'Tick-borne viruses of West Pakistan: I. Isolation and general characteristics', *American journal of epidemiology*. Oxford University Press, 92(3), pp. 180–191.

Bell-Sakyi, L. (1991) 'Continuous cell lines from the tick Hyalomma anatolicum anatolicum', *Journal of parasitology*, 77(6), pp. 1006–1008. doi: 10.2307/3282757.

Belov, G. A., Fogg, M. H. and Ehrenfeld, E. (2005) 'Poliovirus proteins induce membrane association of GTPase ADP-ribosylation factor.', *Journal of virology*, 79(11), pp. 7207–7216. doi: 10.1128/JVI.79.11.7207-7216.2005.

Bente, D. A. *et al.* (2010) 'Pathogenesis and immune response of Crimean-Congo hemorrhagic fever virus in a STAT-1 knockout mouse model.', *Journal of virology*, 84(21), pp. 11089–11100. doi: 10.1128/JVI.01383-10.

Bente, D. A. *et al.* (2013) 'Crimean-Congo hemorrhagic fever: history, epidemiology, pathogenesis, clinical syndrome and genetic diversity.', *Antiviral research*. Netherlands, 100(1), pp. 159–189. doi: 10.1016/j.antiviral.2013.07.006.

Bergeron, E. *et al.* (2010) 'Crimean-Congo hemorrhagic fever virus-encoded ovarian tumor protease activity is dispensable for virus RNA polymerase function.', *Journal of virology*, 84(1), pp. 216–226. doi: 10.1128/JVI.01859-09.

Bergeron, É. *et al.* (2015) 'Recovery of Recombinant Crimean Congo Hemorrhagic Fever Virus Reveals a Function for Non-structural Glycoproteins Cleavage by Furin', *PLoS Pathogens*, 11(5). doi: 10.1371/journal.ppat.1004879.

Bergeron, É., Vincent, M. J. and Nichol, S. T. (2007) 'Crimean-Congo Hemorrhagic Fever Virus Glycoprotein Processing by the Endoprotease SKI-1/S1P Is Critical for Virus Infectivity', *Journal of Virology*, 81(23), pp. 13271 LP – 13276. doi: 10.1128/JVI.01647-07.

Bertolotti-Ciarlet, A. *et al.* (2005) 'Cellular localization and antigenic characterization of crimean-congo hemorrhagic fever virus glycoproteins.', *Journal of virology*, 79(10), pp. 6152–6161. doi: 10.1128/JVI.79.10.6152-

6161.2005.

Bird, B. H., Albariño, C. G. and Nichol, S. T. (2007) 'Rift Valley fever virus lacking NSm proteins retains high virulence in vivo and may provide a model of human delayed onset neurologic disease.', *Virology*. United States, 362(1), pp. 10–15. doi: 10.1016/j.virol.2007.01.046.

Blakqori, G. *et al.* (2007) 'La Crosse Bunyavirus Nonstructural Protein NSs Serves To Suppress the Type I Interferon System of Mammalian Hosts', *Journal of Virology*, 81(10), pp. 4991–4999. doi: 10.1128/JVI.01933-06.

Blondeau, F. *et al.* (2004) 'Tandem MS analysis of brain clathrin-coated vesicles reveals their critical involvement in synaptic vesicle recycling', *Proceedings of the National Academy of Sciences of the United States of America*, 101(11), pp. 3833 LP – 3838. doi: 10.1073/pnas.0308186101.

Bridgen, A. *et al.* (2001) 'Bunyamwera bunyavirus nonstructural protein NSs is a nonessential gene product that contributes to viral pathogenesis', *Proceedings of the National Academy of Sciences*, 98(2), pp. 664 LP – 669. doi: 10.1073/pnas.98.2.664.

Bridgen, A. and Elliott, R. M. (1996) 'Rescue of a segmented negative-strand RNA virus entirely from cloned complementary DNAs', *Proceedings of the National Academy of Sciences*, 93(26), pp. 15400 LP – 15404. doi: 10.1073/pnas.93.26.15400.

Burney, M. I. *et al.* (1980) 'Nosocomial outbreak of viral hemorrhagic fever caused by Crimean Hemorrhagic fever-Congo virus in Pakistan, January 1976.', *The American journal of tropical medicine and hygiene*. United States, 29(5), pp. 941–947. doi: 10.4269/ajtmh.1980.29.941.

Butenko, A. M. *et al.* (1968) *Isolation and Investigation of Astrakhan Strain ('Drozdov') of Crimean Hemorrhagic Fever Virus and Data on Serodiagnosis of this Infection*. NAVAL MEDICAL RESEARCH UNIT NO 3 FPO NEW YORK 09527.

Buttigieg, K. R. *et al.* (2014) 'A novel vaccine against Crimean-Congo Haemorrhagic Fever protects 100% of animals against lethal challenge in a mouse model.', *PloS one*, 9(3), p. e91516. doi: 10.1371/journal.pone.0091516.

Carlton, J. G. *et al.* (2005) 'Sorting nexin-2 is associated with tubular elements of the early endosome, but is not essential for retromer-mediated endosome-to-TGN transport', *Journal of Cell Science*, 118(19), pp. 4527 LP –

4539. doi: 10.1242/jcs.02568.

Carter, S. D. *et al.* (2012) 'Structure, function, and evolution of the Crimean-Congo hemorrhagic fever virus nucleocapsid protein.', *Journal of virology*, 86(20), pp. 10914–10923. doi: 10.1128/JVI.01555-12.

Casanova, J. E. (2007) 'Regulation of Arf Activation: the Sec7 Family of Guanine Nucleotide Exchange Factors', *Traffic*. John Wiley & Sons, Ltd, 8(11), pp. 1476–1485. doi: 10.1111/j.1600-0854.2007.00634.x.

Causey, O. R. *et al.* (1970) 'Congo virus from domestic livestock, African hedgehog, and arthropods in Nigeria.', *The American journal of tropical medicine and hygiene*. United States, 19(5), pp. 846–850. doi: 10.4269/ajtmh.1970.19.846.

Celikbas, A. K. *et al.* (2014) 'Crimean-congo hemorrhagic fever among health care workers, Turkey', *Emerging Infectious Diseases*. doi: 10.3201/eid2003.131353.

Charlton, F. W. *et al.* (2019) 'Cellular cholesterol abundance regulates potassium accumulation within endosomes and is an important determinant in bunyavirus entry.', *The Journal of biological chemistry*, 294(18), pp. 7335–7347. doi: 10.1074/jbc.RA119.007618.

Cherry, S. *et al.* (2006) 'COPI Activity Coupled with Fatty Acid Biosynthesis Is Required for Viral Replication', *PLOS Pathogens*. Public Library of Science, 2(10), p. e102. Available at: https://doi.org/10.1371/journal.ppat.0020102.

Chlanda, P. *et al.* (2016) 'The hemifusion structure induced by influenza virus haemagglutinin is determined by physical properties of the target membranes.', *Nature microbiology*, 1(6), p. 16050. doi: 10.1038/nmicrobiol.2016.50.

Christova, I. *et al.* (2009) 'Crimean-Congo hemorrhagic fever, southwestern Bulgaria.', *Emerging infectious diseases*, pp. 983–985. doi: 10.3201/eid1506.081567.

Chumakov, M. . (1945) 'A new tick-borne virus disease—Crimean hemorrhagic fever. Crimean hemorrhagic fever (acute infectious capillary toxicosis)', *Moscow: Izd Otd Primorskoi Armii*, 13(45).

Claude, A. *et al.* (1999) 'GBF1: A novel Golgi-associated BFA-resistant guanine nucleotide exchange factor that displays specificity for ADP-ribosylation factor 5.', *The Journal of cell biology*, 146(1), pp. 71–84.

Collett, M. S. (1986) 'Messenger RNA of the M segment RNA of Rift Valley fever virus.', *Virology*. United States, 151(1), pp. 151–156. doi: 10.1016/0042-

6822(86)90114-5.

Collins, P. L. *et al.* (1995) ‘Production of infectious human respiratory syncytial virus from cloned cDNA confirms an essential role for the transcription elongation factor from the 5’ proximal open reading frame of the M2 mRNA in gene expression and provides a capability for vaccine’, *Proceedings of the National Academy of Sciences of the United States of America*, 92(25), pp. 11563–11567. doi: 10.1073/pnas.92.25.11563.

Connolly-Andersen, A.-M., Magnusson, K.-E. and Mirazimi, A. (2007) ‘Basolateral entry and release of Crimean-Congo hemorrhagic fever virus in polarized MDCK-1 cells.’, *Journal of virology*, 81(5), pp. 2158–2164. doi: 10.1128/JVI.02070-06.

Conzelmann, K. K. and Schnell, M. (1994) ‘Rescue of synthetic genomic RNA analogs of rabies virus by plasmid-encoded proteins.’, *Journal of virology*, 68(2), pp. 713–719.

Crotty, S. *et al.* (2000) ‘The broad-spectrum antiviral ribonucleoside ribavirin is an RNA virus mutagen’, *Nature Medicine*, 6(12), pp. 1375–1379. doi: 10.1038/82191.

Devignot, S. *et al.* (2015) ‘A virus-like particle system identifies the endonuclease domain of Crimean-Congo hemorrhagic fever virus.’, *Journal of virology*, 89(11), pp. 5957–5967. doi: 10.1128/JVI.03691-14.

Deyde, V. M. *et al.* (2006) ‘Crimean-Congo hemorrhagic fever virus genomics and global diversity.’, *Journal of virology*, 80(17), pp. 8834–8842. doi: 10.1128/JVI.00752-06.

Dias, A. *et al.* (2009) ‘The cap-snatching endonuclease of influenza virus polymerase resides in the PA subunit.’, *Nature*. England, 458(7240), pp. 914–918. doi: 10.1038/nature07745.

Donaldson, J. G. and Jackson, C. L. (2000) ‘Regulators and effectors of the ARF GTPases.’, *Current opinion in cell biology*. England, 12(4), pp. 475–482. doi: 10.1016/s0955-0674(00)00119-8.

Donnelly, M. L. L. *et al.* (2001) ‘The “cleavage” activities of foot-and-mouth disease virus 2A site-directed mutants and naturally occurring “2A-like” sequences.’, *The Journal of general virology*. England, 82(Pt 5), pp. 1027–1041. doi: 10.1099/0022-1317-82-5-1027.

Drosten, C. *et al.* (2003) ‘Molecular diagnostics of viral hemorrhagic

fevers.', *Antiviral research*. Netherlands, 57(1–2), pp. 61–87. doi: 10.1016/s0166-3542(02)00201-2.

Duclos, C., Lavoie, C. and Denault, J. B. (2017) 'Caspases rule the intracellular trafficking cartel', *FEBS Journal*, 284(10), pp. 1394–1420. doi: 10.1111/febs.14071.

Dunn, E. F. *et al.* (1995) 'Transcription of a recombinant bunyavirus RNA template by transiently expressed bunyavirus proteins', *Virology*. Elsevier, 211(1), pp. 133–143.

Durden, L. A. *et al.* (1993) 'Experimental vector incompetence of a soft tick, Ornithodoros sonrai (Acari: Argasidae), for Crimean-Congo hemorrhagic fever virus.', *Journal of medical entomology*. England, 30(2), pp. 493–496. doi: 10.1093/jmedent/30.2.493.

Ekert, P. G. and Vaux, D. L. (1997) 'Apoptosis and the immune system.', *British medical bulletin*. England, 53(3), pp. 591–603. doi: 10.1093/oxfordjournals.bmb.a011632.

Elmore, S. (2007) 'Apoptosis: A Review of Programmed Cell Death', *Toxicologic Pathology*, 35(4), pp. 495–516. doi: 10.1080/01926230701320337.

Ergönül, O. (2006) 'Crimean-Congo haemorrhagic fever.', *The Lancet. Infectious diseases*. United States, 6(4), pp. 203–214. doi: 10.1016/S1473-3099(06)70435-2.

Erickson, B. R. *et al.* (2007) 'N-linked glycosylation of Gn (but not Gc) is important for Crimean Congo hemorrhagic fever virus glycoprotein localization and transport.', *Virology*. United States, 361(2), pp. 348–355. doi: 10.1016/j.virol.2006.11.023.

Estrada-Peña, A. and de la Fuente, J. (2014) 'The ecology of ticks and epidemiology of tick-borne viral diseases.', *Antiviral research*. Netherlands, 108, pp. 104–128. doi: 10.1016/j.antiviral.2014.05.016.

Fernandez-Garcia, M. D. *et al.* (2014) 'European survey on laboratory preparedness, response and diagnostic capacity for Crimean-Congo haemorrhagic fever, 2012.', *Euro surveillance : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin*. Sweden, 19(26). doi: 10.2807/1560-7917.es2014.19.26.20844.

Ferré-D'Amaré, A. R. and Scott, W. G. (2010) 'Small self-cleaving ribozymes.', *Cold Spring Harbor perspectives in biology*, 2(10), p. a003574. doi:

10.1101/cshperspect.a003574.

Filippone, C. *et al.* (2013) ‘Molecular diagnostic and genetic characterization of highly pathogenic viruses: application during Crimean-Congo haemorrhagic fever virus outbreaks in Eastern Europe and the Middle East.’, *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases*, 19(2), pp. E118-28. doi: 10.1111/1469-0691.12075.

Filippova, M. *et al.* (2007) ‘The large and small isoforms of human papillomavirus type 16 E6 bind to and differentially affect procaspase 8 stability and activity.’, *Journal of virology*, 81(8), pp. 4116–4129. doi: 10.1128/JVI.01924-06.

Flick, K. *et al.* (2003) ‘Rescue of Hantaan virus minigenomes.’, *Virology*. United States, 306(2), pp. 219–224. doi: 10.1016/s0042-6822(02)00070-3.

Flick, R. *et al.* (2003) ‘Reverse genetics for crimean-congo hemorrhagic fever virus.’, *Journal of virology*, 77(10), pp. 5997–6006. doi: 10.1128/jvi.77.10.5997-6006.2003.

Flick, R. and Pettersson, R. F. (2001) ‘Reverse genetics system for Uukuniemi virus (Bunyaviridae): RNA polymerase I-catalyzed expression of chimeric viral RNAs.’, *Journal of virology*, 75(4), pp. 1643–1655. doi: 10.1128/JVI.75.4.1643-1655.2001.

Fontana, J. *et al.* (2008) ‘The unique architecture of Bunyamwera virus factories around the Golgi complex.’, *Cellular microbiology*, 10(10), pp. 2012–2028. doi: 10.1111/j.1462-5822.2008.01184.x.

Foulke, R. S., Rosato, R. R. and French, G. R. (1981) ‘Structural polypeptides of Hazara virus.’, *The Journal of general virology*. England, 53(Pt 1), pp. 169–172. doi: 10.1099/0022-1317-53-1-169.

Frias-Staheli, N. *et al.* (2007) ‘Ovarian tumor domain-containing viral proteases evade ubiquitin- and ISG15-dependent innate immune responses.’, *Cell host & microbe*, 2(6), pp. 404–416. doi: 10.1016/j.chom.2007.09.014.

Garrison, A. R. *et al.* (2013) ‘Crimean-Congo hemorrhagic fever virus utilizes a clathrin- and early endosome-dependent entry pathway.’, *Virology*. United States, 444(1–2), pp. 45–54. doi: 10.1016/j.virol.2013.05.030.

Gazina, E. V *et al.* (2002) ‘Differential requirements for COPI coats in formation of replication complexes among three genera of Picornaviridae.’,

Journal of virology, 76(21), pp. 11113–11122. doi: 10.1128/jvi.76.21.11113-11122.2002.

Gerlach, P. *et al.* (2015) 'Structural Insights into Bunyavirus Replication and Its Regulation by the vRNA Promoter.', *Cell*, 161(6), pp. 1267–1279. doi: 10.1016/j.cell.2015.05.006.

Guo, Y. *et al.* (2012) 'Crimean-Congo hemorrhagic fever virus nucleoprotein reveals endonuclease activity in bunyaviruses.', *Proceedings of the National Academy of Sciences of the United States of America*, 109(13), pp. 5046–5051. doi: 10.1073/pnas.1200808109.

Haferkamp, S. *et al.* (2005) 'Intracellular localization of Crimean-Congo Hemorrhagic Fever (CCHF) virus glycoproteins.', *Virology journal*, 2, p. 42. doi: 10.1186/1743-422X-2-42.

Han, J. *et al.* (1996) 'The E1B 19K protein blocks apoptosis by interacting with and inhibiting the p53-inducible and death-promoting Bax protein.', *Genes & development*. United States, 10(4), pp. 461–477. doi: 10.1101/gad.10.4.461.

Hansford, K. M. *et al.* (2019) 'Hyalomma rufipes on an untraveled horse: Is this the first evidence of Hyalomma nymphs successfully moulting in the United Kingdom?', *Ticks and tick-borne diseases*. Netherlands, 10(3), pp. 704–708. doi: 10.1016/j.ttbdis.2019.03.003.

Hawman, D. W. *et al.* (2018) 'Favipiravir (T-705) but not ribavirin is effective against two distinct strains of Crimean-Congo hemorrhagic fever virus in mice.', *Antiviral research*. Netherlands, 157, pp. 18–26. doi: 10.1016/j.antiviral.2018.06.013.

Henderson, S. *et al.* (1993) 'Epstein-Barr virus-coded BHRF1 protein, a viral homologue of Bcl-2, protects human B cells from programmed cell death.', *Proceedings of the National Academy of Sciences of the United States of America*, 90(18), pp. 8479–8483. doi: 10.1073/pnas.90.18.8479.

Hinkula, J. *et al.* (2017) 'Immunization with DNA Plasmids Coding for Crimean-Congo Hemorrhagic Fever Virus Capsid and Envelope Proteins and/or Virus-Like Particles Induces Protection and Survival in Challenged Mice.', *Journal of virology*, 91(10). doi: 10.1128/JVI.02076-16.

Honig, J. E., Osborne, J. C. and Nichol, S. T. (2004) 'Crimean-Congo hemorrhagic fever virus genome L RNA segment and encoded protein.', *Virology*. United States, 321(1), pp. 29–35. doi: 10.1016/j.virol.2003.09.042.

Hoogstraal, H. (1966) ‘Ticks in Relation to Human Diseases Caused by Viruses’, *Annual Review of Entomology*. Annual Reviews, 11(1), pp. 261–308. doi: 10.1146/annurev.en.11.010166.001401.

Hoogstraal, H. (1979) ‘The epidemiology of tick-borne Crimean-Congo hemorrhagic fever in Asia, Europe, and Africa.’, *Journal of medical entomology*, 15(4), pp. 307–417. doi: 10.1093/jmedent/15.4.307.

Hover, S. *et al.* (2016) ‘Modulation of Potassium Channels Inhibits Bunyavirus Infection.’, *The Journal of biological chemistry*, 291(7), pp. 3411–3422. doi: 10.1074/jbc.M115.692673.

Inoue, H. and Randazzo, P. A. (2007) ‘Arf GAPs and their interacting proteins.’, *Traffic (Copenhagen, Denmark)*. England, 8(11), pp. 1465–1475. doi: 10.1111/j.1600-0854.2007.00624.x.

Jackson, C. L. and Casanova, J. E. (2000) ‘Turning on ARF: the Sec7 family of guanine-nucleotide-exchange factors.’, *Trends in cell biology*. England, 10(2), pp. 60–67. doi: 10.1016/s0962-8924(99)01699-2.

James, T. W. *et al.* (2011) ‘Structural basis for the removal of ubiquitin and interferon-stimulated gene 15 by a viral ovarian tumor domain-containing protease.’, *Proceedings of the National Academy of Sciences of the United States of America*, 108(6), pp. 2222–2227. doi: 10.1073/pnas.1013388108.

Jameson, L. J. *et al.* (2012) ‘Importation of Hyalomma marginatum, vector of Crimean-Congo haemorrhagic fever virus, into the United Kingdom by migratory birds.’, *Ticks and tick-borne diseases*. Netherlands, 3(2), pp. 95–99. doi: 10.1016/j.ttbdis.2011.12.002.

Jeeva, S. *et al.* (2019) ‘Crimean-Congo hemorrhagic fever virus nucleocapsid protein harbors distinct RNA-binding sites in the stalk and head domains.’, *The Journal of biological chemistry*, 294(13), pp. 5023–5037. doi: 10.1074/jbc.RA118.004976.

Karlberg, H., Tan, Y. J. and Mirazimi, A. (2011) ‘Induction of caspase activation and cleavage of the viral nucleocapsid protein in different cell types during Crimean-Congo hemorrhagic fever virus infection’, *Journal of Biological Chemistry*, 286(5), pp. 3227–3234. doi: 10.1074/jbc.M110.149369.

Karti, S. S. *et al.* (2004) ‘Crimean-Congo hemorrhagic fever in Turkey.’, *Emerging infectious diseases*, 10(8), pp. 1379–1384. doi: 10.3201/eid1008.030928.

Kawamoto, K. *et al.* (2002) 'GBF1, a Guanine Nucleotide Exchange Factor for ADP-Ribosylation Factors, is Localized to the cis-Golgi and Involved in Membrane Association of the COPI Coat', *Traffic*. John Wiley & Sons, Ltd, 3(7), pp. 483–495. doi: 10.1034/j.1600-0854.2002.30705.x.

Kinsella, E. *et al.* (2004) 'Sequence determination of the Crimean-Congo hemorrhagic fever virus L segment.', *Virology*. United States, 321(1), pp. 23–28. doi: 10.1016/j.virol.2003.09.046.

Kleinfelter, L. M. *et al.* (2015) 'Haploid Genetic Screen Reveals a Profound and Direct Dependence on Cholesterol for Hantavirus Membrane Fusion.', *mBio*, 6(4), p. e00801. doi: 10.1128/mBio.00801-15.

Klemm, C. *et al.* (2013) 'Systems to establish bunyavirus genome replication in the absence of transcription.', *Journal of virology*, 87(14), pp. 8205–8212. doi: 10.1128/JVI.00371-13.

Kohl, A. *et al.* (2003) 'Bunyamwera Virus Nonstructural Protein NSs Counteracts Interferon Regulatory Factor 3-Mediated Induction of Early Cell Death', *Journal of Virology*, 77(14), pp. 7999–8008. doi: 10.1128/JVI.77.14.7999-8008.2003.

Kohl, A. *et al.* (2006) 'Genetic elements regulating packaging of the Bunyamwera orthobunyavirus genome.', *The Journal of general virology*. England, 87(Pt 1), pp. 177–187. doi: 10.1099/vir.0.81227-0.

Kormelink, R. *et al.* (1991) 'The nonstructural protein (NSs) encoded by the ambisense S RNA segment of tomato spotted wilt virus is associated with fibrous structures in infected plant cells.', *Virology*. United States, 181(2), pp. 459–468. doi: 10.1016/0042-6822(91)90878-f.

Lasecka, L. and Baron, M. D. (2014) 'The molecular biology of nairoviruses, an emerging group of tick-borne arboviruses', *Archives of Virology*, 159(6), pp. 1249–1265. doi: 10.1007/s00705-013-1940-z.

Lawrence, C. P. *et al.* (2006) 'The cathepsin B inhibitor, z-FA-FMK, inhibits human T cell proliferation in vitro and modulates host response to pneumococcal infection in vivo.', *Journal of immunology (Baltimore, Md. : 1950)*. United States, 177(6), pp. 3827–3836. doi: 10.4049/jimmunol.177.6.3827.

Leblebicioglu, H. *et al.* (2015) 'Consensus report: Preventive measures for Crimean-Congo Hemorrhagic Fever during Eid-al-Adha festival.', *International journal of infectious diseases : IJID : official publication of the International*

Society for Infectious Diseases. Canada, 38, pp. 9–15. doi: 10.1016/j.ijid.2015.06.029.

Lipatov, A. S. *et al.* (2008) 'The role of the N-terminal caspase cleavage site in the nucleoprotein of influenza A virus in vitro and in vivo.', *Archives of virology*. Austria, 153(3), pp. 427–434. doi: 10.1007/s00705-007-0003-8.

Liu, X. *et al.* (1996) 'Induction of apoptotic program in cell-free extracts: requirement for dATP and cytochrome c.', *Cell*. United States, 86(1), pp. 147–157. doi: 10.1016/s0092-8674(00)80085-9.

Liu, Z. *et al.* (2017) 'Systematic comparison of 2A peptides for cloning multi-genes in a polycistronic vector', *Scientific Reports*, 7(1), p. 2193. doi: 10.1038/s41598-017-02460-2.

Lowen, A. C. *et al.* (2004) 'Efficient bunyavirus rescue from cloned cDNA', *Virology*, 330(2), pp. 493–500. doi: 10.1016/j.virol.2004.10.009.

Lumley, S. *et al.* (2014) 'Non-fatal case of Crimean-Congo haemorrhagic fever imported into the United Kingdom (ex Bulgaria), June 2014.', *Euro surveillance : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin*. Sweden, 19(30). doi: 10.2807/1560-7917.es2014.19.30.20864.

Maag, D. *et al.* (2001) 'Hepatitis C virus RNA-dependent RNA polymerase (NS5B) as a mediator of the antiviral activity of ribavirin.', *The Journal of biological chemistry*. United States, 276(49), pp. 46094–46098. doi: 10.1074/jbc.C100349200.

Maes, P. *et al.* (2018) 'Taxonomy of the family Arenaviridae and the order Bunyavirales: update 2018.', *Archives of virology*. Austria, 163(8), pp. 2295–2310. doi: 10.1007/s00705-018-3843-5.

Makwana, D. *et al.* (2015) 'First confirmed case of Crimean-Congo haemorrhagic fever from Sirohi district in Rajasthan State, India', *Indian Journal of Medical Research*. doi: 10.4103/0971-5916.169221.

Maltezou, H. C. *et al.* (2010) 'Crimean-Congo hemorrhagic fever in Europe: current situation calls for preparedness.', *Euro surveillance : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin*. Sweden, 15(10), p. 19504.

Mamuchishvili, N. *et al.* (2015) 'Notes from the field: Increase in reported Crimean-Congo hemorrhagic fever cases--country of Georgia, 2014', *MMWR.*

Morbidity and mortality weekly report. U.S. Centers for Disease Control, 64(8), pp. 228–229. Available at: https://pubmed.ncbi.nlm.nih.gov/25742385.

Mardani, M. *et al.* (2003) 'The efficacy of oral ribavirin in the treatment of crimean-congo hemorrhagic fever in Iran.', *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*. United States, 36(12), pp. 1613–1618. doi: 10.1086/375058.

Martin, M. L. *et al.* (1985) 'Distinction between Bunyaviridae genera by surface structure and comparison with Hantaan virus using negative stain electron microscopy', *Archives of Virology*, 86(1), pp. 17–28. doi: 10.1007/BF01314110.

Matsumoto, Y. *et al.* (2019) 'A Minigenome Study of Hazara Nairovirus Genomic Promoters.', *Journal of virology*, 93(6). doi: 10.1128/JVI.02118-18.

Le May, N. *et al.* (2004) 'TFIIH transcription factor, a target for the Rift Valley hemorrhagic fever virus.', *Cell*. United States, 116(4), pp. 541–550. doi: 10.1016/s0092-8674(04)00132-1.

Mazzola, L. T. and Kelly-Cirino, C. (2019) 'Diagnostic tests for Crimean-Congo haemorrhagic fever: a widespread tickborne disease.', *BMJ global health*, 4(Suppl 2), p. e001114. doi: 10.1136/bmjgh-2018-001114.

Mega, D. F. *et al.* (2020) 'Mutagenic analysis of Hazara nairovirus non-translated regions during single and multi-step growth identifies both attenuating and functionally-critical sequences for virus replication', *Journal of Virology*, p. JVI.00357-20. doi: 10.1128/JVI.00357-20.

Misselwitz, B. *et al.* (2011) 'RNAi screen of Salmonella invasion shows role of COPI in membrane targeting of cholesterol and Cdc42.', *Molecular systems biology*, 7, p. 474. doi: 10.1038/msb.2011.7.

Morales, J. *et al.* (2014) 'Review of poly (ADP-ribose) polymerase (PARP) mechanisms of action and rationale for targeting in cancer and other diseases.', *Critical reviews in eukaryotic gene expression*, 24(1), pp. 15–28. doi: 10.1615/critreveukaryotgeneexpr.2013006875.

Morin, B. *et al.* (2010) 'The N-terminal domain of the arenavirus L protein is an RNA endonuclease essential in mRNA transcription', *PLoS pathogens*. Public Library of Science, 6(9).

Mukhopadhyay, A. *et al.* (1997) 'Sequential actions of Rab5 and Rab7 regulate endocytosis in the Xenopus oocyte.', *The Journal of cell biology*, 136(6),

pp. 1227–1237. doi: 10.1083/jcb.136.6.1227.

Nagata, T. *et al.* (1999) ‘Tissue tropism related to vector competence of Frankliniella occidentalis for tomato spotted wilt tospovirus.’, *The Journal of general virology*. England, 80 (Pt 2), pp. 507–515. doi: 10.1099/0022-1317-80-2-507.

Niu, T.-K. *et al.* (2005) ‘Dynamics of GBF1, a Brefeldin A-sensitive Arf1 exchange factor at the Golgi.’, *Molecular biology of the cell*, 16(3), pp. 1213–1222. doi: 10.1091/mbc.e04-07-0599.

Van Noorden, C. J. (2001) ‘The history of Z-VAD-FMK, a tool for understanding the significance of caspase inhibition.’, *Acta histochemica*. Germany, pp. 241–251. doi: 10.1078/0065-1281-00601.

Novoa, R. R. *et al.* (2005) ‘Virus factories: associations of cell organelles for viral replication and morphogenesis.’, *Biology of the cell*, 97(2), pp. 147–172. doi: 10.1042/BC20040058.

Nuttall, P. A. *et al.* (1994) ‘Adaptations of arboviruses to ticks.’, *Journal of medical entomology*. England, 31(1), pp. 1–9. doi: 10.1093/jmedent/31.1.1.

Papa, A. *et al.* (2002) ‘Genetic characterization of the M RNA segment of Crimean Congo hemorrhagic fever virus strains, China.’, *Emerging infectious diseases*, 8(1), pp. 50–53. doi: 10.3201/eid0801.010087.

Patterson, J. L. and Kolakofsky, D. (1984) ‘Characterization of La Crosse virus small-genome transcripts.’, *Journal of virology*, 49(3), pp. 680–685.

Peng, R. *et al.* (2020) ‘Structural insight into arenavirus replication machinery’, *Nature*, 579(7800), pp. 615–619. doi: 10.1038/s41586-020-2114-2.

Peter, B. J. *et al.* (2004) ‘BAR domains as sensors of membrane curvature: the amphiphysin BAR structure’, *Science*. American Association for the Advancement of Science, 303(5657), pp. 495–499.

Pflug, A. *et al.* (2014) ‘Structure of influenza A polymerase bound to the viral RNA promoter.’, *Nature*. England, 516(7531), pp. 355–360. doi: 10.1038/nature14008.

Phalen, T. and Kielian, M. (1991) ‘Cholesterol is required for infection by Semliki Forest virus.’, *The Journal of cell biology*, 112(4), pp. 615–623. doi: 10.1083/jcb.112.4.615.

Punch, E. K. *et al.* (2018) ‘Potassium is a trigger for conformational change in the fusion spike of an enveloped RNA virus’, *Journal of Biological Chemistry*,

293(26). doi: 10.1074/jbc.RA118.002494.

Racaniello, V. R. and Baltimore, D. (1981) 'Cloned poliovirus complementary DNA is infectious in mammalian cells.', *Science (New York, N.Y.)*. United States, 214(4523), pp. 916–919. doi: 10.1126/science.6272391.

Reguera, J., Weber, F. and Cusack, S. (2010) 'Bunyaviridae RNA polymerases (L-protein) have an N-terminal, influenza-like endonuclease domain, essential for viral cap-dependent transcription.', *PLoS pathogens*, 6(9), p. e1001101. doi: 10.1371/journal.ppat.1001101.

Rodrigues, R. *et al.* (2012) 'Crimean-Congo Hemorrhagic Fever Virus-Infected Hepatocytes Induce ER-Stress and Apoptosis Crosstalk', *PLoS ONE*. Edited by D. Goletti, 7(1), p. e29712. doi: 10.1371/journal.pone.0029712.

Rodriguez, S. E. *et al.* (2019) 'Vesicular Stomatitis Virus-Based Vaccine Protects Mice against Crimean-Congo Hemorrhagic Fever.', *Scientific reports*, 9(1), p. 7755. doi: 10.1038/s41598-019-44210-6.

Rojek, J. M. and Kunz, S. (2008) 'Cell entry by human pathogenic arenaviruses.', *Cellular microbiology*. England, 10(4), pp. 828–835. doi: 10.1111/j.1462-5822.2007.01113.x.

Rose, J. K. (2003) 'Optimization of transfection.', *Current protocols in cell biology*. United States, Chapter 20, p. Unit 20.7. doi: 10.1002/0471143030.cb2007s19.

Ruusala, A. *et al.* (1992) 'Coexpression of the membrane glycoproteins G1 and G2 of Hantaan virus is required for targeting to the Golgi complex.', *Virology*. United States, 186(1), pp. 53–64. doi: 10.1016/0042-6822(92)90060-3.

Sakahira, H., Enari, M. and Nagata, S. (1998) 'Cleavage of CAD inhibitor in CAD activation and DNA degradation during apoptosis.', *Nature*. England, 391(6662), pp. 96–99. doi: 10.1038/34214.

Salanueva, I. J. *et al.* (2003) 'Polymorphism and structural maturation of bunyamwera virus in Golgi and post-Golgi compartments.', *Journal of virology*, 77(2), pp. 1368–1381. doi: 10.1128/jvi.77.2.1368-1381.2003.

Salata, C. *et al.* (2018) 'The DEVD motif of Crimean-Congo hemorrhagic fever virus nucleoprotein is essential for viral replication in tick cells', *Emerging Microbes & Infections*, 7(1), p. 190. doi: 10.1038/s41426-018-0192-0.

Sanchez, A. J. *et al.* (2006) 'Crimean-congo hemorrhagic fever virus glycoprotein precursor is cleaved by Furin-like and SKI-1 proteases to generate

a novel 38-kilodalton glycoprotein.', *Journal of virology*, 80(1), pp. 514–525. doi: 10.1128/JVI.80.1.514-525.2006.

Sanchez, A. J., Vincent, M. J. and Nichol, S. T. (2002) 'Characterization of the Glycoproteins of Crimean-Congo Hemorrhagic Fever Virus', *Journal of Virology*, 76(14), pp. 7263 LP – 7275. doi: 10.1128/JVI.76.14.7263-7275.2002.

Schnell, M. J., Mebatsion, T. and Conzelmann, K. K. (1994) 'Infectious rabies viruses from cloned cDNA.', *The EMBO journal*, 13(18), pp. 4195–4203.

Scholte, F. E. M. *et al.* (2017) 'Crimean-Congo Hemorrhagic Fever Virus Suppresses Innate Immune Responses via a Ubiquitin and ISG15 Specific Protease', *Cell Reports*, 20(10), pp. 2396–2407. doi: 10.1016/j.celrep.2017.08.040.

Schwarz, T. F., Nsanze, H. and Ameen, A. M. (1997) 'Clinical features of Crimean-Congo haemorrhagic fever in the United Arab Emirates', *Infection*. doi: 10.1007/BF01740819.

Seaman, M. N. J. *et al.* (1997) 'Endosome to Golgi retrieval of the vacuolar protein sorting receptor, Vps10p, requires the function of the VPS29, VPS30, and VPS35 gene products', *The Journal of cell biology*. The Rockefeller University Press, 137(1), pp. 79–92.

Shepherd, A. J. *et al.* (1987) 'Field and laboratory investigation of Crimean-Congo haemorrhagic fever virus (Nairovirus, family Bunyaviridae) infection in birds.', *Transactions of the Royal Society of Tropical Medicine and Hygiene*. England, 81(6), pp. 1004–1007. doi: 10.1016/0035-9203(87)90379-8.

Shepherd, A. J. *et al.* (1989) 'Experimental studies on the replication and transmission of Crimean-Congo hemorrhagic fever virus in some African tick species.', *The American journal of tropical medicine and hygiene*. United States, 40(3), pp. 326–331. doi: 10.4269/ajtmh.1989.40.326.

Shi, X. *et al.* (2016) 'Bunyamwera orthobunyavirus glycoprotein precursor is processed by cellular signal peptidase and signal peptide peptidase', *Proceedings of the National Academy of Sciences*, 113(31), pp. 8825 LP – 8830. doi: 10.1073/pnas.1603364113.

Simon, M. *et al.* (2009) 'Microtubule-dependent and microtubule-independent steps in Crimean-Congo hemorrhagic fever virus replication cycle.', *Virology*. United States, 385(2), pp. 313–322. doi: 10.1016/j.virol.2008.11.020.

Simon, M., Johansson, C. and Mirazimi, A. (2009) 'Crimean-Congo

hemorrhagic fever virus entry and replication is clathrin-, pH- and cholesterol-dependent.', *The Journal of general virology*. England, 90(Pt 1), pp. 210–215. doi: 10.1099/vir.0.006387-0.

Skaletskaya, A. *et al.* (2001) 'A cytomegalovirus-encoded inhibitor of apoptosis that suppresses caspase-8 activation.', *Proceedings of the National Academy of Sciences of the United States of America*, 98(14), pp. 7829–7834. doi: 10.1073/pnas.141108798.

Smirnova, S. E. *et al.* (1977) 'Experimental Hazara Virus infection in mice.', *Acta virologica*. Slovakia, 21(2), pp. 128–132.

Smith, D. R. *et al.* (2018) 'Attenuation and efficacy of live-attenuated Rift Valley fever virus vaccine candidates in non-human primates', *PLOS Neglected Tropical Diseases*. Public Library of Science, 12(5), p. e0006474. Available at: https://doi.org/10.1371/journal.pntd.0006474.

Spengler, J. R. *et al.* (2019) 'Heterologous protection against Crimean-Congo hemorrhagic fever in mice after a single dose of replicon particle vaccine.', *Antiviral research*, 170, p. 104573. doi: 10.1016/j.antiviral.2019.104573.

Spengler, J. R., Bergeron, É. and Rollin, P. E. (2016) 'Seroepidemiological Studies of Crimean-Congo Hemorrhagic Fever Virus in Domestic and Wild Animals', *PLoS Neglected Tropical Diseases*. doi: 10.1371/journal.pntd.0004210.

Steinhauer, D. A., Domingo, E. and Holland, J. J. (1992) 'Lack of evidence for proofreading mechanisms associated with an RNA virus polymerase.', *Gene*. Netherlands, 122(2), pp. 281–288. doi: 10.1016/0378-1119(92)90216-c.

Stenmark, H. *et al.* (1994) 'Distinct structural elements of rab5 define its functional specificity.', *The EMBO journal*, 13(3), pp. 575–583.

Stennicke, H. R. *et al.* (2000) 'Internally quenched fluorescent peptide substrates disclose the subsite preferences of human caspases 1, 3, 6, 7 and 8.', *The Biochemical journal*, 350 Pt 2(Pt 2), pp. 563–568.

Surtees, R. *et al.* (2015) 'The crystal structure of the Hazara virus nucleocapsid protein', *BMC Structural Biology*, 15, p. 24. doi: 10.1186/s12900-015-0051-3.

Swanepoel, R. *et al.* (1987) 'Epidemiologic and clinical features of Crimean-Congo hemorrhagic fever in southern Africa.', *The American journal of tropical medicine and hygiene*. United States, 36(1), pp. 120–132. doi: 10.4269/ajtmh.1987.36.120.

Szemiel, A. M., Failloux, A.-B. and Elliott, R. M. (2012) ‘Role of Bunyamwera Orthobunyavirus NSs protein in infection of mosquito cells.’, *PLoS neglected tropical diseases*, 6(9), p. e1823. doi: 10.1371/journal.pntd.0001823.

Taylor, R. C., Cullen, S. P. and Martin, S. J. (2008) ‘Apoptosis: controlled demolition at the cellular level.’, *Nature reviews. Molecular cell biology*. England, 9(3), pp. 231–241. doi: 10.1038/nrm2312.

Tezer, H. and Polat, M. (2015) ‘Diagnosis of Crimean-Congo hemorrhagic fever.’, *Expert review of anti-infective therapy*. England, 13(5), pp. 555–566. doi: 10.1586/14787210.2015.1021782.

Thornberry, N. A. (1997) ‘The caspase family of cysteine proteases.’, *British medical bulletin*. England, 53(3), pp. 478–490. doi: 10.1093/oxfordjournals.bmb.a011625.

Vanhomwegen, J. *et al.* (2012) ‘Diagnostic assays for Crimean-Congo hemorrhagic fever.’, *Emerging infectious diseases*, 18(12), pp. 1958–1965. doi: 10.3201/eid1812.120710.

Vasilenko, S. (1976) *On the etiology, epidemiology and the specifc immunoprevention of Crimean hemorrhagic fever in Bulgaria*, *PhD Theses*. NCIPD, Sofia.

Vincent, M. J. *et al.* (2003) ‘Crimean-Congo hemorrhagic fever virus glycoprotein proteolytic processing by subtilase SKI-1.’, *Journal of virology*, 77(16), pp. 8640–8649. doi: 10.1128/jvi.77.16.8640-8649.2003.

Walter, C. T. and Barr, J. N. (2011) ‘Recent advances in the molecular and cellular biology of bunyaviruses.’, *The Journal of general virology*. England, 92(Pt 11), pp. 2467–2484. doi: 10.1099/vir.0.035105-0.

Wang, W. *et al.* (2015) ‘Structural and Functional Diversity of Nairovirus-Encoded Nucleoproteins’, *Journal of Virology*, 89(23), pp. 11740–11749. doi: 10.1128/JVI.01680-15.

Wang, Y. *et al.* (2012) ‘Structure of Crimean-Congo Hemorrhagic Fever Virus Nucleoprotein: Superhelical Homo-Oligomers and the Role of Caspase-3 Cleavage’, *Journal of Virology*, 86(22), pp. 12294–12303. doi: 10.1128/JVI.01627-12.

Watts, D. M. *et al.* (1988) ‘The arboviruses: epidemiology and ecology’, *Boca Raton*, 177, p. 222.

Welch, S. R. *et al.* (2017) ‘Identification of 2’-deoxy-2’-fluorocytidine as a

potent inhibitor of Crimean-Congo hemorrhagic fever virus replication using a recombinant fluorescent reporter virus.', *Antiviral research*. Netherlands, 147, pp. 91–99. doi: 10.1016/j.antiviral.2017.10.008.

Welch, S. R. *et al.* (2019) 'Fluorescent Crimean-Congo hemorrhagic fever virus illuminates tissue tropism patterns and identifies early mononuclear phagocytic cell targets in Ifnar-/- mice', *PLOS Pathogens*. Public Library of Science, 15(12), p. e1008183. Available at: https://doi.org/10.1371/journal.ppat.1008183.

Wilson, M. L. *et al.* (1990) 'Distribution of Crimean-Congo hemorrhagic fever viral antibody in Senegal: environmental and vectorial correlates.', *The American journal of tropical medicine and hygiene*. United States, 43(5), pp. 557–566. doi: 10.4269/ajtmh.1990.43.557.

Witkowski, J. T. *et al.* (1972) 'Design, synthesis, and broad spectrum antiviral activity of 1- -D-ribofuranosyl-1,2,4-triazole-3-carboxamide and related nucleosides.', *Journal of medicinal chemistry*. United States, 15(11), pp. 1150–1154. doi: 10.1021/jm00281a014.

Wolff, S., Becker, S. and Groseth, A. (2013) 'Cleavage of the Junin Virus Nucleoprotein Serves a Decoy Function To Inhibit the Induction of Apoptosis during Infection', *Journal of Virology*, 87(1), pp. 224–233. doi: 10.1128/JVI.01929-12.

Xiao, X. *et al.* (2011) 'Identification of a putative Crimean-Congo hemorrhagic fever virus entry factor.', *Biochemical and biophysical research communications*, 411(2), pp. 253–258. doi: 10.1016/j.bbrc.2011.06.109.

Yang, J.-S. *et al.* (2011) 'COPI acts in both vesicular and tubular transport.', *Nature cell biology*, 13(8), pp. 996–1003. doi: 10.1038/ncb2273.

Zeller, H. G., Cornet, J. P. and Camicas, J. L. (1994) 'Crimean-Congo haemorrhagic fever virus infection in birds: field investigations in Senegal.', *Research in virology*. France, 145(2), pp. 105–109. doi: 10.1016/s0923-2516(07)80012-4.

Zhirnov, O. P. *et al.* (1999) 'Caspase-dependent N-terminal cleavage of influenza virus nucleocapsid protein in infected cells.', *Journal of virology*, 73(12), pp. 10158–10163.

Zhou, Q. *et al.* (1998) 'Interaction of the baculovirus anti-apoptotic protein p35 with caspases. Specificity, kinetics, and characterization of the caspase/p35

complex', *Biochemistry*, 37(30), pp. 10757–10765. doi: 10.1021/bi980893w.

Zivcec, M. *et al.* (2015) 'Assessment of Inhibitors of Pathogenic Crimean-Congo Hemorrhagic Fever Virus Strains Using Virus-Like Particles.', *PLoS neglected tropical diseases*, 9(12), p. e0004259. doi: 10.1371/journal.pntd.0004259.

Appendix

1.1. Nucleotide sequences

1.1.1. Hazara S segment

TCTCAAAGACAAACATGCCGCAGACGCCCCACGTTTTCATCCTTTTGAGAG
CAAAACCGGTCGCCTCACAACATCAGCGAGAATGGAGAACAAGATTGTTG
CCAGTACTAAGGAAGAGTTCAACACCTGGTACAAGCAGTTTGCTGAGAAAC
ACAAACTGAACAACAAGTACACGGAATCTGCATCCTTTTGTGCCGAAATTC
CTCAGCTCGACACCTACAAGTACAAGATGGAGCTAGCCAGCACGGACAAT
GAGAGGGACGCCATCTACAGCTCAGCACTCATTGAGGCCACCCGATTCTG
TGCTCCTATCATGGAGTGTGCATGGGCTTCCTGCACTGGGACAGTTAAAA
GAGGTCTGGAATGGTTCGACAAAAACAAAGACTCTGACACCGTGAAGGTC
TGGGATGCCAACTACCAAAAGCTAAGAACAGAGACACCGCCTGCTGAGGC
ATTACTGGCATACCAGAAAGCCGCCCTCAACTGGAGAAAGGATGTCGGCT
TCTCCATTGGGGAGTACACCAGTATACTAAAAAAGGCAGTGGCGGCAGAG
TATAAGGTGCCAGGGACAGTGATCAATAACATCAAAGAAATGCTCAGCGAT
ATGATTCGACGCAGGAACAGGATCATCAACGGTGGCTCCGATGATGCGCC
AAAGAGAGGGCCTGTTGGGCGTGAGCATCTTGACTGGTGCAGAGAGTTTG
CCTCAGGCAAATTTTTGAATGCTTTCAACCCTCCTTGGGGAGAGATTAACA
AGGCAGGTAAATCTGGGTACCCTCTCTTGGCCACGGGATTAGCAAAGCTG
GTTGAGCTAGAGGGGAAAGACGTGATGGACAAGGCCAAGGCAAGCATTG
CACAACTTGAAGGATGGGTCAAAGAGAACAAAGACCAAGTCGACCAAGAC
AAGGCAGAAGACCTCCTAAAAGGGGTGAGAGAAAGCTACAAGACTGCCCT
GGCTCTGGCAAAGCTGTCGAACGCCTTCCGTGCCCAGGGAGCTCAAATAG
ACACTGTGTTCAGCAGCTACTACTGGCCTTGGAAGGCAGGAGTGACACCA
GTCACCTTCCCATCTGTGTCCCAATTCCTTTTCGAGCTGGGTAAGAACCCC
AAGGGACAGAAAAAAATGCAAAAGGCACTCATCAACACCCCTCTCAAATGG
GGAAAAAGGCTCATTGAACTGTTTGCTGATAACGACTTCACAGAGAACAGA
ATCTACATGCACCCCTGTGTGCTAACATCTGGAAGAATGTCCGAGCTTGGA
ATATCGTTTGGAGCCGTGCCCGTCACCAGTCCCGATGATGCCGCCCAAGG
GTCCGGGCACACAAAGGCAGTCCTCAACTATAAGACCAAGACCGAGGTCG
GCAACCCGTGTGCCTGCATCATTTCAAGCCTCTTTGAGATCCAAAAAGCAG
GTTATGACATTGAGAGCATGGACATAGTCGCCTCTGAACATCTTCTCCATC

AATCTCTTGTGGGGAAGCGCTCGCCCTTCCAAAACGCCTACCTCATTAAAG
GCAATGCCACCAACATCAACATCATCTAGGATTGCGGCTAGGTTCACTGCC
AACACATCTCTACTGCTGTTTTTCTTCTGCATTGCTTTGCTTCTATTCTTACT
TTCATTCCTATTCTATTCTGTTTTGTAAAATTTGGGGCTGTGCGGCAACGAT
ATCTTTGAGA

1.1.2. Hazara M segment

TCTCAAAGACAGACTTGCGGCACACACAAAAGGAAACTCCAGTGGTCTCG
CATTGACTTGTCTGAAGATGGAAGGGTCCTATTGGTGGCTATCCCTATTGG
CCCTATTGGCCTGGGGAGCAAACGGAGAGAGCACCAGCCCGGCTGAGAC
CTCTCCAGCCCCAACCACTCCCAACCCCCCGGTGGTAAACCCGAGTCTTA
GAAGAAAGATCGTGAACCAGAGAATCCTCTCAGCCATGGGCATGGACTCT
GATCCCAGCAACGAAGCACTCAATGGTGTTTGCCAGAGCATACACTCCAAT
GGCTGCAATGCTAATGAACTCAAACTGAGATTAGCAGACTTTTTCATTGAC
ACAAACAGCTCTCAGTGTTATGATGAGATACTAGTGAAAAAACCCTGCAGC
AGCCTAACCCCTGCTCACAACTCACATTGGGTCCCGAGGGGTCTTGATAA
GAGTGAGGTCGACAAGATCTTTGACACAAAACTTAAACTATTCTTCTCACA
GTCAAGAAAGGTGACCTGCCTCTCTGCCAGTGCACTCAATCCTTCACAGTT
TGTGAAACACTTTCAGGTTAAAATACAAGAAACGAGTGGCCCTGCCAAACA
AAGCCTCAGATCTCTACACTGTGTGAATCTAGTCTGGTCACACAGCCACAA
AGGAGAGAAGGAGGTTGTTCATGTCCTGCAGTCAGCAGTGCCAGTTAAGT
TGAAAAACTGCTTGGCAATGCTGAACTTTAGGCAGTGCTACTACAACCAGC
AATCTGAAGGACCGGTTGTGGTACCAAGCTACCAGCATAACGGGGAAAAG
TGGGTGACAGGAGCCTACACCATGACCGTTGAGGTTGACAAACATGCTGA
TGGGCCGTGTGAAATCTCGACAACATGCATCACAGAAGGCAGCGAAATAA
AACCCGGAGTCCACAGCCTCAGAGGATTTAAAACCACACTAGTCATACATG
GGAAAAGAAACACCGGGAGGAGGCTTCTCTCAAGCAGCAATGCACGCCA
GGAATGCAGCTCGGGCACATTTTTGGGAGAGGGGGGCAGTGCTCAAGTG
GTTGGCCCAAAAAATGACGGACCCGGTGACCATATCACCTTTTGCAATGGT
TCTGTGGTCACCAAGATCAGGCTTGGACAAGAACACGGGTGCTACACTGT
CAGAAGGATTAAAACATACAGAAACTGCCGGCCAGAAGAAGGATCCTCCG
CCTGCGAAGTGGACGATGAGCTCAAGCCCTGTGGTGCGCAGAAATGCATG
AATGTTCATCTAAGTGTCAAAGGCCTAGTCAAAACGAGCAGGGGGAGCAA

CGTTCAAGTGCACTCCTGTGACAAAGACTGTCTAATTCAGATACCAGAGGG
CTTTGGTGACATTCAGATCGACTGTCCAGGAGGGACACAGCACTACCTAG
AGTCTAATGTGCTTGATGTAGACTGTCCCATGTACAACAGGCTCGGTGGTC
TGATGCTATATTTCTGTAGGATGTCTCACAGACCCAGGACATGCCTCGCCC
TCTTCATCTGGTTAGGGGCAGGCTATGGCATCACATGCATAGCAGGTTACA
TGGTGTACTATGCCATCCTTGCTTTAAGCATGTTGACCAGGTGTCTCAAGC
GCAAATACATGGTAAAGGGAGACTTCTGCCTAAAGTGTGAACAGAAATGTG
TCACCAGTCTAGATCAGACCCTACACGATGAAAGTTGCTCCTATAACATCT
GCCCTTATTGTGGTAACAGGCTTCCAGAAGAAGGTCTAAGGAGGCATGTC
CCAAGCTGTCCCAAAAGGAAGCAAAGATTAGAAGAAATTGACTTATATCTA
GACTATCTGTTAGTCCCATGCCCCCTACACTTTGCTCTCAGCACAGCAGTG
AAGCTGGGCACTTTGCTTAAGAGGTTGAGCTGGGTTACTGTCTTCCTCTGC
TTGTTCCTTACTGCTATCGCGCCTGTGCAAGGACAGGTCACCACCAGCCC
AGTACTCCCGAGCAATCAGAGCACCGAGTGCACGCTATTGCCTCCTCCTG
TGTTTCTCATCTTTTCCGCAGTTCTCATGTCCAAGACTCTGAAGAGGATGG
GTCCTGTAAACAAGGTGGGTGCGGCCGGACACAGTGCCAGAAGGACAAA
CAGCCCAAAAAATTTGTACAAGAGCAAACAGATTGCCAACACCAAAAGTGG
CCCCAGAGAGCCACGCCGGCGCGTAGTGGTCAAAGCATTACTCATCCTAA
CAGCATCATCAGCCCTCCAGTCCATTCATCTGGCCCAAGCATTTGACTCTG
GCTCTCTGCCAGAAGGAGCATGGGAGGAAGAGATGCAGCTAGTTCAAGGA
TGCAACCAAGAGTGCTCGCTAGAAGAAGATGAATGCTCCTGTCCTGATGG
CCAGTCAATGACACGGAAGTTACTTTTCTTTAAGGGTCTGAATTCGGCAGC
ATCCAAAATGGCCTCTTCTCATAGATTGCTCACTAGTGTGTCTATCGATACA
CCATGGGGGGCAATCAAAGTGGAATCAACTTACAAGCCGAGACTCGCCTC
CTCCAACATACAGCTTGCCTGGAATTCTATCGAAGAGCAGGGGGACAAAG
TGATTCTTTCAGGGAAGTCTACTTCTATCATCAAGCTAGAAGAAAAAACGG
GCATGCAGTGGTCTCTAGGCTCTGAGTCAGCAGCAGAGGAGAAGAGGCT
CCTGGTGTCCATACTGGATTACACACAAGTATATTCGAGCACGTTTCAGTA
CATCACCGGGGATAGAACCGTCTCTGAGTGGCCAAAAGCCACCTGTACTG
GAGACTGCCCTGACCGGTGCGGGTGCAGCACATCCTCCTGCCTTTACAAG
TCCTGGCCACACAGCCGTAACTGGAGATGCAACCCAACATGGTGCTGGGG
GGTAGGAACAGGTTGTACATGCTGTGGGGTTGACATTCTTAGACCCTTTAA
CAAATACTTTGTCACCAAATGGACCACTGAGTATGTAAGAACAGATGTGCT
GGTCTGTGTTGAGCTAACAGACCAAGAGAGGCACTGTGATGTTGTGGAGG

CAGGCAGTCAATTTGTCATTGGGCCTGTGAGGGTGGTAGTCTCTGACCCC
CAGAATGTTCAGACTAAGCTTCCATCAGAGATCCTGACCATCCAAAAACTT
GAAGGAAATCAGGTTGTGGATATAATGCATGCCACCTCCATTGTCTCCGCC
AAGAACGCATGCAAACTGCAGAGCTGCACGCACGGCAGCCCAGGTGACA
TGCAGATCCTCCATACAGACAACCTCATACAGCACAGTCATGATGGTGGTT
TAAATCTTGCCGACCTTAACCCACTTGTGAACTCAACCTGGATGTCCTGGG
AAGGATGTGATCTAGATTACTACTGCACTACGGGTAGTTGGCCGAGCTGC
ACCTACACTGGCATAAACTCAGAAAATACAGAGAGCTTTGACAACCTCTTA
AACACTGAATCCAACCTATGTGAGCGGTTCCACTTTCACTCAAAAAGAATA
TCTGCATCAGGCTCAACTCTTCAGATGGACCTTAAGGGAAGGCCCAACAG
TGGAGGTGGTGAGCTATCTGTCTTAGTAGATGTGAAGGGTCTTGAGTTGC
ATTCAAAGAAAATATCTCTGAAGGGACTATCATTCAAAACTCTGTCTTGCTC
TGGCTGCTATGCGTGCAGCTCTGGCTTATCCTGCACTGTGGAGGTGAGGA
TTGAGAGACCGGATGAGTTCACCGTCCACCTTAGAAGTGTCAGTCCAGAT
ATAGCGGTGGCGGAAGGCAGCATTATCGCAAGAAGGATGACAGGTGGTC
CACTCAGCAGGCTCAGAGCCTTTGCAGTTAGAAAGGTTAAGAAGATCTGCT
TTGAGATTGTTGAAAAGTCATACTGCAAAGACTGCAAGAATGAAGACACCA
CAAAGTGCATTGAGGTGGAGCTCCAACCACCAAAGGACATTCTCCTCGAG
CATAAGGGTACAATAATAAAGCGACAAAATGAAACTTGTGTTTCAGGACTT
CAGTGCTGGACAGAATCTGCTTCTAGCTTTGTGTCCGGGGTTGGTAGTTTT
TTTAGAAACTACCTAGGCAGCATCACTCTCGGTATAGTTTTAACACTCCTAC
CAGTGGCTGTGGTCCTCCTCTTCTTCTGTTACGGAGACAAGTTATTTAAGC
TTTGCAGTTGCTTTCGATGTTGTAGAGGACTATCAAGGGGCAAAGTCAGGA
AGGAACTCGACGAAGATGAGCTGAGGAACAAGTTAAAAAAATTCAGCAAA
GAAGGTGAGCTGTTTGGGAAAGAAAAAAAAGACGCCAGAACCATAGCTCT
GCTGCTATCTGGCAAAGGGAAGAACTACAAGGAACTTGTTTGACTGCCCC
AAACCCCGCCAGACTGCAACCCCACAGGGCAAGGCCACAGTCAACTCTAT
AGCTTCGCCCGCCGCTCTTCCCCTGCCTCCATGCTGACCATAGCCAGCTA
GAGCATGTCTACCGATATGTTAACCCTTTGTGTAACTCTTTCACTGTGAGTT
TACTAATCAATTGTCATGTATGTTTTAGTGCTAGTTATTATGTGTTTATAAGT
TTAGGGTGTGCCGCCACGATATCTTTGAGA

1.1.3. Hazara L segment

TCTCAAAGACATCATCCCCCTTATCCCCAAGTTAACATGGACTTTCTTGAAG
GAATAACGTGGGATAGTGTCTCTGATATCCAATCGGTTAGCAATCCGAGCT
TCACCATTACTGATTACTTTGAAGTGGTTAGACAGCCTGCGGATGGCAACT
GTTTTTACCACAGCCTCGCAGAGCTCTACATTCCCAACAAATCTGACCACG
CATACAGGTTAGTCAAAAATGAGCTGAGGGAAGCTGCCGAGAAGTATTTC
CCGACAGAGCCTGAAGCTGCTGCCACTGGCATGAGGCTCGATGAGTATTT
GGACACTGCCCTACGGGACAATGAGTGGGGTGGAAGTCTTGAGGCAGCA
ATGCTCTCACGCCATCTGGGACTCACTGTGGTTATTTGGCTGGTAGATGG
CAGCAATAGAGTCGTTGGTGCTACAAGGTTTGGTAAGGGTTCACTCAAAAC
GGCATTGCACCTGTTACACAGTGGGTTGACACACTTTGATGCTCTGCGGC
TACTAGCCACGGAAGAAGACCCCCAGCAGGAGACAATGACACTGGTAGAG
AAGATGGAGCTTGTGGAGAGATTTACATTGACTGAGGGAGAGGAATGTTT
GCAGGAGGAGGAACTGCTCCTAGACAGCGAAACAGATACTGCTGCTCCG
GAGGAGCCCTCCACCAGCGAGCCTAGGCTCAGATCGCAGGCCATTTTACT
GCATAGACTAGTTAGACATGGTGAAAATATCCCTGTCAGAGTTGGTAGGGT
CCTGGACTGTCTCTTCAACTGCAAGCTATGTGTGGAGGTTAGTCAAGAACT
GCTCATTTTGAAGCCAGACCGGAAGGAGGCCACCGCAGAGTCAATGAGTC
TCAGACAGCTAGGACACAAGATGCTAACAAGAGACAAGCAGCTAAAGAGC
GAGTTTTCAAGGCATAAGCTATACCTCACCAAGGACCTTCTTGACCATCTT
GATGTTGGAGGCCTTCTCCGATCAGCTTTCCCAGGTAAAGGCCTTGAAAG
AAACTTGTCTCTCCTTCACTCAGAGTTAGTGCTGGACATTTGCACCGTTGT
TCTGGGGACACTTCTCTCTACTTTCTTGTATGGATCAAACAACAAAAACAAA
CGAAGATTCATCACCAACTGCCTGCTTGGCACCAGCCTCTCAGGTAAGCG
AGTTTTTAAATCTCTCAGCAAGTTGACATCCAACTTGTTGTACAGGAGCCC
ACGCAGGGCTGTTTCTATCATTTGCAATGATCTCTACGGTAGGCTCATCCA
AAAGCTGAGCAACTGTTTCGCACTGATGAATCCCATTAGCCTTCTCGCACT
TAGAAACCTTGACTGCGATAACATGGAGCTCAAAGACTACATCGACATGGT
CGTTGAGATGTCAAAGCTAGACAATTCTGATGTTGATTTTACACACAGAGA
AATTGCTGATATAAATCAGCTGACGGATAGACTTCAGATATTGGCTGGTAG
GAAAAATCCTGATGTACTACTTGAATGGTACAAAAAGGAGGAGTTACATAA
GAGGGCCCTAAGGGATATTGCGGGCGCCCAGGAGCACCTGATCAGCGAC
TTCTTTAGGCGGAAAGACATCATGAAGTTCATCAGCACGTCTGGAAAGGCC
TCAAGTGCTGGAAGCCTAGGCAATGTACTATCATATGCCCATAACCTTTAT
TTAAGCAAAGAAAGCCTCCGGATGTCCAGTGAGGATGTAACCCAGCTGTT

AATTGAGATTAAGAGGCTATACAAGCTCCAGGGCGACCAGAGTGTTGAGC
CAATTGCTCTCATCTGCGATCGGTTAGAAGACAGTTTCAGGCGTCTCTGCA
GGGAACTACCTGCAGATTGTGCACAAGAGTGTGTGACTCTCTTTGAGGAC
ATTAGAAACTCAACCAGCCACAGTACCGCTTGGAAGCACGCGCTGAGGCT
GAAGGGCACTGCCTACGAAGGTATGTTTTCCCGTCAATACAACTGGAGGT
ACACCCCAGAAGACATTAAGCCAAGCCTGACTATGCTGATCCAGACTCTGT
TTCCCGAGAAGTTTGAAATGTTTTTGGACAGGACACAGCTGCACCCTGAGT
TCAGGGATCTTACACCAGACTACGCACTAACACAGAAAGTGTACTTCAAAA
GAAATCAGATAGTGGAGGTTGTGAGCCACCAAATTTCAATTGGGGAAACAC
TTGATGAGTCTGTTGATGCTATACCCCTTGAAGAAAAGAAAATGTTCCCTCT
GCCAGAGACACCTGTTGGCGAAGTGTACTCCATCCAAAGCATCCTGAAGA
ACTTCCAGGACAAGATAGATAGATGCAGGGATGACTCAAAGGCCAGTGCT
GAAGAGGGAAAAGCATCCGACTCGGCTGGTGAACAGCTTGACTACAATAG
CCGAATCATTATAGATAAAAATAACATTGAAATATCTGATGAAGAGGAGCTC
ATCCGGCGCCAACTGTTGCTTGTGGAGGTTGGATATCAGACAGATGTTGA
CAGCAAAATCACAACTGACTACAAGAAATGGAAAGACATTCTCAGACTTCT
GGAGATGCTAGACATCAAGTGTTCCTTTGTGGCTTGTGCAGACTGTTCCTC
AACACCTTCAGACAGCTGGTGGATTTCTGAGGATAAGGTAAGGTTGTTGAA
GAATTCCATAAGCCATCTCTTCAGTTGTCTGACCAAGAACTCACCATCTGA
TGTTACTGACATTGTGGTGGGCTCTATCAGCACACAGAAGGTAAGAAGTTA
TCTGAAATCCGGTTCAGCAACTAAAACACCAATATCCAGTAAAGATGTGCA
GGAAACATGGGCCAGACAGCAAGAGTACATTATCAACCGTCCGACTGGGA
TCTCTATTCCAAAAAGGCTGGCCGATGCAATGAAACAAGGCTTTGTCGATG
GTGTAGTCATGTCAGCAGACTCAAGCAAGGAGTGTGTGGCTAACATAAAG
AAGAATGCAGAGAGGCTGACGGATGAGTATGAGAGAACAAAGTTCAAGCA
CGAACTGAACTACAGCAGAGTCACTAGTGAGAAACTACTCTTAGGATGGTT
GGGCGAAGATCTACAAGGTATAAGATGTGATAATTGTCTACATACAATCAA
AGAGACTGTTGAACAGATGAATGAAAACTGTGACAGGCTTGAGTATCTTGC
CAGCAGCTGCCTCTTGAGCAATCACTGCAGTGGTTGCCATCCCAGCGGGA
TAGCCCTAAATAACCAAACAAATGTGCAGAAAAGGCTACCAGAGATGGGA
CTCTTGAAGCATTCAGAGAACAAAGGCTTTGAAGATACTAATGAGGCCATT
ACTGACTTAGACAAGCTCGTAAGGCTCACACTGCCAGGTAAAACTGAGAAA
GAGAGGAGAGTCAAAAGGAATGTGGAACTGCTGATAAGACAAATGATGCA
GCAATCTGGCATTGAATGCATAAAGCTGCCATCAGGCCAGATAGTCACACA

CAGACTAACAAGAAAAAACAAGCAGGCACCTAGCACTCAAGAGTGTGAAA
GGATTGAGGAGAGAGTAACAAAGTTAAAAAAGGAACTATCTGAAAAGAAGC
TCTCTCAATACTCAAAGCACATCAATTCCACTATCGCACACTCACTGGAAC
GCTTGGACAGACAAACAGGCTCCAGGTGTTCTGTCCCGAAGGAGTGGCTG
GAGAAGCTATTGAGGGATCTGAAAGTGCCTACCAAGGATGAGGACATTCT
GATAGGGATTCAAAGGTCCATGCAAGAGAAAGTTGGGTTCACAGTCAACA
ATGACAAACTACTGATAAGGGATGAAGACGACCTAGTACGGTTCATAGAGA
GTAGAAGCAAATCATTACTGGAAACTAATGAGGAAGGTATTTTCCAATCTG
ACTGCGTACTCTTCAAGGAGGTGGTAGCAGAGGCTATGTGCAGGTACACA
TCCACGCCATACCAAGAAATACCCGAGACCTTAGTGAAATTGATTAACTTG
CTGTGCAAGTTTGTCTGGTTCCAGGAATGCATCCTGTACGGCAAGGTGTG
TGAAACTTTCCTCCGGTGTTGCACAGAGTTTAGCAGGTCTGGTATCAAATT
AGTCAAAATAAGACACTGTGACGCTAACCTGGCTATAAAGCTACCTTCAAA
CAAAAAAGAGAATATGCTGTGCTGCATATACAGCAAGGACATGGAGCTTAT
CAAGGGACCTTTCTTTCTCAATCGAAGGCAGGCTATTTTGGGAGCAGCATA
TCCTTACATTCTAATTACAACCTATGTGCAGGTGCTGCAACAGCATCGTTG
CCTAGAGGTACTAAACAATCATGGGCCTCGAATACTCGAGAATATAAGTAG
GTGCACAAAGACACTCCTGGAAACCGCTACTAAGGAGCTGTCATTCACACT
TAGAGGGCTTTTTGAGAAAGCCTACGAAACTAGGACCAAACAGTGCCAGC
TCGGAGGCAACTTCCTCAGCAGAAGCAGCAGAGACCACTTTGTCTCAGTG
ATCTCAGGGCTGAATCTTGTTTACGGACTGTTGATCAGAGACAACCTTTTA
GCAAATTCACAACAACAAAACAAGCAACTGCAGATGCTAAGGTTTGGCATG
CTCTGTGGCCTAAGCCGTCTTTCCTGTCCTAAAGAGCTAGGAAAAAAGTTT
TCAGCAAGCTGTCGCAGGATGGAAGACAATGTTATGAGGCTTTACCTACA
GTCCACTGTATACTCTGCCAACAGAGACTGTGAGATGAACGTCTCCAACTG
GAAGCTAAAAGACCTTTGTCCAGAGGTCACCATACCCTGCTTTTCAGTCTA
TGGGCTCTTTGTAAATAGTGATAGACAGTTGATTTATGACATTTATAATGTG
CACATCTACAACAAGGAGATGGACAACTTTGATGAAGGGTGTATAAATGTC
CTAGAGGAAACTGCAGAACGCCACATGATGTGGGAGCTCAATCTGCTTGA
AACCCTAAACCCAAACACTAAAGATGACCGTACAGCAAGGCTGCTTCTAGG
TTGTCCCAATATTAGGAAGTGCACTGGCAAGGATGGCAGAAAGATAAGAC
CACTATTACATGACTCAGAGAGTGCAGACAGCAGCAGTGAAAGCTCAACC
ATCTCGGGGAGGCGGTCCTACGGCTCTAGCAAAGGCAAGATACAGAGCAT
GTTTGGGAGATACAACTCAAACAAGAAACCTTTTGAACTCAAGCCGGGTCT

TGAAGTGAGCAACGATCCGCTTCATGACTTTCAACAGGTCGTGACTGGAG
GTAGTGCGTATTCAGAGTATTCACCGAACCAGGATAGTTTGCTGAAAGACT
ATATCCAAATAATAAGGAGCAACCCTGGCTACACAATGGGGTCATTTGAGC
TTATACAGGCAGTGACTGAATTTGCCAGGACCAAGTTTCCTGCAGAAGGCA
TTGAGAAGGCAAGACGAGACCCAAAGAACTGGGTAAGCATCTCTGAGGTT
ACTGAAACAACCAGTATAGTGGCCACTCCGAAGGTCCACATGATGTTGAAA
GACTGCTACAAGGTTCTCTTAGGTACTGAAAATAAGAAAATTGTAAAGATG
CTAAGAGGGAAGCTAAAGAAGCTTGGTGCCATAAGCACGGATGTGGAGAT
AGGGAAGAAGGACTGCCTGGACCTCCTGAACACGGTTGAAGGCCTATCCG
AGGAGCAGAAGAAGAACATAGTCAATGGCATCTTCGAGCCTTCCAAATTAT
CCTTTTACCACTGGAGAGAACTCATCAAGAAAGATCTGTATGAGGTACTTC
TGACGGATGACGGCAATTACATCTTCTGCTGGCTCAAAACTCTATCCAGTG
CAATAAAGGGAAGGCTCAAGCGGGATCTAAGATTCATGAATCAAAACAGC
CAGACCAGCAGTCAGACAGACCTCTTTAGCGAAGAAGAGTATGAGGAATT
ACTTAAAATGAAGCAGGTTATCTGCTCACAATCAGATGCTGAGGATGAACT
GAATGTTGATGTTCTCCTTGACGCATGGGTGAAGTGCGCCCACAAACCCA
GAGATGCCTCCTCCATAATAAACGAGGGGATGAGTAGAGTCCTACCAATCT
CAGAGATTCTATTTGAGCTAAGGATGCAGCACCTAGAGCTTACAAAGCTGA
AGAAAGACAATCCTTCTGTGAGCTTTACCAAGGAAGAAGTCACCGTTAAAA
GGATGGAAAAGCAGTTCCTGGCTAAACACAATTTGGACATAATGCACCTTA
CGAACCTCATATTCTACTGTGCCCTTGCTGCTCCCTGGTGTGTTCATTACA
AAGCTCTTGAAGCCTACCTCGTCAGACACCCAGAAATACTGGAGTTCAGTG
GAAGTGCGCCCACCGAAAGCAAGGTGCTAGATCTTTCAGTGGCTGCATTG
ATTATAAAGATGACAGAAAACTATAGGGATGACACATCAGACGGAGTAGAA
GTGAAGGTCAGGTTCCTTGTCAGATACATCATAACTCTATTTACTGCAAATG
GAGAGCCATTTTCCCTCAGTCTGAGCGACGGCGGCCTAAACGAGGACTTG
CAGAAGACCACAGACGAAAAACTGCTACATCAGACCAAAGTTGTCTTTGCC
AAAATTGGGTTGTCCGGGAAAAACTATGACTTCATCTGGACAGTACAGATG
ATTGCAAACAGCAATTTCAACGTGTGTAAAAGGCTCACAGGGAGGTCGAC
AGGCGAGAGACTACCTAGAAGCGTCAGAAGCAAGGTGATCTACGAGATGG
TCAAGCTAGTGGGAGAGACGGGGATGGCAATACTTCAGCAACTTGCCTTC
GCCCAGGCTTTAAACTACGACCACAGATTTTATGCTGTCTTGGCCCCAAAA
GCCCAGCTAGGGGGAAGTAGAGATCTTTTGGTACAAGAAACTGGCACCAA
GGTTATACATGCCACCACAGAAATGTTTAGTCGCAATCTCCTCAAGACAAC

CAAAGACGACGGGCTAACCAACCCTCACTTGAAGGAAACCATTCTCAATG
CAGGGCTTGAAGCTCTTCAAACTATGAGATTGGTAGACGGAAAGCCTGCT
GCTGAGGGCAGCAGCCTAGTAACGTTCTACAAGGTGGTCTGCATATCCGG
CGACAATACGAAATGGGGTCCAATTCACTGTTGCTCCTTTTTCTCTGGCAT
GATGCAGCAGTTGCTTAAAGATGTGCCTGACTGGTGTTCATTCTACAAGTT
AACTTTCGTAAAGAACTTGTGCAGACAGGTAGAAATTCCAACAGCAAGCAC
AAAAAAAATCCTCAATGTTCTAAGGTTTTATCTGAGCGACAAGGGAGGAGT
TGAGAGACTAAGTGAAGAGGAAATTAGGAACAGGCTCTGTGAAACCCTAG
ATCTCTGGGGCGGAAATGACATAGTGAAGTTCCTTATTACTACTTACCTGA
GCAAGGGCATTATGGCGATGAATAGTTACAATCACATGGGGCAAGGCATT
CATCATGCAACCTCTTCAGTGCTCACATCTGTCATGGCAGAACTGTTCGAA
GAACTAACTGTTGATTACTATAAGAAGCATTATCCAAACCTTTCTGTATCGG
TGACACATGCTGGTAGCTCAGATGACTATGCAAAGTGCATTGTTGTCACTG
GGCTGCTGTCAAAAGATCTGTTTGACAAGTACTCTGAAACATTCTGGATGC
ATACCTGCAGACTTAAAAATTTCACAGCTGCAGTGCAGAGATGCTGTCAGA
TGAAAGATAGTGCCAAAACACTGGTTGGTGACTGCTTTTTGGAGTTCTATA
GTGAGTTCATGATGGGTTACAGAGTCACCCCTGCAGTGATTAAGTTTATTT
TTACTGGGCTTATAAACAGCTCAGTGACTTCGCCGCAGAGTCTCTCTCAGG
CGTGCCATGTGTCATCGCAACAGGCGATGTATAACAGTGTTCCCCTGCTTA
CAAATGCAACCTTCACTCTTTTGAGGCAACAGGTTTTCTTCTCACACGTTGA
AGACTTCATTAGGAGGTATGGACTTTTGACGCTTGGTTCACTTTCACCGTT
TGGAAGGCTCTTTGTACCAACTTACTCAGGCCTTGTTAGTTCTGCTGTGGC
TCTAGAAGACAGTGAAACCATAGCAAGATCCTCCTCTACACTGGTGGAGAA
CAGCATCTTCCTTGAAACCAGCAGCTTATCAATAATTGATCAAATCTCTTCT
AGCTCTAGCTCTGAAGGCGAAGGATTTAGCACTGCAAGCAGCACAACTGT
TGAGTCAACACATTCTGCAAGTTCATCTTCGAGCTTCACATTCGAGCTGAA
CAGACCTCTCTCGGAGACTGAGTTACAGTTCCTTAAAACCTTGAGAGACAA
CAGCAGGCAGACCTCTAGTGAACACATTCAGGATAAGATCACAGAACTTTA
TTCTACTTCAAAGGAGGGACCTCTTGACAAGTACTACTTACTCTACAGCAG
TAAAATAGTTGACTCGTGCTCGTGGCTCAAAAAGGGTCGTGAGAAAGGCC
CTATTGAATGTGCAAAAAGGCTTCAATGCATACTAAATGTGCTGATTGCCG
GCTATTATCGCTCTTTTGGCAGTGATGGCACTGAGAAACAAGTCAAGGCTT
GCCTTAACAGGGACGACAACAGAGTCATAGAAGATCCCATGATTCAGCTC
ATTCCAGAGAAGTTAAGGAGAGAGTTAGAGAGGTTAGGGGTCTCCAGGAT

GGAGGTGGATGAGTTGATGCCGGCCACATCTCCTGATGATACCCTATGTC
AATTAGTTGCAAAAAAGCTGATCAGCCTCAACGTGTCCACAGAAGAGTACT
CGGCAGAGGTTTCCAGACTTAAGCAGACCCTCACAGCTAGGAACGTGTTA
CACGGGCTAGCTGGGGGTATAAAAGAACTGTCACTACCCATTTATACTATC
TTCATGAAGTCTTACTTCTTCAAAGACAATGTGTTCCTTGACCTAGATGACC
GCTGGAGCACAAAGCATAGCACAAACTATAGGGACAGCACAGGCAGATTG
TTGACAGGCAGAGTAATCACTAAGTACTCACACTGGTTGGACAACTTTCTA
AATTGTAAAGTCAGTGTCGATAGAGTCCAAGAGGTCAGAGACTGTTCTCTC
TTCAACCCTGACCTTCGGTGTGTCAATTTGCTGATCGGTGAGAACAAGGTT
AGAGAGCTCTCCATCGTAGTAAGCCACTTGAAAGTGTTTGCGAGAGAGTTT
GACAACCTGAATCTGCAATTCTCTGATTTGAACCGTCAAAAGCTAAAAATTG
TAGAGTCAAGGCCACCAGAGTCGGAGCTAGAGGCAAACAAAGTAGTAATA
GTCAAGTCCAAACTCTTCAGTGCAACTGAACATGTTAGACTTTCAAACAAC
CCTGCAGTGGTGATGGGTTACCTTCTAGAAGAGTCCTCAATTTCTGAGGTC
AAGCCTACCAAAGTGGACTACTCCAACCTGCTCAAGGACAGGTTCAAACT
GATGCAATTCTTCCCATCTGTCTTTTCACTCTTGAGGACACTCCAGGTTGA
ATCACGAGAAACAGAGAAGTTGGGTGACCCGGTTGACATGAATCTGGTCT
CTAGGTACTCTAACCATCTGACACTGCTGTGCAGGATGATTCAACAAGCAA
GGCCGTCATTGACTGTGTTCTACATGCTCAAAAGCACTCACCTCGCAACAG
AGCCAACTGTGTCTGAGCTGGTTAGTTTTGGAGTGAAAGAAGGCCGTTATC
TAAGGCTGTCTGACTCAGGGCTGGATGCCAGCACTTACTCGGTAAAGTAC
TGGAAAATTCTACACTGCATCTCAGCTATAGGTGAGCTCCCTCTAAGCCCC
AAGGATAAAACAACGCTGCTTATGAGCTTTTTAAATTGGAAAGTTGACCTTG
AATCCTGCGAGCAGGACTGCCCGCTTTATAAGAACGAACTCAGTGTTCTCA
GCGAGTTCTCAGGTCAGGTTATAATAAACACTTTAGCTAGCGAGTTGAGCT
CTGTCCGAAAGGATGGAGAGAGGGATAGTCTGACTGACCTCATTGACTAT
GTCAATTCACCAAGTGAATTGCTGAAGAAGAAGCCCTACTTAGGAACAACG
GCAAAATTCTCTTCCTGGGGTGATTCGAATAAGAGTGGGAAGTTCACATAT
AGCAGTAGGTCAGGTGAAGCAATTGGCATCTTCATAGGGGGCAAGCTGCA
CATACATATTTCAGAAGAGTCAACAGGGCTACTATGCGAGGTGGAGCGCT
GTGTGCTAAGCTGGCTGTCTAGACGTAGAACCGACATCATTACGAAGGAA
CAGCATGGTCAATTTCTGGCTTTTTTGCCTACACTGTCAGAGGTTGCACAG
AAGAACAGAGATGGAGGTGTGCAGGGTGTCTGTGTTGACCCTAGCAACCC
AAGACTACTTAGGTTCACTGTTGCCAAAAGGCAGAGCCCAGTGATAAAGGT

AAAGAAGCAGATTCTGACAGTGAAGAAGCAGATCACATATGATGCAGAAAG
TGAACCTAGACTACAGTGGGGACATGGCAGCTTAGCCATTGTGTATGATG
AATGTGAGACTCAGACAACGTACCATGAAAACATAATTAAGATCAAACAAC
TAATTGACAACACAGTGGACAAAGGGAAAATCCTACCGCAGTCCGTGTTCT
CGGACACGAGGATAATATTGGCAAAGATTAGGTTTAAGAGCGACTTGCTCC
TTAATTCTCTTTGCCTGCTACATGCCTTCTTGCGCCACACAGCTACTTACG
CAGTGATGGAGGTAGAGTCCAAGAGTCAGCTACTGGAAAAGTTTCTCAGA
TCAGGCGGTGTACAGTTTAGAAGCTGTTGCACTAGCATAAAAGACAAACTA
AAGACTAAGGAGTTAGACGGGTTAATCACTCAGACATTAGATGAAGAAGTT
GCAGTTTGCGATGAGCTAAACAAGGTCTTTTCTGAAGCACAGGTTCCCCTT
AGTAGCTGGTCAGAGGTGCAAACCTACATCGAGGAGGTTGGCTTCAACAA
CGTCCTGGTGAACGTAGACAAAAGTCCTGCCAAGAGTGAGTTGCTATGGA
GGTTCTCACTGGATAGCCATGTGAGCAATCTGGGCATGATTAGGGATCTC
AGGTCACTTGTTGGCTATGTCAGCACCGAAGCTGTTCCAAAGTTTCTGCTC
CCTTTCATATTCTTCGAAGGGCTACTTGTTAGCATCATTGGCAAGTGCAAG
ACTCTAAAGGAGTTGATAAATTCCACTGGCACCACAGATAGAGACATAGAC
ATCCTGCTCTGCCTTATTTTATTCTGCTTTCAGAATGACTCCTTTGCCAGAG
AGGGGCCTAGATGCAGTGCCTCAGCCCTTAAACAGCTGCAGTGTTGTAAT
GTAGTCAGAGTTAACAACCGAATCAACATAGAGTTAGTATCCGAAGGAAGC
AATGTAAGCTTGAAAGTCGTCATTGTCTTGATTGATGCGACTGAGAGCACC
CTGGAAAGATCGAAGCGAGTGAAGGTAGCCAAGAGGAACCTGAGCAGTTC
GCTGGGGTTAATGTTCTTTGATAAGTCTATTGATGTCACTGGCCTTAAGAG
GGTTGCATCCAGAGTCAAGGTGTCGAGTGACAAAGAAAAAGAGTTCTTGG
ACTTTGTTCTTCCATCACAGTCTGCTTCTGAGGTTGACTATCAAACCATATT
AGACTTGACCATAGACAAAGCAAAAAAGGGCAGGACTGTTCACGGGATTG
AAGATTTGCTCCTTACGCTCATGGGTAAAGCAGGCACACCAAAGTCAGAG
GATGAAGAAGACCTGATACAGGATAGTACAGAGGGACACCTCTGCTTAGA
AGACCTGCTGGAGGATACCAGTGTTTCAGATCCCAGGGCCGAGTCTGAAG
ATGAAGAACAGATCATGAAGCGTGGTTTTAAGTTCAACTGGGACAGTGACT
GACGGTGTGGAGTGTTGCTGATCTCCAAACACCATAAACTCACCTGTGTCT
CCAAACCCACAACTACCAAAAGTGAAAACTTTAGCAAATGAGGCATCTGAA
TCCCCCACTGTAAAATACTTTCTTATATTAGAAGATTAAAATTGGGGTGTGG
GGGGAACGATATCTTTGAGA

1.1.4. EGFP-P2A gene insert

GAATTCATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCA
TCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTC
CGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTC
ATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCA
CCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAG
CAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCG
CACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGA
AGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGA
CTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACA
ACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAG
GTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGC
CGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTG
CCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAA
CGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGG
ATCACTCTCGGCATGGACGAGCTGTACAAGTAAGGAAGCGGAGCTACTAA
CTTCAGCCTGCTGAAGCAGGCTGGAGACGTGGAAGAGAACCCTGGACCG
CGGCCGC

1.2. Primers

Primer	Orientation	Sequence
HindIII insertion	Forward	GGA TTA GCA AAG CTT GTT GAG CTA GAG GGG
HindIII insertion	Reverse	CGT GGC CAA GAG AGG GTA CCC AGA TTT AC
HindIII sequencer	Forward	TCG GCT TCT CCA TTG GGG AGT ACA C
HindIII sequencer	Reverse	GCA TCA TCG GGA CTG GTG ACG G
S segment EcoRI insertion	Forward	TCG TCC GAA TTC TCT CGC TGA TGT TGT GAG G
S segment NotI insertion	Forward	TGC TGC GGC CGC ATG GAG AAC AAG ATT GTT GC
S segment EcoRI deletion	Forward	ATG GTG AGC GGC GAG
S segment EcoRI deletion	Reverse	TCT CGC TGA TGT TGT GAG G

S segment NotI deletion	Forward	ATG GAG AAC AAG ATT GTT GC
S segment NotI deletion	Reverse	CGG TCC AGG GTT CTC TTC
S segment DQVE substitution	Forward	GTC AAA GAG AAC AAA GAC CAA GTC GAA CAA GAC AAG GCA
S segment DQVE substitution	Reverse	
S segment ENKE substitution	Forward	GTC AAA GAG AAC AAA GAA CAA GTC GAC CAA G
S segment ENKE substitution	Reverse	
S segment AQVA substitution	Forward	AGA CCA AGT CGC CCA AGA CAA GG
S segment AQVA substitution	Reverse	TTG TTC TCT TTG ACC CAT CC
S segment DVMA substitution	Forward	AGA CGT GAT GGC CAA GGC CAA GG
S segment DVMA substitution	Reverse	TTC CCC TCT AGC TCA ACC
S segment LFAA substitution	Forward	ACT GTT TGC TAA CGA CTT CAC
S segment LFAA substitution	Reverse	TCA ATG AGC CTT TTT CCC

1.3. Membrane trafficking screen targets

Gene Symbol	Full Gene Name	Sense siRNA Sequence	Antisense siRNA Sequence
AP2A1	adaptor-related protein complex 2, alpha 1 subunit	GGCGGAAAUUAAGAGAAUCtt	GAUUCUCUUAAUUUCCGCCtc
AP2A2	adaptor-related protein complex 2, alpha 2 subunit	CCACUUUAGCACAGAAGAAtt	UUCUUCUGUGCUAAAGUGGtg
AP1B1	adaptor-related protein complex 1, beta 1 subunit	GCUUAUCUGCUACAUCGGCtt	GCCGAUGUAGCAGAUAAGCtc
AP2B1	adaptor-related protein complex 2, beta 1 subunit	GCGGUAAAAGUCCUAAUGAtt	UCAUUAGGACUUUUACCGCtg
AP1G1	adaptor-related protein complex 1, gamma 1 subunit	CGGCGUGCAAUGGAAUUGAtt	UCAAUUCCAUUGCACGCCGtt
ARF1	ADP-ribosylation factor 1	GGAACUGGUACAUUCAGGCtt	GCCUGAAUGUACCAGUUCCtg
CACNA1A	calcium channel, voltage-dependent, P/Q type, alpha 1A subunit	GGCGGCAACAACAGAUUGAtt	UCAAUCUGUUGUUGCCGCCtc
CACNA1B	calcium channel, voltage-dependent, N type, alpha 1B subunit	GUGGCCUCCAUUCGAGUAUtt	AUACUCGAAUGGAGGCCACtc
CACNA1C	calcium channel, voltage-dependent, L type, alpha 1C subunit	GGAGCACAUUCGAUAACUUtt	AAGUUAUCGAAUGUGCUCCtc
CACNA1D	calcium channel, voltage-dependent, L type, alpha 1D subunit	GGACCAACUUCUCAGCCGAtt	UCGGCUGAGAAGUUGGUCCtt
CACNA1E	calcium channel, voltage-dependent, R type, alpha 1E subunit	GGUAUUCUAGAAGGAUUUGtt	CAAAUCCUUCUAGAAUACCtg

CACNA1F	calcium channel, voltage-dependent, L type, alpha 1F subunit	GGAUCUUUAAGGUCACCAGtt	CUGGUGACCUUAAAGAUCCtg
CACNA1S	calcium channel, voltage-dependent, L type, alpha 1S subunit	GGGUCUCCCAGACAAGUCAtt	UGACUUGUCUGGGAGACCCtt
CAV1	caveolin 1, caveolae protein, 22kDa	GGGACACACAGUUUUGACGtt	CGUCAAAACUGUGUGUCCCtt
CAV2	caveolin 2	GCACACAACGAUUAUAGUAtt	UACUAUAAUCGUUGUGUGCtt
CAV3	caveolin 3	GGAGAUGUGACGAAGGAACtt	GUUCCUUCGUCACAUCUCCtg
CDKN2A	cyclin-dependent kinase inhibitor 2A (melanoma, p16, inhibits CDK4)	CGUAGAUAUAUGCCUUCCCtt	GGGAAGGCAUAUAUCUACGtt
LYST	lysosomal trafficking regulator	CCUACCGCUCUCAGCAGAUtt	AUCUGCUGAGAGCGGUAGGtt
AP2M1	adaptor-related protein complex 2, mu 1 subunit	GGAAAACAUCAAGAACAAUtt	AUUGUUCUUGAUGUUUUCCtc
AP1S1	adaptor-related protein complex 1, sigma 1 subunit	CAAGGAACGGAAGAAGAUGtt	CAUCUUCUUCCGUUCCUUGtc
AP2S1	adaptor-related protein complex 2, sigma 1 subunit	GGUCUUAAACGAAUAUUUCtt	GAAAUAUUCGUUUAAGACCtc
AP3S1	adaptor-related protein complex 3, sigma 1 subunit	GGUAACUGGACUUCUAAAGtt	CUUUAGAAGUCCAGUUACCtt
CLTA	clathrin, light chain (Lca)	GGAAAGUAAUGGUCCAACAtt	UGUUGGACCAUUACUUUCCtg
CLTB	clathrin, light chain B	GGUCACGGAACAGGAAUGGtt	CCAUUCCUGUUCCGUGACCtt
CLTC	clathrin, heavy chain (Hc)	GCCACAGCUGGAAUAAUUGtt	CAAUUAUUCCAGCUGUGGCtt
COPA	coatomer protein complex, subunit alpha	GGCAUUGACUUCCAUAAGCtt	GCUUAUGGAAGUCAAUGCCtc
COPB1	coatomer protein complex, subunit beta 1	CCAACAUGGUUGAUUUAAAtt	UUUAAAUCAACCAUGUUGGtg
DNM1	dynamin 1	GGGAAAGAAAUUCACCGACtt	GUCGGUGAAUUUCUUUCCCtt
DNM2	dynamin 2	GCUCAUCUUCUCAAAAACAtt	UGUUUUUGAGAAGAUGAGCtg

ENG	endoglin (Osler-Rendu-Weber syndrome 1)	GGUGUCAGCAAGUAUGAUCtt	GAUCAUACUUGCUGACACCtg
STX2	syntaxin 2	GGUGGAGGAGAUUAGAAACtt	GUUUCUAAUCUCCUCCACCtg
HPS1	Hermansky-Pudlak syndrome 1	CGGAAAAUGGCAACUUCCUtt	AGGAAGUUGCCAUUUUCCGtg
HSPA8	heat shock 70kDa protein 8	CGGAAAAGUCGAGAUAAUUtt	AAUUAUCUCGACUUUUCCGtg
ITGA3	integrin, alpha 3 (antigen CD49C, alpha 3 subunit of VLA-3 receptor)	GGAGAUGCCACUUCUCACUtt	AGUGAGAAGUGGCAUCUCCtc
LMAN1	lectin, mannose-binding, 1	GGACAGAAUCGUAUUCAUCtt	GAUGAAUACGAUUCUGUCCtt
RAB3A	RAB3A, member RAS oncogene family	GGUCAAGACCAUCUAUCGCtt	GCGAUAGAUGGUCUUGACCtt
RASA1	RAS p21 protein activator (GTPase activating protein) 1	GGCGUGUACGAGCUAUUCUtt	AGAAUAGCUCGUACACGCCtt
SNAP25	synaptosomal-associated protein, 25kDa	GGGUAACAAAUGAUGCCCGtt	CGGGCAUCAUUUGUUACCCtg
STX1A	syntaxin 1A (brain)	GGUCAUGUCGGAGUACAACtt	GUUGUACUCCGACAUGACCtc
STX3	syntaxin 3	CGGCUUUUAUGGACGAGUUtt	AACUCGUCCAUAAAAGCCGtg
STX4	syntaxin 4	GCAGCUUGAACGCAGUAUUtt	AAUACUGCGUUCAAGCUGCtg
STX5	syntaxin 5	GGACAUCAAUAGCCUCAACtt	GUUGAGGCUAUUGAUGUCCtg
STXBP1	syntaxin binding protein 1	GGCUAUGAGUUAUGAUCUGtt	CAGAUCAUAACUCAUAGCCtg
STXBP2	syntaxin binding protein 2	GGUGCUUAUCAUGGAUCACtt	GUGAUCCAUGAUAAGCACCtt
STXBP3	syntaxin binding protein 3	GGUGUAUACUCUUGAUGUAtt	UACAUCAAGAGUAUACACCtg
VAMP1	vesicle-associated membrane protein 1 (synaptobrevin 1)	CCAGUAACAGACGACUACAtt	UGUAGUCGUCUGUUACUGGtc

VAMP2	vesicle-associated membrane protein 2 (synaptobrevin 2)	GCAUGUUAAUUAGCGUAUCtt	GAUACGCUAAUUAACAUGCtg
VAMP7	vesicle-associated membrane protein 7	GGUGACAGAGCAGAUUCUGtt	CAGAAUCUGCUCUGUCACCtc
SYT1	synaptotagmin I	GCCAUGCAUUCCUAACACAtt	UGUGUUAGGAAUGCAUGGCtg
SYT4	synaptotagmin IV	GGGAGUUGAUAUUUACCCUtt	AGGGUAAAUAUCAACUCCCtt
SYT5	synaptotagmin V	GGUGCAGCCAGAAGUAGAGtt	CUCUACUUCUGGCUGCACCtt
AP3B2	adaptor-related protein complex 3, beta 2 subunit	CGCCAUCCCUAAACUCUACtt	GUAGAGUUUAGGGAUGGCGtg
STX7	syntaxin 7	CCUGAAUUGAGGCAACAGUtt	ACUGUUGCCUCAAUUCAGGtg
AP3B1	adaptor-related protein complex 3, beta 1 subunit	GGGAAAAAUGCAUCUGAACtt	GUUCAGAUGCAUUUUUCCCtt
VAMP8	vesicle-associated membrane protein 8 (endobrevin)	GGAUCUGGAAGCCACAUCUtt	AGAUGUGGCUUCCAGAUCCtc
VAMP4	vesicle-associated membrane protein 4	GGACAAAUCAGAAAGCUUAtt	UAAGCUUUCUGAUUUGUCCtg
STX16	syntaxin 16	GAAGCUGUUCAGUAUUAUAtt	UAUAAUACUGAACAGCUUCtg
STX11	syntaxin 11	GGAUGAAAACCAGCUGCUGtt	CAGCAGCUGGUUUUCAUCCtg
STX10	syntaxin 10	CCAUCGGUAUAGUGGAAGCtt	GCUUCCACUAUACCGAUGGtc
GBF1	golgi brefeldin A resistant guanine nucleotide exchange factor 1	GCCCAAUGUAUUCCUUCGAtt	UCGAAGGAAUACAUUGGGCtc
SNAP23	synaptosomal-associated protein, 23kDa	CCAACAGAGAUCGUAUUGAtt	UCAAUACGAUCUCUGUUGGtg
AP1S2	adaptor-related protein complex 1, sigma 2 subunit	GCACUGUCUUAUCACAUCGtt	CGAUGUGAUAAGACAGUGCtc
AP1G2	adaptor-related protein complex 1, gamma 2 subunit	GCCUGGCUCUGGUAAAUAGtt	CUAUUUACCAGAGCCAGGCtt
AP1M1	adaptor-related protein complex 1, mu 1 subunit	GCCUUUGAUAUGGAUCGAGtt	CUCGAUCCAUAUCAAAGGCtt

CACNA1I	calcium channel, voltage-dependent, T type, alpha 1I subunit	GGAGCAGACAGCAAUACUUtt	AAGUAUUGCUGUCUGCUCCtg
CACNA1H	calcium channel, voltage-dependent, T type, alpha 1H subunit	GUGCCUUUGUCAGGAACAAtt	UUGUUCCUGACAAAGGCACtg
CACNA1G	calcium channel, voltage-dependent, T type, alpha 1G subunit	CCGAUGCUUCCUACCUGAGtt	CUCAGGUAGGAAGCAUCGGtt
AP3D1	adaptor-related protein complex 3, delta 1 subunit	CCUUCAAGCGAAUUGGCUAtt	UAGCCAAUUCGCUUGAAGGtg
SYT7	synaptotagmin VII	CGAGACCUUCCUCUUUGAAtt	UUCAAAGAGGAAGGUCUCGtt
AP4M1	adaptor-related protein complex 4, mu 1 subunit	AUCCGGGUCGAUGAAGUCUtt	AGACUUCAUCGACCCGGAUtc
VAPB	VAMP (vesicle-associated membrane protein)-associated protein B and C	GGGAAACCAUGAGUAAUGCtt	GCAUUACUCAUGGUUUCCCtt
COPB2	coatomer protein complex, subunit beta 2 (beta prime)	GCGGAUGACAUGCAGAUUAtt	UAAUCUGCAUGUCAUCCGCtc
VAMP3	vesicle-associated membrane protein 3	GGUGGUGGACAUAAUGCGAtt	UCGCAUUAUGUCCACCACCtc
SNAP29	synaptosomal-associated protein, 29kDa	CCCACACCUUCGAGCCUAUtt	AUAGGCUCGAAGGUGUGGGtt
STX8	syntaxin 8	CCUCUUGGAUGAUCUUGUAtt	UACAAGAUCAUCCAAGAGGtt
RGS6	regulator of G-protein signaling 6	GGUACUGACAUUGUGCAGUtt	ACUGCACAAUGUCAGUACCtg
RICS	Rho GTPase-activating protein	GGUUCUUCAAAGCUGCACAtt	UGUGCAGCUUUGAAGAACCtg
IQSEC1	IQ motif and Sec7 domain 1	GCAAGUCUGAGUCGGACUAtt	UAGUCCGACUCAGACUUGCtc
AP1M2	adaptor-related protein complex 1, mu 2 subunit	GCGAAAUCGUCGGUACCAUtt	AUGGUACCGACGAUUUCGCtc
HUWE1	HECT, UBA and WWE domain containing 1	GCAAUGGCUGCCAGAAUUAtt	UAAUUCUGGCAGCCAUUGCtg

PREB	prolactin regulatory element binding	GGCAUAAAGAAUGGCGUGCtt	GCACGCCAUUCUUUAUGCCtg
M6PRBP1	mannose-6-phosphate receptor binding protein 1	GGCUCUCAAAACGGGCAUCtt	GAUGCCCGUUUUGAGAGCCtt
STX6	syntaxin 6	GCCUUCAUUACAAGUACUCtt	GAGUACUUGUAAUGAAGGCtt
AP3S2	adaptor-related protein complex 3, sigma 2 subunit	CCUGGAAACUUUGUGUCUGtt	CAGACACAAAGUUUCCAGGtc
SEC24B	SEC24 related gene family, member B (S. cerevisiae)	GCUCUUCUGUUGCGUCUCAtt	UGAGACGCAACAGAAGAGCtg
SEC23A	Sec23 homolog A (S. cerevisiae)	GGAAUCAGUUUCCACCUAGtt	CUAGGUGGAAACUGAUUCCtt
AP4B1	adaptor-related protein complex 4, beta 1 subunit	GCGAGUGAUUAGGUACAUGtt	CAUGUACCUAAUCACUCGCtg
EHD1	EH-domain containing 1	GGUCCAUAAAGACUGAGCGtt	CGCUCAGUCUUUAUGGACCtt
AP3M2	adaptor-related protein complex 3, mu 2 subunit	GCCAGUGAUCAAAGACAAUtt	AUUGUCUUUGAUCACUGGCtc
AP4S1	adaptor-related protein complex 4, sigma 1 subunit	CGAGAUGGCUAUUUAUGAAtt	UUCAUAAAUAGCCAUCUCGtt
HPS5	Hermansky-Pudlak syndrome 5	GCAUAGCUGUGUCUCGGAAtt	UUCCGAGACACAGCUAUGCtc
AP1GBP1	AP1 gamma subunit binding protein 1	GCUCCUAUUCCAACUUUAAtt	UUAAAGUUGGAAUAGGAGCtg
COPE	coatomer protein complex, subunit epsilon	GGAGCUGAAGAGAAUGCAGtt	CUGCAUUCUCUUCAGCUCCtt
COPZ1	coatomer protein complex, subunit zeta 1	GCAUAUUUAGAUAAUAGGGtt	CCCUAUUAUCUAAAUAUGCtg
COPG1	coatomer protein complex, subunit gamma 1	GGAGAUGUCUUGCAUUGCAtt	UGCAAUGCAAGACAUCUCCtt
SEC31A	SEC31 homolog A (S. cerevisiae)	GAUGAGCACCUCAUUCUAAtt	UUAGAAUGAGGUGCUCAUCtg
SYT11	synaptotagmin XI	GGUGGACGUAGGAACCUGUtt	ACAGGUUCCUACGUCCACCtt
AP4E1	adaptor-related protein complex 4, epsilon 1 subunit	GCUGUUCUGGCAUUAUACAtt	UGUAUAAUGCCAGAACAGCtt
STX12	syntaxin 12	CCUAUGGAGACAGUAAUUAtt	UAAUUACUGUCUCCAUAGGtc
DNM3	dynamin 3	GGGACAGUUGCUCUCCAUAtt	UAUGGAGAGCAACUGUCCCtg

VPS33B	vacuolar protein sorting 33 homolog B (yeast)	GCAACACGAAGUAGACAAGtt	CUUGUCUACUUCGUGUUGCtt
COPG2	coatomer protein complex, subunit gamma 2	CCAGGGUGAACACUUUGGAtt	UCCAAAGUGUUCACCCUGGtt
AP3M1	adaptor-related protein complex 3, mu 1 subunit	GGAAUUACAGUGACAGUUCtt	GAACUGUCACUGUAAUUCCtt
STXBP6	syntaxin binding protein 6 (amisyn)	GGUCAAGGCGAAUAUUUAAtt	UUAAAUAUUCGCCUUGACCtc
RACGAP1	Rac GTPase activating protein 1	GGACUUUGAGGAUUUCCGUtt	ACGGAAAUCCUCAAAGUCCtt
EHD4	EH-domain containing 4	GGGAAUGCUUUAGUCGUGGtt	CCACGACUAAAGCAUUCCCtg
EHD3	EH-domain containing 3	CCUCAUACUAUUCUCAGAAtt	UUCUGAGAAUAGUAUGAGGtc
EHD2	EH-domain containing 2	CCAUUCCAUAUUUAAGUGGtt	CCACUUAAAUAUGGAAUGGtg
SAR1B	SAR1 homolog B (S. cerevisiae)	GGAAAAACAACAUUGCUACtt	GUAGCAAUGUUGUUUUUCCtg
COPZ2	coatomer protein complex, subunit zeta 2	CCUCUUCCUAUACGUGGUGtt	CACCACGUAUAGGAAGAGGtc
SYT17	synaptotagmin XVII	GGAUUUCGAGUCUUGAGUCtt	GACUCAAGACUCGAAAUCCtc
STX18	syntaxin 18	GCGACUAAUUGGUGAAAUGtt	CAUUUCACCAAUUAGUCGCtg
STX17	syntaxin 17	GGAUGACCUAGUACUUCUGtt	CAGAAGUACUAGGUCAUCCtt
C4orf16	chromosome 4 open reading frame 16	GGCAUUAUGAUUCCAUUGCtt	GCAAUGGAAUCAUAAUGCCtt
SAR1A	SAR1 gene homolog A (S. cerevisiae)	GGCAAAACCACUCUUCUUCtt	GAAGAAGAGUGGUUUUGCCtg
FAM62B	family with sequence similarity 62 (C2 domain containing) member B	CCCACCUUAGCACCUUUAGtt	CUAAAGGUGCUAAGGUGGGtg
SYT13	synaptotagmin XIII	GGAGAGUUGUCUGUUAUGUtt	ACAUAACAGACAACUCUCCtt
TBC1D15	TBC1 domain family, member 15	CCAAUGACCAAGACGGCUUtt	AAGCCGUCUUGGUCAUUGGtc
ARHGAP10	Rho GTPase activating protein 10	GGCAGUCGAAGAGACAAGGtt	CCUUGUCUCUUCGACUGCCtt
HPS6	Hermansky-Pudlak syndrome 6	CGUUCUACUCAUCUGGAGCtt	GCUCCAGAUGAGUAGAACGtg

PTGES2	prostaglandin E synthase 2	CCGAGUUCGGCAAUAAGUAtt	UACUUAUUGCCGAACUCGGtc
SYT15	synaptotagmin XV	GGCAUUGUCAGUGUGUUUGtt	CAAACACACUGACAAUGCCtc
FAM62C	family with sequence similarity 62 (C2 domain containing), member C	CGGUGUCAAGGCACACACUtt	AGUGUGUGCCUUGACACCGtt
SYT16	synaptotagmin XVI	GGAGAUAUGUUAUCUGCUUtt	AAGCAGAUAACAUAUCUCCtg
DTNBP1	dystrobrevin binding protein 1	GGCGAGUUUUGAGGAGGUAtt	UACCUCCUCAAAACUCGCCtc
SYT3	synaptotagmin III	CGUUUCAAUUCUCGGUGCCtt	GGCACCGAGAAUUGAAACGtc
HPS3	Hermansky-Pudlak syndrome 3	GCGAGGCUGGAGAUUAUUUtt	AAAUAAUCUCCAGCCUCGCtg
HPS4	Hermansky-Pudlak syndrome 4	GCUAGUUGGAUUCUUUAAUtt	AUUAAAGAAUCCAACUAGCtg
SYT8	synaptotagmin VIII	CUCCUGAAUGCACCACAUGtt	CAUGUGGUGCAUUCAGGAGtg
SYT12	synaptotagmin XII	GGUACCUUCAGCAGAAGUAtt	UACUUCUGCUGAAGGUACCtg
STX1B	syntaxin 1B	GGUUCGGUCCAAAUUGAAAtt	UUUCAAUUUGGACCGAACCtt
SNX26	sorting nexin 26	GGCAACAAAGUGAUGGGAGtt	CUCCCAUCACUUUGUUGCCtt
SYT2	synaptotagmin II	GAGUCAGGAGGACAUGUUUtt	AAACAUGUCCUCCUGACUCtc
AP1S3	adaptor-related protein complex 1, sigma 3 subunit	AUAGAAAAUCAGGACAAUGtt	CAUUGUCCUGAUUUUCUAUtg
STXBP5	syntaxin binding protein 5 (tomosyn)	GGCUACGUCAUUAUGUGGAtt	UCCACAUAAUGACGUAGCCtg
SYT9	synaptotagmin IX	GGAUCGUUAUACGUCUUCAtt	UGAAGACGUAUAACGAUCCtt
SYT6	synaptotagmin VI	CCCUGAUUGUGCGUAUCCUtt	AGGAUACGCACAAUCAGGGtc
VPS13B	vacuolar protein sorting 13 homolog B (yeast)	GCUCGAGUUAAAGUUGGAUtt	AUCCAACUUUAACUCGAGCtt
STXBP4	syntaxin binding protein 4	GCCCUGAGACAGCAAGUACtt	GUACUUGCUGUCUCAGGGCtt
SYT14	synaptotagmin XIV	CCACCAUAUCAGGAUGACAtt	UGUCAUCCUGAUAUGGUGGtg

SYT10	synaptotagmin X	GCAAACCUGUGACUUCCAAtt	UUGGAAGUCACAGGUUUGCtt
STX19	syntaxin 19	GCAAGCUGUUAUUUAUGAAtt	UUCAUAAAUAACAGCUUGCtg
AP2A1	adaptor-related protein complex 2, alpha 1 subunit	GGAGACAAAGCCUUGGAUGtt	CAUCCAAGGCUUUGUCUCCtt
AP2A2	adaptor-related protein complex 2, alpha 2 subunit	GGCCAUCAGCUUAAUCAUUtt	AAUGAUUAAGCUGAUGGCCtc
AP1B1	adaptor-related protein complex 1, beta 1 subunit	GGACGAGGAUCCAUAUGUGtt	CACAUAUGGAUCCUCGUCCtt
AP2B1	adaptor-related protein complex 2, beta 1 subunit	GGACUGUGAAGAUCCUAAUtt	AUUAGGAUCUUCACAGUCCtt
AP1G1	adaptor-related protein complex 1, gamma 1 subunit	GGCCAAUGAUUUAUUGGAUtt	AUCCAAUAAAUCAUUGGCCtg
ARF1	ADP-ribosylation factor 1	GGCAGUUUCUGGUACUCCUtt	AGGAGUACCAGAAACUGCCtc
CACNA1A	calcium channel, voltage-dependent, P/Q type, alpha 1A subunit	GGAUGCGUUUCUACAUCCGtt	CGGAUGUAGAAACGCAUCCtc
CACNA1B	calcium channel, voltage-dependent, N type, alpha 1B subunit	GGCAGGGAUCAAAAUCAUCtt	GAUGAUUUUGAUCCCUGCCtc
CACNA1C	calcium channel, voltage-dependent, L type, alpha 1C subunit	GGAUGACGAAAAUCGGCAAtt	UUGCCGAUUUUCGUCAUCCtg
CACNA1D	calcium channel, voltage-dependent, L type, alpha 1D subunit	GGCAAACGAAAUACUAGCAtt	UGCUAGUAUUUCGUUUGCCtt
CACNA1E	calcium channel, voltage-dependent, R type, alpha 1E subunit	GGCAAGGAGUCCUCUUCAGtt	CUGAAGAGGACUCCUUGCCtt
CACNA1F	calcium channel, voltage-dependent, L type, alpha 1F subunit	GCGAAGAAACCAGCACAGCtt	GCUGUGCUGGUUUCUUCGCtt

CACNA1S	calcium channel, voltage-dependent, L type, alpha 1S subunit	GGUGCUACAUUCCCAAAAAtt	UUUUUGGGAAUGUAGCACCtc
CAV1	caveolin 1, caveolae protein, 22kDa	CGAAAUACUGGUUUUACCGtt	CGGUAAAACCAGUAUUUCGtc
CAV2	caveolin 2	GGAUUGAAUACUUGGACCCtt	GGGUCCAAGUAUUCAAUCCtg
CAV3	caveolin 3	CGAAGGAACAAGGUCUAAUtt	AUUAGACCUUGUUCCUUCGtc
CDKN2A	cyclin-dependent kinase inhibitor 2A (melanoma, p16, inhibits CDK4)	CCAGAGAGGCUCUGAGAAAtt	UUUCUCAGAGCCUCUCUGGtt
LYST	lysosomal trafficking regulator	GCAGACGACAGCGUAAAAUtt	AUUUUACGCUGUCGUCUGCtt
AP2M1	adaptor-related protein complex 2, mu 1 subunit	GGCACAGCUGAUGAAACAAtt	UUGUUUCAUCAGCUGUGCCtt
AP1S1	adaptor-related protein complex 1, sigma 1 subunit	GGAACGGAAGAAGAUGGUGtt	CACCAUCUUCUUCCGUUCCtt
AP2S1	adaptor-related protein complex 2, sigma 1 subunit	GGUUUACACGGUCGUGGACtt	GUCCACGACCGUGUAAACCtt
AP3S1	adaptor-related protein complex 3, sigma 1 subunit	CCAUGCAGUUGUUUACCAAtt	UUGGUAAACAACUGCAUGGta
CLTA	clathrin, light chain (Lca)	GGAGCUAGAAGAAUGGUAUtt	AUACCAUUCUUCUAGCUCCtt
CLTB	clathrin, light chain B	GGCUUUCGUGAAGGAAUCCtt	GGAUUCCUUCACGAAAGCCtc
CLTC	clathrin, heavy chain (Hc)	GGUGGUAAUCAUUGAUAUGtt	CAUAUCAAUGAUUACCACCtg
COPA	coatomer protein complex, subunit alpha	CCUUGGAUCCUGACUAGUUtt	AACUAGUCAGGAUCCAAGGtc
COPB1	coatomer protein complex, subunit beta 1	GGAUCACACUAUCAAGAAAtt	UUUCUUGAUAGUGUGAUCCtg
DNM1	dynamin 1	GGUCAUGCUUCUCAUCGAUtt	AUCGAUGAGAAGCAUGACCtg
DNM2	dynamin 2	GCAGUCCUACAUCAACACGtt	CGUGUUGAUGUAGGACUGCtc
ENG	endoglin (Osler-Rendu-Weber syndrome 1)	GGUGACAUAUACCACUAGCtt	GCUAGUGGUAUAUGUCACCtc
STX2	syntaxin 2	GGAGAUUAGAAACAGUAUUtt	AAUACUGUUUCUAAUCUCCtc

HPS1	Hermansky-Pudlak syndrome 1	GGUCCUCUUCUACUGGACAtt	UGUCCAGUAGAAGAGGACCtc
HSPA8	heat shock 70kDa protein 8	CCUAAAUUCGUAGCAAAUUtt	AAUUUGCUACGAAUUUAGGtc
ITGA3	integrin, alpha 3 (antigen CD49C, alpha 3 subunit of VLA-3 receptor)	CGAGAUGUGCAAUAGCAACtt	GUUGCUAUUGCACAUCUCGtt
LMAN1	lectin, mannose-binding, 1	GGAAAUCUCUAAAAGAGGAtt	UCCUCUUUUAGAGAUUUCCtc
RAB3A	RAB3A, member RAS oncogene family	GGAAUCCUUCAAUGCAGUGtt	CACUGCAUUGAAGGAUUCCtc
RASA1	RAS p21 protein activator (GTPase activating protein) 1	CCUAAUAGGUUAUUACAGUtt	ACUGUAAUAACCUAUUAGGtc
SNAP25	synaptosomal-associated protein, 25kDa	GGUUGUACAUAGUGGUCAUtt	AUGACCACUAUGUACAACCtt
STX1A	syntaxin 1A (brain)	CCAAAGGUCUUGGUACAACtt	GUUGUACCAAGACCUUUGGtg
STX3	syntaxin 3	GCGGCAGCUCGAAAUUACUtt	AGUAAUUUCGAGCUGCCGCtg
STX4	syntaxin 4	CGCAGUAUUCGUGAGCUGCtt	GCAGCUCACGAAUACUGCGtt
STX5	syntaxin 5	GCGCAAGUCCCUCUUUGAUtt	AUCAAAGAGGGACUUGCGCtt
STXBP1	syntaxin binding protein 1	GGACUGUAUGAAGCAUUACtt	GUAAUGCUUCAUACAGUCCtc
STXBP2	syntaxin binding protein 2	GGGCAUCACCAUUGUUGAAtt	UUCAACAAUGGUGAUGCCCtc
STXBP3	syntaxin binding protein 3	GGGUAAAACUCAUUCACAGtt	CUGUGAAUGAGUUUUACCCtt
VAMP1	vesicle-associated membrane protein 1 (synaptobrevin 1)	GGAGGUGGUGGACAUCAUAtt	UAUGAUGUCCACCACCUCCtc
VAMP2	vesicle-associated membrane protein 2 (synaptobrevin 2)	CCUCUUGUACCCUCCUCUCtt	GAGAGGAGGGUACAAGAGGtt
VAMP7	vesicle-associated membrane protein 7	UGAGAGAACAAGGAGUUAAtt	UUAACUCCUUGUUCUCUCAta

SYT1	synaptotagmin I	GGAAAGGAAGAUGCAUUUUtt	AAAAUGCAUCUUCCUUUCCtt
SYT4	synaptotagmin IV	GGUUUUCAAGAGAUGAUAUtt	AUAUCAUCUCUUGAAAACCtg
SYT5	synaptotagmin V	GGCCAAAUAAAUAUUCAGCtt	GCUGAAUAUUUAUUUGGCCtc
AP3B2	adaptor-related protein complex 3, beta 2 subunit	GGAGAUGCUGGACACCAACtt	GUUGGUGUCCAGCAUCUCCtt
STX7	syntaxin 7	GCAGCAGUAUACUAACCAGtt	CUGGUUAGUAUACUGCUGCtt
AP3B1	adaptor-related protein complex 3, beta 1 subunit	GGACCCAAACCAACUAAUUtt	AAUUAGUUGGUUUGGGUCCtt
VAMP8	vesicle-associated membrane protein 8 (endobrevin)	GGGACAACCCUCCAUAAAUtt	AUUUAUGGAGGGUUGUCCCtg
VAMP4	vesicle-associated membrane protein 4	GCUUAUCGGAUAAUGCAACtt	GUUGCAUUAUCCGAUAAGCtt
STX16	syntaxin 16	GCAGAAGAUGAAAGAAUUGtt	CAAUUCUUUCAUCUUCUGCtt
STX11	syntaxin 11	CGUGUUUUCCGAGAACUUGtt	CAAGUUCUCGGAAAACACGtc
STX10	syntaxin 10	GGACCAUAUGGUCAGCCCAtt	UGGGCUGACCAUAUGGUCCtt
GBF1	golgi brefeldin A resistant guanine nucleotide exchange factor 1	CCACCUUAUCAUCUAACCUtt	AGGUUAGAUGAUAAGGUGGtt
SNAP23	synaptosomal-associated protein, 23kDa	GGAAUCAAGACCAUCACUAtt	UAGUGAUGGUCUUGAUUCCtg
AP1S2	adaptor-related protein complex 1, sigma 2 subunit	GGUCAGUUUCUCUUAAUCUtt	AGAUUAAGAGAAACUGACCtt
AP1G2	adaptor-related protein complex 1, gamma 2 subunit	GGUUUGCUCCAACCAAACGtt	CGUUUGGUUGGAGCAAACCtc
AP1M1	adaptor-related protein complex 1, mu 1 subunit	GGUGGUGCAGGUGUUUUCCtt	GGAAAACACCUGCACCACCtt
CACNA1I	calcium channel, voltage-dependent, T type, alpha 1I subunit	CCAGAGCUCAUCCAACAUAtt	UAUGUUGGAUGAGCUCUGGtc

CACNA1H	calcium channel, voltage-dependent, T type, alpha 1H subunit	GGCUGGAUUUCUUCAUCGUtt	ACGAUGAAGAAAUCCAGCCtg
CACNA1G	calcium channel, voltage-dependent, T type, alpha 1G subunit	GGAGAAGGCACUAGUAGAGtt	CUCUACUAGUGCCUUCUCCtt
AP3D1	adaptor-related protein complex 3, delta 1 subunit	GGAUAUUGAUCGAGGACUCtt	GAGUCCUCGAUCAAUAUCCtt
SYT7	synaptotagmin VII	GGAACAUCAUACAUAGUAGtt	CUACUAUGUAUGAUGUUCCtc
AP4M1	adaptor-related protein complex 4, mu 1 subunit	CUAUGGCUAUGUACAGACCtt	GGUCUGUACAUAGCCAUAGtc
VAPB	VAMP (vesicle-associated membrane protein)-associated protein B and C	CCGGAAGACCUUAUGGAUUtt	AAUCCAUAAGGUCUUCCGGtt
COPB2	coatomer protein complex, subunit beta 2 (beta prime)	GCCUUUCAUUCUAACUAGCtt	GCUAGUUAGAAUGAAAGGCtg
VAMP3	vesicle-associated membrane protein 3	GGAAAUAUUGGUGGAAGAAtt	UUCUUCCACCAAUAUUUCCtc
SNAP29	synaptosomal-associated protein, 29kDa	GCCUAUCACCAGAAGAUCGtt	CGAUCUUCUGGUGAUAGGCtc
STX8	syntaxin 8	CCUGCUCUCAAUAAAUUCCtt	GGAAUUUAUUGAGAGCAGGtt
RGS6	regulator of G-protein signaling 6	GGACCCAGUUGAAGCAAUAtt	UAUUGCUUCAACUGGGUCCtc
RICS	Rho GTPase-activating protein	GGACAUCCAUUCUGUGGGUtt	ACCCACAGAAUGGAUGUCCtg
IQSEC1	IQ motif and Sec7 domain 1	GGUCUCAGUGACUAACGACtt	GUCGUUAGUCACUGAGACCtg
AP1M2	adaptor-related protein complex 1, mu 2 subunit	GGUCCACUUCCUAUGGAUCtt	GAUCCAUAGGAAGUGGACCtg
HUWE1	HECT, UBA and WWE domain containing 1	GAAUUAAGCAAAUCAAGCCtt	GGCUUGAUUUGCUUAAUUCtg
PREB	prolactin regulatory element binding	GCUGGCCUAAAGAUGCAAUtt	AUUGCAUCUUUAGGCCAGCtg
M6PRBP1	mannose-6-phosphate receptor binding protein 1	CCCUAUCAGUCAGUAGAGAtt	UCUCUACUGACUGAUAGGGtg

STX6	syntaxin 6	GGCAUUAGCUGAAAGAAAAtt	UUUUCUUUCAGCUAAUGCCtg
AP3S2	adaptor-related protein complex 3, sigma 2 subunit	GGGUUAGGGACUGCAUUGAtt	UCAAUGCAGUCCCUAACCCtt
SEC24B	SEC24 related gene family, member B (S. cerevisiae)	GGGAUUUCCCUCUACUUGUtt	ACAAGUAGAGGGAAAUCCCtg
SEC23A	Sec23 homolog A (S. cerevisiae)	GGUUCAUGAACUUGGAUGUtt	ACAUCCAAGUUCAUGAACCtg
AP4B1	adaptor-related protein complex 4, beta 1 subunit	GCACUUAUUUGGCUACUUGtt	CAAGUAGCCAAAUAAGUGCtt
EHD1	EH-domain containing 1	GGCUCUGAAGAACCAUGAGtt	CUCAUGGUUCUUCAGAGCCtt
AP3M2	adaptor-related protein complex 3, mu 2 subunit	GCCUCCUACCAUCCUUCGAtt	UCGAAGGAUGGUAGGAGGCtt
AP4S1	adaptor-related protein complex 4, sigma 1 subunit	GGAACCAAUUGAUGAACUUtt	AAGUUCAUCAAUUGGUUCCtg
HPS5	Hermansky-Pudlak syndrome 5	GCCCUGAUCUGAAAGUGAGtt	CUCACUUUCAGAUCAGGGCtc
AP1GBP1	AP1 gamma subunit binding protein 1	GCCGAUAGUGUAUCACCACtt	GUGGUGAUACACUAUCGGCtg
COPE	coatomer protein complex, subunit epsilon	GGAGAACGACUUUGACAGGtt	CCUGUCAAAGUCGUUCUCCtt
COPZ1	coatomer protein complex, subunit zeta 1	GCCAGUUCCUAAAACUAAAtt	UUUAGUUUUAGGAACUGGCtt
COPG1	coatomer protein complex, subunit gamma 1	GCAGCCUAACAAAAGACAUtt	AUGUCUUUUGUUAGGCUGCtg
SEC31A	SEC31 homolog A (S. cerevisiae)	CAUUUGAGGAUCUUAUUCAtt	UGAAUAAGAUCCUCAAAUGtg
SYT11	synaptotagmin XI	CCAUAUACUUGCAUGUCUAtt	UAGACAUGCAAGUAUAUGGtt
AP4E1	adaptor-related protein complex 4, epsilon 1 subunit	GGAAGUUAUAGCUAAGCUCtt	GAGCUUAGCUAUAACUUCCtc
STX12	syntaxin 12	GGGUUGCUGGGCAUUUGAAtt	UUCAAAUGCCCAGCAACCCtt
DNM3	dynamin 3	GGUUACCAGUGAUUGUAUAtt	UAUACAAUCACUGGUAACCtt
VPS33B	vacuolar protein sorting 33 homolog B (yeast)	GCUAUACAAGGUGGAGAACtt	GUUCUCCACCUUGUAUAGCtt
COPG2	coatomer protein complex, subunit gamma 2	GGAUGGGUAUGAUGAUGAGtt	CUCAUCAUCAUACCCAUCCtc

AP3M1	adaptor-related protein complex 3, mu 1 subunit	GCCUCAACAUACAGUUUAAtt	UUAAACUGUAUGUUGAGGCtc
STXBP6	syntaxin binding protein 6 (amisyn)	GGAGGACAAAGAAAAAGAUtt	AUCUUUUUCUUUGUCCUCCtc
RACGAP1	Rac GTPase activating protein 1	GGAUUUCCGUAAAAAGUGGtt	CCACUUUUUACGGAAAUCCtc
EHD4	EH-domain containing 4	GGAUCAUCCUGCUCUUUGAtt	UCAAAGAGCAGGAUGAUCCtg
EHD3	EH-domain containing 3	GGAACAGGACCUAUUCAGGtt	CCUGAAUAGGUCCUGUUCCtc
EHD2	EH-domain containing 2	GGAGAACAAGAAGAAGCAGtt	CUGCUUCUUCUUGUUCUCCtt
SAR1B	SAR1 homolog B (S. cerevisiae)	GGCUGUUAGAGUCAAAAGAtt	UCUUUUGACUCUAACAGCCtt
COPZ2	coatomer protein complex, subunit zeta 2	GGCUGUUUUCAUCCUAGAUtt	AUCUAGGAUGAAAACAGCCtt
SYT17	synaptotagmin XVII	GCUGGUGCAUGGACUCAAAtt	UUUGAGUCCAUGCACCAGCtg
STX18	syntaxin 18	GGAUGACAGACACAGAACGtt	CGUUCUGUGUCUGUCAUCCtc
STX17	syntaxin 17	GGUCGACUACAUAAUGGAGtt	CUCCAUUAUGUAGUCGACCtg
C4orf16	chromosome 4 open reading frame 16	GCUUUAUACGCUGCACAGCtt	GCUGUGCAGCGUAUAAAGCtt
SAR1A	SAR1 gene homolog A (S. cerevisiae)	GGAAAGCUCUCCAACCAUGtt	CAUGGUUGGAGAGCUUUCCtt
FAM62B	family with sequence similarity 62 (C2 domain containing) member B	GGAUCAAUGGUGUUAAGGUtt	ACCUUAACACCAUUGAUCCtg
SYT13	synaptotagmin XIII	GGAUGUCUGUGUCAUGGAGtt	CUCCAUGACACAGACAUCCtc
TBC1D15	TBC1 domain family, member 15	GGAUGCCGAAGUAAUAGUGtt	CACUAUUACUUCGGCAUCCtt
ARHGAP10	Rho GTPase activating protein 10	GGAUGUACAAACCUCCAGGtt	CCUGGAGGUUUGUACAUCCtc
HPS6	Hermansky-Pudlak syndrome 6	GGAAGGUCCUAAGUACAGAtt	UCUGUACUUAGGACCUUCCtc
PTGES2	prostaglandin E synthase 2	GGGCUGAGAUCAAGUUCUCtt	GAGAACUUGAUCUCAGCCCtg
SYT15	synaptotagmin XV	CGAGCACUUCAUCUUUCAGtt	CUGAAAGAUGAAGUGCUCGtc

FAM62C	family with sequence similarity 62 (C2 domain containing), member C	GGCACACACUAAUACGUGCtt	GCACGUAUUAGUGUGUGCCtt
SYT16	synaptotagmin XVI	GGACUUAGAUAAUAUUCAGtt	CUGAAUAUUAUCUAAGUCCtg
DTNBP1	dystrobrevin binding protein 1	GGACUGUUCCAUUUUUGCCtt	GGCAAAAAUGGAACAGUCCtg
SYT3	synaptotagmin III	GGAGUUCAAUGACCGAAUCtt	GAUUCGGUCAUUGAACUCCtg
HPS3	Hermansky-Pudlak syndrome 3	GCACCUGCUCUAUAGACGUtt	ACGUCUAUAGAGCAGGUGCtg
HPS4	Hermansky-Pudlak syndrome 4	GCCAUUAGUCUCCACGAGUtt	ACUCGUGGAGACUAAUGGCtt
SYT8	synaptotagmin VIII	ACCAUGCAGGGUAGAAAGAtt	UCUUUCUACCCUGCAUGGUtt
SYT12	synaptotagmin XII	GGCAGUCUCAGCAUUGAGGtt	CCUCAAUGCUGAGACUGCCtt
STX1B	syntaxin 1B	GGUAAUGACCGAAUAUAACtt	GUUAUAUUCGGUCAUUACCtc
SNX26	sorting nexin 26	GGGUCCUUGUACAGAAAUGtt	CAUUUCUGUACAAGGACCCtc
SYT2	synaptotagmin II	GUCAGGAGGACAUGUUUGCtt	GCAAACAUGUCCUCCUGACtc
AP1S3	adaptor-related protein complex 1, sigma 3 subunit	CAAGGGAAAUUACGGCUACtt	GUAGCCGUAAUUUCCCUUGtc
STXBP5	syntaxin binding protein 5 (tomosyn)	GCCAUUGAACUGUCAUCUAtt	UAGAUGACAGUUCAAUGGCtt
SYT9	synaptotagmin IX	GCUGCCAGGAUUUCAUUUAtt	UAAAUGAAAUCCUGGCAGCtg
SYT6	synaptotagmin VI	GGCUAUUCAGAUCCCUAUGtt	CAUAGGGAUCUGAAUAGCCtg
VPS13B	vacuolar protein sorting 13 homolog B (yeast)	GGAACUGAAAUUACCAUUCtt	GAAUGGUAAUUUCAGUUCCtg
STXBP4	syntaxin binding protein 4	GGAUGGUCGUUUGAAGCCAtt	UGGCUUCAAACGACCAUCCtt
SYT14	synaptotagmin XIV	CCUUGUUCUUCUACCUAUAtt	UAUAGGUAGAAGAACAAGGtg
SYT10	synaptotagmin X	CGUGUGCAAAGACAAAUUAtt	UAAUUUGUCUUUGCACACGtg
STX19	syntaxin 19	GGAACUUGUUAAUUUGGAGtt	CUCCAAAUUAACAAGUUCCtt

AP2A1	adaptor-related protein complex 2, alpha 1 subunit	GGCUGUGAAUCUGUUGAGUtt	ACUCAACAGAUUCACAGCCtc
AP2A2	adaptor-related protein complex 2, alpha 2 subunit	GCCACAAGUCUGAUCACCAtt	UGGUGAUCAGACUUGUGGCtg
AP1B1	adaptor-related protein complex 1, beta 1 subunit	GGUGCUGAUGAAGUUCAUGtt	CAUGAACUUCAUCAGCACCtt
AP2B1	adaptor-related protein complex 2, beta 1 subunit	GGAACAUCAACUUAAUUGUtt	ACAAUUAAGUUGAUGUUCCtc
AP1G1	adaptor-related protein complex 1, gamma 1 subunit	GGAUUUUCCUAAAGUUACUtt	AGUAACUUUAGGAAAAUCCtc
ARF1	ADP-ribosylation factor 1	CCUCAGAACUGGUCUAUUUtt	AAAUAGACCAGUUCUGAGGtt
CACNA1A	calcium channel, voltage-dependent, P/Q type, alpha 1A subunit	GGCUGGAAUUAAAAUCAUUtt	AAUGAUUUUAAUUCCAGCCtc
CACNA1B	calcium channel, voltage-dependent, N type, alpha 1B subunit	GGGAAGUUUUUCUACUGCAtt	UGCAGUAGAAAAACUUCCCtt
CACNA1C	calcium channel, voltage-dependent, L type, alpha 1C subunit	GGAAGAAGAGUGUUUACAAtt	UUGUAAACACUCUUCUUCCtc
CACNA1D	calcium channel, voltage-dependent, L type, alpha 1D subunit	GGACACUGCAUGACAUUGGtt	CCAAUGUCAUGCAGUGUCCtt
CACNA1E	calcium channel, voltage-dependent, R type, alpha 1E subunit	GGUAUUUUGACUAUGUGUUtt	AACACAUAGUCAAAAUACCtc
CACNA1F	calcium channel, voltage-dependent, L type, alpha 1F subunit	GUACGUAUUCCUGGUGAUUtt	AAUCACCAGGAAUACGUACtc
CACNA1S	calcium channel, voltage-dependent, L type, alpha 1S subunit	GCUUCUAAUGUCAUUCAAAtt	UUUGAAUGACAUUAGAAGCtc
CAV1	caveolin 1, caveolae protein, 22kDa	GGAGUUAGUGGAUUACUGCtt	GCAGUAAUCCACUAACUCCtt

CAV2	caveolin 2	GGUCUUUUCAAAUAAUGCUtt	AGCAUUAUUUGAAAAGACCtg
CAV3	caveolin 3	GGAACAAGGUCUAAUUUGUtt	ACAAAUUAGACCUUGUUCCtt
CDKN2A	cyclin-dependent kinase inhibitor 2A (melanoma, p16, inhibits CDK4)	GCUUUUAAAAAUGUCCUGCtt	GCAGGACAUUUUUAAAAGCtc
LYST	lysosomal trafficking regulator	GCCCUAAUAAAUAGUGUGAtt	UCACACUAUUUAUUAGGGCtt
AP2M1	adaptor-related protein complex 2, mu 1 subunit	CGUUAAGCGGUCCAACAUUtt	AAUGUUGGACCGCUUAACGtg
AP1S1	adaptor-related protein complex 1, sigma 1 subunit	GGGACCUCAAAGUUGUCUAtt	UAGACAACUUUGAGGUCCCtc
AP2S1	adaptor-related protein complex 2, sigma 1 subunit	GGUGCUGAAACAGCUGCUGtt	CAGCAGCUGUUUCAGCACCtt
AP3S1	adaptor-related protein complex 3, sigma 1 subunit	GGAGGAUUAUUAAUUGGAGtt	CUCCAAUUAAUAAUCCUCCtt
CLTA	clathrin, light chain (Lca)	GCUACAGAAAACAAAAGCAtt	UGCUUUUGUUUUCUGUAGCtg
CLTB	clathrin, light chain B	UGCUGCAUCUAAGGUCACGtt	CGUGACCUUAGAUGCAGCAtc
CLTC	clathrin, heavy chain (Hc)	GGGCAAUUCAAAGAUAUCAtt	UGAUAUCUUUGAAUUGCCCtt
COPA	coatomer protein complex, subunit alpha	GGGCACAACCAUUAUGUGAtt	UCACAUAAUGGUUGUGCCCtg
COPB1	coatomer protein complex, subunit beta 1	GGAUCUUCAACAUCCUAAUtt	AUUAGGAUGUUGAAGAUCCtt
DNM1	dynamin 1	GGAGAACUGCCUCAUCCUGtt	CAGGAUGAGGCAGUUCUCCtt
DNM2	dynamin 2	GGGCUUCAUGUCCAACAAGtt	CUUGUUGGACAUGAAGCCCtt
ENG	endoglin (Osler-Rendu-Weber syndrome 1)	GGCAUCCAAGCAAAAUGGCtt	GCCAUUUUGCUUGGAUGCCtg
STX2	syntaxin 2	GGGUGAAAUGAUCAACAACtt	GUUGUUGAUCAUUUCACCCtg
HPS1	Hermansky-Pudlak syndrome 1	GGAGUACUUCACACCAGCUtt	AGCUGGUGUGAAGUACUCCtc
HSPA8	heat shock 70kDa protein 8	GGAGGUGUCUUCUAUGGUUtt	AACCAUAGAAGACACCUCCtc

ITGA3	integrin, alpha 3 (antigen CD49C, alpha 3 subunit of VLA-3 receptor)	GGAAACAGCUACAUGAUUCtt	GAAUCAUGUAGCUGUUUCCtt
LMAN1	lectin, mannose-binding, 1	GGCUCAGUGUGGACAAAGAtt	UCUUUGUCCACACUGAGCCtc
RAB3A	RAB3A, member RAS oncogene family	GAUUCUCAUCAUCGGCAACtt	GUUGCCGAUGAUGAGAAUCtt
RASA1	RAS p21 protein activator (GTPase activating protein) 1	GGAGAUAUGUUCAUUGUUCtt	GAACAAUGAACAUAUCUCCtt
SNAP25	synaptosomal-associated protein, 25kDa	CCCGUGUUCUCCUCCAAAUtt	AUUUGGAGGAGAACACGGGtg
STX1A	syntaxin 1A (brain)	GGAGAUUCGAGGCUUCAUUtt	AAUGAAGCCUCGAAUCUCCtc
STX3	syntaxin 3	GGGUGAGAUGUUAGAUAACtt	GUUAUCUAACAUCUCACCCtg
STX4	syntaxin 4	GGAGUUCUUCCACAAGGUCtt	GACCUUGUGGAAGAACUCCtc
STX5	syntaxin 5	GGAUUCCUACAUCCAGAGUtt	ACUCUGGAUGUAGGAAUCCtg
STXBP1	syntaxin binding protein 1	GGGCAUAACGAUUGUGGAAtt	UUCCACAAUCGUUAUGCCCtc
STXBP2	syntaxin binding protein 2	GGAGCUGAAUAAGUAUUCUtt	AGAAUACUUAUUCAGCUCCtt
STXBP3	syntaxin binding protein 3	GGGUUUAGAGAUUCUUACUtt	AGUAAGAAUCUCUAAACCCtc
VAMP1	vesicle-associated membrane protein 1 (synaptobrevin 1)	GGCAGGAGCAUCACAAUUUtt	AAAUUGUGAUGCUCCUGCCtg
VAMP2	vesicle-associated membrane protein 2 (synaptobrevin 2)	UUUGCGCCAUCAUCCUCAUtt	AUGAGGAUGAUGGCGCAAAtc
VAMP7	vesicle-associated membrane protein 7	CGGUUCAAGAGCACAGACAtt	UGUCUGUGCUCUUGAACCGta
SYT1	synaptotagmin I	GGAAGAUGCAUUUUCUAAGtt	CUUAGAAAAUGCAUCUUCCtt
SYT4	synaptotagmin IV	GGUAUACAUGUUUGAAGAAtt	UUCUUCAAACAUGUAUACCtc

SYT5	synaptotagmin V	GGACUGUCAGAUCCAUACGtt	CGUAUGGAUCUGACAGUCCtc
AP3B2	adaptor-related protein complex 3, beta 2 subunit	GGAUCAGCUGAUAGAAGUCtt	GACUUCUAUCAGCUGAUCCtt
STX7	syntaxin 7	GCUAUUGUAUAAAGGAUGGtt	CCAUCCUUUAUACAAUAGCtt
AP3B1	adaptor-related protein complex 3, beta 1 subunit	GGUUUACACUUGUCAACUUtt	AAGUUGACAAGUGUAAACCtt
VAMP8	vesicle-associated membrane protein 8 (endobrevin)	GGAGAGAAAAAAGAGAGAUtt	AUCUCUCUUUUUUCUCUCCtt
VAMP4	vesicle-associated membrane protein 4	GAAAUCUUUUGGAAGAUGAtt	UCAUCUUCCAAAAGAUUUCtc
STX16	syntaxin 16	CCAGCGAUUUGAUUUAUUGtt	CAAUAAAUCAAAUCGCUGGtg
STX11	syntaxin 11	GCUCAACGUACAAAAGACGtt	CGUCUUUUGUACGUUGAGCtc
STX10	syntaxin 10	CCAUAUGGUCAGCCCAACAtt	UGUUGGGCUGACCAUAUGGtc
GBF1	golgi brefeldin A resistant guanine nucleotide exchange factor 1	GGCUCAGUGAGUUAUUGAGtt	CUCAAUAACUCACUGAGCCtc
SNAP23	synaptosomal-associated protein, 23kDa	GGCUUGGACCAAAUAAAUAtt	UAUUUAUUUGGUCCAAGCCtt
AP1S2	adaptor-related protein complex 1, sigma 2 subunit	CGUAGUGUUCUUGAAGAAAtt	UUUCUUCAAGAACACUACGtg
AP1G2	adaptor-related protein complex 1, gamma 2 subunit	GAUUCACAGACAAGAGGGUtt	ACCCUCUUGUCUGUGAAUCtg
AP1M1	adaptor-related protein complex 1, mu 1 subunit	GGAUGUGAAAUUUUUCCGUtt	ACGGAAAAAUUUCACAUCCtt
CACNA1I	calcium channel, voltage-dependent, T type, alpha 1I subunit	GGAGAACUUCACCAUACAAtt	UUGUAUGGUGAAGUUCUCCtc
CACNA1H	calcium channel, voltage-dependent, T type, alpha 1H subunit	GGUCAUCACACACAAGAUGtt	CAUCUUGUGUGUGAUGACCtt

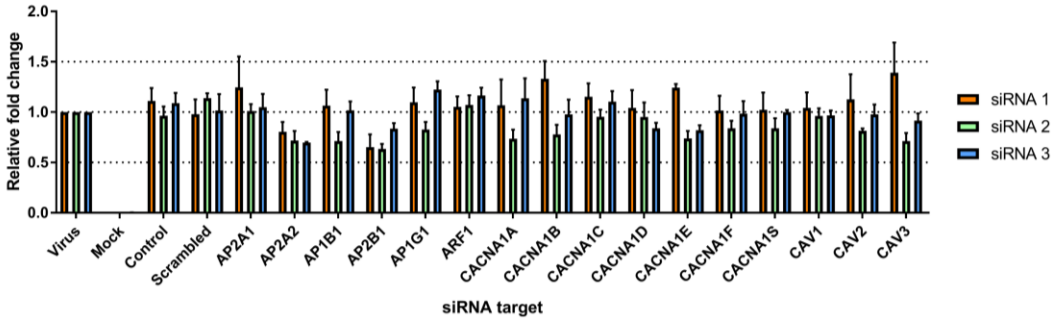
CACNA1G	calcium channel, voltage-dependent, T type, alpha 1G subunit	GGAACAUCACCAAUAAAUCtt	GAUUUAUUGGUGAUGUUCCtg
AP3D1	adaptor-related protein complex 3, delta 1 subunit	GGUCAUUUGUUGCGUUGAAtt	UUCAACGCAACAAAUGACCtc
SYT7	synaptotagmin VII	CCAACUUGUAUCGAGUGGGtt	CCCACUCGAUACAAGUUGGtt
AP4M1	adaptor-related protein complex 4, mu 1 subunit	GAGAUUGUCUGUACUGAUAtt	UAUCAGUACAGACAAUCUCtc
VAPB	VAMP (vesicle-associated membrane protein)-associated protein B and C	GGUUAUGGAAGAAUGUAAGtt	CUUACAUUCUUCCAUAACCtt
COPB2	coatomer protein complex, subunit beta 2 (beta prime)	GCUCUUAAUUAGAACUGGAtt	UCCAGUUCUAAUUAAGAGCtc
VAMP3	vesicle-associated membrane protein 3	GGUUCUGGAAAGAGACCAGtt	CUGGUCUCUUUCCAGAACCtt
SNAP29	synaptosomal-associated protein, 29kDa	GCUUAUCGAACUAGACAUUtt	AAUGUCUAGUUCGAUAAGCtt
STX8	syntaxin 8	GGUGAAAAGGCACCAAAGCtt	GCUUUGGUGCCUUUUCACCtt
RGS6	regulator of G-protein signaling 6	GGUUAAAAAGUCCGUGUAUtt	AUACACGGACUUUUUAACCtt
RICS	Rho GTPase-activating protein	GGCUGAUAAAAAUCCACGAtt	UCGUGGAUUUUUAUCAGCCtt
IQSEC1	IQ motif and Sec7 domain 1	GGGAGUCCAGUACCUCAUCtt	GAUGAGGUACUGGACUCCCtt
AP1M2	adaptor-related protein complex 1, mu 2 subunit	GGAGUACAUCACUCAGCAGtt	CUGCUGAGUGAUGUACUCCtg
HUWE1	HECT, UBA and WWE domain containing 1	GCAAAUCAAGCCUUUGUUAtt	UAACAAAGGCUUGAUUUGCtt
PREB	prolactin regulatory element binding	CCACGAUAAUACCCUGCUUtt	AAGCAGGGUAUUAUCGUGGtt
M6PRBP1	mannose-6-phosphate receptor binding protein 1	GCCUGUAGCUAAUGGACUUtt	AAGUCCAUUAGCUACAGGCtg
STX6	syntaxin 6	GGGAUUGUUUCAGAGAUGGtt	CCAUCUCUGAAACAAUCCCtg
AP3S2	adaptor-related protein complex 3, sigma 2 subunit	GGGUGGAAGUUUGAUUGGUtt	ACCAAUCAAACUUCCACCCtc

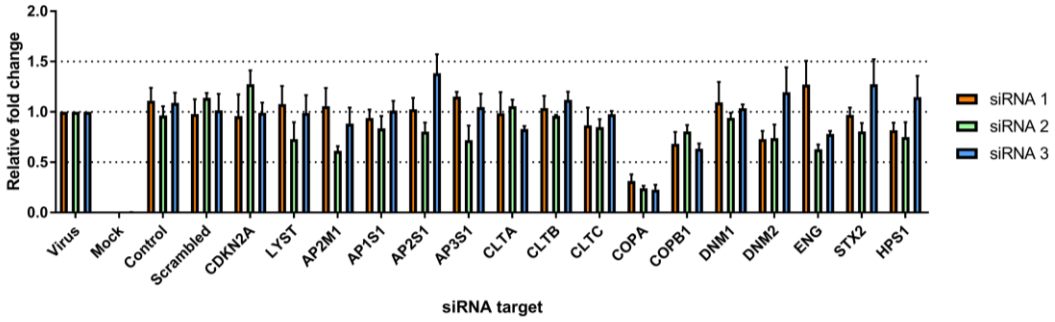
SEC24B	SEC24 related gene family, member B (S. cerevisiae)	GCCCUAUUAUAUACAUCAAtt	UUGAUGUAUAUAAUAGGGCtg
SEC23A	Sec23 homolog A (S. cerevisiae)	GGCAUUUCAAAAAGCUAUGtt	CAUAGCUUUUUGAAAUGCCtt
AP4B1	adaptor-related protein complex 4, beta 1 subunit	GGAAGAACUAUUUGACAUUtt	AAUGUCAAAUAGUUCUUCCtc
EHD1	EH-domain containing 1	GCCUCACCAUUUCUCAAGGtt	CCUUGAGAAAUGGUGAGGCtt
AP3M2	adaptor-related protein complex 3, mu 2 subunit	CCUACAGGAUUCUAAGGUUtt	AACCUUAGAAUCCUGUAGGtt
AP4S1	adaptor-related protein complex 4, sigma 1 subunit	GGUACAGAAAACCUCCACGtt	CGUGGAGGUUUUCUGUACCtt
HPS5	Hermansky-Pudlak syndrome 5	GGAUGUGCUAUUAGUUAAUtt	AUUAACUAAUAGCACAUCCtg
AP1GBP1	AP1 gamma subunit binding protein 1	CCGAUUUUAAAACAGCCGAtt	UCGGCUGUUUUAAAAUCGGtg
COPE	coatomer protein complex, subunit epsilon	GCUGUUCGACGUAAAGAACtt	GUUCUUUACGUCGAACAGCtc
COPZ1	coatomer protein complex, subunit zeta 1	CCUUCCUGAGACUUAAGCAtt	UGCUUAAGUCUCAGGAAGGtt
COPG1	coatomer protein complex, subunit gamma 1	CCAAUCCCUCAAAGUACAUtt	AUGUACUUUGAGGGAUUGGtg
SEC31A	SEC31 homolog A (S. cerevisiae)	GACAAGUUCAGCAUAUUUUtt	AAAAUAUGCUGAACUUGUCtg
SYT11	synaptotagmin XI	GGCAGUCUAUAAUUAACUGtt	CAGUUAAUUAUAGACUGCCtt
AP4E1	adaptor-related protein complex 4, epsilon 1 subunit	CCUUCAUUGUUUGCUAAUAtt	UAUUAGCAAACAAUGAAGGtg
STX12	syntaxin 12	GGGCCCAAAUGCAUAAGUUtt	AACUUAUGCAUUUGGGCCCtt
DNM3	dynamin 3	GGUUACUCGUCUGAAUAUUtt	AAUAUUCAGACGAGUAACCtt
VPS33B	vacuolar protein sorting 33 homolog B (yeast)	GGAUUACCGAUCUCUGAAAtt	UUUCAGAGAUCGGUAAUCCtt
COPG2	coatomer protein complex, subunit gamma 2	GGCUGUCAACAAUAUCAUCtt	GAUGAUAUUGUUGACAGCCtc
AP3M1	adaptor-related protein complex 3, mu 1 subunit	GGUAGUAAAAGGGAACAUUtt	AAUGUUCCCUUUUACUACCtt
STXBP6	syntaxin binding protein 6 (amisyn)	GGACAAAGAAAAAGAUUCCtt	GGAAUCUUUUUCUUUGUCCtc

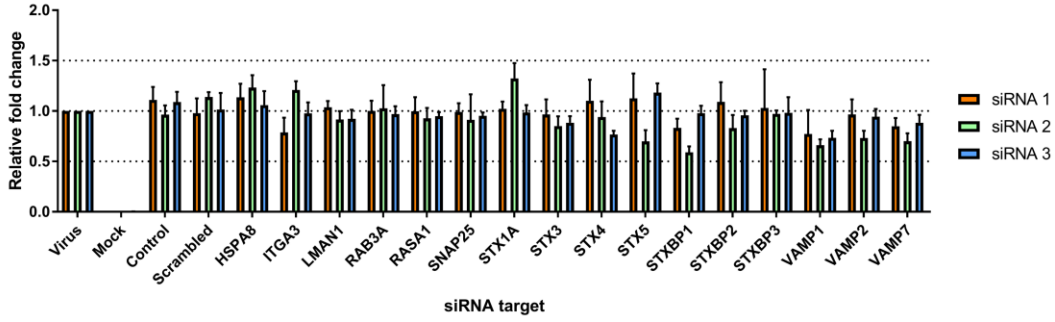
RACGAP1	Rac GTPase activating protein 1	CCACAGACACCAGAUAUUAtt	UAAUAUCUGGUGUCUGUGGtg
EHD4	EH-domain containing 4	GGAUAAAAGCCUUUAUACAtt	UGUAUAAAGGCUUUUAUCCtg
EHD3	EH-domain containing 3	GGUGGUUUUAGUUCAUAGAtt	UCUAUGAACUAAAACCACCtt
EHD2	EH-domain containing 2	GGACAAGUCCAAAUACGACtt	GUCGUAUUUGGACUUGUCCtt
SAR1B	SAR1 homolog B (S. cerevisiae)	GGCUUACUCAGAGAUUUGAtt	UCAAAUCUCUGAGUAAGCCtg
COPZ2	coatomer protein complex, subunit zeta 2	GGAGCAGAUGGUUUUCGAGtt	CUCGAAAACCAUCUGCUCCtt
SYT17	synaptotagmin XVII	GGAUCGUCAUUGGCCAGUAtt	UACUGGCCAAUGACGAUCCtc
STX18	syntaxin 18	GGGAAGAUCAUAGUUAAUAtt	UAUUAACUAUGAUCUUCCCtg
STX17	syntaxin 17	GGAGCUAAGGUUCAUUAGGtt	CCUAAUGAACCUUAGCUCCtg
C4orf16	chromosome 4 open reading frame 16	GCACUUAAGUAUAGCAACAtt	UGUUGCUAUACUUAAGUGCtg
SAR1A	SAR1 gene homolog A (S. cerevisiae)	GGUUAACAUCUGUAGCUUUtt	AAAGCUACAGAUGUUAACCtt
FAM62B	family with sequence similarity 62 (C2 domain containing) member B	GGGUGUUCUAAGGAUACAUtt	AUGUAUCCUUAGAACACCCtt
SYT13	synaptotagmin XIII	GGAGGUGCUUAUUAAAUACtt	GUAUUUAAUAAGCACCUCCtt
TBC1D15	TBC1 domain family, member 15	CGAAAAUGUUGUCAUGACUtt	AGUCAUGACAACAUUUUCGtc
ARHGAP10	Rho GTPase activating protein 10	GGGAACUGUACUACUCGCAtt	UGCGAGUAGUACAGUUCCCtt
HPS6	Hermansky-Pudlak syndrome 6	GCUAUUUUUAAGAUGUGUGtt	CACACAUCUUAAAAAUAGCtc
PTGES2	prostaglandin E synthase 2	GGCAUCUCAUGGAAAAAAAtt	UUUUUUUCCAUGAGAUGCCtg
SYT15	synaptotagmin XV	CCCAGAGCUGUACAAGUUCtt	GAACUUGUACAGCUCUGGGtt
CHR3SYT	family with sequence similarity 62 (C2 domain containing), member C	GGGCUCUUGAAACAGAGUUtt	AACUCUGUUUCAAGAGCCCtt

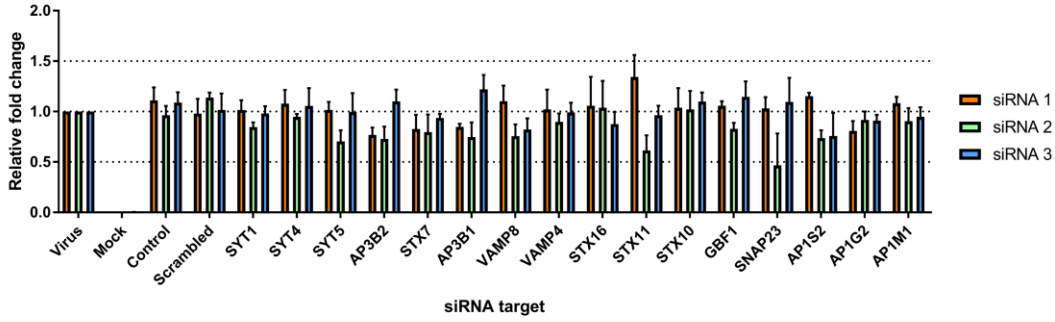
SYT16	synaptotagmin XVI	GGACUCAUGUUCCACAAUGtt	CAUUGUGGAACAUGAGUCCtt
DTNBP1	dystrobrevin binding protein 1	GGUAUGAGGAUACAUGGGCtt	GCCCAUGUAUCCUCAUACCtg
SYT3	synaptotagmin III	GGAAAAAGCCAAACUCAUUtt	AAUGAGUUUGGCUUUUUCCtc
HPS3	Hermansky-Pudlak syndrome 3	CGUCAUUACAAGUAACAACtt	GUUGUUACUUGUAAUGACGtg
HPS4	Hermansky-Pudlak syndrome 4	CCUCAGUUCUAUUACGUUAtt	UAACGUAAUAGAACUGAGGtt
SYT8	synaptotagmin VIII	CGCCGCUGCCUCAGAUCCCtt	GGGAUCUGAGGCAGCGGCGtg
SYT12	synaptotagmin XII	GGAAGAUGAGCAAAAAGAAtt	UUCUUUUUGCUCAUCUUCCtc
STX1B	syntaxin 1B	GGAAGAAAAUCAUGAUCAUtt	AUGAUCAUGAUUUUCUUCCtc
SNX26	sorting nexin 26	GGUUUCUAACUUUGUAACUtt	AGUUACAAAGUUAGAAACCtg
SYT2	synaptotagmin II	GAAAUAUGAGACCAAAGUCtt	GACUUUGGUCUCAUAUUUCtt
AP1S3	adaptor-related protein complex 1, sigma 3 subunit	AUGGUACAUCACUCUCCCUtt	AGGGAGAGUGAUGUACCAUtt
STXBP5	syntaxin binding protein 5 (tomosyn)	GGAGGUUUGUCAUAUGAUAtt	UAUCAUAUGACAAACCUCCtg
SYT9	synaptotagmin IX	GACAACUCAACUUGUCAAAtt	UUUGACAAGUUGAGUUGUCtt
SYT6	synaptotagmin VI	CCCUCGGUUGUGAUUUCAUtt	AUGAAAUCACAACCGAGGGtt
VPS13B	vacuolar protein sorting 13 homolog B (yeast)	GGAUUCAUGUACCAUGGACtt	GUCCAUGGUACAUGAAUCCtc
STXBP4	syntaxin binding protein 4	CCCAGUCUACUUGAAAAGGtt	CCUUUUCAAGUAGACUGGGtg
SYT14	synaptotagmin XIV	GGUCAUGAAAUAGAAAGUUtt	AACUUUCUAUUUCAUGACCtg
SYT10	synaptotagmin X	GCUAUUGAGCCUGCAAUAAtt	UUAUUGCAGGCUCAAUAGCtt
STX19	syntaxin 19	GGAUUUAAGGGAUCUUUUCtt	GAAAAGAUCCCUUAAAUCCtt

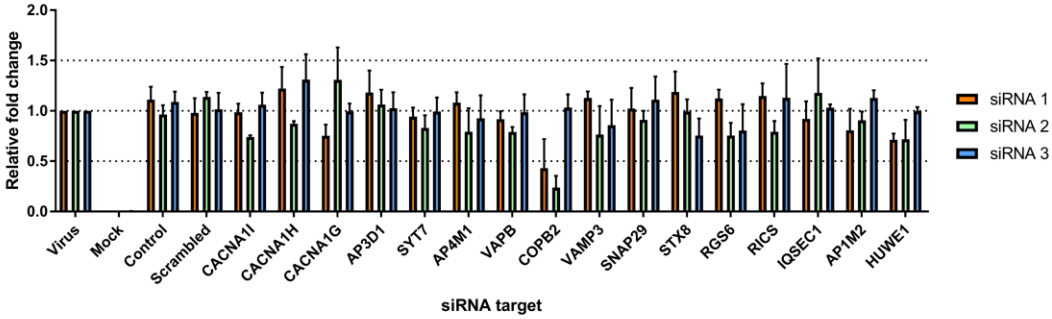
1.4. Full siRNA screen results

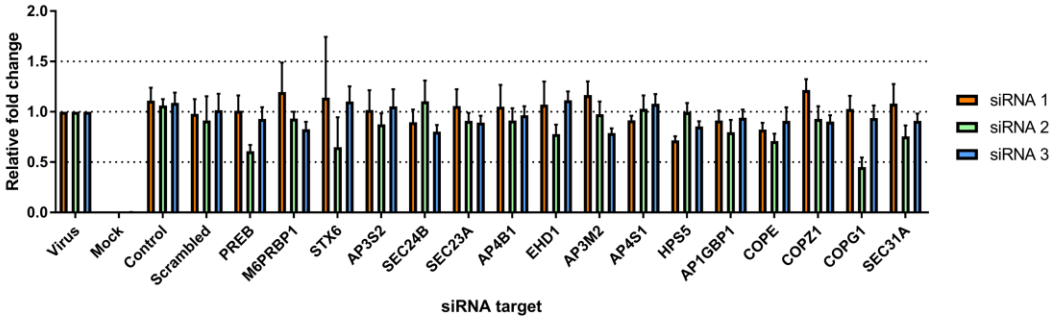


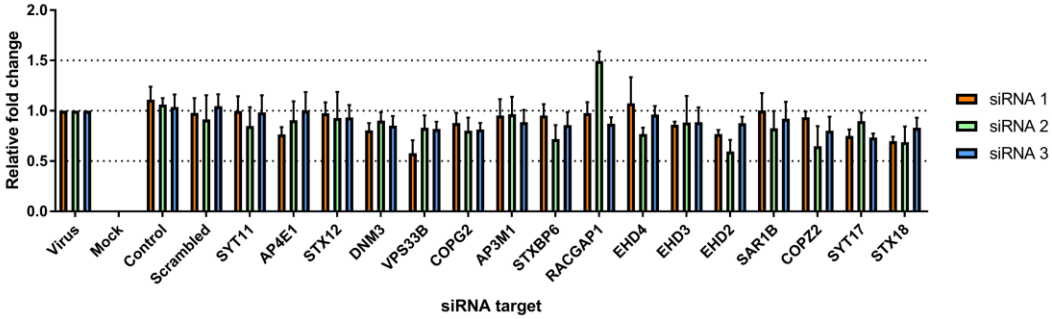


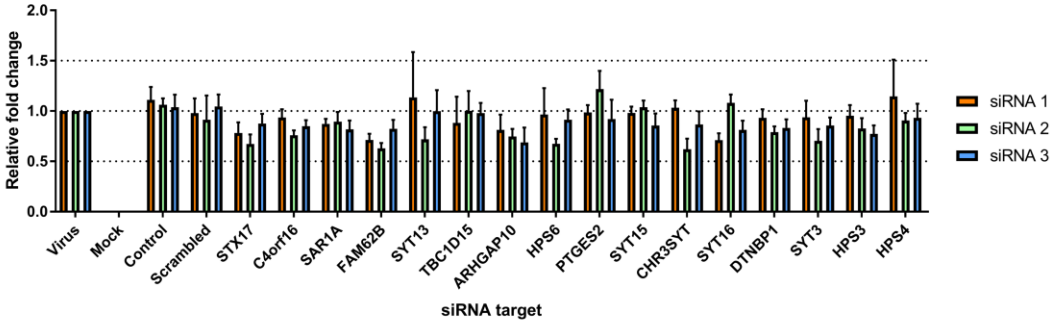


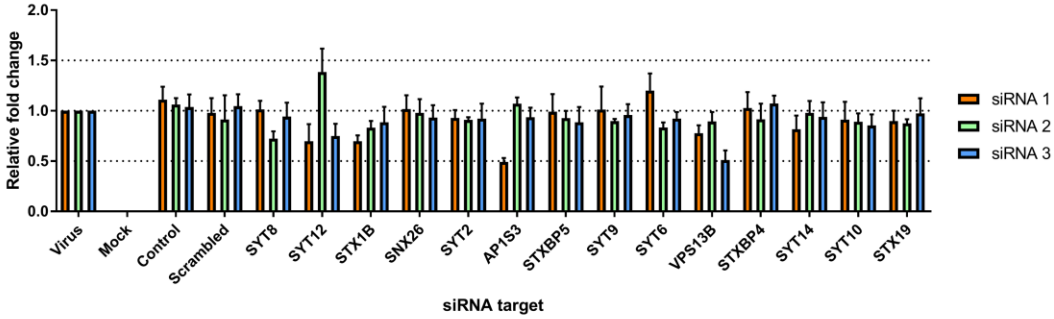



Relative fold change
2.0
1.5
1.0
0.5
0.0
Virus
Mock
Control
Scrambled
CACNA1I
CACNA1H
CACNA1G
AP3D1
SYT7
AP4M1
VAPB
COPB2
VAMP3
SNAP29
STX8
RGS6
RICS
IQSEC1
AP1M2
HUWE1
siRNA target
siRNA 1
siRNA 2
siRNA 3


Relative fold change
2.0
1.5
1.0
0.5
0.0
Virus
Mock
Control
Scrambled
PREB
M6PRBP1
STX6
AP3S2
SEC24B
SEC23A
AP4B1
EHD1
AP3M2
AP4S1
HPS5
AP1GBP1
COPE
COPZ1
COPG1
SEC31A
siRNA target
siRNA 1
siRNA 2
siRNA 3


Relative fold change
2.0
1.5
1.0
0.5
0.0
Virus
Mock
Control
Scrambled
SYT11
AP4E1
STX12
DNM3
VPS33B
COPG2
AP3M1
STXBP6
RACGAP1
EHD4
EHD3
EHD2
SAR1B
COPZ2
SYT17
STX18
siRNA target
siRNA 1
siRNA 2
siRNA 3


Relative fold change
2.0
1.5
1.0
0.5
0.0
Virus
Mock
Control
Scrambled
STX17
C4orf16
SAR1A
FAM62B
SYT13
TBC1D15
ARHGAP10
HPS6
PTGES2
SYT15
CHR3SYT
SYT16
DTNBP1
SYT3
HPS3
HPS4
siRNA target
siRNA 1
siRNA 2
siRNA 3


Relative fold change
2.0
1.5
1.0
0.5
0.0
Virus
Mock
Control
Scrambled
SYT8
SYT12
STX1B
SNX26
SYT2
AP1S3
STXBP5
SYT9
SYT6
VPS13B
STXBP4
SYT14
SYT10
STX19
siRNA target
siRNA 1
siRNA 2
siRNA 3