THE UNIVERSITY *of* EDINBURGH

THE UNIVERSITY of EDINBURGH
The Royal (Dick) School of Veterinary Studies

Identification of potential accessory proteins encoded by the bovine rotavirus strain RF

Ola Diebold

Thesis submitted for the degree of Doctor of Philosophy

The University of Edinburgh

August 2023

Declaration

I declare that this thesis was produced by myself and that the work and illustrations herein are my own, except where otherwise explicitly stated and acknowledged in the text. This work has not been previously submitted for any other degree or professional qualification at this or any other university. No quotation, work or figures from this thesis may be published or cited without proper acknowledgement.

Signed: Date: …31.08.2023…….

Acknowledgements

I have immensely enjoyed every step of my PhD journey. Along the way I have learned many invaluable lessons and this memorable experience would not have been possible without the help and support of the following people whom I will always cherish.

Firstly, I would like to express my sincere gratitude to my primary supervisor Dr Elly Gaunt for giving me the opportunity to do what made me very happy and what I am very passionate about – virology. Thank you for believing in me from the very first time we met and for trusting in my abilities. For providing scientific knowledge and teaching me the skills and techniques of a virologist. For your encouragement to be brave and to question everything. For guiding me and helping to build the foundations of my career journey. I thank you for giving me the chance to grow into the scientist I am today.

I would also like to thank my second supervisor Professor Paul Digard for his knowledge and wisdom that helped to guide this thesis. For challenging and stimulating questions that encouraged me to learn, understand and broaden my knowledge. For sharing your ideas and great suggestions, and for your personal life stories, which I always enjoyed.

I am very grateful to all the members of the Gaunt laboratory, especially Dr Colin Sharp. From PCR to cloning, Colin you have been a great tutor and an inspirational post-doc. Thank you for your training and for all your help throughout my PhD. Thank you to my friend Blanka for all the support and help as well as making the ASV trip so amazing and easy. Thank you to previous and current members of the Digard and Grey laboratories for help and advice, and for small things that made this whole journey an enjoyable ride. In particular, thank you to my friend Hui Min for all the help, support and great food during our lunch times. Thank you to Ploen and Fed, for all their encouragement, kindness and a lot of fun times. A big thanks to Rose for teaching me everything about organoids, for all your help and all the laughs.

Thank you to Dr Samantha Lycett and Dr Barbara Shih for their commitment in teaching me computer science.

Thank you to Dr Bob Dalziel for being a part of my thesis committee, for knowledge sharing and guidance.

A very special thanks to Dr Jo Stevens who really stepped up and provided necessary support during a very hard time in my life, for being understanding as a professional but also as a mother.

My deepest appreciation to Dr Ulrich Desselberger for taking his invaluable time to read and thoroughly comment on this thesis.

I want to thank Dr Abigail Diack, who was crazy enough to hire me as a research assistant before my PhD, giving me the chance to prove myself. I have learned a lot during that time, thank you for teaching me and encouraging me to spread my wings.

A very big thank you to the Microbiology Society for the career development opportunities and for financial support to attend national and international conferences.

A huge thanks to my friend Lizzy for always being there for me even when things got tough, for your support, advice and awesome gossip dinners.

I would also like to extend my thanks to the following people who played a big part in my journey of completion of this thesis:

Rute - was I lucky to have met you! You have been like a sister to me. From the moment we met I think two stars collided somewhere. Thank you for all your help and support, not just with science but for late chats and food. I am still hoping our Bora Bora idea will happen someday.

An Peter und Renate, vielen Dank für eure Unterstützung, die es mir ermöglicht hat, die ersten Schritte auf meiner beruflichen Reise zu gehen. Ich bin sehr froh, dass ihr ein Teil meines Lebens seid und ich werde euch auf ewig dankbar sein.

Дорогая моя мама, я хочу от всей души поблагодарить тебя за то, что ты дала мне силы, научила меня настойчивости и стремлению, помогала и поддерживала в самое трудное время говорила: “что все будет хорошо”.

To my daughter, Amy, I want to thank you for showing me the importance of family. Family provides the foundation and stability from which you can build your future. It gives you confidence to flourish, grow and develop. Family’s love and support is irreplaceable and I will continue to always be your family, to love you and support you in anything you chose to do.

And finally, to my husband, Marius, without whom this whole journey just would not have been possible. There will never be enough words to express my gratitude. Thank you for everything, for supporting me immensely not just during the PhD but in the last 10 years. For encouraging me to pursue my dreams, for small things like keeping my clothes warm and making dinner when I came home very late (which happened very often). For your patience and understanding, for standing by my side, for teaching me and for believing in me that I could do anything. I want to thank you for being a wonderful husband and a stepfather, for your continuing unfailing support and your unconditional love, you kept me going through it all. For that, I will and continue to be eternally grateful.

Abstract

Rotaviruses (RVs) are a major cause of acute gastroenteritis in infants and young children worldwide, accounting for ~215,000 deaths annually, mostly in developing countries. Likewise, RV-associated enteritis in young calves and piglets has a significant economic impact on livestock production as a result of the high morbidity and mortality caused.

The RV genome comprises 11 segments of dsRNA encoding structural and non-structural proteins that are essential for virus replication. However, the coding capacity of individual RV genes varies between strains with evidence demonstrating expression of accessory proteins from alternative open reading frames. Work in this thesis aimed to identify accessory proteins encoded by the bovine RV strain RF and investigate their role in pathogenicity. Additionally, lack of a plasmid-only reverse genetics system to artificially engineer and mutate an infectious RV has limited studies of its genome. Thus, a plasmid-only reverse genetics system was established for the RF strain that was used in proof-of-principle experiments to tag viral protein NSP3 with SARS-CoV-2 spike peptides.

In vitro transcription and translation assays led to the identification of an unknown polypeptide of ~100 kDa in the RF gene segment 1 (VP1). Bioinformatic analyses identified highly conserved AUG codons corresponding to putative alternative translation initiation sites. Specifically, in the cell free system, mutation of AUG12 (M130) and AUG14 (M146) decreased the expression of the unknown polypeptide, suggesting that both AUGs were utilised for downstream translation initiation. In contrast, in cells transfected with VP1-EGFP tagged constructs, only AUG12 (M130) was utilised for the expression of the unknown polypeptide. Thus, a novel VP1 isoform named VP1-N129 was identified.

In the context of RV infection, only the leucine substitution at M130 residue (M130L) affected virus production whereas substitution of methionine to valine (M130V) had no effect on the viral rescue nor on replication. Neither M146L nor M146V mutation affected the viral rescue or replication kinetics. To assess the impact of M130 and M146 mutations on the viral polymerase activity, attempts were made to develop a ‘mini-genome’ assay. Preliminary results from this assay showed no effect on the polymerase function, although further experiments are needed to improve the sensitivity of the assay.

In conclusion, a novel N-terminally truncated isoform of VP1 has been identified but further studies are required to confirm its role as a potential accessory protein in RV infection.

Lay Summary

Rotaviruses (RVs) are a leading cause of diarrheal-associated deaths in infants and young children, particularly in developing countries. RVs also affect livestock, impacting economic productivity due to illness and high mortality in young dairy calves and newborn piglets.

The genetic material of RVs consists of 11 segments that encode proteins crucial for virus replication. Some RV segments have the ability to make more than one protein from different start sites. These proteins, called accessory proteins, are not essential but have an important role in viral pathogenesis. ‘Reverse genetics’ is a tool used to investigate the biological effects of introducing mutations in the genes of viruses, which can be used to find these accessory proteins. This thesis aimed to develop the reverse genetics system for the bovine RV strain RF, which was then used to identify and characterise the accessory proteins with a focus on understanding their role in causing disease.

Using computer analyses and various experiments, frequently observed alternative start sites were identified in the RF gene segment 1 at amino acid positions 130 and 146. Specifically, mutating the start site at amino acid position 130 affected the expression of a truncated form of the protein from segment 1 in cells. Thus, a novel version of a viral protein expressed from RF gene segment 1 called VP1-N129 was discovered. Using reverse genetics, different mutations were introduced at amino acid positions 130 and 146 that affected virus production. Segment 1 codes for a viral polymerase which is an enzyme that makes copies of the viral genetic material. An experiment was designed to test whether the amino acid changes made using reverse genetics affected the function of the viral polymerase. Preliminary results from this experiment showed no impact of segment 1 mutations on the viral polymerase activity, thus requiring further investigation.

In conclusion, an alternative start site for a potential accessory protein was identified in RF segment 1. The role of this protein in viral pathogenicity and its mechanism of expression is yet to be determined.

Table of Contents

List of Figures

List of Tables

Abbreviations

Abbreviation	Full name
3D	Three dimensional
ACE2	Angiotensin converting enzyme 2
ANOVA	Analysis of Variance
ASFV	African swine fever virus
BHK-21	Baby Hamster Kidney-21 cells
bp	Base pair
BSA	Bovine Serum Albumin
BSR-T7	BHK-21 clone expressing phage T7 ribonucleic acid polymerase
Ca^{2+}	Calcium ion
CCMB	Calcium manganese-based
cDNA	Complementary Deoxyribonucleic Acid
CoV	Coronavirus
COVID-19	Coronavirus disease 2019
CPE	Cytopathic effect
CSU	Central Services Unit
DAPI	4′,6-Diamidino-2-Phenylindole
DIP	Defective Interfering Particle
DLP	Double-layered particle
DMEM	Dulbecco's Modified Eagle Medium
DMSO	Dimethyl Sulfoxide
DNA	Deoxyribonucleic Acid
dNTP	Deoxyribonucleotide triphosphate
DOS	Duplex-open state
DRiP	Defective Ribosomal Product
dsDNA	Double-stranded DNA
dsRNA	Double-stranded RNA
DVG	Defective viral genomes
EDTA	Ethylenediaminetetraacetic Acid
EGFP	Enhanced Green Fluorescent Protein
EIF2	Eukaryotic Initiation Factor 2
eIF4GI	Eukaryotic translation initiation factor 4GI
ER	Endoplasmic Reticulum
FAST	Fusion-Associated Small Transmembrane protein

FBS	Fetal Bovine Serum
FMDV	Foot and mouth disease virus
FUT2	Functioning Fucosyltransferase 2 enzyme
GFP	Green Fluorescent Protein
GMEM	Glasgow's Minimum Essential Medium
GTAse	Guanylyltransferase
HA	Haemagglutinin
HBGA	Histo-Blood Group Antigens
HDV	Hepatitis Delta Virus
HEK293T	Human Embryonic Kidney 293 cells carrying SV40 T antigen
HIT	Histidine triad
HLH	Helix-loop-helix
hpi	Hours post infection
hsc70	Human heat shock cognate protein 70
IAV	Influenza A Virus
IF	Immunofluorescence
IFN	Interferon
IRES	Internal Ribosome Entry Site
IRF3	Interferon regulatory factor 3
IVT	*In Vitro* transcription/translation
kDa	Kilo Dalton
LAV	Live-Attenuated Vaccine
LB	Luria-Bertani
MA104	Monkey African Green kidney 104 cells
MCS	Multiple Cloning Site
MEM	Minimum Essential Medium
MERS-CoV	Middle East Respiratory Syndrome Coronavirus
MOI	Multiplicity of infection
mRNA	Messenger Ribonucleic Acid
MTase	Methyltransferase
NBF	Neutral Buffered Formalin
NCBI	National Center for Biotechnology Information
NCDV	Nebraska calf diarrhoea virus
NSP	Non-Structural Protein
ORF	Open Reading Frame
P2A	Porcine teschovirus-1 2A peptide
PABP	Poly(A)-binding protein

PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
PFU	Plaque Forming Units
PKR	Protein Kinase R
RBD	Receptor Binding Domain
RdRp	RNA-dependent RNA-polymerase
RF	Rotavirus France
RNA	Ribonucleic Acid
rRF	Rescued rotavirus France
RRV	Rhesus rotavirus
rSA11	Rescued Simian Agent 11
RTPase	RNA triphosphatase
RT-PCR	Reverse transcription polymerase chain reaction
RV	Rotavirus
SA11	Simian Agent 11
SARS-CoV	Severe Acute Respiratory Syndrome Coronavirus
SARS-CoV-2	Severe Acute Respiratory Syndrome Coronavirus 2
SDS-PAGE	Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis
SFM	Serum Free Medium
SOB	Super Optimal Broth
ssRNA	Single-Stranded RNA
T2A	Thosea asigna virus 2A peptide
T7P	T7 Promoter
T7T	T7 Terminator
TAE	Tris-acetate EDTA
TBE	Tris-borate-EDTA
TBS-T	Tris-buffered saline with Tween20
TEMED	Tetramethylethylenediamine
TES	Transcript-elongated state
TLP	Triple-Layered Particle
TURBS	Termination upstream ribosome-binding site
UTR	Untranslated Region
VGM	Virus Growth Medium
VP	Viral Protein
VV	Vaccinia virus
WHO	World Health Organisation
WT	Wild-Type

Chapter 1: Introduction

1.1 General introduction to Rotaviruses

Rotaviruses (RVs) are the leading cause of acute gastroenteritis in infants and young children under the age of five accounting for ~128,500 deaths annually, particularly in low-income countries [1]. Human RVs were first discovered in 1973 when Bishop *et al.,* (1973) identified virus particles in electron micrographs of duodenal biopsy samples from children with acute diarrheal disease [2]. Several animal viruses described during the previous 10 years were later found to be RVs based on the shared characteristics with and similar morphology to human RVs [3-5]. Accordingly, RV-associated enteritis in young calves and piglets has a significant economic impact on livestock production as a result of the high morbidity and mortality caused as well as the high cost of treatment for infected animals [6-9].

1.2 Classification

RVs comprise the genus *Rotavirus* within the subfamily *Sedoreovirinae* of the *Sedoreoviridae* family in the order of *Reovirales* [10]. There are ten distinct species of RVs (A–J) that have been classified according to the antigenic variability of the viral protein 6 (VP6) [11-13]. Additionally, novel species K and L, identified in common shrews, have been recently defined, mainly based on their genome sequence identities [14, 15]. RV species A–C have been found in animals and humans whereas species D–J have only been identified in animals. Species A are extensively studied and individual RV strains are often classified by a binary serotype system GxP[x] analogous to the HxNx designation of influenza virus strains [16]. The two outer capsid proteins VP7 (G antigen - glycosylated) and VP4 (P antigen protease–sensitive) are used to determine the strain serotype based sequence differences and reactivity with neutralising antibodies [10]. To date, 41 G-types and 57 P-types have been identified [17]. In 2008, a comprehensive nucleotide sequence-based classification system was established for species A RV assigning a specific genotype to each of the 11 gene segments [18]. According to the updated guidelines from the Rotavirus Classification Working Group, nomenclature for individual strains was assigned to be: RV species/species of origin/country of identification/common name/year of identification/G- and P-type [19]. Thus, the prototype bovine strain RF is designated RVA/Cow/France/RF/1975/G6P6 [20]. RV species A are the most predominant accounting for over 90% of infections in humans and animals, and thus will be the focus of this thesis [21].

1.3 Virion structure

The mature infectious RV virion is a non-enveloped triple-layered particle (TLP) that exhibits a T = 13 icosahedral symmetry and contains 11 segments of double-stranded RNA (dsRNA) as its genome (Figure 1.1) [22, 23]. Based on the TLP morphology detailed by electron microscopy, Flewett *et al.,* (1974) proposed the name 'rotavirus' from the Latin *'rota'* meaning *'wheel'* [24]. The TLP is ~100 nm in diameter and comprises the outer layer VP7 embedded with VP4 spikes, the intermediate VP6 layer and the core shell made of VP2 [23]. The TLP structure also includes 132 large aqueous channels grouped into three types (type I, type II and type III) based on their locations in the T = 13 icosahedral lattice [25, 26]. The core shell has 120 VP2 molecules arranged as asymmetrical dimers organised in a unique T = 1 icosahedral symmetry encompassing the dsRNA genome [23, 27]. One isoform of VP2, VP2-A, converges around the fivefold axes forming a star-shape complex, while the VP2-B isoform interlocks between the adjacent VP2-A [27]. Inside the core shell, the polymerase complex, composed of the RNA-dependent RNA polymerase (RdRp) (VP1) and the RNA capping enzyme (VP3), is anchored at the fivefold axes through simultaneous interactions with multiple subdomains of VP2-A and VP2-B [28, 29]. The current view is that each genome segment interacts with one specific polymerase complex [27, 30]. This capsid architecture found in RVs and orbiviruses [31] contrasts with the turret structure found in the reoviruses [32], aquareovirus [33] and cypovirus [34] of the *Sedoreoviridae* family, where the capping enzyme is not located within the core but protrudes at the fivefold axis.

The core shell is surrounded by an intermediate layer of 260 trimers of VP6 forming the transcriptionally active, non-infectious double-layered particle (DLP) [10]. Binding of VP6 stabilises the core allowing capped transcripts to exit through the type I channels at the fivefold axis [35]. The DLP in turn is covered by 260 trimers of glycoprotein VP7, and 60 spikes of VP4 trimers, forming a mature TLP [36]. The VP4 (88 kDa) on the TLP is proteolytically cleaved into VP8* (28 kDa) and VP5* (60 kDa) subdomains, which represent the amino- and carboxyl-terminal regions of the protein respectively [37, 38]. Proteolytic cleavage is required for efficient RV infectivity as a series of conformational changes in VP4 facilitate entry of RV into cells [39, 40]. VP8* mediates initial binding of the virus to cells whilst VP5* plays a role in direct cell membrane penetration [38]. The lectin domains of the VP8* fragment form the tips of the spikes that cover the hydrophobic loop of the VP5* β barrel domains [36, 38]. The C-terminal of the VP5* forms the foot domain, which is buried in the particle and anchors the spike through its interaction with both the VP6 and the VP7 layers [23]. VP8* and VP5* remain non-covalently associated with each other on the mature virion [38].

The outer layer of the capsid shields the virus from the harsh environment of the digestive system [10]. Furthermore, VP7 and VP4 are targets of both homotypic and cross-reactive heterotypic neutralising antibodies [36]. Specifically, antibodies that target the VP8* subdomain block RV attachment, thus the VP8* subunit protein has been used in several vaccine platforms [41-44].

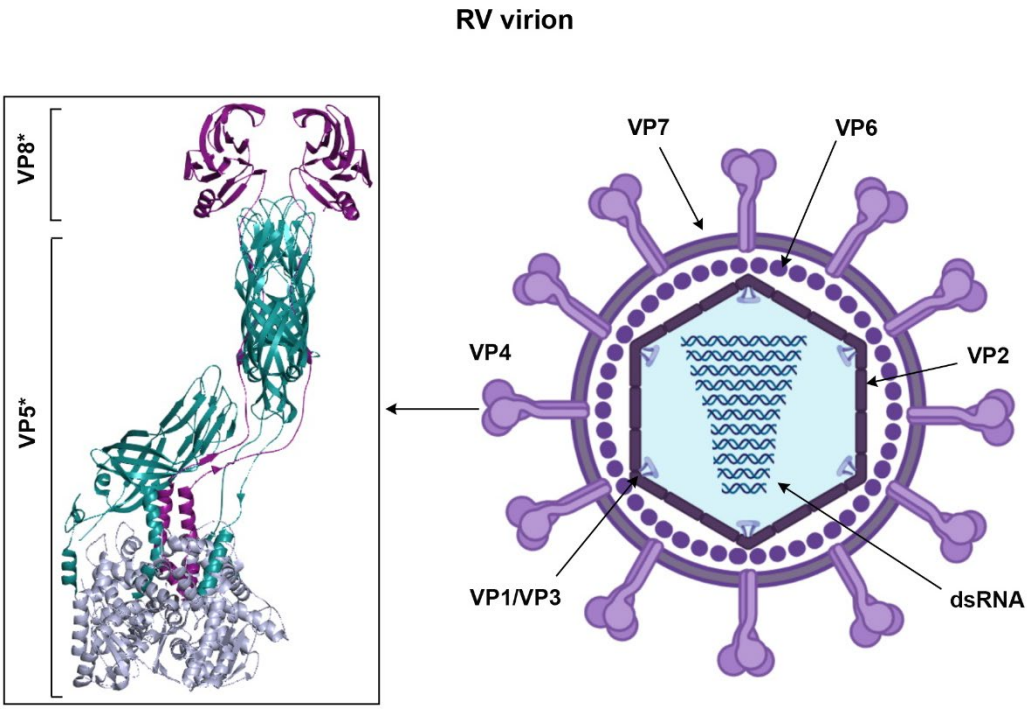


Figure 1.1 Schematic of RV virion.
The virion is composed of three concentric layers: VP4 spikes and outer layer VP7, intermediate layer VP6 and inner layer VP2. A transcription complex of VP1 and VP3 is located inside the VP2 layer at the fivefold axes of symmetry. The viral dsRNA genome comprises 11 segments of different sizes. VP4 is proteolytically cleaved into an N-terminal fragment, VP8* (28 kDa), and a C-terminal fragment, VP5* (60 kDa). Ribbon representation of VP4 is shown in box. VP8* domain – magenta, VP5* β-barrel domains – teal and VP5* foot domain – light blue. Schematic not to scale. *(Adapted from Herrmann et al., 2020 [38]; created with BioRender.com; crystal structure annotated in PyMOL [330] using PDB ID: 6WXE).*

1.4 Genome structure and organisation

The RV genome is ~18,500 base pairs (bp) in size and encodes six structural (VP1 - VP4, VP6 and VP7) and depending on the strain, five or six non-structural proteins (NSP1 - NSP5 ± NSP6) [10]. All of the gene segments are considered monocistronic except segment 11 which encodes two proteins, NSP5 and NSP6 [45]. The gene segments range in size between 667 (segment 11) and 3302 (segment 1) bp in length [46]. Species A RV transcripts are capped, lack a 3' poly(A) tail, are A+U rich (60–70%) and contain short conserved 5' (GGC) and 3' (UGUGACC) terminal sequences that are part of the untranslated regions (UTRs) [10, 47]. The lengths of the 5' and 3' UTRs vary for different segments ranging between 9 – 49 nucleotides at the 5' end and 17 – 182 nt at the 3' end (Figure 1.2) [46]. The 3' conserved consensus sequences contain important *cis*-acting signals for gene expression, packaging and replication [48, 49]. *In silico* analyses revealed conserved secondary structures in the positive strand RNAs ((+)RNAs), including long-range interactions at the 5' and 3' terminal regions of all segments [50]. Complementary base pairing of 5' and 3' regions of each segment are predicted to facilitate RNA circularisation by forming panhandle structures with the 3'-GACC conserved terminal sequence extending as a single-stranded tail [49-51]. The interaction of this extended RNA with different RV proteins is thought to regulate the dsRNA synthesis [51]. Similar features of the RNA termini (5' cap and 3' conserved sequences) are found in the genomes of reoviruses and orbiviruses of the *Sedoreoviridae* family as well as in other virus families such as *Orthomyxoviridae* and *Bunyaviridae* [52-55].

A summary of the major features of the RV genes are highlighted in Figure 1.2. The RV genome structure and organisation based on bovine RV RF prototype is shown in Figure 1.3 with the functions of their respective proteins summarised in Table 1.

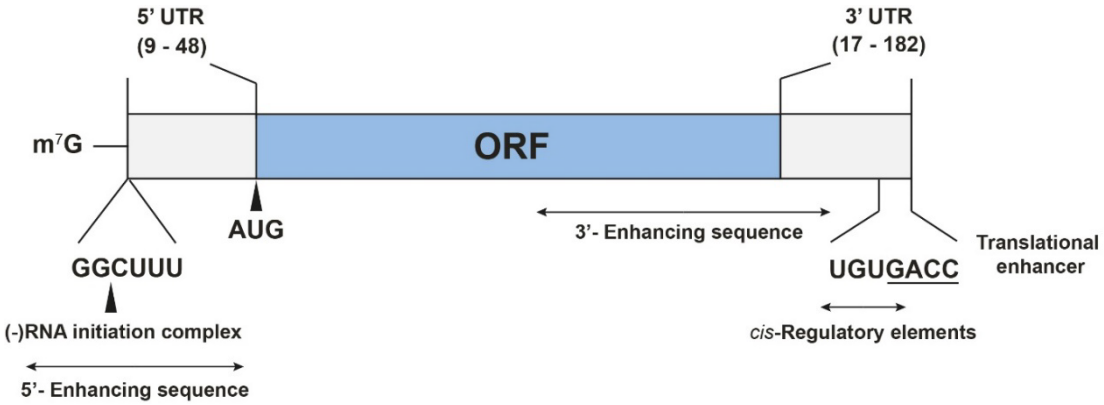


Figure 1.2 Schematic of general features of RV gene structure.
All 11 RV genes are capped at the 5′ end (represented by m^7G) and lack poly(A) tail. The main open reading frame (ORF) is in blue and the AUG codon (black arrowhead) denotes the canonical start site. The 5′- and 3′-UTRs of the mRNA vary in length (nucleotides in brackets) and are shown in white. The RdRp interacts with the 5'-end conserved sequence to initiate negative strand RNA synthesis designated as (-)RNA initiation complex (black arrowhead). Enhancing sequences (long arrows) present at 5'- and 3'-end base-pair and form panhandle structures from which the 3' consensus sequence extends as a single stranded tail. The 3' *cis*-regulatory elements required for RNA replication and transcription are shown (short arrow). NSP3 interacts with the 3′-end (5′-GACC-3') translational enhancer sequence (underlined) and acts as a poly(A) tail. Schematic not to scale. *(Adapted from Crawford et al., 2023 [10])*.

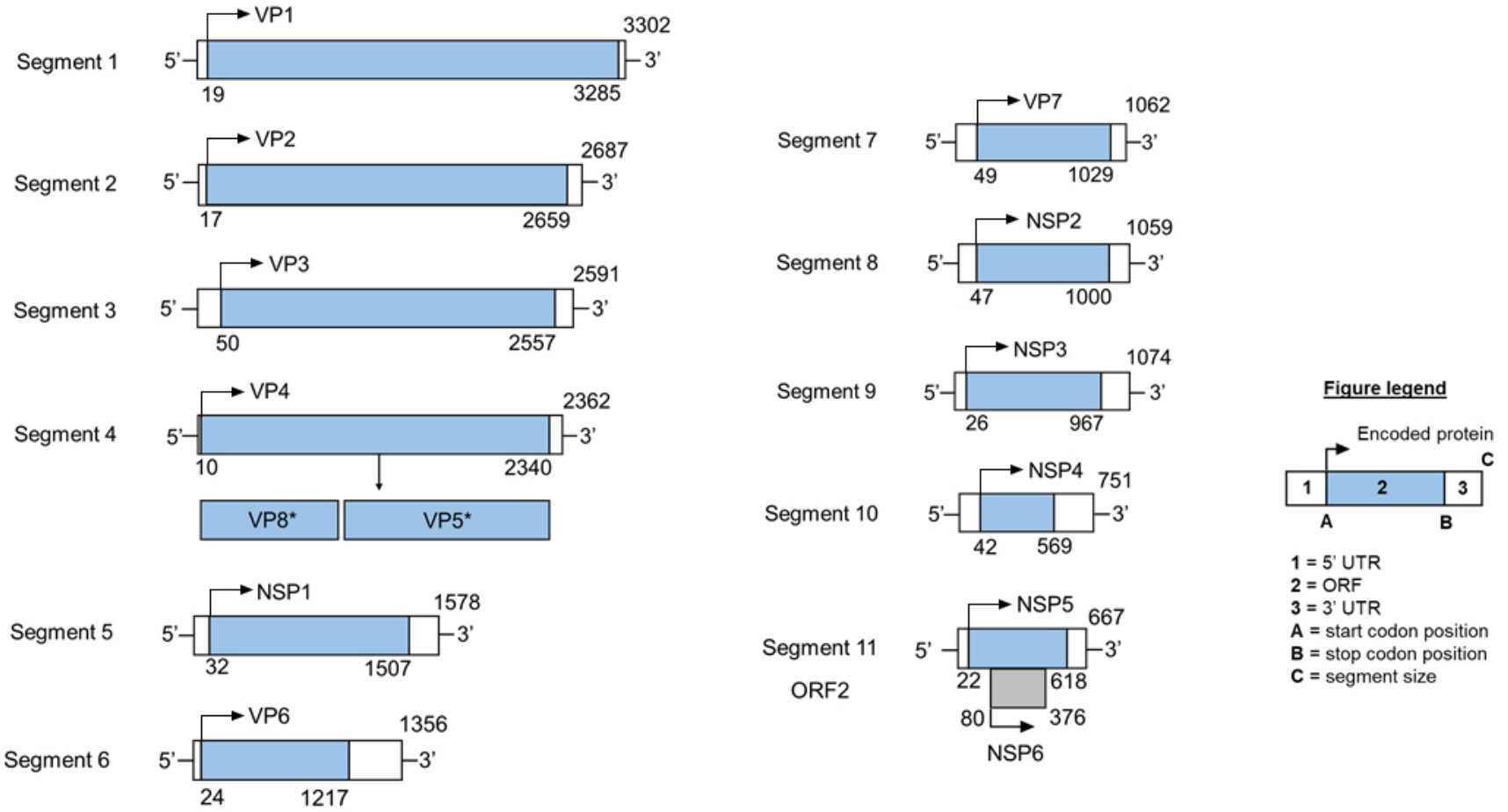


Figure 1.3 RV gene products.
RNA segments illustrated in the positive sense with their encoded proteins (size shown in nucleotides). Coding regions from the 11 segments are represented by boxes with colours defining different reading frames (blue - primary products, frame 1; grey - secondary products, frame 2). Information is based on bovine RV strain RF (strain RVA/Cow/France/RF/1975/G6P6[1]) [20]. The gene segments range in size from 667 (segment 11) to 3,302 (segment 1). Schematic not to scale. *(Adapted from Crawford et al., 2023 [10]).*

Table 1 RV gene products, localization and respective functions
Information is based on bovine RV strain RF (RVA/Cow/France/RF/1975/G6P6[1]) [20] downloaded from NCBI and UniProtKB. *(Adapted from Crawford et al., 2023 [10] and Richards et al., 2013 [263]).*

Gene segment	Accession #	UniProtKB ID	Encoded protein	Segment size (bp)	Gene size (bp)	Gene size (aa)	Mw (kDa)	5' UTR (bp)	3' UTR (bp)	Location in virion and cell	Functions
1	KF729687	P17468	VP1	3302	3267	1088	125	18	17	Core - attached inside the inner capsid	RNA-directed RNA polymerase (RdRp); interacts with VP3; RNA binding
2	KF729642	P12472	VP2	2687	2643	880	102.5	16	28	Core - inner capsid	RNA binding; protects dsRNA from degradation; required for RdRp activity
3	KF729645	Q8BB04	VP3	2591	2508	835	98	49	34	Core - attached inside the inner capsid	mRNA capping; affinity for ssRNA; mRNA methyltransferase and nucleotidyltransferase activity
4	KF729650	Q96802	VP4	2362	2331	776	86.8	9	22	Outer capsid; host ER and cell membrane	Cleaved into VP5 and VP8; virulence and host range restriction; host-virus interaction; attachment and entry
5	KF729656	P12475	NSP1	1578	1476	491	58.7	31	71	Host cytoskeleton	IFN antagonist; RNA binding
6	KF729658	P04509	VP6	1356	1194	397	44.8	23	139	Intermediate capsid	Required for transcriptional activity of DLPs
7	KF729664	W0G1I0	VP7	1062	981	326	37.1	48	33	Outer capsid; host ER lumen	Defines the G serotype of species A RVs; calcium dependent; attachment and entry
8	KF729667	H9N1A1	NSP2	1059	954	317	36.7	46	59	Host cytoplasm in viroplasms	Essential for viroplasm formation; ssRNA binding; ATP binding; RTPase and NTPase activity
9	KF729675	Q86504	NSP3	1074	942	313	36.4	25	107	Host cytoplasm	Initiates capped mRNA translation; binds to viral (+)ssRNA, host eIF4G and Hsp90; displaces PABP
10	KF729677	Q6YLV5	NSP4	751	528	175	20.4	41	182	Host rough ER membrane	Viroporin; participates in RNA replication; modulates intracellular calcium levels; enterotoxin; interacts with viroplasms and host autophagy pathway

11	KF729684	Q9E8F2	NSP5	667	597	198	21.6	21	49	Host cytoplasm in viroplasms	Essential for viroplasm formation; interacts with NSP2; participates in RNA replication
ORF 2 of Seg11	KF729684	Q9E8F1	NSP6	667	297	98	11.5	79	291	Host cytoplasm in viroplasms and host mitochondrion	Interacts with NSP2 and NSP5

1.5 Viral accessory proteins

Viral accessory proteins are defined as non-essential proteins encoded by some virus strains [56]. Typically, accessory proteins play an important role in viral pathogenicity or spread as demonstrated by the accessory proteins of many RNA viruses, such as the PA-X protein of influenza A virus (IAV) or the XP protein of astrovirus [57, 58]. Expression of viral accessory proteins typically involves non-canonical gene expression mechanisms that diversify the coding repertoire of the virus genome [59].

To date, studies have documented the expression of potential accessory proteins in some RV species.

The NSP6 protein (92 aa; 12 kDa) of RV is expressed from an overlapping open reading frame (ORF) in segment 11 (possibly by ribosomal leaky scanning [60]) is the most intensively studied accessory protein, yet its precise function has not been fully characterized [45]. NSP6 has been proposed to localize to viroplasms through interaction with NSP5, and to mitochondria via its conserved N-terminal signal sequence [61-64]. Komoto *et al.*, (2017) demonstrated that NSP6 is not essential for viral replication in cell culture, but may be important for viral replication *in vivo* [65]. This is supported by sequence analysis of segment 11 from different RV strains which revealed that some culture-adapted strains lack the NSP6 protein [62]. However, a recent study showed that NSP6 was not required for replication and pathogenicity *in vivo* and a recombinant NSP6-deficient virus was able to induce diarrhoea in suckling mice [66]. Two human virus isolates Mc323 and 512-C do not have the NSP6 start codon, while the lapine Alabama strain and the porcine OSU strain both possess a truncated NSP6 ORF [67-70]. ORF coding for full-length NSP6 is also missing in all species C RVs [61]. Interestingly, the NSP4-encoding segment of species D RV strain Ch-49 isolated from a chicken possesses two partially overlapping ORFs, analogous to RV species A NSP5 [71]. In infected cells, the NSP6 protein is expressed at low levels and has a high turnover rate with the majority of protein degraded within 2 hr of synthesis [62]. Nonetheless, the presence of the NSP6 ORF is highly conserved among RV species A isolates from natural infections suggesting that further studies are required to confirm its role in pathogenicity [65, 66].

Of the two in-phase initiation codons in gene segment 7 (VP7) of RV simian agent 11 (SA11) strain, only the second codon has an optimal Kozak context for translation initiation [72, 73]. *In vitro* translation of segment 7 mRNA using wheat germ extracts produced two distinct polypeptide species of different molecular weights (37 kDa and 35.5 kDa) [72]. Both products were also detected in infected cells and in purified virus particles of other RV strains such as canine (CU-1), bovine (NCDV) and rhesus (RRV) [72]. Further digestion with Endo H

enzyme showed that both products were glycosylated but that only one product contained a cleavable signal sequence [72]. Whether the two products are functionally distinct remains unknown.

Studies of species B RVs showed that gene segment 5 (NSP1) contains two overlapping ORFs and was predicted to encode putative proteins NSP1-1 and NSP1-2 [74, 75]. Diller *et al.*, (2019) demonstrated that NSP1-1 expression mediated fusion and syncytium formation of cultured cells, thus enhancing viral replication [75]. Based on the nucleotide sequences flanking the initiation codons, translation of NSP1-2 was predicted to be more efficient than NSP1-1, although its function remains unknown [74, 75].

Segment 9 (NSP3) of the porcine species C RV (Cowden strain) has been shown to produce three proteins [76, 77]. Expression of the full-length NSP3 protein (45 kDa) *in vitro* resulted in two proteolytically cleaved polypeptide products with molecular weight of 35 kDa and 8 kDa [76]. Coupled *in vitro* transcription and translation reactions demonstrated that cleavage of NSP3 occurred cotranslationally with its synthesis or immediately after, since the two products were detected prior to detection of the full-length NSP3 [76]. Functional analysis revealed that the smallest polypeptide (8 kDa) specifically binds to dsRNA while the 35 kDa polypeptide might be involved in inhibition of interferon (IFN) induced protein kinase R (PKR) [76].

Lastly, gene segment 7, coding for NSP3, of species A RVs contains two in-frame AUG codons 25 and 32 nucleotides from the 5’ end, with second AUG having an optimal Kozak motif [47, 78]. The first AUG was shown to be mutated in an avian RV and a human isolate M of species A [79, 80], whereas a functional protein was produced when translation started at the second AUG [78, 81]. Additionally, conserved RNA structural elements such as long-range interactions have been detected in some segments of species A RVs (VP4, VP6, NSP2, NSP5 and NSP3), raising the possibility that translation initiation from downstream AUG in NSP3 could be mediated through an internal ribosomal entry site (IRES) [50].

Overall, evidence demonstrates the potential for alternative translation initiation in some segments of the RV genome although the mechanisms of expression as well as the exact functions of identified polypeptide species have not yet been defined. Many RNA viruses employ non-canonical translation initiation mechanisms to regulate gene expression in order to maximise their coding capacity, making RVs no exception.

1.6 Replication cycle

RVs primarily infect mature enterocytes in the middle and on the tip of the intestinal epithelium and replicate exclusively in the cytoplasm [10]. At a high multiplicity of infection (MOI), the complete replication cycle in cell culture takes between 10 to 12 hr at 37°C [82]. The RV replication cycle can be summarised by the following steps:

- Attachment and cell entry
- Positive-strand RNA ((+)RNA) synthesis
- Translation of viral proteins
- Viroplasm formation, packaging and dsRNA replication
- Virion maturation and release

A schematic representation of the RV life cycle is shown in Figure 1.6.

1.6.1 Attachment and cell entry

The VP8* lectin domains of the outer capsid VP4 mediate RV attachment and depending on the strain bind to different host cell receptors such as sialic-acid glycans (gangliosides GM1 and GD1a) and histo-blood group antigens (HBGAs) [83-85]. Animal RV strains such as RRV and SA11 as well as human strains Wa and DS-1 bind sialic acid or sialic acid-containing gangliosides GM3 or GM1 [83, 86, 87]. However, various human strains of the P[6], P[11] and P[14] serotypes revealed binding specificity for HBGA, while interaction of the P[8] serotype with HBGA depends on the presence of a functioning fucosyltransferase 2 (FUT2) enzyme [88]. However, studies using human intestinal organoids showed that functional FUT2 did not affect susceptibility to RV infection of the P[8] serotype [89]. Nevertheless, epidemiological studies have demonstrated that children with less severe gastroenteritis had genetic polymorphisms affecting expression of the FUT2 enzyme [90]. Interestingly, major human RV serotypes P[8], P[4] and P[6] as well as norovirus, another pathogen causing acute gastroenteritis in children, all use polymorphic HBGA receptors for cellular attachment [91, 92].

Additionally, various cell surface molecules can act as potential co-receptors for RV cell entry. All co-receptors including integrins (α2β1, αVβ3, αXβ2 and α4β1) and human heat shock cognate protein 70 (hsc70) are associated with lipid-rafts found on the cellular plasma membrane, providing a platform with which RV TLPs can associate [85, 93, 94]. Hewish *et al.,*

(2000) showed that integrins expression in poorly permissive cells increased SA11 attachment and infectivity [95]. For RRV, following attachment to sialic-acids and integrins, the VP5* domain, specifically 642 and 659 residues were shown to interact with hsc70 [96]. A synthetic peptide mimicking this VP5* region (peptide KID) and antibodies to hsc70 block RV entry but not cell binding, suggesting that VP5* interacts with hsc70 at a post-attachment step [93, 96]. Additionally, Gualtero *et al.*, (2007) reported interaction between DLPs of the RRV, YM and Wa RV strains with hsc70, suggesting an indirect role of VP6 in RV cell entry [97].

After the initial attachment to cells, the VP8* lectin domains dissociate and expose the hydrophobic loops of the VP5* β-barrel domains [23, 37, 38]. The VP5* subunits fold back onto themselves and the interaction of the VP5* hydrophobic loops with the lipid bilayer leads to perforation of the plasma membrane by the VP5* foot domain, facilitating RV entry [37, 98, 99]. In its post-fusion conformation, VP5* remains bound to the TLP through contact with the VP7 subunits [38, 100]. The refolding of the V5* subunit is analogous to the membrane fusion mechanism of IAV where the proteolytically primed hemagglutinin (HA) interacts with the target membrane and refolds to insert the hydrophobic peptide into the host cell membrane [101].

Different RV strains enter cells through different endocytic pathways. Both receptor- and clathrin-mediated endocytosis as well as direct plasma membrane penetration have been suggested as mechanisms of TLP internalisation [102]. During endocytosis, a decrease in the calcium ion (Ca^{2+}) concentration in the endosome leads to the dissociation of both VP4 and VP7 from the particle mediating the release of the transcriptionally active DLP into the cytoplasm [103].

1.6.2 Positive-strand RNA ((+)RNA) synthesis

The polymerase complex, VP1 and VP3, synthesises capped, non-polyadenylated, (+)RNA transcripts from the negative strand of the genomic RNA, that are extruded out of the DLP into the cytoplasm [10]. The (+)RNAs function as mRNAs for viral protein translation by cellular ribosomes, and also as templates for synthesis of new dsRNA genomes in newly formed DLPs after interaction with VP1 [29, 104]. The VP1 has a cage-like structure with four tunnels leading to a catalytic core (residues 333–778) enclosed between the N-terminal (residues 1–332) and C-terminal domains (residues 779–1088) [105]. Two of the tunnels allow entry of the RNA template and NTPs whereas two others function as RNA exit tunnels [105, 106]. The cap-binding site of the N-terminal domain of VP1 splits the dsRNA genome through its interaction with the 5’ conserved m7GpppGGC residue of (+)RNA present in all RV

segments [49]. The N-terminus of RV VP1 has a neighbouring helix-loop-helix (HLH) subdomain (residues 39–69) which further separates the genomic duplex [107]. After a short part of the helix is unwound, the unpaired negative strand RNA ((-)RNA) traverses towards the active site of RdRp and immediately pairs with complementary NTPs within the core that form a backbone of the nascent RNA [107]. The dsRNA genome is pushed along by the newly synthesised nascent RNA backbone until it reaches the C-terminal domain of VP1 where the coding strand reanneals with the template and reforms the dsRNA genome [107]. The presence of distinct exit tunnels ensures that the nascent RNA is released into the cytoplasm while the (-)RNA is reused in subsequent rounds of (+)RNA synthesis [105, 108].

Proposed model of RV transcription is shown in Figure 1.4.

(+)RNAs are capped at the 5' end by VP3 that exhibits guanylyltransferase (GTase) and methyltransferase (MTase) activities using GTP and *S*-adenosyl methionine as substrates respectively [109]. In the leading proposed model for capping the endogenous transcript, the phosphodiesterase domains of VP3 tetramers are positioned between two VP1 molecules at each fivefold axis, facing the transcript exit channels [110]. The γ-phosphate of 5′-pppGp is removed by the RNA triphosphatase domain of VP3 and the emerging transcript transits through the GTase, MTase and 2'-*O*-methytransferase domains of VP3 for subsequent capping [110]. The capped transcript then exits through the type I channel system at the fivefold axis in the VP2 layer of the DLP [105, 111]. Periz *et al.,* (2013) showed that each channel of the DLP specialises in extruding one specific segment at any given time [104]. The proposed location of VP3 parallel to the inner surface of the VP2 layer is consistent with previous work by Estrozi *et al.,* (2013) [28]. Lawton *et al.,* (2001) showed that the transcription initiation and capping is followed by a pause in elongation after five to six nucleotides have been transcribed, suggesting that VP3 is in close proximity to VP1 during endogenous transcription [112]. *In vitro* and *in situ* reconstructions of DLPs have demonstrated that VP1 is attached to the innermost layer of the core formed by VP2 at fivefold axes, whereas the location of VP3 was not identified [105, 107]. Thus, the precise mechanism of VP3 capping requires further analyses.

At a later stage of virus replication, primary transcription must be inhibited to allow RNA replication to proceed and virus assembly to be completed. Inhibition of transcription is not fully understood but cryo-EM studies revealed that addition of VP6 trimers can inhibit elongation and translocation of transcripts [113, 114].

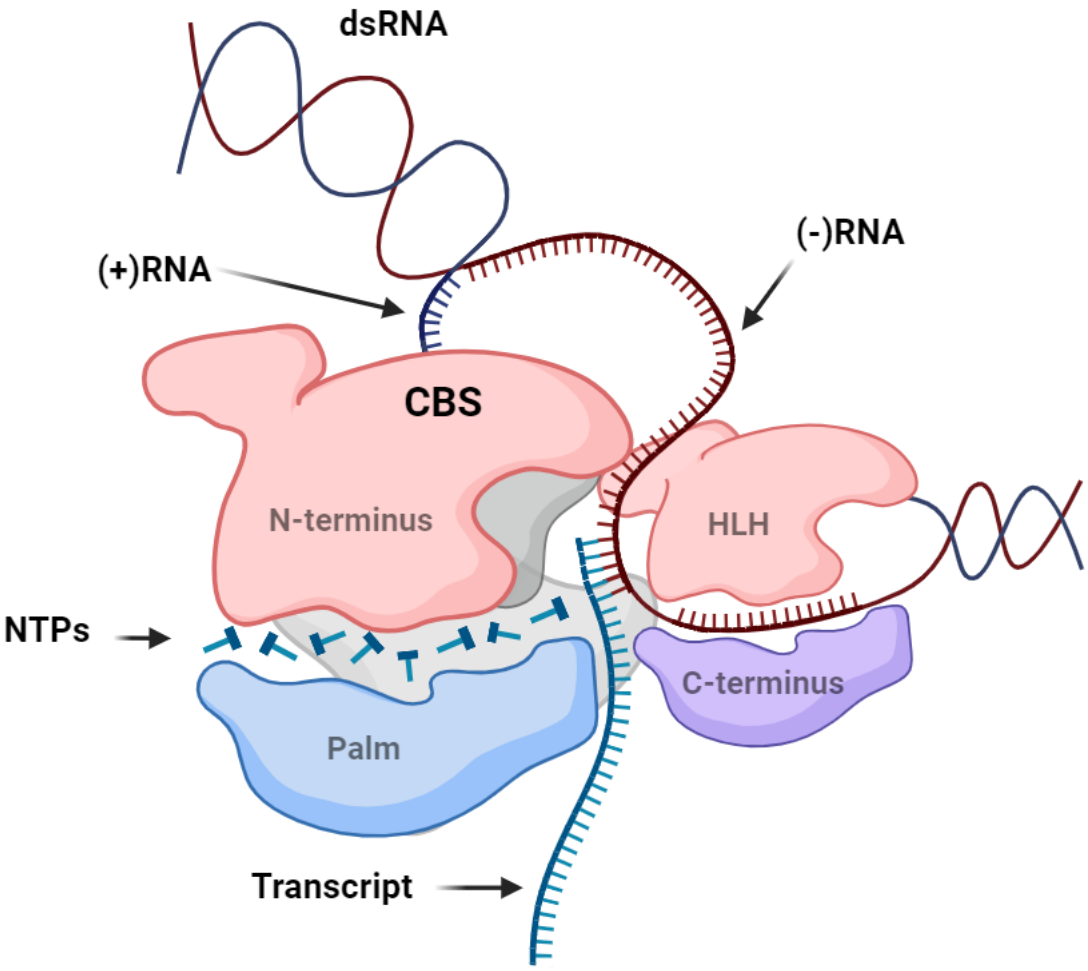


Figure 1.4 Proposed model of RV transcription.
Schematic showing RdRp with its corresponding domains. For diagram simplicity only the N-terminal domain with its HLH subdomain (pink), palm subdomain (blue) and C-terminal domain (purple) are labelled. Fingers and thumbs of the core are shown in grey. During transcription, the cap-binding site (CBS) of the N-terminal domain binds to the genomic (+)RNA and separates the dsRNA while the HLH subdomain further guides the (-)RNA into the core. Unpaired (-)RNA pairs with complementary NTPs that form the backbone of the nascent transcript. The transcript then leaves the polymerase core through one of the aqueous channels while the template stand reanneals with the (+)RNA and reforms the dsRNA genome. Diagram not to scale. *(Adapted from Ding et al., 2019 [104] and created with BioRender.com).*

1.6.3 Translation of viral proteins

Viral mRNAs are capped but not polyadenylated and the translation of viral proteins is facilitated by NSP3 [115, 116]. The N-terminus of NSP3 interacts with the 3'-GACC conserved consensus sequence of the viral mRNAs, while the C-terminus of NSP3 interacts with the eukaryotic translation initiation factor 4GI (eIF4GI) [78, 117]. The simultaneous interaction of NSP3 with the 3' end of viral mRNA and eIF4GI results in circularisation of viral mRNA and translation of RV proteins [118, 119]. Through these interactions, NSP3 evicts the cellular poly(A)-binding protein (PABP) from its interaction with eIF4G, causing PABP to accumulate in the nucleus [120, 121]. The depletion of PABP from the cell cytoplasm reduces the translation of polyadenylated mRNAs, thereby suppressing host cell protein synthesis [122-124].

A second mechanism by which RV inhibits translation of cellular proteins involves phosphorylation of eIF2α early after infection [125]. The continuous phosphorylated status of eIF2α preferentially translates viral mRNAs over those of the host and depends on the synthesis of VP2, NSP2 and NSP5 [125]. The PKR has also been implicated in this phosphorylation event, triggered by viral dsRNA in the cytoplasm outside of viroplasms, although the exact mechanism has not yet been determined [126].

RV proteins are synthesised on free ribosomes in the cytoplasm. The viral glycoproteins VP7 and NSP4 are co-translationally inserted into the endoplasmic reticulum (ER) membrane due to the presence of signal sequences at their N-termini [73, 127].

1.6.4 Viroplasm formation, packaging and dsRNA replication

1.6.4.1 Viroplasm formation

Following viral protein translation, two non-structural proteins NSP2 and NSP5 together form the cytoplasmic inclusion bodies known as ‘viroplasms’ where RV replication and assembly occur [128, 129]. In the absence of other RV proteins, co-expression of NSP2 and NSP5 is necessary and sufficient for the formation of viroplasm-like structures in uninfected cells [130]. Silencing of either NSP2 or NSP5 by RNA interference or using temperature-sensitive mutants prevents viroplasm formation, genome replication and virion assembly [131-133]. After RV infection, the number of viroplasms decrease with time, whereas the area of each viroplasm increases, suggesting viroplasms fusion [130, 134]. Viroplasms associate with lipid droplets and interfering with their interaction decreases the number and size of viroplasms thereby reducing production of infectious virus [135-137]. Functional proteasomes and elements of the autophagy pathway are also required for early assembly of viroplasms and RV replication [138-140].

Viroplasm assembly is regulated by phosphorylation and involves interaction of two forms of NSP2 with hypo- and hyper-phosphorylated forms of NSP5 [141-143]. Phosphorylation of the C-terminal helix domain of NSP2 at its S313 residue by the cellular kinase CK1α results in two forms of NSP2: the cytoplasmic or ‘diffuse’ (dNSP2) and viroplasm-specific (vNSP2) NSP2 [142]. The dNSP2 associates with the hypophosphorylated form of NSP5 to initiate viroplasm assembly [142, 143]. The dNSP2 also forms complexes with the VP1, the N-terminal hub of VP2 and alpha/beta tubulin [142, 144]. Studies suggest that dNSP2 may traffic viroplasm-associated components to viroplasm sites through its association with microtubules [134, 144]. The vNSP2 form is detected only in mature viroplasms that are present later during infection and accumulates concurrently with their increase in size throughout the replication cycle [142]. Cryo-EM structures show that single-stranded RNA (ssRNA) and NSP5 compete for binding in the vNSP2 octamer grooves [145]. The vNSP2 octamers, in association with hyperphosphorylated NSP5, form a lattice structure for viroplasm assembly, likely on lipid droplets [135].

Recently, studies demonstrated that the viroplasms represent condensates formed via liquid–liquid phase separation of NSP5 and NSP2 [129, 146]. Depending on the stage of infection, the material properties (e.g., fluidity) of these condensates change concomitant with NSP5 phosphorylation [146]. Additionally, interaction with the host components, including lipid

bilayers, microtubules and tubulin, can promote nucleation of condensates and spatially regulate the kinetics of their formation in cells [146].

1.6.4.2 Packaging and dsRNA replication

Based on structural and biochemical studies, RV replication including genome encapsidation and DLP assembly occurs simultaneously with the formation of core particles [147, 148]. The 11 (+)RNA segments interact with viral core proteins and are packaged in equimolar amounts [148]. Although the precise mechanism of coordinated packaging is not well understood, several models have been proposed some of which are described below [36, 48, 149].

During RV replication, (+)RNA serves as a template for (-)RNA synthesis, yielding dsRNA products [150]. The (+) strand of the dsRNA is capped, while the (-) strand is uncapped, lacking the 5′ γ-phosphate (γP) [151]. Studies showed that the NSP2 octamer relaxes the secondary structures of viral RNA templates for dsRNA synthesis by VP1, and assists with their translocation into pre-virion cores [152]. Analysis of the NSP2 structure revealed a deep cleft within the NSP2 octamer containing nucleotide-binding histidine triad (HIT)-like motifs associated with the nucleoside-triphosphatase and RNA triphosphatase (RTPase) activities of the protein [151, 153, 154]. After the (-) strand synthesis initiation, NSP2 interacts with the 5′ consensus sequence of the nascent RNA product in such a way that its 5′ end enters the HIT motif, where it engages the catalytic site [152]. Thus, the RTPase activity of NSP2 motif may account for the absence of γP on the (-) strand of dsRNA genome segments [151, 153-155]. Furthermore, the interaction of NSP2 with the 5′ end of the (-) strand, and the interaction of the RdRp with the 3′ end of the (-) strand may be responsible for initial circularisation of the genome segments prior to core formation [151].

Core assembly is initiated by specific interactions of VP1 with the conserved 3’ terminal sequence on all 11 mRNAs containing *cis*-acting signals important for transcription [156]. The 3’ consensus sequence of (+)RNA differs between RV strains with species A (5’-UGUGACC-3’) and C (5’-UGUGGCU-3’) sharing the same ‘UGUG’ motif, while the 3’ termini of (+)RNA in species B have a different consensus sequence (5’-ACCC-3’) [108, 156]. This interaction is catalytically inactive and requires the N-terminal domain of VP2 for (-)RNA synthesis and for encapsidation of VP1 and VP3 into progeny cores [157, 158]. Interaction of VP1 with VP2 leads to conformational changes in the priming loop within the catalytic core of VP1 stabilising the initiating nucleotide in the priming site of RdRp [159, 160]. This correct alignment results in the formation of the first phosphodiester bond of the (-)RNA product [161]. For the polymerase to transition from initiation to elongation, the amphipathic α-helix of VP2 locks the HLH subdomain of VP1 and displaces the C-terminal plug from the exit tunnel allowing

threading of dsRNA driven by nucleotide hydrolysis [107, 108]. Simultaneously, the priming loop retracts allowing elongation of the dsRNA product out of the polymerase [105].

During the core assembly, inter-molecular RNA-RNA interactions between non-translating viral transcripts are favoured by the presence of the chaperone NSP2 within viroplasms [162-164]. According to Borodavka *et al.*, (2017) (+)ssRNAs undergo remodelling upon binding to NSP2 and due its high affinity for ssRNA, NSP2 may remain associated with the ssRNA prior its replication [162]. Previous studies showed that loss of NSP2 results in inhibition of replication and virion assembly [165]. Recently, Strauss *et al.*, (2023) have demonstrated that clustering of VP1-bound non-translating viral transcripts within viroplasms for RNA genome assortment and assembly was facilitated through its interaction with NSP2 and NSP5 [163]. Additionally, interaction of VP1 with the N-terminal ‘arms’ of VP2 could aid the formation of the VP2 lattice around the assorted 11 distinct RNA transcripts [30, 158].

The interaction of the polymerase complex (VP1 and VP3) with the VP2 dimers and 11 dsRNA segments results in the formation of the inner capsid layer, providing a platform for the subsequent addition of VP6 trimers to form the DLP [27].

Proposed model of genome assortment and packaging is shown in Figure 1.5.

At later stages of infection, a logarithmic increase in mRNA abundance indicates that the newly synthesised DLPs in viroplasms have become transcriptionally active, initiating a secondary wave of transcription and replication [166, 167]. Studies showed that silencing the expression of the viral proteins that constitute the DLP (VP1, VP2, VP3 and VP6) ablates the logarithmic increase in the RNA synthesis [166].

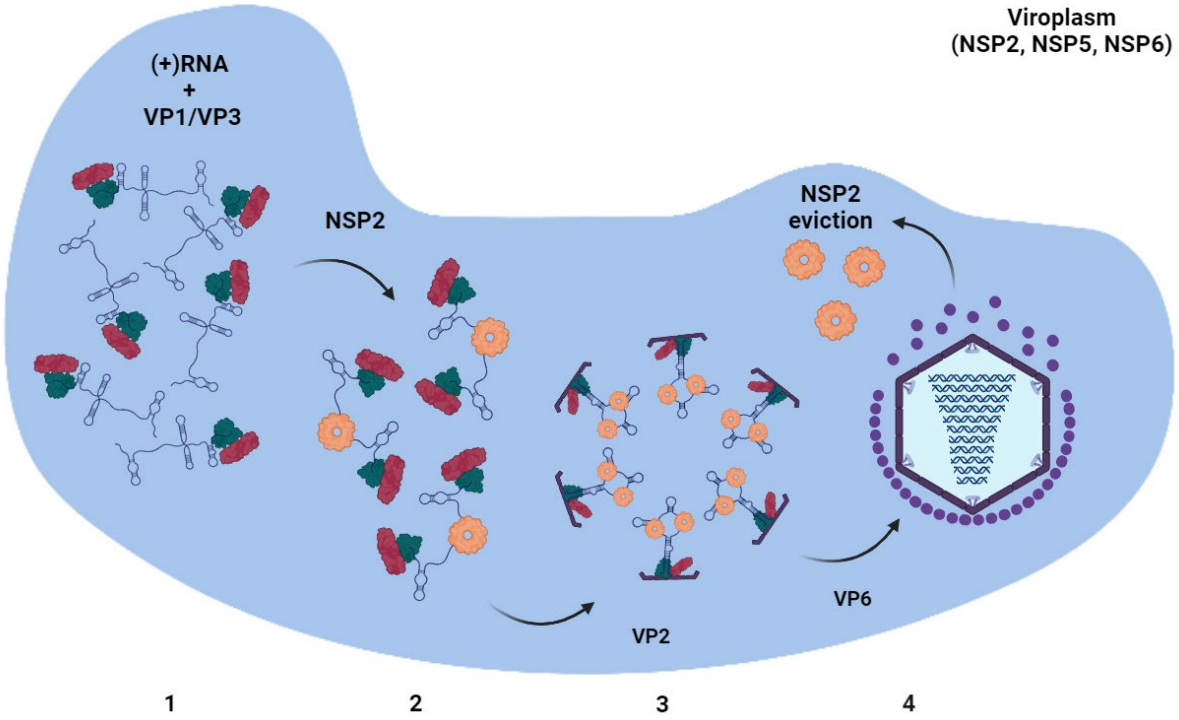


Figure 1.5 Proposed model of genome assortment and packaging.
1. Within viroplasms, (+)RNAs bind VP1 (green) and VP3 (red), forming RNA complexes. **2.** Binding of the NSP2 octamer (yellow) causes structural remodelling of (+)RNAs, promoting interaction between different RNA transcripts – a process representing RNA assortment. **3.** The assorted RNA complexes are predicted to interact with VP2 to facilitate core assembly, providing a platform for the addition of VP6 that results in the displacement of NSP2 (**4**). Schematic not to scale. *(Adapted from Borodavka et al., 2018 [148] and created with BioRender.com).*

1.6.4.3 Virion maturation and release

Virion assembly requires NSP4-mediated increases in cytoplasmic calcium and the cellular process of autophagy [168, 169]. NSP4 is a transmembrane glycoprotein located in the ER and acts as a viroporin releasing the ER calcium into the cytoplasm [170-172]. Cytoplasmic increase in calcium levels activates calcium signalling pathways involving calcium/calmodulin-dependent protein kinase kinase 2 and 5′ adenosine monophosphate-activated protein kinase to trigger autophagy [168]. Coat protein complex II vesicles traffic NSP4, the outer capsid VP7 and the autophagy marker protein LC3 II to the viroplasms [169, 173]. The C-terminal cytoplasmic domain of NSP4 acts as an intracellular receptor that binds VP6 on the newly synthesised DLPs and mediates budding of maturing viral particles into the viroplasm-associated membranes [174, 175]. Assembled DLPs leaving the viroplasms are transiently enveloped but this envelope is removed by an unknown mechanism and the outer capsid proteins, VP7 and VP4, are assembled onto the particle [173]. Previously, DLPs were thought to bud through the ER membranes [36].

Virion maturation is a calcium-dependent process and studies showed that in the absence of calcium, the budding of virus particles into the viroplasm-associated membranes is not observed [176]. Structural studies of VP7 showed that calcium stabilises the trimeric interactions between VP6 and VP7 whereas in the absence of calcium VP7 is retained in ER [147, 177]. Calcium was shown to be important for oligomerisation of NSP4, VP4, and VP7 in the viroplasm-associated membranes as well as for proper folding of VP7 and outer capsid assembly [178]. NSP4 also has a binding site for the VP5* domain of VP4 that may play a role in removing of the transient envelope [179]. Cuadras *et al.*, (2006) showed that following budding of DLPs into the viroplasm-associated membranes, VP4 and VP7 are assembled onto the particle and VP7 is involved in the removal of transient envelope [180]. Furthermore, *in vitro* reconstitution studies of TLPs showed that the DLPs require addition of VP4 before VP7 to produce infectious virus [23, 181, 182].

Alterations in the permeability of the plasma membrane of infected cells results in the release of mature TLPs as well as cellular and viral proteins [10]. Despite host cell lysis, most DLPs and many TLPs remain associated with the cellular debris [183]. VP4 can remodel microfilaments through its interaction with actin and lipid rafts, destabilising the brush borders of cell membranes and facilitating the release of RV from infected cells [184-186]. Although it is possible that the RV particles are simply trapped by the cytoskeleton.

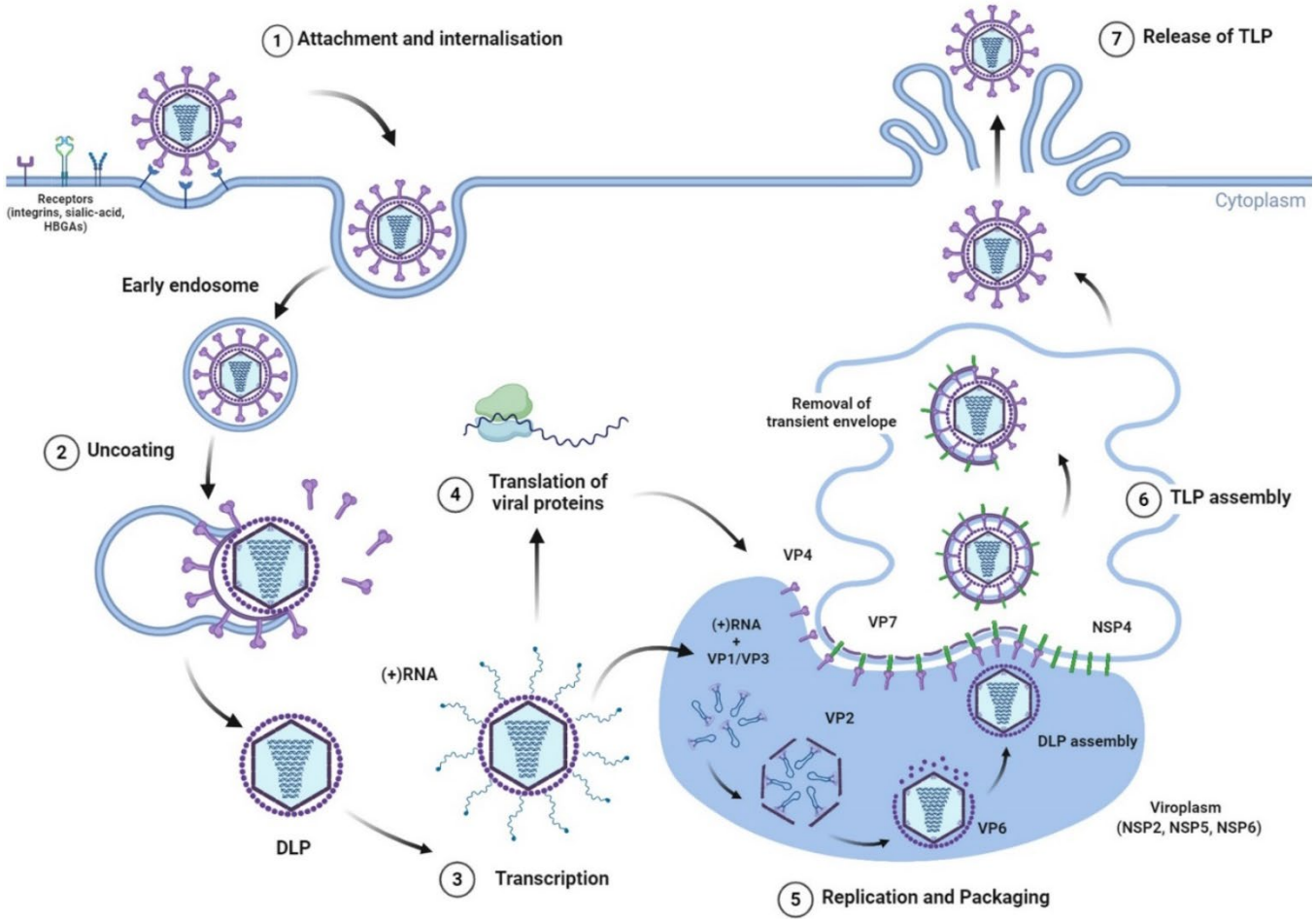


Figure 1.6 Schematic of the RV replication cycle.
Major features of the RV replication cycle are detailed in **section 1.6**. **1**. Following attachment, conformational change of VP4 results in internalisation of TLP by receptor-mediated endocytosis. **2**. Uncoating results in the removal of the outer capsid proteins VP4 and VP7 due to low calcium within the endosome, causing release of DLPs into the cytoplasm. **3**. Transcription of (+)RNA. **4**. Translation of viral proteins**. 5**. Replication and packaging of viral proteins into new DLPs occurs in the viroplasm. **6**. DLPs mature into TLPs in rough ER where the outer capsid layer is assembled**. 7.** The new virion particle is released via lysis or endocytosis. *(Adapted from Crawford et al., 2023 [10] and created with BioRender.com)*.

1.7 Prevention of RV disease

1.7.1 Natural protection

Primary RV infection usually occurs during the first year of life resulting in severe gastroenteritis. The majority of RV infections (>90%) are caused by viruses with the most common G-types (G1-4, G9 and G12) and P-types (P[4], P[6] and P[8]) [187-189]. Correspondingly, epidemiological studies showed that following a natural infection, children developed both homotypic and heterotypic immunity [190, 191]. Natural RV infection did not provide sterilising immunity against re-infections but reduced disease severity due to the presence of mucosal RV-specific antibodies (IgA) in the gut lumen [192]. Increased titres of serum RV-specific IgM, IgG and IgA as well as secretory IgA were also observed after infection and provided a good correlate of protection [193]. However, the level of protection induced by primary and subsequent re-infections was lower in developing countries than in the developed world, possibly exacerbated by comorbidities [194-196].

1.7.2 Vaccines

In the 1980s, the initial oral monovalent live-attenuated vaccine (LAV) involved a cell-culture passaged bovine RV (NCDV strain, G6P[1], RIT 4237) that delivered over 80% protection against disease in infants and young children in two separate efficacy trials in Finland [197, 198]. However, initial promising results were not reproduced by trials in several middle- and low-income countries [199, 200]. Consequently, further work on this vaccine candidate was not pursued.

The simian RV vaccine (RRV strain, G3P[3], MMU 18006) was studied extensively as a monovalent *Jennerian* vaccine candidate in several countries during the late 1980s [201]. Results from efficacy trials were encouraging and it appeared to be more antigenic than the RIT 4237 vaccine candidate, although in some settings it failed to induce high levels of protection [202-204]. The reason for this variable protective efficacy was not determined, but the vaccine appeared to work best against homologous G3 strains and less well against heterologous strains [205, 206].

Failure of the monovalent RRV vaccine to reproducibly induce high levels of protection against heterotypic RV strains led to the development of a quadrivalent (RV4) simian RRV-based vaccine [207]. RV4 incorporated the most important human genotypes at the time (G1

– G4) genetically coupled to the simian RRV attenuated in human [207]. Due to the segmented nature of the RV genome, these viruses reassorted at high frequency during mixed infection [208]. Single human RV gene substitution reassortants were isolated that possessed the human G1, G2, G4 or simian G3 VP7 genes on a background of 10 simian RV genes and these were multiplexed to create a quadrivalent RV4 vaccine [209]. Based on several large successful Phase III trials the RV4 vaccine was licenced and produced under the name RotaShield (Wyeth Lederle Vaccines) [210, 211].

During 1998/99 as part of a universal mass vaccination programme, RV4 was administered to almost 1 million young children in the USA [46]. However, cases of intussusception (estimated 1:2500) were reported during the first 2 weeks after the first vaccine dose [212]. This correlated risk led to voluntary discontinuation of vaccine production by the manufacturer, although it was found that age at vaccination was a confounding factor in the rare occurrence of intussusception following the RV4 use [213, 214]. Thus, avoiding the age period (3-4 months through to 9 months) when gut intussusception was most prevalent and implementing the initial vaccination during the first 2 to 3 months of life might have substantially reduced or eliminated the risk of intussusception [215].

Following withdrawal of the RV4 vaccine, extensive efforts were made to develop alternative vaccines that resulted in two separate but parallel approaches.

One such approach involved a monovalent LAV vaccine (RV1) derived from the virulent human RV isolate 89–12, of serotype G1P[8] [216]. Unlike the *Jennerian* candidates, this vaccine candidate was attenuated by serial passage in tissue culture [216]. In a large Phase III trial (>60,000 recipients), it provided high levels of protection (>90%) against disease by homotypic G1P[8] strains and reasonable protection against G heterotypic, P homotypic G3P[8], G4P[8], and G9P[8] strains [217]. Additionally, RV1 was found to be efficacious against entirely heterotypic G2P[4] strains, although to a lesser extent [218]. Given the high level of efficacy and safety profile, RV1 vaccine was manufactured by GlaxoSmithKline under the trade name Rotarix, and has been licensed for use around the world since 2005 [217].

While Rotarix was under development, a *Jennerian* bovine RV vaccine was used with the Wistar calf strain (WC3) G6P[5] as the genetic backbone to create a pentavalent vaccine, RV5 [219, 220]. Originally, the monovalent WC3 vaccine candidate had demonstrated variable protective efficacy in several trials from different countries, but it was withdrawn from further study [221, 222]. Five mono-WC3 reassortants were formulated to incorporate human-derived isolate VP7 G1–G4 genes and a single human-derived P[8] VP4 gene [223]. This approach was similar to that taken for the development of RotaShield except that the genetic backbone was derived from a bovine RV rather than a simian strain. A large Phase III safety trial that

included more than 68,000 infants showed high levels of protection (98%) against severe disease and no risk of intussusception was identified [224]. The pentavalent vaccine named RotaTeq, manufactured by Merck, was licensed for use throughout the world from 2006.

Several other regional RV vaccines are currently available and alternative vaccine strategies are being evaluated, such as inactivated virus-like particles and subunit vaccines [225-228]. Due to an increase in incidence of norovirus-associated diarrhoea in children, combination RV vaccines are also being trialled [44, 229].

1.7.3 Vaccine efficacy

The RV5 and RV1 vaccines have substantially reduced the disease burden of RV-associated gastroenteritis in infants and young children worldwide [230, 231]. Phase IV post-licensure studies in high-income and moderate-income countries confirmed the high efficacy and safety of the Phase III trials [232-234]. A substantial decrease in hospitalizations from gastroenteritis was observed in children of all age groups including those not eligible for vaccination, suggesting “herd immunity” induced by mass vaccination programmes [235-237].

In contrast to the high effectiveness of RV vaccination in developed countries, in some middle- and low-income regions, such as India and Sub-Saharan Africa, the vaccine efficacy was substantially lower (30–50%) [238-240]. Similar to RV1 and RV5, a variety of orally administered vaccines, such as those targeting polio and some cholera vaccines, performed less well in poorer countries [196]. The reasons behind the suboptimal performance of the RV vaccines remain unknown but may include factors such as malnutrition, presence of maternal antibodies and/or HBGA differences [195, 241]. A very low rate of vaccine associated intussusception (<1:50,000) is still associated with the two currently licenced vaccines (RV1 and RV5), although the benefits of the vaccine outweigh the risks based on the currently available data [242-244]. Additionally, the live-attenuated nature of RV1 and RV5 limits their administration to immunocompromised or prematurely born children due to risk of vaccine-mediated illness [245]. LAVs also carry a risk of reassortment with wild-type (WT) human RVs and there is limited information on protection against heterotypic strains by RV1 and RV5 [246-248]. Furthermore, human–animal RV reassortants have been detected in human isolates as certain animal RVs share a neutralization antigen with some human RV strains [249]. However, zoonotic transmission and frequency of human/animal RV reassortment in the field remains unclear. Despite these drawbacks, since the mortality from RV-associated gastroenteritis is highest in low-income countries, the World Health Organisation recommends that all countries routinely vaccinate young children against RV [250].

Similarly, as RV infects a wide variety of livestock, zoonotic transmission of RVs has been reported occasionally, posing a risk to humans, as current vaccines may not be able to protect against novel strains [251-255]. RV infections cause significant economic loss to the livestock industry due to declines in production and associated treatment costs [256-258]. RV vaccines are available for cattle, pigs and horses, and are mainly protective against species A RVs [8, 259, 260]. A major obstacle for the development of a multivalent RV vaccine is that non-species A RVs are hard to propagate to high titres [261]. Currently available RV vaccines for animal use rely on immunisation of pregnant animals thereby relying on maternal transfer of antibodies via colostrum [8, 260]. However, passive immunisation varies and its success can be affected by the duration of colostrum administration, health status of animal and previous exposure to RV [260]. Thus, there is a need for development of animal RV vaccines that are more immunogenic such as subunit vaccines that will negate the need to propagate non-species A RVs.

The combination of all these factors highlights the need for the development of next-generation RV vaccines, especially for low- and middle-income countries where RV-associated mortality remains high. The key challenge in designing a LAV is a means by which to effectively attenuate virulence to avoid symptomatic disease while retaining strong immunogenicity [262]. For this reason, current LAVs targeting other pathogens are developed so that they replicate sufficiently to induce a strong immune response (e.g. FluMist, influenza viruses), while others require multiple doses and may induce some mild disease (e.g. MMRVaxPro, measles, mumps and rubella virus) [262]. The recent development of plasmid-only reverse genetics systems for RVs *(see Reverse genetics)* has provided a new approach in engineering next-generation vaccines.

1.8 Reverse genetics

The reverse genetics technique permits controlled manipulation of the viral genome to characterise the importance and functions of viral genetic traits. Reverse genetics of RNA viruses typically involves manipulation of their genomes at the DNA level, followed by transfection of plasmid(s) or RNA(s) into cells to produce infectious virus.

The development of a reverse genetics system to artificially engineer and mutate an infectious RV has proven challenging. These challenges were attributed to inefficient translation of transcripts from transfected mRNAs or cDNAs, which was often associated with cytotoxic effects that decreased viral rescue efficiency [46, 263]. Historically, researchers relied on inefficient and difficult gene reassortment experiments involving co-infection of the same cells with two parental RV strains to generate viral mutants [208, 209, 264, 265]. However, reassortant viruses were difficult to generate due to the requirement of laborious strong selection conditions and/or plaque isolation [266]. Subsequently, various strategies, such as helper-virus dependent and plasmid-based reverse genetics systems were employed to generate recombinant RVs to study the function of viral proteins and viral replication in the context of infection. For other viruses of the *Sedoreoviridae* family such as orthoreoviruses [267-269] and orbiviruses [270-272], plasmid- or RNA transcript- based reverse genetics systems are well established.

1.8.1 Helper-virus-dependent reverse genetics system for RVs

Prior to the development of an efficient plasmid-only reverse genetics system, methods to rescue recombinant RVs using helper-viruses were reported, some of which are described below. This strategy was based on the previously developed helper-virus-dependent reverse genetics system for IAV [273].

The first RV system was based on transfection of cells expressing T7 RNA polymerase driven by the vaccinia virus (VV) with plasmids encoding the simian VP4 gene [274, 275]. Plasmids were constructed so that the gene was flanked by a T7 polymerase promoter (T7P) and a hepatitis delta virus (HDV) ribozyme sequence, followed by a T7 RNA polymerase terminator (T7T). The T7P was used to initiate transcription at a defined guanosine residue corresponding to the 5'-end of an RV segment whilst self-cleavage of the HDV produced RNA with the native RV 3' terminus [274]. Cells were then infected with the human RV KU strain and the culture supernatant was passaged in MA104 cells in the presence of antibodies that specifically neutralised VP4 of the KU helper virus [274, 275]. The selective pressure favoured

the rescue of a recombinant virus containing the VP4 segment of different RV strains in the backbone of the WT helper virus [274, 275]. However, the rescue efficiency was very low and limited to segments encoding VP4 and VP7 as these genes were specifically targeted by the antibodies used [276].

A similar strategy with a different selective pressure technique was used to generate an infectious RV carrying an exogenous NSP3 segment with a partial head-to-tail duplication [277]. Sequence duplications have been identified in a variety of RV segments and were shown to be preferentially packaged into viral progeny over the WT genes [263, 278-281]. Thus, cDNAs containing the rearranged NSP3 gene under T7P were transfected into the cells and, upon infection with a helper RV and T7 RNA polymerase driven by the VV, the WT segment was replaced by the partially duplicated one [277]. However, the recombinant RVs were only recovered after 18 passages, highlighting the inefficiency of this system [277].

Another approach exploited a dual-selection mechanism, or a ‘two-hit’ strategy, to rescue recombinant RVs with chimeric NSP2 genes [282]. The first selection method involved the use of temperature-sensitive mutants followed by the use of RNA interference [282]. On infection of cells expressing a temperature-sensitive helper virus mutant, only the recombinant RVs carrying the temperature-resistant NSP2 gene were able to grow [282]. The same cells were simultaneously transfected with an small interfering RNA targeting the NSP2 gene of the helper virus. This resulted in successful rescue of recombinant RVs, although its low rescue efficiency restricted its application.

Development of an improved plasmid based reverse genetics system for mammalian orthoreoviruses showed that orthoreovirus rescue was possible using the engineered baby hamster kidney (BHK) cell line constitutively expressing T7 RNA polymerase (BHK-T7 cells) [268]. Subsequently, Johne *et al*., (2016) generated an avian-mammalian reassortant RV using the BSR5/T7 cells (synonymous to BHK-T7 cells) eliminating the use of VV as source of T7 RNA polymerase [283].

Despite the advantage of generating recombinant RVs, the above-described strategies involved the need of a helper-virus and/or specific-selection systems, highlighting the limitations for studying particular functions of the viral proteins and incapability for targeted mutagenesis.

1.8.2 Plasmid-only reverse genetics system for RVs

In 2017, a breakthrough was made when Kanai *et al.*, (2017) developed a plasmid-only reverse genetics system for the simian SA11 strain of RV [284]. This system did not require selection pressure since no helper virus was involved. Based on a previously validated approach with orthoreoviruses [267, 268], in this system SA11 RV was rescued following co-transfection of BHK-T7 cells with 14 plasmids [284]. These included 11 plasmids corresponding to each of the 11 SA11 RV gene segments (designed as described in *section 1.8.1*) and three polymerase-II-promoter-driven expression plasmids encoding the Nelson Bay orthoreovirus fusion-associated small transmembrane (FAST) protein and two subunits of the VV capping enzyme (D1R and D12L) [284]. FAST proteins are non-structural proteins encoded by fusogenic members of the *Aquareovirus* and *Orthoreovirus* genera of the *Sedoreoviridae* family, and promote viral replication and pathogenesis *in vivo* [285, 286]. Species B RVs encode a FAST homolog whereas species A lack this protein [75]. The aim of co-expressing FAST was to improve the virus rescue by increasing virus spread through cell-cell fusion and syncytium formation [285, 286]. The VV capping enzyme was co-expressed as a mechanism to stabilize the viral (+)ssRNAs produced in the cytoplasm by adding an authentic 5' cap thereby increasing the rate of viral protein translation [268, 284]. Following transfection, BHK-T7 cells were co-cultured with MA104 cells (highly permissive for RV infection, but poorly transfectable), followed by inoculation of lysates onto fresh MA104 cells to rescue SA11 viruses [284].

A follow-up study from the same laboratory demonstrated that co-transfection of the expression plasmids encoding the VV capping enzymes did not provide the additional advantage as initially expected [65]. Furthermore, the possible toxicity of the FAST proteins through their potential to promote apoptosis and cell death has limited their use [287]. A further modification to the system included increasing the quantities of the two plasmids harbouring the RV NSP2 and NSP5 genes [288]. The increased rescue efficiency observed under these conditions was likely due to these non-structural proteins being essential for formation of the viroplasms that are required for virus replication [130]. Accordingly, following transfection of BHK-T7 cells and co-cultured with CV-1 cells, the rescue efficiency of RVs was greatly improved [288, 289].

To further enhance the rescue efficiency, Philip *et al.*, (2019) used a different capping enzyme, NP868R of African swine fever virus (ASFV), fused to the T7 RNA polymerase together with the 11 RV segments [290]. The use of this fusion plasmid (C3P3-G1) increased

the capping and stability of the viral transcripts, improving the rescue efficiency of RV and reovirus [291].

A more recent modification included the use of genetically modified MA104 cells, which are routinely used to produce and passage infectious RVs. The engineered MA104 cell line (MA104N*V) stably expressed the N protease of bovine viral diarrhoea virus, targeting interferon regulatory factor 3 (IRF3) for degradation, and the V protein of parainfluenza virus 5, targeting STAT1 for degradation [292]. As a result, the antiviral innate immune response was significantly attenuated. This optimized reverse genetics system allowed the recovery of recombinant simian, human, and murine-like RVs [292].

1.8.3 Application of reverse genetics systems

Using the helper virus-dependent reverse genetics system, Komoto *et al.*, (2011) generated a recombinant SA11 RV with a polybasic cleavage site in the VP4 gene [293]. The proteolytic cleavage of VP4 results in conformation changes that mediate viral entry (*see 1.6.1 Attachment and cell entry*) [40]. Introducing a furin cleavage site in VP4 would allow RV to undergo multiple rounds of replication without the need of exogenous trypsin, as shown for IAV [294]. However, the rescued virus had lower replication titres than those of the WT RV possibly due to disordered conformational changes in VP4.

To investigate the capacity of the RV genome to accommodate additional genetic information, a dual selection method was used to mutate the SA11 NSP2 gene [295]. The recombinant RVs had sequence duplications of the NSP2 gene and heterologous sequences (FLAG tag, the hepatitis C virus type 2 E2 epitope or the cricket paralysis virus IRES) at the 3’ UTR [295]. However, the rearranged NSP2 gene containing 400- and 800-nucleotide duplications failed to rescue, suggesting the presence of *cis*-acting signals important for RNA packaging [295].

Following its development, the plasmid-based reverse genetics system has been utilised to study the role of several non-structural and structural proteins during RV infection. Initially, Kanai *et al.*, (2017) showed that the C-terminal 103 residues of NSP1 were required to inhibit the innate immune response by inducing degradation of IRF3 [284]. NSP1 ORF was further modified to express fluorescent proteins such as NanoLuc-luciferase (N-Luc) and split-green fluorescent protein (GFP), highlighting its tolerance for large insertions without a significant effect on RV replication *in vitro* [284, 296]. In the presence of the porcine teschovirus-1 2A (P2A) sequence, recombinant RVs were generated carrying fluorescent

proteins in the NSP1 gene, such as N-Luc, mCherry and enhanced green fluorescent protein (EGFP) [288]. Presence of the 2A peptide allowed co-expression of two protein products as a result of ribosomal skipping, therefore lowering the risk of interfering with the function of the NSP1 gene [297]. The replication-competent RV carrying the N-Luc gene was useful for dose-dependent screening of antivirals [284].

The optimised plasmid-only reverse genetics system was used to generate a repertoire of recombinant RVs expressing fluorescent reporter proteins from the NSP3 ORF. The C-terminus of the NSP3 ORF was fused with one of six fluorescent reporter proteins (UnaG, mKate, mRuby, TagBFP, CFP, YFP) in the presence of the P2A element [298]. This further demonstrates the ability of RV to package additional foreign sequences as part of the viral genome. Indeed, recent studies showed that RVs could be utilised as vaccine vectors by modifying the NSP3 gene to express different fragments of the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) spike protein (*see Chapter 4*) [299, 300].

Through reverse genetics, the crucial role of NSP5 hyperphosphorylation for viroplasm formation was demonstrated following the construction of recombinant RV with a non-functional transcript of NSP5, grown in an NSP5 *trans*-complemented cell line [301]. NSP5 phosphorylation was required for the assembly of fully functional and round-shaped viroplasms, whereas mutants with Ser67Ala mutation assembled spindle like viroplasms and replicated more slowly than the WT RV [301-303]. In contrast, NSP6, which is expressed from the +1 ORF of the NSP5 segment, was not essential for virus replication in a cell culture in this system (*see section 1.5 Viral accessory proteins*) [65].

With the help of reverse genetics, the NSP1, NSP2, NSP4 and NSP5 genes have also been modified and incorporated into recombinant RVs *in vivo* [167]. MA104 cells were engineered to express CRISPR-Csy4 nuclease fused to NSP5 that upon infection was recruited to viroplasms and specifically cleaved the (+)ssRNA intermediates during dsRNA genome replication [167, 304]. As a result, the viral progeny contained transcripts encoding fluorescent or other tagged reporters, demonstrating that the secondary transcription step contributed to the overall production of viral proteins in infected cells [167].

Some new functional elements of structural RV proteins were also discovered through the plasmid-only reverse genetics system. Song *et al.*, (2020) showed that a virus harbouring mutations in the C-terminal domain of VP3 was attenuated due to the absence of a virulence domain that could effectively neutralize the RNase L signalling pathway [305]. The study

highlighted that the VP3 mutants could be further explored as potential next-generation vaccine candidates for the control of RV infection in humans [305].
In a mouse model, RRV infection results in inflammation and obstruction of the bile ducts, a condition known as biliary atresia [306, 307]. Studies reported that arginine at position 446 of the VP4 protein is essential for viral entry and mutation of this residue resulted in significantly reduced symptoms and mortality in neonatal mice [307]. Thus, the murine model of biliary atresia could be used to explore the important role of VP4 in viral pathogenesis.

The recent success of the plasmid-only reverse genetics system has opened the possibility to study functions of RV proteins in more detail, generate stable RV reporter expression systems and use RV as an expression vector for the development of novel bivalent vaccines. Reverse genetics has also been used to generate many reassortant viruses containing genome segments from different species, thus highlighting the importance of this technique for rapid generation of vaccine candidates in the event of emergence of novel RV strains of clinical significance.

1.9 Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)

In December 2019, an outbreak of severe pneumonia cases of unknown aetiology was reported in the city of Wuhan, Hubei province, China [308]. Patients presented with high fever and difficulty breathing with most of the cases (66%) epidemiologically relating to the Huanan seafood wholesale market [309]. Following full genome sequencing, the causative agent was identified as SARS-CoV-2 of the *Sarbecovirus* subgenus, which was responsible for the outbreak of the coronavirus disease 2019 (COVID-19) [308]. By February 2020, the virus was circulating in several continents and on the 11 March 2020, the WHO declared COVID-19 a pandemic.

Coronaviruses (CoVs) are enveloped viruses containing a large (26–32kb) single-stranded, (+)RNA genome consisting of 14 ORFs that encode 27 proteins [310]. The CoV virion comprises the main structural proteins: spike, membrane, envelope and nucleocapsid [310]. CoVs replicate in the cytoplasm and enter the host cells by either direct fusion of the viral envelope with the host cell membrane, or by membrane fusion within the endosome after endocytosis [311]. CoVs of the *Alpha-* and *Betacoronavirus* genus (OC43, HKU1, 229E, NL63) continuously circulate in the human population and cause mild respiratory infections [312]. Before SARS-CoV-2, SARS-CoV (2002 – 2003) and Middle Eastern respiratory syndrome CoV (MERS-CoV) (from 2012) have caused zoonotic infections and epidemics [313, 314]. Following the outbreak of SARS-CoV, development of preclinical and Phase I vaccine candidates was stopped as the virus was eradicated from the human population [315-317]. Until recently, no vaccines against CoVs were licenced for use in humans since they were considered a low priority.

SARS-CoV-2 was of high priority due to its high transmissibility and mortality affecting all ages of the population [318]. Following the emergence of COVID-19, antiviral drugs, monoclonal antibodies and non-pharmaceutical interventions such as physical distancing and lockdowns were deployed in an attempt to control the spread and impact of SARS-CoV-2. As the world stood still, researchers worldwide raced to develop an effective vaccine while medical staff worked tirelessly to save lives. Through prior studies of vaccine development for animal and human CoVs, the spike glycoprotein of SARS-CoV-2 was identified as the major antigenic target for neutralising antibodies and vaccine development [319]. As a result, various vaccine strategies were utilised leading to the development of >180 vaccine candidates [22, 320, 321]. In some cases, the production processes were simply adapted from those of existing vaccines or vaccine candidates, whilst relying on preclinical and toxicological data from related vaccines [320].

To assist with vaccine development research, we pivoted our existing research to explore the potential of using RV as an expression vector for SARS-CoV-2 spike peptides, thereby utilising the established RV vaccine platform to deliver a safe and effective vaccine.

1.10 Aims

The presence of a polycistronic segment 11 in RV species A, as well as the identification of alternative ORFs in other RV species, led us to hypothesise that transcription is initiated from alternative AUGs in other RV segments. The recent success of the plasmid-only reverse genetics system for RVs has led to the discovery of new functions of viral proteins, generation of reporter and attenuated viruses as well as candidates for next-generation RV vaccines. Thus, we hypothesised that the precise viral gene manipulation by the reverse genetics system could help identify potential accessory proteins encoded by the bovine RV strain RF.

- **The main aims of this thesis were:**

1. To establish a plasmid-only reverse genetics system for the bovine strain RF and compare the viral characteristics of rescued viruses to the parental isolate and the simian strain SA11 **(*Chapter 3*).**

2. Use the established reverse genetics system to express SARS-CoV-2 spike peptides from the NSP3 gene and test their immunogenicity **(*Chapter 4*).**

3. Identify potential accessory proteins encoded by the bovine strain RF genome and analyse their sequence conservation using bioinformatic analyses **(*Chapter 5*).**

4. Investigate the effects of mutation on the expression of identified accessory proteins in the context of virus rescue and replication *in vitro* **(*Chapter 5*).**

Chapter 2: Materials and Methods

2.1 Materials

2.1.1 General reagents

General plasticware for molecular experiments and tissue culture was supplied by Corning Incorporated, VWR, Sarstedt, Greiner Bio-One and Thermo Scientific.
The Roslin Institute Central Services Unit (CSU) prepared and provided distilled water and phosphate buffered saline (PBS) (Oxoid).

Table 2 Reagents and kits with their respective suppliers and catalogue numbers

Name	Supplier	Catalogue #
10x Tris/Glycine/SDS	Bio-Rad	1610732
1kb DNA Ladder	Promega	G5711
2-Mercaptoethanol	Sigma-Aldrich	M3148-25ML
40% acrylamide:bisacrylamide (19:1)	Sigma-Aldrich	A9926-100ML
4-20% Novex™ Tris-Glycine Mini Protein Gels (1 mm)	Invitrogen	XP04205BOX
Agarose (UltraPure™)	Invitrogen	16500500
Bovine Serum Albumin Fraction V	Merck	10735086001
Cellulose (colloidal, microcrystalline)	Sigma-Aldrich	435244-250G
EasyTag™ EXPRESS^{35}S Protein Labeling Mix, [^{35}S]-, 2mCi	Perkin Elmer	NEG772002MC
Lipofectamine 2000 transfection reagent	Invitrogen	11668019
MEGAscript™ T7 Transcription Kit	Invitrogen	AM1333
MinElute Gel Extraction Kit	Qiagen	28604
Monarch® PCR and DNA Cleanup Kit	New England Biolabs	T1030S
Neutral Buffered Formalin (NBF)	CellPath	BAF-0010-10A
Nitrocellulose membrane (0.2 µm)	Fisher Scientific	15249794
Nuclease Free Water	Qiagen	129114
Precision Plus Protein Dual Color Standard	Bio-Rad	1610374
ProLong™ Gold Antifade Mountant	Invitrogen	P36930
ProtoGel Resolving Buffer	Scientific Laboratory Supplies	NAT1268
ProtoGel Stacking Buffer	Scientific Laboratory Supplies	NAT1270

Qubit™ dsDNA Broad Range	Invitrogen	Q32850
Qubit™ RNA Broad Range	Invitrogen	Q10210
Qubit™ RNA High Sensitivity	Invitrogen	Q32852
QuikChange II Site-Directed Mutagenesis Kit	Agilent Technologies	200523
RNeasy® Mini Kit	Qiagen	74104
Silver Stain Plus Kit	Bio-Rad	1610443
Superscript™ III First-Strand synthesis super mix	Invitrogen	18080044
SYBR DNA gel stain	Invitrogen	S33102
Tetramethylethylenediamine (TEMED)	PanReac AppliChem	A1148, 0100
TNT® T7 Coupled Reticulocyte Lysate System	Promega	L4610
Toluidine Blue O	Sigma-Aldrich	89640-25G
Trypsin from porcine pancreas Type IX-S	Sigma-Aldrich	T0303-1G
Tween 20	Sigma-Aldrich	P9416-100mL
Western Blotting Filter Paper	Thermo Scientific	10675935
Whatman Grade 3MM Chr Blotting Paper	Scientific Laboratory Supplies	CHR1130
X-ray films	Fisher Scientific	10752067

2.1.2 Enzymes

The following enzymes were used according to manufacturer's instructions unless otherwise stated.

Table 3 Enzymes

Enzymes	Supplier	Catalogue #
DNA restriction enzymes	New England Biolabs Promega	Various
DpnI	Agilent Technologies	200523
GoTaq® G2 DNA polymerase	Promega	M7845
PfuUltra High-Fidelity DNA polymerase	Agilent Technologies	200523
Platinum™ *Taq* polymerase	Invitrogen	10966018
Q5® High-Fidelity DNA polymerase	New England Biolabs	M0491S
RQ1 RNase-Free DNase	Promega	M6101
Shrimp Alkaline Phosphatase (rSAP)	New England Biolabs	M0371S
T4 DNA ligase	Promega	M1801

2.1.3 Bacterial cells

Subcloning Efficiency™ DH5α Competent cells	Invitrogen (18265017)
DH10β Competent cells	Gift from Dr Christine Tait-Burkard
XL1-Blue Competent cells	Agilent Technologies (200523)

2.1.4 Bacterial media

- Ampicillin sodium salt (working concentration 100 µg/mL) — Cambridge BioScience (14417-25g-CAY)
- Carbenicillin disodium salt (working concentration 100 µg/mL) — FORMEDIUM LTD (CAR0005)
- Kanamycin sulphate salt (working concentration 50 µg/mL) — Scientific Laboratory Supplies (60615-5G)

Luria-Bertani Broth (LB) (Sigma-Aldrich, L3022-1kg), Luria-Bertani agar (LA) (Miller) (FORMEDIUM, LMM0202) and Hanahan's Super Optimal Broth (SOB) (Sigma-Aldrich, H8032-500g) were prepared and provided by the Roslin Institute CSU according to the following recipes:

- LB: 10 g/L tryptone, 5 g/L yeast extract, 5 g/L sodium chloride, pH 7.0.
- LA: 10 g/L tryptone, 5 g/L yeast extract, 5 g/L sodium chloride, 15 g/L agar pH 7.0.
- SOB: 0.186 g/L potassium chloride, 2.4 g/L magnesium sulphate, 0.5 g/L sodium chloride, 20 g/L tryptone, 5 g/L yeast extract, pH 7.0.

2.1.5 Mammalian cell lines

Table 4 Mammalian cell lines

Cell line	Source	Reference
Baby hamster kidney fibroblasts (BHK-21 cells) clone BSR-T7 expressing T7 RNA polymerase	Gift from Professor Massimo Palmarini (MRC-University of Glasgow Centre for Virus Research, UK)	(Buchholz et al., 1999) [322]
African Green monkey kidney epithelial (MA104) cells	Gift from Professor Richard Randall (University of St Andrews, UK)	-
Human Embryonic Kidney 293T cells, containing SV40 T-antigen (HEK293T)	American Type Culture Collection (ATCC)	(DuBridge *et al.*, 1987) [323]

2.1.6 Cell culture reagents

Table 5 Cell culture reagents

Name	Supplier	Catalogue #
Dulbecco's Modified Eagle Medium (DMEM)	Sigma-Aldrich	D5796
Foetal Bovine Serum (FBS)	Gibco	
Geneticin (G-418 solution)	Scientific Laboratory Supplies	4727878001
Glasgow's Minimum Essential Medium (GMEM)	Gibco	G5154-500ML
IntestiCult™ Organoid Growth Medium (Mouse)	STEMCELL Technologies	06005
L-Glutamine	Sigma-Aldrich	59202C-100ML
Corning® Matrigel	Scientific Laboratory Supplies	354234
Opti-MEM reduced serum medium	Gibco	11058021
Penicillin/Streptomycin solution (10,000 U/mL)	Gibco	15140122
Tryptose phosphate broth	Life Technologies	18050039
0.25% Trypsin-EDTA	Gibco	T3924-100ML

2.1.7 Drugs and inhibitors

• MG-132 (10 mM in DMSO)	Sigma-Aldrich (M7449-200µL)
• ROCK pathway inhibitor	Cambridge Biosciences (Y-27632)
• Galunisertib	Cambridge Biosciences (LY2157299)
• p38 inhibitor	Enzo Life Sciences (SB202190)

2.1.8 Cell culture media composition

- Complete medium: DMEM supplemented with 10% (v/v) heat-inactivated FBS and 1% (v/v) Penicillin/Streptomycin.
- Serum free medium (SFM): DMEM supplemented with 1% (v/v) Penicillin/Streptomycin.
- Virus growth medium (VGM): DMEM supplemented with 1% (v/v) Penicillin/Streptomycin and 0.5 µg/mL porcine pancreatic trypsin type IX.
- Organoid Growth Medium: IntestiCult™ Organoid Growth Medium supplemented with 10 µM ROCK pathway inhibitor, 10 µM Galunisertib and 55 nM p38 inhibitor.

2.1.9 Antibodies and dyes

Table 6 Primary antibodies

Name	Clonality	Host species	Dilution	Application	Supplier	Catalogue #
Anti-Rotavirus VP6 [A2], IgG2b	Monoclonal	Mouse	1:1000	WB, IF	Abcam	ab181695
Alpha-Tubulin, IgG2a	Monoclonal	Rat	1:5000	WB	Novus Biologicals	NB600-506
Anti-beta Actin, IgG	Polyclonal	Rabbit			Abcam	ab8227
Anti-GFP (JL8)	Monoclonal	Mouse	1:1000		Takara Bio Europe	632381

Table 7 Secondary antibodies

Name	Conjugate	Dilution	Application	Supplier	Catalogue #
Donkey anti-mouse IgG (H+L)	Alexa Fluor® 488	1:1000	IF	Invitrogen	A-21202
Donkey anti-mouse IgG (H+L)	IRDye® 800CW	1:10,000	WB	LI-COR	926-32212
Goat anti-rat IgG (H+L)	IRDye® 680RD				926-68029
Goat anti-rabbit IgG (H+L)	IRDye® 680LT				926-68071

Table 8 Fluorescent dyes

Name	Dilution	Application	Supplier	Catalogue #
DAPI (Nucleus)	1:5000	IF	Invitrogen	D3571
Hoechst (Nucleus)	1:5000			H3570
Phalloidin (F-actin)	1:100			A12379

2.1.10 General buffers and homemade gels

The following buffers and gels were prepared in-house:

- 1X SDS-PAGE running buffer: 25 mM Tris-HCl, 192 mM glycine, 0.1% (w/v) SDS, pH 8.3.
- Protein transfer buffer: 25 mM Tris-HCl, 192 mM glycine, 20% (v/v) methanol.
- Gel fixing solution: 50% (v/v) methanol, 10% (v/v) acetic acid.
- Washing buffer: Tris-buffered saline (20 mM Tris-HCl, 150 mM NaCl, pH 7.6) / 0.1% (v/v) Tween 20 (TBS-T).
- Blocking buffer: TBS-T, 5% (w/v) skimmed milk.
- Antibody diluent: 5% BSA/TBS-T.
- Tris-borate-EDTA (TBE) buffer: 89 mM Tris-borate, 2 mM EDTA, pH 8.3.
- 2X Laemmli sample buffer: 65.8 mM Tris-HCl [pH 6.8], 100 mM DTT [pH 6.8], 2.1% sodium dodecyl sulphate (SDS), 26.3% (w/v) glycerol, 0.01% bromophenol blue.
- Calcium manganese-based (CCMB) buffer: 10 mM potassium acetate, 10% (v/v) glycerol, 80 mM $CaCl_2$, 20 mM $MnCl_2$, 10 mM $MgCl_2$, pH 6.4.
- 10% resolving polyacrylamide gel: 3.3 mL 40% acrylamide:bisacrylamide (19:1), 2.5 mL 4X resolving buffer, 4.2 mL water, 100 µL 10% (w/v) Ammonium persulfate, 10 µL TEMED.
- 4% stacking polyacrylamide gel: 1.3 mL 40% acrylamide:bisacrylamide (19:1), 2.5 mL 4X stacking buffer, 6.2 mL water, 50 µL 10% (w/v) Ammonium persulfate, 10 µL TEMED.

Tris-acetate EDTA (TAE) buffer was prepared and supplied by the Roslin Institute CSU according to the following recipe:

- TAE buffer: 40 mM Tris, 20 mM acetic acid, 1 mM EDTA.

2.1.11 Plasmids

SA11 pT7 plasmids were constructed by Takeshi Kobayashi [284] and were purchased from Addgene with the corresponding ID numbers listed in **Table 9**.

Table 9 SA11 RV plasmids

Gene	Plasmid Name	Addgene ID#
VP1	pT7-VP1SA11	#89162
VP2	pT7-VP2SA11	#89163
VP3	pT7-VP3SA11	#89164
VP4	pT7-VP4SA11	#89165
VP6	pT7-VP6SA11	#89166
VP7	pT7-VP7SA11	#89167
NSP1	pT7-NSP1SA11	#89168
NSP2	pT7-NSP1SA11	#89169
NSP3	pT7-NSP1SA11	#89170
NSP4	pT7-NSP1SA11	#89171
NSP5	pT7-NSP1SA11	#89172

Eleven RF T7 plasmids were constructed to express the individual full-length RF gene segments flanked by the T7 promoter and HDV ribozyme, followed by the T7 RNA polymerase terminator sequences. The constructs were synthesised by Invitrogen GeneArt using either pMK-RQ (KanR), pMA-RQ (AmpR) or pMA-T (AmpR) vectors. Plasmids were transformed into chemically competent *E.coli* DH5α or DH10β cells (***section 2.2.8***) and purified using the QIAGEN® Plasmid Midi Kit according to the manufacturer's protocol (***section 2.2.9***).

2.1.12 Viruses

The Bovine RV strain RF (G6P6[1]) was a kind gift from Dr Ulrich Desselberger, University of Cambridge.

2.1.13 Oligonucleotides

Table 10 Oligonucleotides used for sequencing of plasmids and viruses

Primer Name	Sequence (5' – 3')	Application
RF_FW_VP1whole	GACATGCTAGCCAATATGA	Sequence RF_VP1 gene
RF_REV_VP1whole	CTGGATAACTCTAGCATCA	
RFVP1_FW_whole#2	GGAAGGAGAGATGTACCAGGA	
T7_FW_sequencing	TAATACGACTCACTATAGG	Sequence 5' end of all RF plasmids
HDV_REV	TAGCCATCCGAGTGGACGA	Sequence 3' end of all RF plasmids
pEGFP-N1_FW	GGGCTAGCGGCTATTAAAGCTATACAATGGGG	Clone VP1 amplicon into pEGFP-N1
pEGFP-N1_REV	GGACCGGTGCAACTGACATCAGCATTTCAATTGAA	
SARS_CoV2_RFNSP3_FW	GCTTTTCAGTGGTTGATGCT	Sequence RF_NSP3 mutant viruses
SARS_CoV2_RFNSP3_REV	TCATAGAGGGTCATGTGAAG	

Table 11 Oligonucleotides used for site-directed mutagenesis of RF_VP1

Primer Name	Mutated nucleotide (amino acid change)	Sequence (5' – 3')
RF_VP1_F_406_ATG - CTG	A406C (M130L)	TGATTCATTGCTGGATCCAGC
RF_VP1_R_406_ATG - CTG		GCTGGATCCAGCAATGAATCA
RF_VP1_F_454_ATG - CTG	A454C (M146L)	AAATGCAGTTCTGTTCTGGTTG
RF_VP1_R_454_ATG - CTG		CAACCAGAACAGAACTGCATTT
RF_VP1_F_604_ATG - CTG	A604C (M196L)	ATATGAAGTACTGAAAGATAAGCCG
RF_VP1_R_604_ATG - CTG		CGGCTTATCTTTCAGTACTTCATAT
RF_VP1_F_655_ATG - CTG	A655C (M213L)	CTTCAATTGAACTGCTGATGTCA
RF_VP1_R_655_ATG - CTG		TGACATCAGCAGTTCAATTGAAG
RF_VP1_F_406_ATG - GTG	A406G (M130V)	TGATTCATTGGTGGATCCAGC
RF_VP1_R_406_ATG - GTG		GCTGGATCCACCAATGAATCA
RF_VP1_F_454_ATG - GTG	A454G (M146V)	AAATGCAGTTGTGTTCTGGTTG
RF_VP1_R_454_ATG - GTG		CAACCAGAACACAACTGCATTT
RF_VP1_FW_196_ATG - CTG	A196C (M66L)	TAGCGATGTTCTGGAGAATGC
RF_VP1_REV_196_ATG - CTG		GCATTCTCCAGAACATCGCTA
FW_T30G_stop_codon_RF_VP1	Introducing stop codon after canonical AUG	TGGGGAAGTAGAATCTAATCTTG
REV_T30G_stop_codon_RF_VP1		CAAGATTAGATTCTACTTCCCCA

2.2 Molecular techniques

2.2.1 Polymerase Chain Reaction (PCR)

PCR was used for cloning, colony screening and sequencing. Different size fragments were amplified using primers listed in **Table 10**. PCR reactions were set up according to the manufacturer's protocol and carried out in a T100™ Thermocycler (Bio-Rad).

Each 25 µL PCR reaction included: 1X reaction buffer (1.5 - 2 mM $MgCl_2$), 10 µM of each primer, 10 mM dNTP mix and 1.25 units of either Q5® High-Fidelity or GoTaq® G2 DNA Polymerase. PCR conditions for Q5® High-Fidelity DNA polymerase included an initial denaturing step of 30 sec at 98°C, followed by 30 cycles of 10 sec at 98°C, 21 sec at 65°C and 30 sec at 72°C (20-30 sec/kb), finishing with a 2 min incubation at 72°C. PCR conditions for GoTaq® G2 DNA Polymerase included a 2 min denaturing step at 94°C, followed by 30 cycles of 18 sec at 94°C, 21 sec at 50°C and 1 min at 72°C (1 min/kb), finishing with a 5 min incubation at 72°C. The PCR product length was confirmed by 0.8-1% (w/v) agarose gel electrophoresis (***section 2.2.3***).

2.2.2 Site-directed mutagenesis

Primers for site-directed mutagenesis were designed using New England Biolabs NEBaseChanger online tool (https://nebasechangerv1.neb.com/) with the following parameters: primer length 18-24 bp, 40-60% GC content, T_m of 50–65°C, start and end with 1-2 G/C pairs. Self-dimerization, hairpin formation and self-annealing of the primer pairs was checked using an online oligonucleotide properties calculator http://biotools.nubic.northwestern.edu/OligoCalc.html [324]. Primers purified by high performance liquid chromatography were ordered from Integrated DNA technologies or Merck.

Site-directed mutagenesis was performed on the RF VP1 plasmid using the QuikChange II Site-Directed Mutagenesis Kit with primers listed in **Table 11** according to the manufacturer's instructions but using half volume reactions. A 25 µL reaction mix contained: 2.5 µL 10X Reaction Buffer, 50 ng of DNA template, 0.6 µL (10 µM) of each primer, 0.5 µL dNTP mix, 0.5 µL *PfuUltra* High-Fidelity DNA polymerase made up to 25 µL with H_2O. Thermal cycling parameters were: 2 min denaturing at 95°C, followed by 18 cycles of 30 sec denaturing at 95°C, 1 min primer annealing at 55°C, 6 min elongation at 68°C, with final 10 min elongation at 68°C.

PCR products were digested using 1 µL of DpnI restriction enzyme to remove parental methylated DNA. Products were visualised using gel electrophoresis (***section 2.2.3***) before transformation into competent *E.coli* cells (***section 2.2.8***).

2.2.3 Gel electrophoresis

Gel electrophoresis was used to separate DNA on agarose gels at final concentrations between 0.8 – 1%. Agarose was dissolved in 1X TAE buffer and 1X SYBR Safe DNA Gel Stain was added to cooled molten agarose. Agarose was poured into a gel tray containing well combs and left to solidify. DNA samples (plasmid DNA = 200 ng each, PCR = 5 µL of reaction) were mixed with DNA loading dye (1X final concentration) and loaded into the set gel alongside a 1kb DNA Ladder. Gels were run at 100V for 30 min, visualised using the Odyssey® XF imaging system (LI-COR) and analysed with Image Studio™ Lite software (LI-COR).

2.2.4 Cloning

PCR amplicons or 3 µg of plasmid DNA were digested using 1 unit of restriction enzyme. Enzymatic reactions were prepared according to the manufacturer's protocol using an appropriate buffer and incubation temperature, and digested for at least 3 hr. To prevent religation of linearised plasmids, 1 unit of rSAP enzyme was added to the mixture to dephosphorylate the 5' and 3' ends of DNA for the duration of digestion according to the manufacturer's protocol. Digested products were purified using the Monarch® PCR and DNA clean-up kit (***section 2.2.5***) and ligation reactions were set up with the T4 DNA ligase enzyme according to the manufacturer's protocol. The ligation reaction mix included: 100 ng of linearised vector DNA, insert DNA at a molar ratio of 3:1 (insert:vector), 1X ligase buffer and 1 unit of T4 DNA ligase enzyme in a total volume of 10 µL. The reaction mixes were left overnight at 16°C before 3 µL of the ligation reaction was transformed into in-house made DH5α *E.coli* strain competent cells (***section 2.2.8***).

2.2.5 DNA and PCR clean-up

PCR amplicons or restriction digested DNA fragments were purified using the Monarch® PCR and DNA clean-up kit according to the manufacturer's protocol. DNA was eluted in 10 µL nuclease free water and purified DNA was quantified and checked for purity by spectrophotometer (DeNovix, DS-11 FX).

2.2.6 DNA extraction from agarose gels

PCR amplicons or digested plasmid DNA fragments (insert or vector) were separated by gel electrophoresis (***section 2.2.3***), visualised with a UV transilluminator and excised with a scalpel. DNA was extracted from agarose gel using the MinElute Gel Extraction Kit according

to the manufacturer's protocol. Gel slices were dissolved at 50°C for 10 min in a buffer containing a pH indicator that allows easy determination of the optimal pH for DNA binding. The solution was applied to the MinElute spin column containing the silica-gel membrane which binds nucleic acids under high-salt conditions provided by the buffer. Impurities were washed away by an ethanol-based wash buffer according to manufacturer's instructions and DNA was eluted in 10 µL nuclease free water. DNA concentration and purity were assessed by spectrophotometer (DeNovix, DS-11 FX).

2.2.7 Preparation of chemically competent bacterial cells

DH5α strains of competent *E.coli* cells were plated on an antibiotic free agar plate and grown overnight at 37°C. One colony was inoculated into 150 mL of antibiotic free LB and grown overnight in a 37°C shaker (180 rpm). Overnight culture was diluted 1:100 in either SOB or LB (600 µL into 60 mL) and incubated in a 37°C shaker for a further 3 hr until OD_{550} reached ~0.3. OD was read in duplicate using 100 µL of culture in a CLARIOstar plate reader (BMG Labtech). After a 10 min incubation on ice, bacterial cells were harvested by centrifugation at 2857 x *g* for 10 min at 4°C. Cell pellets were resuspended in either cold sterile CCMB (SOB cultures) or 100 mM $CaCl_2$ with 15% glycerol (LB cultures) in ¼ of the initial volume (15 mL). Following a 2 hr incubation on ice, bacterial cells were pelleted as before and resuspended again in either cold sterile CCMB (SOB cultures) or 100 mM $CaCl_2$ with 15% glycerol (LB cultures) this time in 1/12 of the starting volume (5 mL). 100 µL cell suspensions were aliquoted into prechilled Eppendorf tubes, snap frozen using a methanol-dry ice bath and stored at -80°C. Cells were tested for transformation competency before use (***section 2.2.8***).

2.2.8 Bacterial transformation

Commercial (XL1-Blue competent cells) or in-house DH5α or DH10β competent cells were used for transformation with plasmid DNA. Briefly, 250 ng of plasmid DNA or 1 µL of site-directed mutagenesis product was mixed with 50 µL of competent cells and incubated on ice for 30 min. Cells were heat shocked for 45 sec at 42°C to facilitate the uptake of DNA by the bacteria, followed by a 2 min incubation on ice. 250 µL of pre-warmed SOB (without antibiotic) was added to cells and incubated for 1 hr in a 37°C shaker (180 rpm) allowing bacteria to recover and grow. Cells were then plated onto the appropriate antibiotic selection agar plates using a sterile spreader and incubated at 37°C overnight. Single colonies were selected and grown in LB containing the appropriate antibiotic overnight in a 37°C shaker (180 rpm).

2.2.9 Plasmid DNA extraction and quantification

Plasmid DNA was prepared using the QIAGEN plasmid purification kits (mini-prep or midi-prep) according to the manufacturer's protocol. Small-scale plasmid DNA purification was prepared from 3 mL of overnight *E.coli* cultures (mini-prep), whereas for the larger-scale plasmid DNA purification 50 mL of overnight cultures were used (midi-prep). DNA concentration was determined using the Qubit™ DNA Broad Range kit according to the manufacturer's protocol. Purity of the DNA was assessed by spectrophotometer (DeNovix, DS-11 FX) through analysis of the A_{260}/A_{280} and A_{260}/A_{230} ratios which indicate possible contamination with proteins, chaotropic salts or non-ionic detergents.

2.2.10 DNA sequencing

Following plasmid DNA purification, cloning, site-directed mutagenesis or PCR, DNA was sent for Sanger sequencing at Genewiz (Germany). In a total volume of between 10-20 µL, plasmid (30-100 ng/µL) or PCR amplicons (10-50 ng/µL) were sequenced with primers designed to target the specific region as listed in **Table 10**. Sequencing results were analysed in SSE v1.3 software [325].

2.3 Cell culture

2.3.1 Subculturing of cell lines

MA104 and HEK293T cells were cultured in DMEM supplemented with heat inactivated 10% FBS and 1% penicillin-streptomycin. BSR-T7 cells were cultured in GMEM supplemented with 1% tryptose phosphate broth, heat inactivated 10% FBS, 1% L-glutamine and 1% penicillin-streptomycin. At every fifth passage, the G-418 selection drug (1 mg/mL) was added to BSR-T7 cell media.

Subculturing of all cell lines was performed twice a week and cells were maintained at 37°C, 5% CO_2. Following removal of the complete medium, adherent cells were washed with PBS and incubated with 0.25% Trypsin-EDTA at 37°C, 5% CO_2 until cells detached from flask (5 mL PBS and 3 mL trypsin for 75 cm^2 flasks; 10 mL PBS and 5 mL trypsin for 175 cm^2 flasks). Trypsin was inactivated by adding fresh complete medium and a fraction of the cell suspension was transferred to a clean flask containing fresh complete medium (total volumes: 14 mL for 75 cm^2 and 25 mL for 175 cm^2).

2.3.2 Cell counting

Cell concentration was calculated using the Neubauer improved haemocytometer counting chamber (Hawksley). Cell suspension was mixed with 0.4% trypan blue at a 1:1 ratio and 10 µL was applied to the haemocytometer, filing one side of the chamber underneath the glass coverslip. Using a light microscope, the number of live cells (unstained) were counted in each set of 16 squares (four sets in total) with a 10X objective. The number of viable cells/mL in the original cell suspension was calculated using the following formula: average cell count from four sets of 16 squares x dilution factor x 10,000.

2.3.3 Plasmid transfection of mammalian cells

At 70-80% confluency, HEK293T cells in a 24-well plate were transfected with plasmid DNA using the Lipofectamine 2000 reagent according to the manufacturer's protocol. 500 ng of plasmid DNA and 1 µL of Lipofectamine reagent were separately diluted in 50 µL Opti-MEM™. After a 5 min incubation at room temperature, diluted mixes were combined and incubated for a further 25 min. During this time, complete cell culture medium was changed to Opti-MEM™ (200 µL/well). The transfection mix was added to cells dropwise and cells were incubated for 48 hr at 37°C, 5% CO_2. Following the 48 hr incubation, supernatant was removed and lysates harvested in 100 µL 2X Laemmli buffer were used to analyse protein expression by western blotting (***section 2.5.2***).

2.4 Virus rescue, titration and growth

2.4.1 Virus rescue using reverse genetics

Viruses were rescued using protocols described by Kanai *et al.,* (2017) [284] and Komoto *et al.*, (2018) [288] with slight modifications. At 70% confluency, BSR-T7 cells in 6-well plates were transfected with 11 plasmids each corresponding to one RV genome segment (2.5 µg for plasmids encoding NSP2 and NSP5; 0.8 µg for the remaining plasmids) along with plasmids encoding two VV capping enzyme subunits (pCAG-D1R and pCAG-D12L - 0.8 µg each) using 16 µL Lipofectamine 2000 per transfection reaction in a total volume of 200 µL of Opti-MEM. After 24 hr incubation at 37°C 5% CO_2, transfected cells were co-cultured with MA104 cells ($1x10^5$ cells/well) for three days in VGM. Co-cultured cells were then lysed three times by freeze/thaw and lysates were incubated for 30 min at 37°C 5% CO_2 with porcine pancreatic trypsin type IX at a final concentration of 10 µg/mL. Lysates were added to fresh confluent MA104 cells in 25 cm^2 flasks and incubated at 37°C 5% CO_2 for 1 hr. After adsorption, MA104 cells were washed with SFM and cultured in VGM for up to seven days or until complete cytopathic effect (CPE) was observed. When CPE was observed, cells were lysed three times by freeze/thaw, cellular debris was precipitated by centrifugation at 2737 x *g* for 30 min at 4°C and virus containing supernatants (P1 stocks) were aliquoted and stored at -80°C then titrated by plaque assay (***section 2.4.2***).

2.4.2 Virus titration by plaque assay

MA104 cells were seeded in 12-well ($1x10^5$ cells/well) or 6-well plates ($3x10^5$ cells/well) using 1-2 mL of complete medium and incubated overnight at 37°C 5% CO_2. Confluent monolayers of MA104 cells were washed with SFM and infected with 400 µL (12-well plates) or 800 µL (6-well plates) of ten-fold serially diluted virus for 1 hr at 37°C 5% CO_2. Following virus adsorption, 1 mL/well (12-well plates) or 2 mL/well (6-well plates) overlay medium was added (1:1 ratio of 2.4% cellulose and VGM) and incubated undisturbed for three (SA11 RV) or four (RF RV) days. Cells were fixed with 1 mL/well of 10% NBF for 1 hr and stained with 0.1% Toluidine blue (dissolved in H_2O) for 1 hr. The plaque-forming unit (PFU) was calculated according to the formula below:

$$\frac{\textit{Average number of plaques}}{\textit{Dilution factor x Volume of diluted virus (mL)}} = \textit{PFU/mL}$$

2.4.3 Measurement of plaque diameter

Plaque assay plates were scanned using an Epson Perfection V750 Pro Scanner at a resolution of 1200 dpi. In Photoshop v24.5.0, images were converted to greyscale and a level 4 posterisation was applied to each image. The area of 40 plaques per virus rescue was measured using ImageJ analysis software.

2.4.4 Viral RNA extraction, RT-PCR and sequencing

RNA was purified from P1 viral stocks using the spin protocol of the QIAamp Viral RNA Kit followed by RQ1 RNase-Free DNase treatment to remove possible DNA contamination according to the manufacturer's protocol. The extracted RNA was dissolved in 30 µL nuclease-free water and stored at -80°C.

cDNA was synthesised with Superscript™ III First-Strand synthesis super mix containing SuperScript™ III Reverse Transcriptase using 5 µL RNA and 1 µL random hexamer primers (10 µM) according to the manufacturer's protocol.

PCR reactions were set up to amplify the gene of interest using primers listed in **Table 10** and carried out according to the manufacturer's instructions in a T100™ Thermal Cycler (Bio-Rad). Each 25 µL PCR reaction included: 1X PCR buffer (-$MgCl_2$), 1.5 mM $MgCl_2$ (DNA-free), 0.2 µM of each primer, 0.2 mM dNTP mix and 1.25 units of Platinum™ Taq DNA polymerase. PCR conditions included an initial denaturing step of 5 min at 95°C, followed by 30-35 cycles of 30 sec at 95°C, 30 sec at 55°C and 2 min at 72°C, finishing with a 5 min incubation at 72°C. The PCR product length was confirmed by 0.8% (w/v) agarose gel electrophoresis (***section 2.2.3***). PCR product was sent for sequencing at Genewiz (Germany) using primers listed in **Table 10**.

2.4.5 Multi-cycle viral growth kinetics

Confluent monolayers of MA104 cells in 24-well plates were infected in technical triplicate with viruses at an MOI of 0.03 PFU/cell (low MOI) in 500 µL SFM. After 1 hr incubation at 37°C, 5% CO_2 (which allowed virus attachment to the cells) the inoculum was removed, cells were washed with SFM and 500 µL of fresh VGM was added to each well. Immediately thereafter, supernatant from one well was harvested (1 hour post infection (hpi)) to provide a measurement for the amount of residual input virus. At different times post infection, supernatants were harvested and stored at -80°C until titrated by plaque assay (***section 2.4.2***).

2.4.6 Electrophoretic analysis of dsRNA

Virions were purified from RF RV stocks using ultracentrifugation as described previously [82]. Briefly, 25 mL of clarified RV stock was pelleted through a buffered 25% (w/v) sucrose cushion (100 mM NaCl, 10 mM Tris-HCl [pH 7], 1 mM EDTA) using a SW32Ti rotor in a Beckman Coulter Optima Max-E ultracentrifuge at 106,750 x *g* for 135 min at 4°C. Viral pellets were resuspended in 350 µL of RLT buffer containing 3.5 µL 2-Mercaptoethanol and dsRNA was extracted using the RNeasy® Mini Kit according to the manufacturer's protocol. Purified RNA was eluted in 30 µL nuclease free water and 8 µL of RNA was then treated with RQ1 RNase-Free DNase to remove possible DNA contamination according to manufacturer's protocol.

Extracted dsRNA was resolved according to size using 5% urea polyacrylamide gels in 1X TBE buffer. Gels were fixed in gel fixing solution and RNA was visualised using the Silver Stain Plus Kit according to the manufacturer's instructions. Gels were dried onto Whatman Grade 3MM blotting paper and imaged using a Samsung Xpress C480FW scanner.

2.5 Protein detection

2.5.1 SDS-PAGE

Proteins were separated by molecular mass using SDS-PAGE on 4-20% Novex™ Tris-Glycine Mini Protein Gels (1 mm).

Lysate samples were mixed with 2X Laemmli buffer and boiled at 95°C for 10 min, vortexed and briefly centrifuged at 16,602 x *g*. Each lane was loaded with 5-15 µL of lysate and 5 µL of the Precision Plus Protein Dual Color Standard as a guide for protein size. Gels were initially run at a constant voltage of 80V for 20-30 min for proteins to pass through the stacking gel after which the voltage was increased to 100V for the desired length of time.

2.5.2 Western blotting

Following SDS-PAGE (***section 2.5.1***), proteins were transferred onto a 0.2 µm nitrocellulose membrane by wet transfer using an in-house made protein transfer buffer (***section 2.1.10***). Membranes were blocked with 5% milk in TBS-T for 1 hr at room temperature followed by three 5 min washes with TBS-T. Membranes were then probed with primary antibodies (**Table 6**) diluted in 5% BSA/TBS-T overnight at 4°C with agitation. After three 5 min washes with TBS-T, membranes were incubated for 1 hr at room temperature with secondary antibodies diluted in TBS-T (**Table 7**). Following three 5 min washes with TBS-T, membranes were imaged using Odyssey ® XF imaging system (LI-COR). Analysis was performed with Image Studio™ Lite software (LI-COR).

2.6 Radioactive isotope experiments

2.6.1 *In vitro* transcription and translation

Coupled *in vitro* transcription and translation reactions were carried out using the Promega TnT® Coupled Reticulate Lysate System supplemented with radioactive ^{35}S protein labelling mix according to the manufacturer's protocol. Briefly, reactions were set up as follows: 8 µL TnT mix, 1 µCi ^{35}S protein labelling mix, 200 ng plasmid DNA, made up to 10 µL with H_2O. Mixes were prepared on ice before reactions were incubated at 30°C for 90 min. The reactions were denatured in 2X Laemmli buffer and boiled for 10 min at 95°C. Samples were analysed using SDS-PAGE (***section 2.5.1***) and autoradiography (***section 2.6.2***).

2.6.2 Autoradiography of dried polyacrylamide gels

SDS-PAGE was performed as described in ***section 2.5.1*** using homemade gels prepared according to the recipes in ***section 2.1.10***. Polyacrylamide gels were cast using 1 mm spacer plates with appropriate short plates and a Bio-Rad casting apparatus. The resolving gel was made using ProtoGel buffer according to the manufacturer's protocol with varying concentration of 40% acrylamide:bisacrylamide (19:1) depending on the size of the target product. Stacking gels were always cast at 4%. Gels were fixed in gel fixing solution (***section 2.1.10***) on a rocker for 45 min with the gel fixing solution being replaced every 15 min. Fixed gels were placed onto Whatman Grade 3MM blotting paper, covered with cling film and dried in a gel dryer (Model 543, Bio-Rad) by heating up to 80°C for 2 hr under vacuum pressure. Dried gels were placed in a sealed cassette with an X-ray film overnight. X-ray films were developed using an Optimax® 2010 NDT automated X-ray film processor (PROTEC) following the manufacturer's protocol.

2.7 Bovine enteroids

2.7.1 Infection of bovine enteroids

Infection of organoids was carried out as described by Derricott *et al.,* (2019) [326], with modifications. Basal-out three dimensional (3D) bovine organoids were mechanically disrupted into multicellular fragments using a p200 pipette to expose the apical surface of the cells in 80% of the enteroids. The sheared enteroids were diluted to the appropriate concentration in organoid growth medium aliquoted into 15 mL falcon tubes. The enteroid fragment suspension was infected with rescued RF (rRF) at an MOI of 10 (high MOI) and incubated for 1 hr at 37°C 5% CO_2. After 1 hr, enteroid/virus suspension was centrifuged at 40 x *g* for 2 min and washed five times by replacing the supernatant with equal volumes of fresh IntestiCult™ (depending on the volume of virus used). Following washes, the enteroid pellet was resuspended in the appropriate volume of IntestiCult™ (depending on the volume of virus used) and plated at 500 µL onto a Corning® Matrigel coated pre-warmed 24-well plate in the presence and absence of coverslips. After 1 hr, 24 hr and 48 hr post infection, supernatant was collected for plaque assays while the cell lysates were used to test different parameters such as mechanical disruption and freeze/thaw.

2.7.2 Immunofluorescent staining of bovine enteroids

Enteroids were grown on coverslips in 24-well plates and infected with rRF and rSA11 at an MOI of 10. After 24 hr, enteroids were fixed with 4% paraformaldehyde for 1.5 hr at 4°C with agitation. Following three washes with PBS, enteroids were permeabilised with 0.5% Triton-X100 in PBS for 15 min and then blocked with 2% horse serum in PBS for 1 hr. Enteroids were incubated with primary antibodies (**Table 6**) overnight at 4°C with agitation. After 24 hr, enteroids were washed three times with PBS and incubated with secondary antibodies (**Table 7**) and phalloidin (**Table 8**). All antibodies were diluted in 2% horse serum/PBS. Enteroids were then washed with PBS three times with the addition of DAPI (**Table 8**) for the final 10 min. Coverslips were rinsed in water and mounted onto microscope slides with ProLong™ Gold Antifade Mountant and imaged using a Zeiss LSM 710 confocal microscope at x630 magnification. Images were analysed using the Zen Black software and processed using Photoshop v24.5.0.

2.8 Bioinformatic analyses

Bioinformatic analyses were performed in collaboration with Dr Samantha Lycett (The Roslin Institute, University of Edinburgh) and Mr Marius Diebold. VP1 nucleotide sequences were downloaded from the National Center for Biotechnology Information Viral Genome resource [327] and aligned using the SSE v1.3 software [325]. Poor quality and incomplete sequences were disregarded. AUG codon preservation and their respective Kozak context [328] was analysed using R scripts, results from which were further analysed in Microsoft Excel. Initial and final R scripts were written by Dr Samantha Lycett and Mr Marius Diebold respectively (***Appendix A, section 8.1***).

2.9 Statistical analysis

GraphPad Prism v9 was used for all statistical analyses. Data are presented as mean and standard deviation from three independent experiments unless otherwise stated. Specific statistical tests were selected for different analyses and are detailed in each figure legend. *P* values were denoted as follows:

- *p-value* > *0.05* - non-significant (ns)
- *p-value* < *0.05* – significant

2.10 Structural modelling

The previously resolved structure of RV polymerase was obtained from the RCSB Protein Data Bank (references: 6OGY and 6OGZ) [107]. Protein predictions for rRF, M130 and M146 mutants were performed in collaboration with Dr Barbara Shih (Lancaster University) using Alpha fold [329] with specific parameters detailed in ***Appendix B, section 8.2***. All protein models were analysed and visualised using PyMOL (https://pymol.org/2/) [330].

Chapter 3: Establishing a reverse genetics system for the bovine rotavirus strain RF

3.1 Background and Aims

A reverse genetics system is an important technique for studying virus replication and pathogenesis through direct and controlled manipulation of the viral genome. RNA viruses, such as IAV and SARS-CoV-2, have been extensively investigated using the reverse genetics approach which has led to discovery of new functions of viral proteins as well as development of new vaccines [331, 332].

Previously, functional analysis of RV genes relied on the isolation of natural variants and reassortants, synthetic RNA transcripts, or helper-virus dependent reverse genetics systems [263, 274, 277, 282, 295]. In 2017, based on a previously validated approach with orthoreoviruses, Kanai *et al.*, (2017) developed an entirely plasmid-based reverse genetics system for the simian RV strain SA11 [267, 268, 284]. Since then, there have been several reports of modifications and improvements to the original protocol, enabling rescue of different RV strains. To date, these include simian strains (SA11 G3P[2] and RRV G3P[3]), human strains (KU G1P[8], Odelia G4P[8] and CDC-9 G1P[8]) and murine strain D6/2 [284, 289, 292, 333]. The availability of reverse genetics systems for different strains will help to validate and compare any observed phenotypic change arising as a result of gene manipulation.

Utilizing the optimised 11-plasmid system reported by Komoto *et al.*, (2018) [288], this chapter aimed to establish an analogous plasmid-only reverse genetics system for the bovine RV strain RF. This is advantageous as one of the currently licenced vaccines against RV, RotaTeq, utilises a bovine RV backbone reassorted with different human strains [334]. Sequence alignment and analysis of RV gene segment 4 (VP4) was performed to compare the cell culture adapted RF strain to the original RF clinical isolate and to the bovine strain WC3 used in the RotaTeq vaccine (Figure 3.1 and Table 12) [335, 336]. Despite being cell culture adapted, RF displayed high homology to the original clinical isolate, whilst sharing a lower identity and similarity to the WC3 strain at nucleotide and amino acid levels (Figure 3.1 and Table 12). Nonetheless, the RF strain is well characterised, grows easily in culture and has been used as a prototype to study RV pathogenesis *in vitro*. Additionally, a reverse genetics system for the bovine strain will be a more biologically relevant model for understanding RV pathogenesis in livestock.

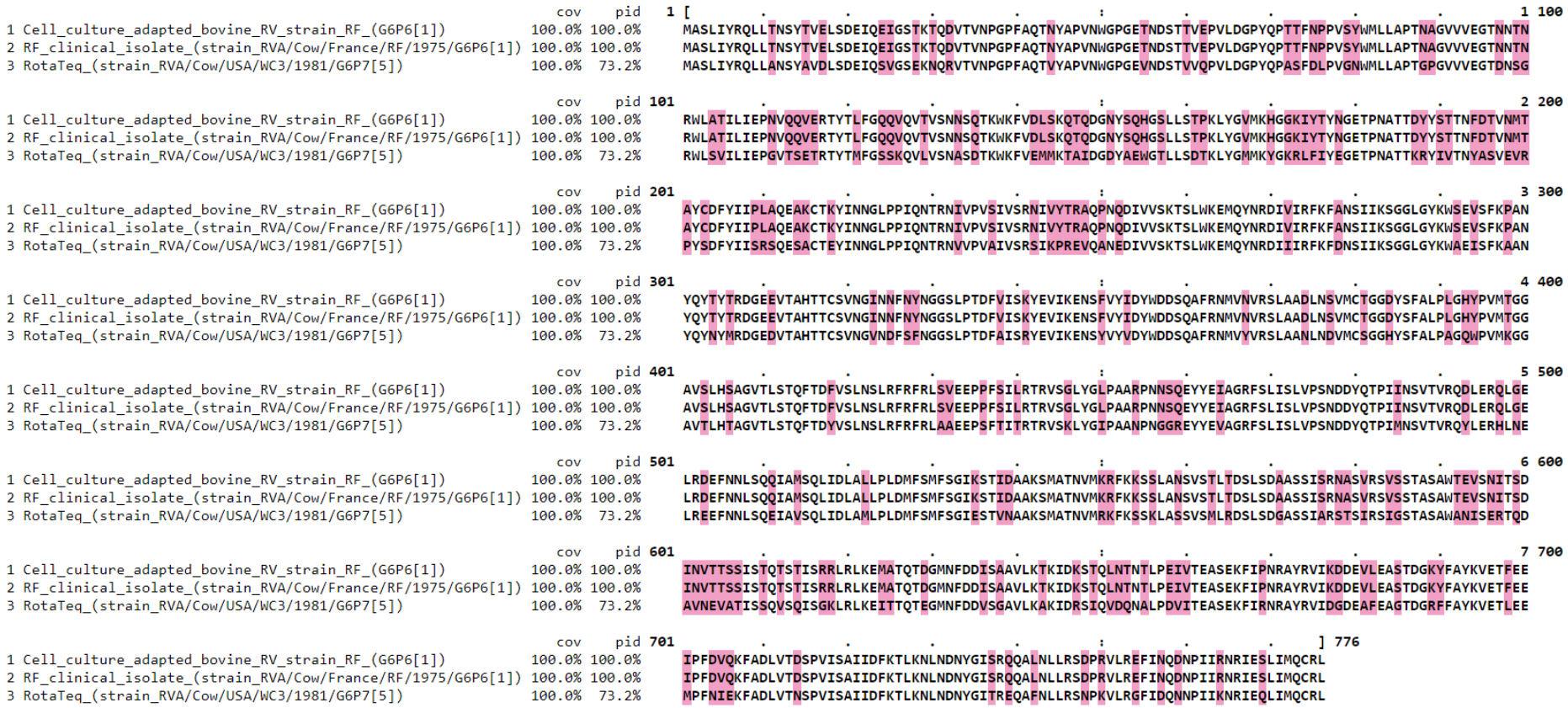


Figure 3.1 Amino acid alignment of RV gene segment 4 (VP4).
Alignment of the cell culture adapted bovine RV strain RF, the RF clinical isolate and the bovine strain used in the RotaTeq vaccine (GenBank accession numbers: KF729650.1, U65924.1 and AY050271.1 respectively). Amino acid residues conserved in all sequences are in black and mismatched residues are highlighted in pink. Width of the alignment is broken down into 100 amino acids per row. Alignment was generated with the Clustal Omega Alignment program and displayed using the MView tool from EMBL-EBI [335, 336]. Percentage coverage (cov) and percent identity (pid) of every sequence with respect to the refence sequence (Cell_culture_adapted_bovine_RV_strain_RF_(G6P6[1]) is shown. Identity was normalised by aligned length.

Table 12 Percentage identity and similarity for RV gene segment 4 (VP4)
The cell culture adapted bovine RF strain was compared to the RF clinical isolate and to the bovine strain used in the RotaTeq vaccine. Percentage identity and similarity at nucleotide (nt) and amino acid (aa) levels are shown. Analysis was performed using the Sequence Manipulation Suite program [337] on aligned sequences generated with the Clustal Omega Alignment program [335].

Gene segment	Encoded protein	Length (bp | aa)	Subject sequence	Target sequences	% nt identity	% nt similarity	% aa identity	% aa similarity
4	VP4	2362 | 776	Cell culture adapted bovine strain RF (G6P6[1])	RF clinical isolate (strain_RVA/Cow/France/RF/1975/G6P6[1])	**64.14**	**71.30**	**100**	**100**
				RotaTeq (strain_RVA/Cow/USA/WC3/1981/G6P7[5])	**54.15**	**66.26**	**73.20**	**83.76**

3.2 Results

3.2.1 Generation of the bovine RV strain RF from cloned cDNA

To establish a reverse genetics system for the bovine RF strain of RV, individual RF gene constructs were designed using complete segment sequences downloaded from the National Centre for Biotechnology Information (NCBI) [263] (for accession numbers see Table 1). Each gene segment was flanked by the sequences of the T7 RNA polymerase promoter (T7P) and the antigenomic hepatitis delta virus (HDV) ribozyme followed by the T7 RNA polymerase terminator (T7T), as carried out previously for the SA11 constructs (Figure 3.2A) [284]. Thus, transcription of the resulting vector would generate full length viral (+) single stranded RNA (ssRNA) transcripts containing native viral 5' and 3' ends [338]. These custom constructs were synthesised by Invitrogen GeneArt using either pMA or pMK plasmid backbones. As a positive control for the reverse genetics, plasmids representing the 11 dsRNA gene segments from SA11 were purchased from Addgene. To ensure that the plasmids contained gene segments of the correct size, 11 plasmids for each RV strain were resolved on an agarose gel according to plasmid size in descending order (Figure 3.2B and C). Synthesis of each RF gene segment was further verified by Sanger sequencing (data not shown).

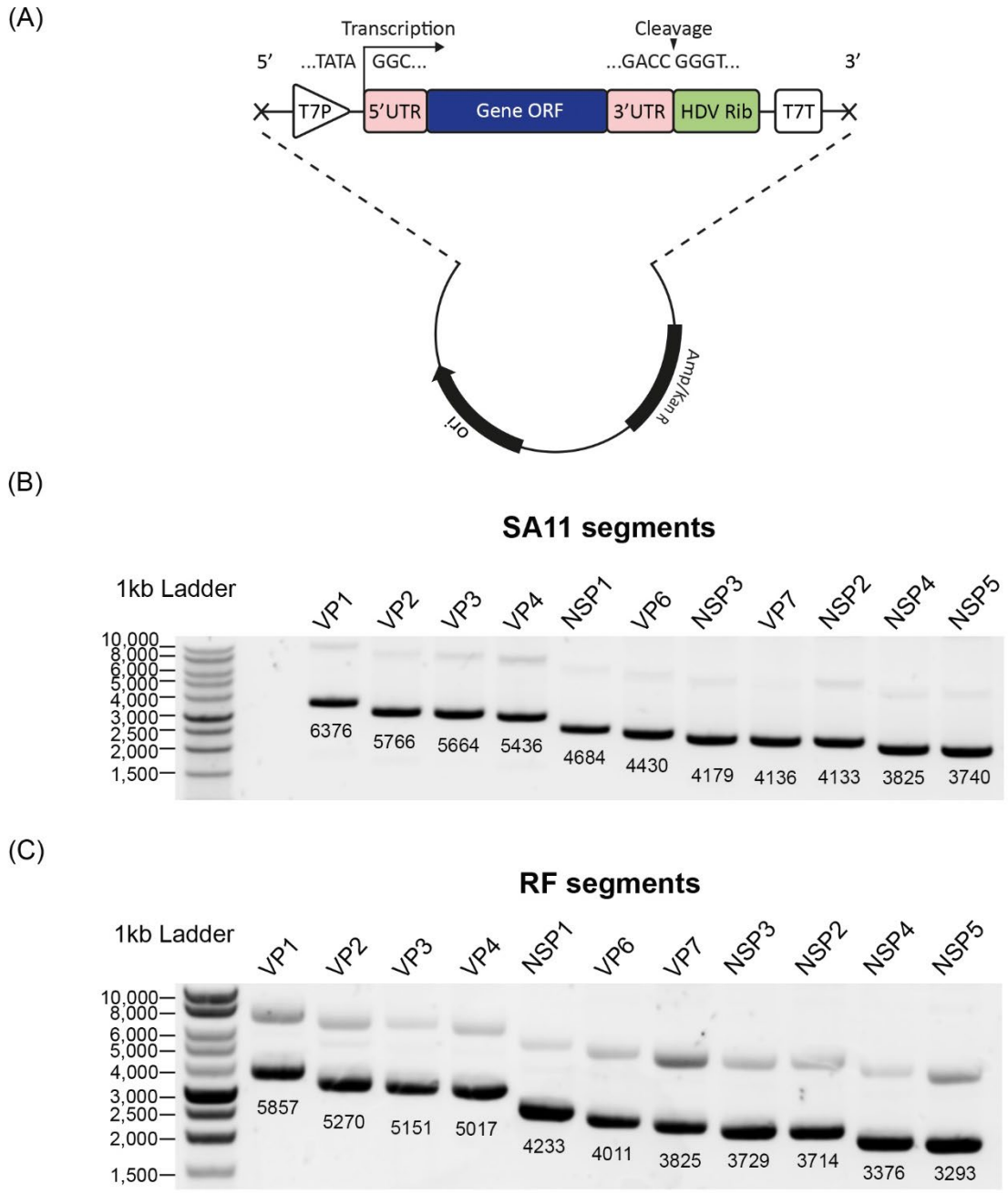


Figure 3.2 Construct design of the bovine RF strain RV for reverse genetics.
(A) Schematic of a plasmid encoding the bovine RF gene (not to scale). Agarose gel of 11 plasmids corresponding to each gene segment of SA11 **(B)** and RF **(C)** RVs. The size of each plasmid and a 1kb DNA ladder (in bp) are shown.

To rescue the bovine RF strain of RV, we utilised a combination of the previously published reverse genetics protocols described by Kanai *et al.*, (2017) and Komoto *et al.*, (2018) with slight modifications [284, 288]. BSR-T7 cells were transfected with 11 plasmids corresponding to each RV genome segment, with increased amounts of plasmids carrying the NSP2 and NSP5 genes (2.5 µg each, compared with 0.8 µg each for the remaining 9 plasmids) (Figure 3.3). Previous studies have shown that RNA transcripts could be efficiently capped by the vaccinia virus (VV) capping enzyme, thereby increasing viral translation [267, 268, 274, 277, 282]. To test whether the virus recovery would benefit from the VV capping enzyme, plasmids encoding the two subunits of the VV capping enzyme (D1R and D12L) [339, 340] were trialled in this protocol (Figure 3.3). Following overnight incubation, transfected cells were co-cultured with MA104 cells for three days in the presence of trypsin. Co-cultured cells were then lysed three times by freeze/thaw and passage zero (P0) virus stock was amplified on fresh MA104 cells to generate passage one (P1) stock. A cytopathic effect (CPE) was observed in MA104 cells suggesting recovery of RF RV (rRF), with no CPE seen in negative control (designated “mock”) in which only 10 of the 11 plasmids were used. SA11 RV was rescued in parallel as a positive control (rSA11). RF rescue was performed in six independent experiments, three of which included the VV capping enzyme, and the virus stock (P1) was titred by plaque assay (Figure 3.4A). The cell-culture adapted bovine RF strain isolate (RF C3E2), was plaqued in triplicate and was used as a positive control. As expected, the mock sample did not produce detectable levels of virus (the limit of detection was 12 PFU/mL) (Figure 3.4A). rRF had comparable titres to RF C3E2 and rSA11 with an average titre of 10^8 PFU/mL. Co-expression of the VV capping enzyme did not affect the viral titres or rescue efficiency.

Virus spread was characterised by assessing the plaque morphology. MA104 cells were infected with 10-fold serial dilutions of P1 stocks from three different rescues and incubated under a semi-solid overlay medium. Plates were scanned and cell-absent areas were calculated in ImageJ. RF C3E2 displayed the largest plaques ranging from approximately 1 mm to 8 mm in diameter (Figure 3.4B and C). In contrast, rescued viruses produced significantly smaller plaques. No difference in plaque phenotype was seen between rRF in the absence and presence of VV. However, in the presence of VV, rRF had significantly smaller plaques than rSA11, possibly suggesting the difference in viral replication between strains (Figure 3.4C).

To compare the replication kinetics of the rRF to RF C3E2 and rSA11, multicycle infections were performed in MA104 cells at a low multiplicity of infection (MOI) of 0.03 (Figure 3.4D). Samples were collected at several times post infection and titred by plaque assay. No

virus was detected in mock infected controls. At 16 hours post infection (hpi), rRF displayed significantly lower titres compared to RF C3E2. All viruses showed similar kinetics, with the titres of rRF being approximately $\log_{10}$ lower at each time point, and endpoint titres reaching 10^6 PFU/mL (Figure 3.4D).

Overall, this data demonstrates that the RF strain was successfully rescued from cloned cDNAs using plasmid only-based reverse genetics system and grew efficiently in cultured cells.

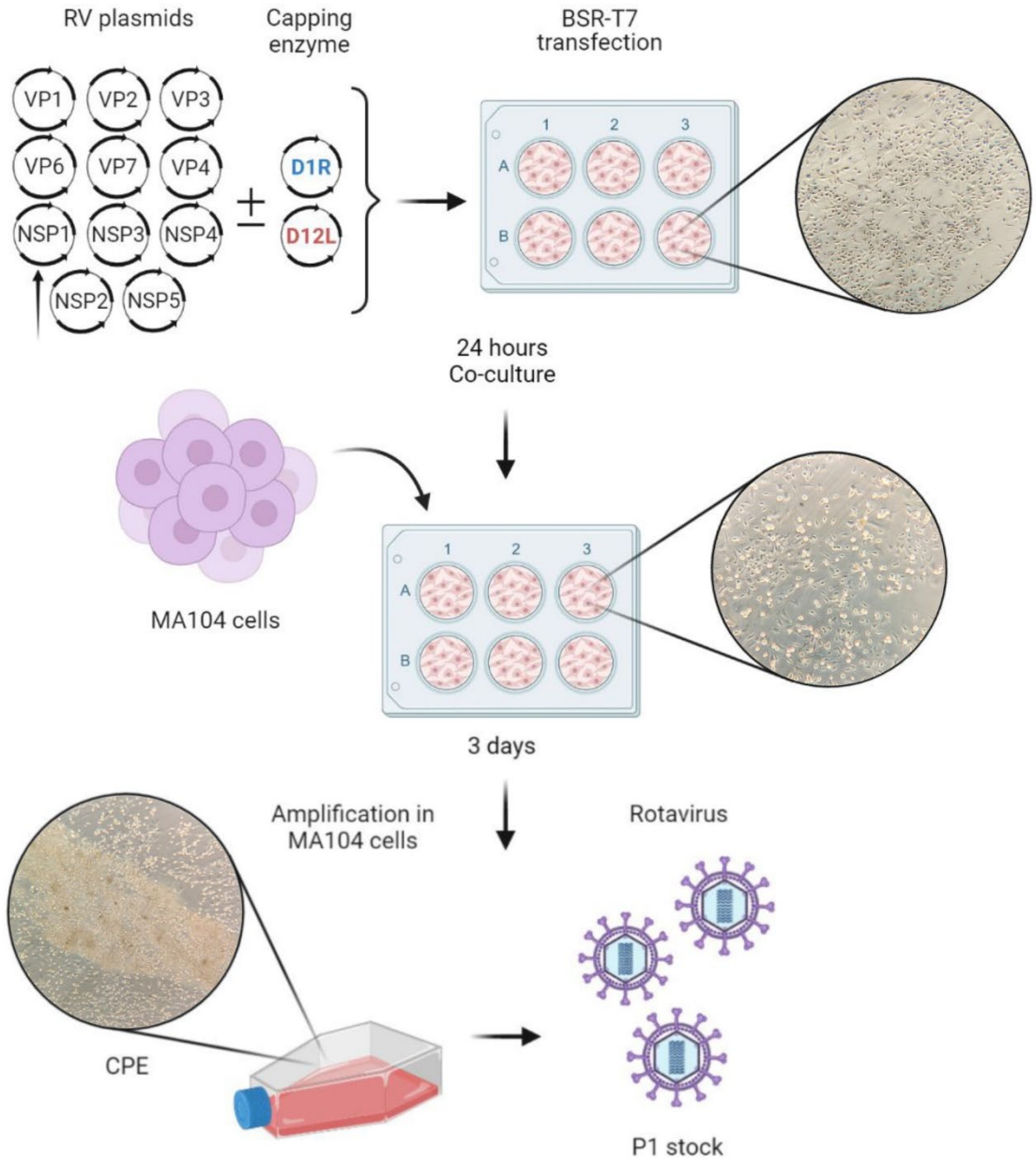


Figure 3.3 Schematic representation of a 13-plasmid reverse genetics system protocol. Cells were visualised by bright-field microscopy at X10 magnification and X20 magnification at co-culture step. (*Adapted from Kanai et al., 2017 [284] and Komoto et al., 2018 [288]; created with BioRender.com*).

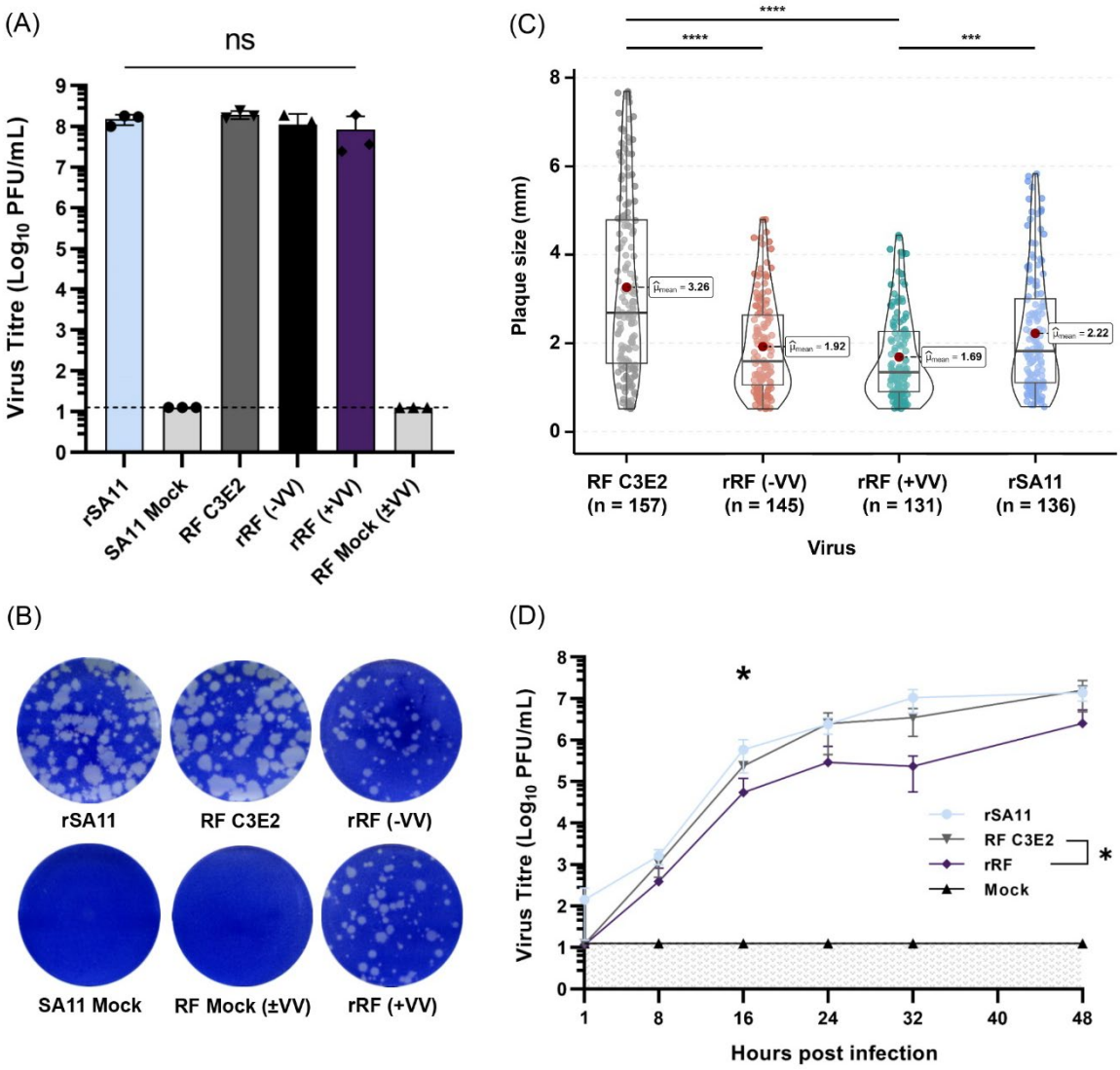


Figure 3.4 Viral characteristics of rRF versus RF C3E2 and SA11 viruses.
(A) Viral stocks (P1) were titred by plaque assay on MA104 cells. Mock samples were included for each strain. RF C3E2 was used as a positive control. Data are mean ± SD (n=3). Statistical annotations are the result of a one-way ANOVA test. Multiple comparison tests were performed between each viral strain. NS: Non-significant. Dashed line represents detection threshold. Samples with and without VV are shown. **(B)** Plaque phenotype of rescued viruses with corresponding mock samples. **(C)** Violin plot of plaque size distribution for each virus. A minimum of 40 plaques were measured for each virus from three independent experiments. Each dot represents a plaque. Statistic annotations are the result of a Welch ANOVA test with Holm's adjustment. *****p-value ≤ 0.0001, ***p-value ≤ 0.001*. The mean values are shown. **(D)** Multi-step growth curves of viruses. MA104 cells were infected at an MOI of 0.03 and samples were collected at the indicated hours post infection. Mock line represents detection threshold. Statistic annotations are the result of a two-way ANOVA test with Tukey's multiple comparisons test performed for each strain at each time-point. **p-value ≤ 0.05*.

3.2.2 Expression of heterologous peptides by rRF NSP3 gene

To exclude the possibility of contamination of rRF by the native RF C3E2 strain, a small tag of eight amino acids corresponding to 'SIINFEKL' was introduced into the C-terminus of rRF NSP3. Previous studies have successfully utilised SA11 NSP3 to express exogenous peptides generating genetically stable recombinant RVs [290, 298, 299]. In this study, rRF NSP3 was tagged with a SIINFEKL peptide in the presence or absence of Thosea asigna virus 2A (T2A) peptide (Figure 3.5A). During translation, the failure of peptide bond formation between 'PG' and 'P' residues in the conserved sequence at the C-terminus of 2A peptides results in production of two proteins [297]. Thus, the T2A sequence was employed in order to lower the risk of the SIINFEKL tag interfering with the function of the NSP3 gene. However, increasing the segment size by incorporating the T2A peptide could negatively affect the virus rescue and/or packaging, so both approaches were trialled. Tagged viruses were assigned the following notations: SIIN for virus without intervening T2A and T2A_SIIN for virus that had an additional 18 amino acids coding for the T2A peptide.

Rescue of RF and tagged viruses was performed in two independent experiments, one of which included the VV capping enzyme. P1 virus stocks were titred by plaque assay on MA104 cells (Figure 3.5B). Mock infected samples did not produce detectable levels of virus (the limit of detection was 12 PFU/mL). The titre of T2A_SIIN virus was lower than that of rRF, while the SIIN virus replicated to an average titre of 10^7 PFU/mL. Notably, the observed decrease in the titre was possibly due to the presence of the T2A peptide.

Subsequently, comparison of plaque morphology between the tagged viruses and rRF was carried out. MA104 cells were infected with 10-fold serial dilutions of P1 stocks from two different rescues and the cell-absent areas were measured in ImageJ. Tagged viruses produced small plaques in the presence and in the absence of VV, with the T2A_SIIN plaques being barely visible in absence of VV (Figure 3.5C). These observations were confirmed by quantification of the plaque size, which showed that, in contrast to rRF, tagged viruses produced smaller plaques ranging between ~0.25–2 mm in size (Figure 3.5D). The VV had no effect on the plaque phenotype of rRF, although the diameter of individual plaques was variable (~0.5-4 mm).

To examine the effect of SIINFEKL peptide insertion on RF infectivity, multicycle infections were performed in MA104 cells at a low MOI of 0.03 (Figure 3.5E). Samples were collected at several times post infection and titred by plaque assay. No virus was detected in

mock infected controls. Although all viruses displayed similar overall kinetics, the replication efficiency of tagged viruses was slightly lower than that of rRF.

To confirm that the tagged viruses contained the expected combination of gene segments, isolated viral dsRNA was resolved on a urea gel and visualised by silver staining. With the exception of gene segment 7 (NSP3), the dsRNA migration patterns were indistinguishable between viruses (Figure 3.6A, segment 7 highlighted in red). Although the lane containing the mock sample broke off from the main gel during drying, only a faint background signal was detected in this sample. The genome profile of SIIN and T2A_SIIN showed that the gene segment 7 migrated as expected, corresponding to its increased size as a result of peptide insertion (Figure 3.6B). Additionally, sequencing of the viral RNA from P3 stocks revealed the expected peptide insertion in segment 7, confirming the plasmid origin of these tagged viruses (Figure 3.6C-D).

Based on these results, the established reverse genetics system for RF allowed efficient rescue of recombinant viruses harbouring heterologous peptides in the NSP3 gene.

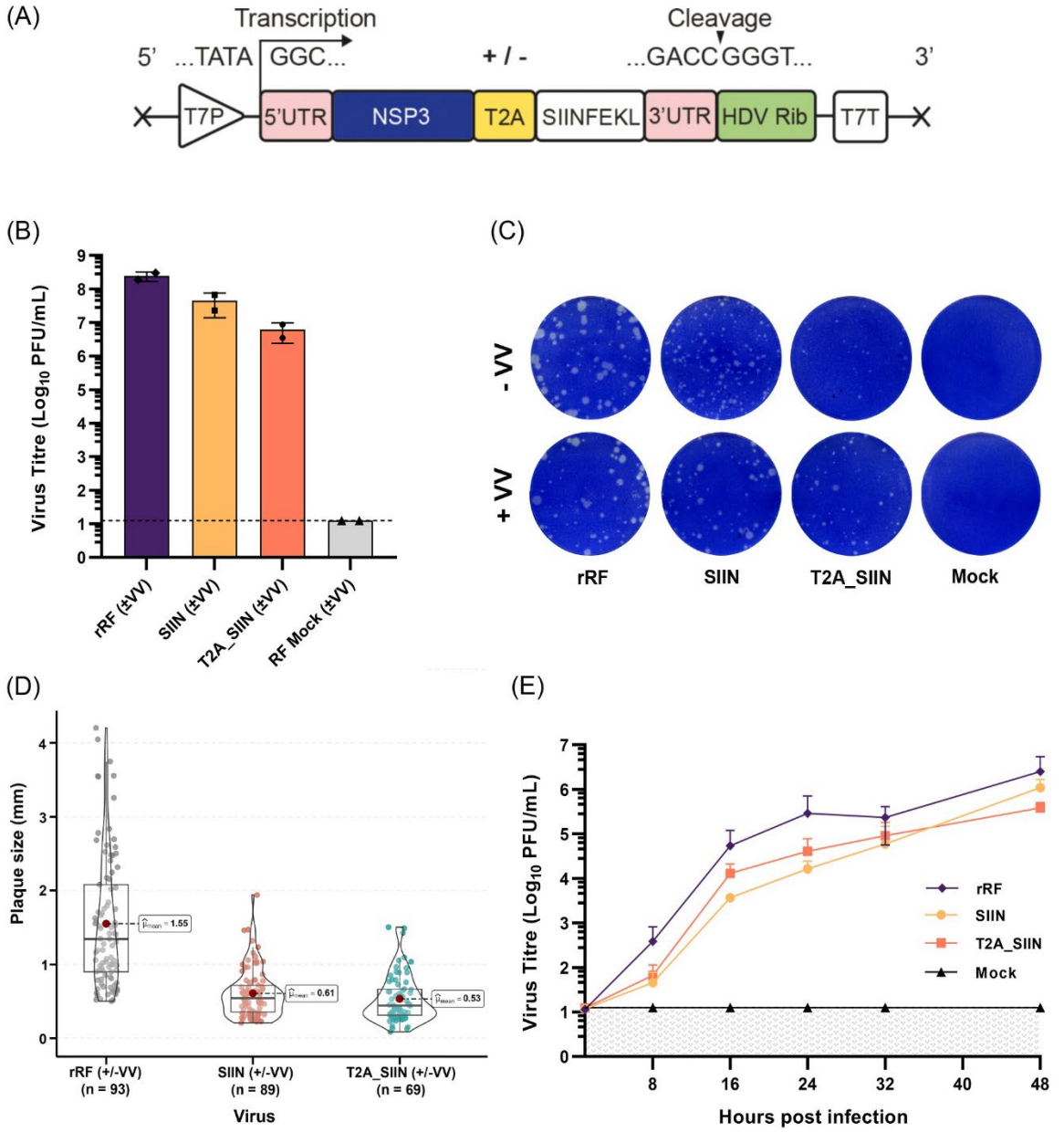
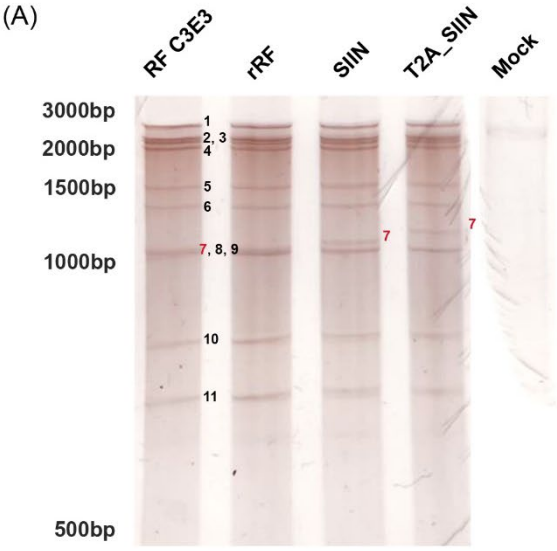


Figure 3.5 Generation of viruses containing heterologous peptides in gene segment 7 (NSP3). **(A)** Prototype of RV gene segment design as in **Figure 3.2A** showing NSP3 fused with a SIINFEKL peptide with or without T2A (yellow box), represented by +/- sign. **(B)** P1 virus stocks were titrated by plaque assay on MA104 cells. Data are mean ± SD (n=2). Dashed line represents detection threshold. Samples including mock, with and without VV, are shown. **(C)** Plaque phenotype of tagged viruses versus rRF with and without VV. **(D)** Violin plot of plaque size distribution for tagged viruses versus rRF. A minimum of 40 plaques were measured for each virus from two independent experiments. Each dot represents a plaque. The mean values are shown. **(E)** Multi-step growth curves of rRF and tagged viruses. Data for rRF is the same as in **Figure 3.4D** only from two independent experiments. MA104 cells were infected at MOI of 0.03 and samples were collected at the indicated hours post infection. Mock line represents detection threshold.

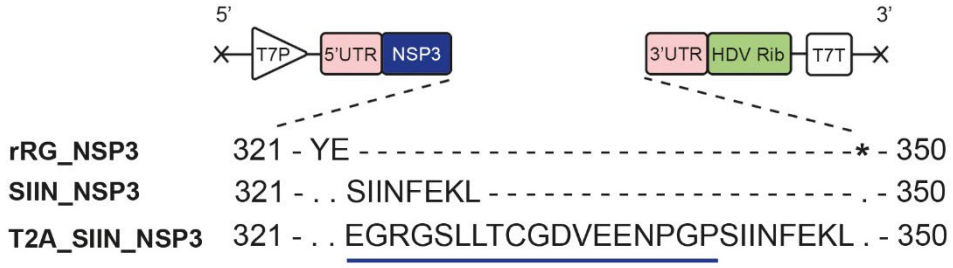
(A)



(B)

Segment	Gene	Segment size (bp)
1	VP1	3302
2	VP2	2687
3	VP3	2591
4	VP4	2362
5	NSP1	1578
6	VP6	1356
7	NSP3	1074
	NSP3_SIIN	1098
	NSP3_T2A_SIIN	1152
8	VP7	1062
9	NSP2	1059
10	NSP4	751
11	NSP5	667

(C)



(D)

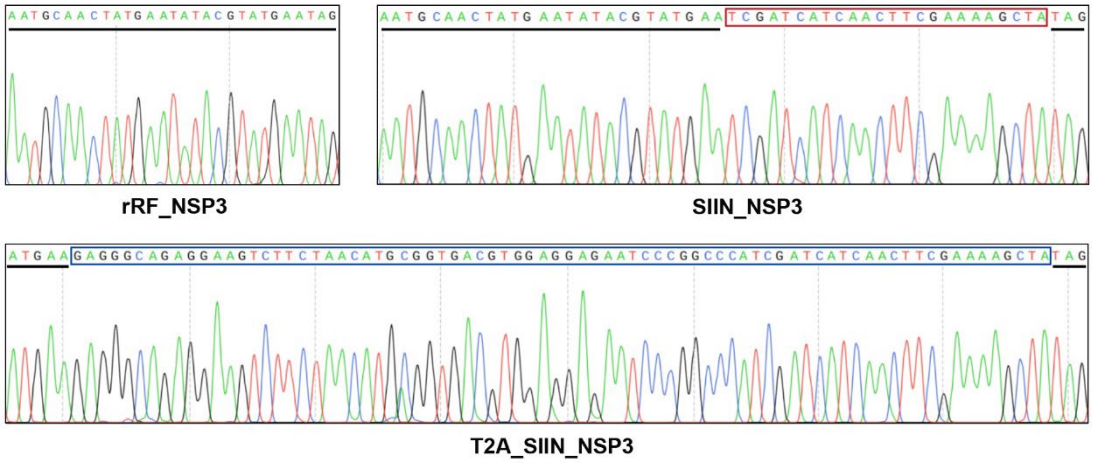


Figure 3.6 Confirmation of foreign peptide expression by gene segment 7 (NSP3). **(A)** Electrophoretic pattern of RNA gene segments for RF C3E3, rRF, tagged viruses and mock infected sample. Viral dsRNA extracted from P3 stocks was separated on a 5% urea gel and visualised by silver staining. Size of the RNA transcripts are shown on the left. Individual RNA segments are labelled in black with gene segment 7 highlighted in red. Lane with the mock sample separated from the main gel during drying. **(B)** Table of all gene segments with their corresponding sizes (in bp). **(C)** Amino acid sequence alignment of rRF and tagged viruses. T2A peptide is underlined in blue. Segment 7 target sequence from total viral RNA was amplified by RT-PCR using primers shown in **Table 10**. Products were sent for sequencing with GENEWIZ. Sequence reads were aligned using SSE v1.3 software. **(D)** Sequence electropherogram of segment 7. rRF sequence is underlined in black. Nucleotides corresponding to 'SIINFEKL' and 'T2A_SIINFEKL' tags are boxed in red and blue respectively.

3.2.3 Susceptibility of bovine enteroids to rRF infection

To explore the feasibility of using a more physiologically representative model for bovine RV infection, three dimensional (3D) bovine enteroids were infected with rRF at a high MOI of ~10 for 24 hr and stained with anti-VP6 antibody (Figure 3.7A). Following 1 hr of adsorption, organoids were thoroughly washed to remove residual virus particles and subsequently embedded in Matrigel for an additional 23 hr. As expected, no signal was detected in mock-infected samples. Secondary only control appeared as mock (data not shown). At 24 hpi, positive staining for VP6 was predominantly observed in the organoid lumen comprising differentiated intestinal epithelial cells, around the nucleus and in puncta surrounding the nucleus. No RV-positive cells were present in the intestinal crypt. This is consistent with previous findings that *in vivo*, RV preferentially infect mature enterocytes at the tips of intestinal villi [84, 89].

Laboratory strain SA11 RV has been used as an experimental model for *in vitro and in vivo* studies of RV pathogenesis, and so the bovine enteroids were also tested for their susceptibility to infection with rSA11 (Figure 3.7B). A faint background VP6 signal was detected in the rSA11 mock sample. In contrast to DAPI staining of other samples, the individual nuclei in the rSA11 mock were not as defined due to the use of a different plane of view image. Similarly to rRF, expression of VP6 was present around the nuclei of cells located within the organoid lumen and in the punctate cytoplasmic inclusion bodies. Additionally, some VP6 staining appeared where no nuclei were present, possibly representing free virus particles released by dead cells or residual inoculum (Figure 3.7B, white arrowhead in bottom panel).

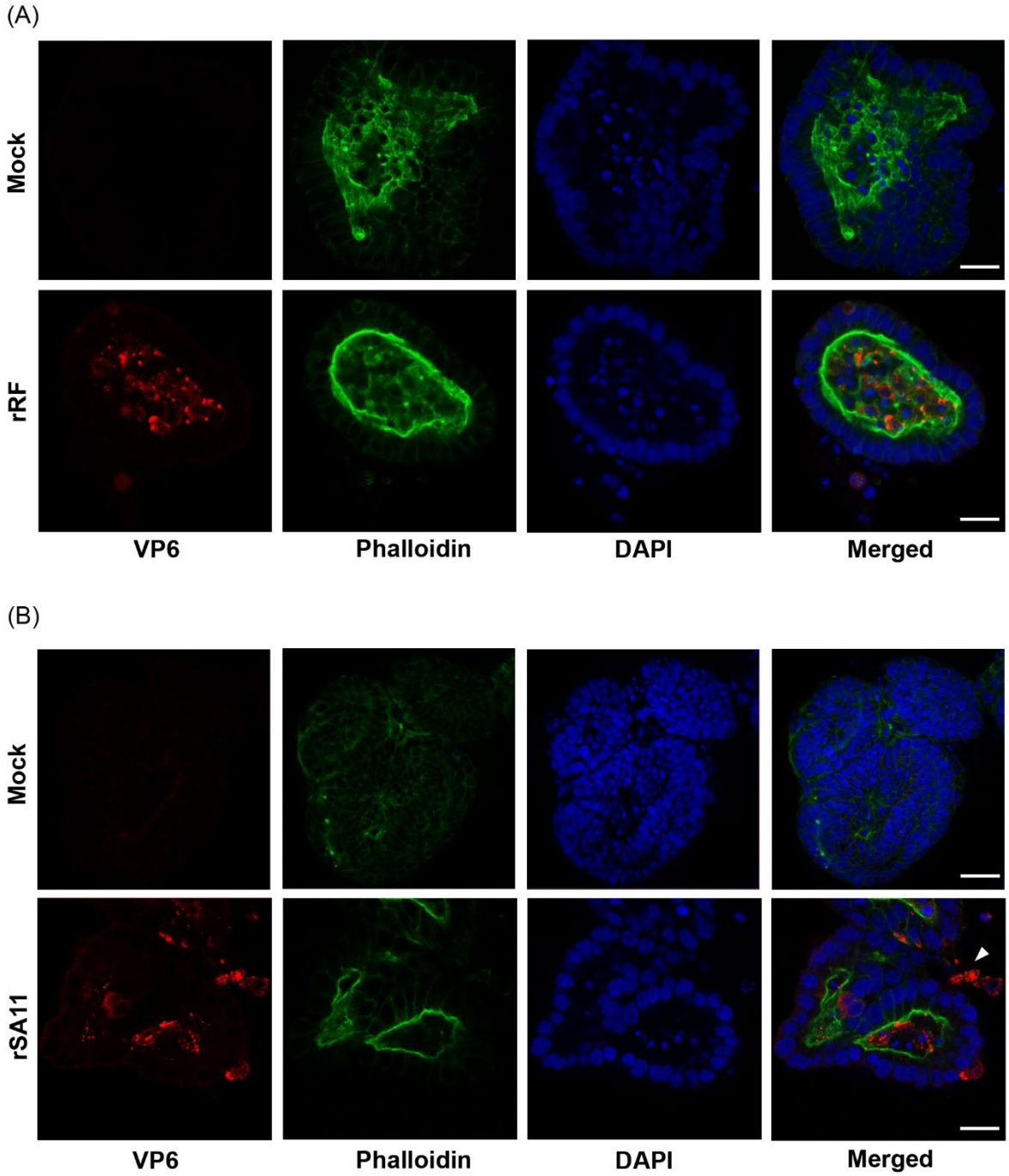


Figure 3.7 Bovine enteroids infected with rRF and rSA11 viruses.
3D bovine enteroids were infected with rRF **(A)** and rSA11 **(B)** viruses, with corresponding mocks at an approximate MOI of 10 PFU per cell and fixed at 24 hpi. Cells were stained for RV VP6, phalloidin (F-actin detection) and DAPI (nuclear stain). Scale bar represents 10µm. White arrowhead in **(B)**, bottom panel, represents free virus particles or residual inoculum.

To assess whether the bovine enteroids supported viral replication, rRF was used to carry out a small pilot study to test the effect of different variables. The effect of adding trypsin was trialled first, as proteolytic cleavage of VP4 is required for membrane interactions and efficient RV infectivity [40, 341]. Bovine enteroids were infected with rRF at high MOI of ~10 PFU/cell in the presence and/or absence of trypsin for 48 hr. Following adsorption and washing to remove unbound residual virus, virus present in the supernatant at 1, 24 and 48 hpi was titred by plaque assay on MA104 cells (Figure 3.8A). No virus was detected in mock-infected samples at each time point (limit of detection was 12 PFU/mL). In the absence of trypsin, after 1 hpi the virus reached an end point titre of 10^4 PFU/mL indicating the slow release of the inoculum. In contrast, in the presence of trypsin no change in the viral titre was observed throughout the 48 hr suggesting no effect of trypsin on the viral replication. These data demonstrate that rRF failed to replicate and produce infectious virions in organoids.

Next, production of infectious progeny was further analysed by mechanically disrupting or lysing the infected bovine enteroids. Following infection with rRF at high MOI of ~10 PFU/cell for 48 hr, virus present in organoids at 1, 24 and 48 hpi was harvested by either freeze/thaw or using a cell disruptor. Released virus was titred by plaque assay on MA104 cells (Figure 3.8B). A $\log_{10}$ increase in the titre was seen between 1 and 24 hpi when the enteroids were lysed by freeze/thaw, although by 48 hpi the titre dipped. In contrast, virus obtained through mechanical disruption of enteroids had consistently higher titres throughout.

Together, these data suggest that the bovine enteroids were susceptible to infection by both rRF and rSA11, but did not support the production of infectious virus progeny in any of the trialled conditions.

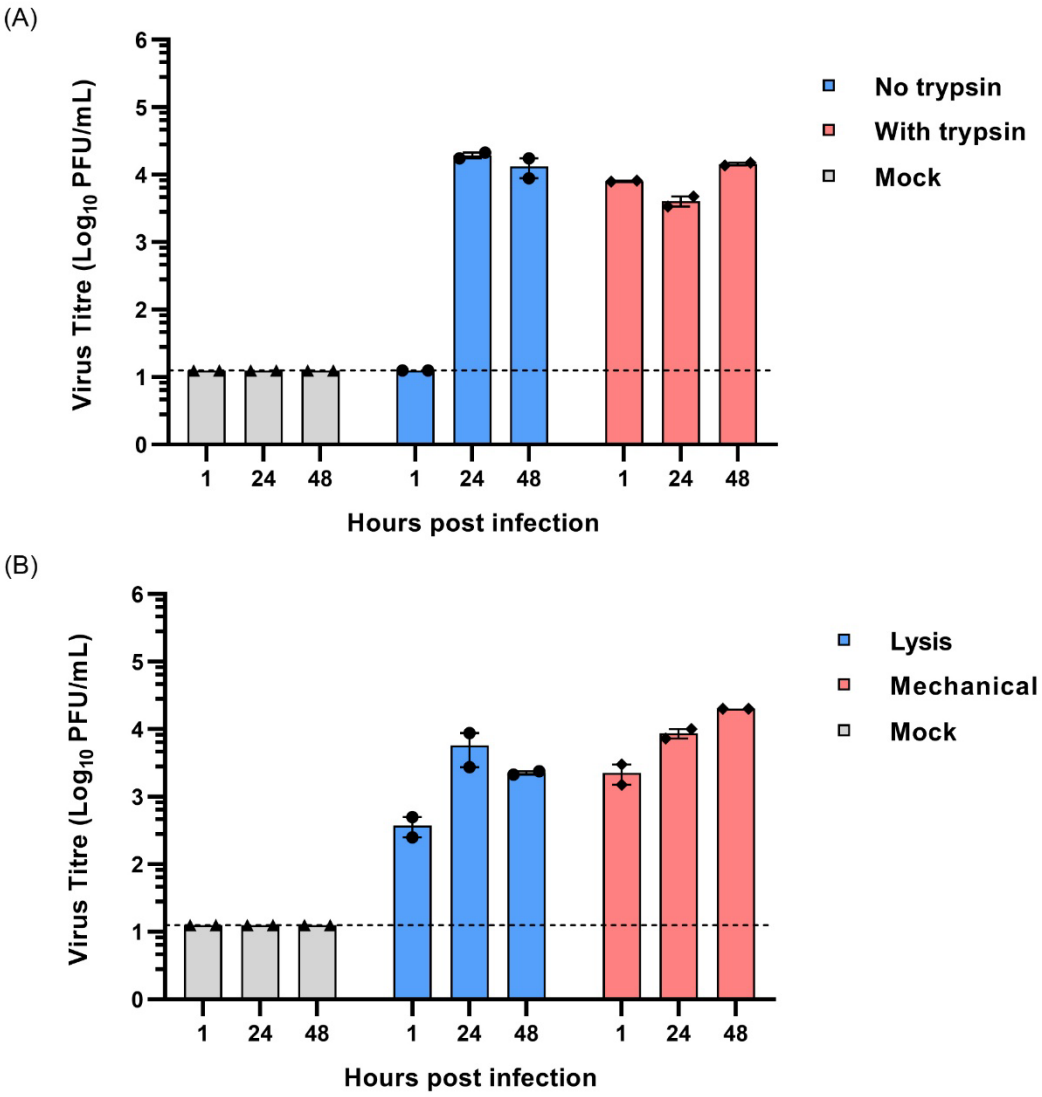


Figure 3.8 Effect of different variables on rRF replication in bovine enteroids. 3D bovine enteroids were infected with rRF at an approximate MOI of 10 PFU per cell for 48 hr. Samples were collected at indicated times post infection and titred by plaque assay on MA104 cells. Effect of trypsin on viral titres in **(A)** and mechanical disruption versus lysis of organoids in **(B)**. Data are the mean ± SEM from two technical experiments. Dashed line represents detection threshold.

3.3 Discussion

In this chapter, a plasmid-based reverse genetics system was established for the bovine RV strain RF. The simian strain SA11 was rescued in parallel as a validation of the previously published optimized 11-plasmid reverse genetics system. Viral characteristics of rRF were subsequently compared to RF C3E2 isolate and rSA11.

The effectiveness of co-expressing the VV capping enzyme in a T7 RNA polymerase system has previously been demonstrated for viruses of the *Sedoreoviridae* family and for hepatitis E virus [267, 268, 284, 342]. Furthermore, another RNA-capping enzyme derived from African swine fever virus (NP868R) or (C3P3-G1) increased the rescue efficiency of simian, human and murine RVs [290, 292]. However, use of plasmids encoding the two subunits of the VV capping enzyme did not provide an additional advantage for RF rescue as anticipated (Figure 3.4A). This supports previous research where co-transfection of two plasmids was not essential for RV rescue [65]. Addition of the VV capping enzyme may be beneficial for future studies involving generation of RV reassortants between strains which are difficult to rescue.

Although the panel of viruses displayed similar overall kinetics, the replication efficiency of rRF was slightly lower than RF C3E2 and rSA11 (Figure 3.4D). This also correlated with the production of significantly smaller plaques (Figure 3.4B-C). Presence of defective viral genomes (DVGs) has previously been suggested to affect viral titres, packaging and replication [343-346]. Some viruses of the *Sedoreoviridae* family were able to package non-canonical segments resulting in gene recombination that contributed to generation of DVGs [263, 280, 344, 347]. Additionally, serially passaging tissue culture-adapted bovine RV at high MOI has been shown to generate viruses with abnormal genome profiles [279]. For RVs, most reported gene recombination events occurred as a result of head-to-tail duplications (concatemers) resulting in longer segments being packaged [67, 80, 278, 347]. Increased levels of DVGs can reduce the production of viral proteins and outcompete the full-length genomes for incorporation into virions thereby producing defective interfering particles (DIPs) [348]. Additionally, accumulation of DVGs within the cell can stimulate pattern recognition receptors leading to upregulation of interferon stimulating genes, which in turn can inhibit viral replication and/or virion production [349-352]. Thus, a heterogeneous population of virions comprising DVGs could explain the observed reduced viral replication and small plaque phenotype of rRF.

Small DVGs exhibit a replicative advantage by retaining their replication and packaging signals [348]. However, in the absence of a standard virus, DVGs are replication incompetent as they require *trans*-complementation provided by the full-length genomes that facilitate replication and packaging into non-infectious DIPs [348]. It is possible that the observed mixed plaque morphology of RF C3E2 and rSA11 (Figure 3.4B and C) could be attributed to acquired compensatory mutations or reversion to WT, which could also explain a more efficient replication of RF C3E2 and rSA11 (Figure 3.4D). Future studies could test for the presence of mutations in rescued viruses by viral plaque purification and subsequent full genome sequencing. Experiments involving segment specific RT-PCR or RNA deep-sequencing would also enable detection of DIPs in viral progeny.

Subsequently, the reverse genetics system for RF was utilised to engineer viruses harbouring foreign peptides in gene segment 7 (encoding NSP3). Previous studies showed that SA11 NSP3 can accommodate large insertions of heterologous peptides generating genetically stable recombinant RVs [290, 298, 299]. As a proof of concept that the RF NSP3 could also be tagged, a small octapeptide the 'SIINFEKL' was used in this study (Figure 3.5A). Many viruses utilise 2A elements to generate multiple proteins from a single ORF including species C RVs where translation of its NSP3 gives rise to two products [76, 353]. Introduction of the T2A element, additional 18 amino acids, upstream of the 'SIINFEKL' peptide revealed that RF NSP3 could also accommodate extra genetic material. Although the viral titres and replication of tagged viruses were comparable to rRF, these viruses displayed significantly smaller plaques (Figure 3.5B-E). The same phenotype was observed for recombinant SA11 RV and RRV expressing fluorescent reporter proteins such as UnaG and fBFP [290, 298]. It is worth noting that UnaG and fBFP were larger (139 and 233 amino acids respectively) than the small tag (8 amino acids) used here. The small plaque phenotype could also be explained by an impaired function of NSP3, although this was not assessed in this study. Presence of the C-terminal tag could impact the ability of NSP3 to efficiently disrupt the eIF4G-PABP complex, thereby reducing the translation of RV proteins [121, 354]. Thus, the reduction of viral protein synthesis could result in decrease of viral progeny. Although no obvious defect in packaging was observed (Figure 3.6A), follow up studies detecting production of viral proteins could provide insight into the functions of NSP3.

In contrast, Montero *et al.*, (2006) showed that blocking NSP3 synthesis increased the yield of viral progeny which correlated with increased synthesis of dsRNA [355]. However, this effect was observed with the RRV strain, while our data originated from RF, which could potentially explain the discrepancy between the two studies. Furthermore, the 3' end of RRV NSP3 differs from the canonical 3' GACC of RF NSP3 [356].

Previous studies have demonstrated that shutdown of host protein synthesis through phosphorylation of eIF4E by the Semliki Forest virus resulted in synthesis of defective ribosomal products (DRiPs) [357]. It is plausible that the mutated NSP3 contributed to the synthesis of DRiPs thereby reducing the production of viral proteins. Thus, the accumulation of DRiPs in the viral progeny could trigger cellular innate immune responses [352] and inhibit viral infectivity as demonstrated by small plaques (Figure 3.5C-D).

The SIINFEKL peptide has previously been used to evaluate the effect of DRiPs on immunosurveillance [358-361]. By introducing SIINFEKL at the C-terminus of IAV gene segment 7, Yang *et al.*, (2016) showed that IAV DRiPs were generated by cytoplasmic non-canonical translation which contributed to antiviral CD8+ T cell immunosurveillance [362]. Hickman *et al.*, (2018) found that the translation of the negative strand, tagged with SIINFEKL, could contribute to DRiP formation and was shown to activate OT-I T cells *in vitro* and *in vivo* [363]. Thus, these tagged viruses could serve as a useful tool for future studies of T-cell mediated immunity to RV infection. Due to the lack of a specific antibody, detection of SIINFEKL was not carried out in this study but could be investigated in the future.

Intestinal organ-like 3D culture systems aim to recapitulate the complex nature and functions of the gut [364]. Both organoids and enteroids have been utilised to study an array of enteric pathogens including human, simian and porcine RVs [89, 365-367]. However, recently developed 3D organoid models of bovine intestinal epithelium have not yet been explored to study RV infection. Therefore, 3D basal-out bovine enteroids established by Hamilton *et al.*, (2018) [368] and further characterised by Blake *et al.*, (2022) [369], were tested to examine their susceptibility to infection by rRF and rSA11. Although VP6 expression was detected following infection with rRF and rSA11, rRF failed to replicate in this system (Figure 3.7 and Figure 3.8). VP6 is a component of a non-infectious DLP, thus it does not accurately represent production of infectious viral progeny [143]. Due to the lack of antibodies, detection of proteins comprising infectious TLPs, such as VP4 and/or VP7, was not assessed in the present study. Staining for specific cell types known to reside in the intestinal crypt, such as Lgr5+ stem cells and Ki67+ proliferative cells [369], could further confirm the localisation of RV virions. Previous work with human intestinal organoids demonstrated that addition of pancreatin and bovine trypsin increased the number of RRV-positive cells [365, 366]. Due to restrictions in organoids availability, different trypsin concentrations and types were not optimised here, but could be assessed in the future.

Upon infection, a drop in Ca^{2+} concentration leads to dissociation of VP4 and VP7 from TLPs, releasing transcriptionally active DLPs into the cytoplasm [181]. Studies have shown

that RV infection induces intracellular calcium waves through adenosine 5'-diphosphate signalling [370]. In human enteroids, inhibition of the calcium/calmodulin-dependent kinase kinase-β signalling pathway significantly reduces the production of infectious virions [168, 371]. Thus, studies exploring the effects of varying the Ca^{2+} concentration may allow further insights into rRF infectivity of bovine enteroids.

Another challenge in working with enteroids is the access to the apical/luminal side where the VP6 expression was mainly detected (Figure 3.7). Several methods have been used to gain access to the interior of enclosed 3D structures including luminal microinjection or mechanical disruption [368, 372]. Additionally, enteroids could be plated as monolayers, allowing a more controlled access to both apical and basolateral surfaces [366, 368]. Directed differentiation of stem cell lines into intestinal-like tissues have also been established and used to study RV infection [365, 373]. Due to resource availability, microinjection proved to be too labour intensive for the scope of this study and was therefore not included. Nonetheless, neither the mechanical disruption nor lysis of organoids resulted in successful replication of rRF in this system (Figure 3.8).

Overall, the methods presented here require optimisation and further characterisation of the bovine enteroids is essential to enhance the reliability of their infection with rRF. As a result, for the reasons stated above and due to data limitations, the results from this section are inconclusive and require further analysis. For future work, it may be beneficial to utilise freshly isolated RVs since both strains used in this study are heavily cell-culture adapted.

The results of this chapter show that the RF strain of RV was successfully rescued using a reverse genetics system. Data presented here demonstrate that RF NSP3 could be used for expression of heterologous peptides (*subject of Chapter 4*). Rescued viruses expressing the SIINFEKL peptide could be utilised to study alternative translation initiation events (*subject of Chapter 5*) as well as T-cell immunity to RV infection in mice. Although *in vitro* cell culture models have provided important insights into RV pathogenesis, they do not capture the complex architecture and cellular heterogeneity present in the intestinal epithelium. Additionally, the current cell models allow only a limited number of well-adapted laboratory strains to be used for experimental research. By identifying optimal culturing conditions of bovine enteroids this model could serve as a valuable tool for studying livestock pathogens, expanding upon the limitations of current cell models and reducing the need for large animal models.

Chapter 4: Rotavirus as a vaccine vector for delivery of SARS-CoV-2 spike peptides

This chapter is based on the published paper:
"Using Species A Rotavirus Reverse Genetics to Engineer Chimeric Viruses Expressing SARS-CoV-2 Spike Epitopes."

Ola Diebold[1], Victoria Gonzalez[1], Luca Venditti[2], Colin Sharp[1], Rosemary A. Blake[1], Wenfang S Tan[1], Joanne Stevens[1], Sarah Caddy[3†,] Paul Digard[1], Alexander Borodavka[2], Eleanor Gaunt[1]

[1]Infection and Immunity Division, Roslin Institute, University of Edinburgh, Easter Bush Campus, Midlothian, EH25 9RG, UK

[2]Department of Biochemistry, University of Cambridge, UK

[3]Cambridge Institute of Therapeutic Immunology and Infectious Disease, Department of Medicine, University of Cambridge, UK

[†]Present address: Baker Institute for Animal Health, Department of Microbiology and Immunology, Cornell University College of Veterinary Medicine, New York, USA

DOI: https://doi.org/10.1128/jvi.00488-22

The original document is included within this chapter along with an introduction and conclusion.

4.1 Background

SARS-CoV-2, the virus responsible for the COVID-19 pandemic, is the third documented coronavirus associated with zoonotic transmission from animal reservoirs in the last two decades [374, 375]. The high morbidity and mortality caused by the virus resulted in several different platforms being utilised for the development of vaccines against COVID-19 [376]. The spike glycoprotein of SARS-CoV-2 was identified as a strong antigenic target based on preclinical studies of vaccines against SARS-CoV and MERS-CoV [377-379]. Specifically, the RBD of the spike protein was able to induce neutralising antibodies and T-cell mediated responses, preventing its interaction with the angiotensin converting enzyme 2 (ACE2) receptor on host cells thereby blocking infection [319, 380]. Thus, the SARS-CoV-2 vaccine candidates that entered Phase III clinical trials expressed either the full length spike protein or the RBD peptide delivered as a protein subunit, RNA molecule or by viral vectors [320, 376].

At the beginning of the COVID-19 pandemic, early vaccine clinical trials excluded vulnerable groups such as children, pregnant women and immunocompromised individuals [320]. Although overall hospitalisation and mortality rates were lower in children than in adults, a high proportion of asymptomatic SARS-CoV-2 infections were reported in children [381, 382]. Notably, children under the age of five carried higher viral loads in the nasopharynx than older children and adults [383-385]. Studies showed that asymptomatic children contributed to SARS-CoV-2 transmission in people who were most at risk from COVID-19 [386]. However, paediatric studies have highlighted that COVID-19-associated deaths in children occurred more frequently in low- and middle-income countries, with case fatality rates exacerbated by health care disparities, poverty and comorbidities [387]. Additionally, vaccine efficacy and accessibility in developing countries was a challenge [388]. Immunisation efforts were therefore needed to target this vulnerable population.

One of the proposed strategies was to introduce COVID-19 vaccines into the existing childhood vaccination programmes. Among the vaccines routinely administered to children in many countries, both developed and developing, are those targeting RV. The pentavalent live-attenuated RV vaccine, RotaTeq, comprises bovine backbone reassorted with the most prevalent human serotypes [334]. RotaTeq is administered orally and induces a strong mucosal immune response with production of neutralising IgA and IgG antibodies [224]. Similarly, orally delivered SARS-CoV-2 vaccines could elicit secretory IgA responses not only to protect the upper respiratory tract but also the intestine, which was shown to be a viral target [389-391]. The ACE2 receptor is highly expressed in the brush border of intestinal epithelium and multiple studies have reported gastrointestinal tract infections by SARS-CoV-2,

suggesting a potential for faecal-oral transmission [391-393]. Previous reports showed that crosstalk between different mucosae sites exists and that immunisation at a single mucosal site results in generation of antigen-specific IgA antibodies at a heterologous site such as the respiratory tract [394, 395]. Based on these considerations, this study aimed to explore the potential of utilising RV as a vaccine vector to express SARS-CoV-2 spike epitopes to induce a robust mucosal immune response. Additionally, using the RV bovine backbone would circumvent the issue of pre-existing immunity against the vector. Thus, RV-spike chimeric vaccine could be produced using the established vaccine pipeline to deliver a safe and easily accessible thermostable vaccine.

4.2 Aims of the study

i) To generate recombinant RVs expressing SARS-CoV-2 spike epitopes.
ii) Validate immunogenicity of SARS-CoV-2 peptides using spike antibodies.

4.3 Statement of authorship

This study was performed using two approaches: tagging of the RF NSP3 with long SARS-CoV-2 spike peptides, and expression of short SARS-CoV-2 spike peptides by the SA11 VP4. The first author performed all experiments for the NSP3 section and supervised MSc student Victoria Gonzalez to perform experiments for the VP4 section, which was in collaboration with Dr Alexander Borodavka and his team at the University of Cambridge. Data generated by Victoria Gonzalez is presented in **Figures 2** and **3**, for which data analysis was performed by the first author.

Ola Diebold: experimental design, performance of experiments, data analysis, investigation, supervision, literature review, first and final draft of the manuscript. **Victoria Gonzalez**: performance of experiments and manuscript editing. **Luca Venditti**: design of the SA11 VP4 constructs expressing SARS-CoV-2 peptides. **Colin Sharp**: experimental design, supervision and manuscript editing. **Rosemary A. Blake**: experimental design and performance of experiments with bovine enteroids and manuscript editing. **Wenfang S. Tan**: generated positive control for ELISA experiments. **Joanne Stevens**: supervision of bovine enteroid experiments. **Sarah Caddy**: mouse experiments (data not included) and manuscript editing. **Paul Digard, Alexander Borodavka and Eleanor Gaunt**: concept, funding acquisition, supervision, experimental design, project administration and manuscript editing.

4.4 Original document

AMERICAN SOCIETY FOR MICROBIOLOGY | Journal of Virology®

CELLULAR RESPONSE TO INFECTION

Using Species a Rotavirus Reverse Genetics to Engineer Chimeric Viruses Expressing SARS-CoV-2 Spike Epitopes

Ola Diebold,[a] Victoria Gonzalez,[a] Luca Venditti,[b] Colin Sharp,[a] Rosemary A. Blake,[a] Wenfang S. Tan,[a] Joanne Stevens,[a] Sarah Caddy,[c] Paul Digard,[a] Alexander Borodavka,[b] Eleanor Gaunt[a]

[a]Infection and Immunity Division, Roslin Institute, University of Edinburgh, Easter Bush Campus, Midlothian, United Kingdom
[b]Department of Biochemistry, University of Cambridge, Cambridge, United Kingdom
[c]Cambridge Institute of Therapeutic Immunology and Infectious Disease, Department of Medicine, University of Cambridge, Cambridge, United Kingdom

ABSTRACT Species A rotavirus (RVA) vaccines based on live attenuated viruses are used worldwide in humans. The recent establishment of a reverse genetics system for rotoviruses (RVs) has opened the possibility of engineering chimeric viruses expressing heterologous peptides from other viral or microbial species in order to develop polyvalent vaccines. We tested the feasibility of this concept by two approaches. First, we inserted short SARS-CoV-2 spike peptides into the hypervariable region of the simian RV SA11 strain viral protein (VP) 4. Second, we fused the receptor binding domain (RBD) of the SARS-CoV-2 spike protein, or the shorter receptor binding motif (RBM) nested within the RBD, to the C terminus of nonstructural protein (NSP) 3 of the bovine RV RF strain, with or without an intervening Thosea asigna virus 2A (T2A) peptide. Mutating the hypervariable region of SA11 VP4 impeded viral replication, and for these mutants, no cross-reactivity with spike antibodies was detected. To rescue NSP3 mutants, we established a plasmid-based reverse genetics system for the bovine RV RF strain. Except for the RBD mutant that demonstrated a rescue defect, all NSP3 mutants delivered endpoint infectivity titers and exhibited replication kinetics comparable to that of the wild-type virus. In ELISAs, cell lysates of an NSP3 mutant expressing the RBD peptide showed cross-reactivity with a SARS-CoV-2 RBD antibody. 3D bovine gut enteroids were susceptible to infection by all NSP3 mutants, but cross-reactivity with SARS-CoV-2 RBD antibody was only detected for the RBM mutant. The tolerance of large SARS-CoV-2 peptide insertions at the C terminus of NSP3 in the presence of T2A element highlights the potential of this approach for the development of vaccine vectors targeting multiple enteric pathogens simultaneously.

IMPORTANCE We explored the use of rotaviruses (RVs) to express heterologous peptides, using SARS-CoV-2 as an example. Small SARS-CoV-2 peptide insertions (<34 amino acids) into the hypervariable region of the viral protein 4 (VP4) of RV SA11 strain resulted in reduced viral titer and replication, demonstrating a limited tolerance for peptide insertions at this site. To test the RV RF strain for its tolerance for peptide insertions, we constructed a reverse genetics system. NSP3 was C-terminally tagged with SARS-CoV-2 spike peptides of up to 193 amino acids in length. With a T2A-separated 193 amino acid tag on NSP3, there was no significant effect on the viral rescue efficiency, endpoint titer, and replication kinetics. Tagged NSP3 elicited cross-reactivity with SARS-CoV-2 spike antibodies in ELISA. We highlight the potential for development of RV vaccine vectors targeting multiple enteric pathogens simultaneously.

KEYWORDS rotavirus, NSP3, VP4, reverse genetics

Species A rotaviruses (RVAs) are a leading cause of severe acute gastroenteritis in infants and young children worldwide, accounting for ~128,500 deaths annually (1–3). Likewise, rotavirus (RV)-associated enteritis in young calves and piglets has a

Editor Tom Gallagher, Loyola University Chicago

Address correspondence to Eleanor Gaunt, Elly.Gaunt@ed.ac.uk, or Alexander Borodavka, ab2677@cam.ac.uk.

The authors declare no conflict of interest.

Received 22 March 2022
Accepted 31 May 2022
Published 27 June 2022

significant economic impact on livestock production as a result of the high morbidity and mortality caused (4–7). Two human live attenuated RV vaccines, Rotarix and RotaTeq, have proven effective in reducing the incidence of RV-related hospitalization and mortality internationally (2, 8–10). Vaccination strategies for livestock rely on induction of active or passive immunity using animal RV vaccines (11–14).

RVA is a double-stranded RNA (dsRNA) virus with 11 genome segments encoding six structural viral proteins (VP1–VP4, VP6, and VP7) and depending on the strain, 5 or 6 nonstructural proteins (NSP1–NSP5 ± NSP6) (3, 15, 16). The mature infectious virion, termed a triple-layered particle (TLP), consists of an outer layer formed by VP4 and VP7. A double-layered particle (DLP), nested within the TLP, contains the intermediate and inner layers of the capsid formed by VP6 and VP2 respectively (3). RV primarily infects mature enterocytes of the intestinal epithelium and replicates exclusively in the cytoplasm (17, 18). Efficient RV cell entry requires proteolytic cleavage of the outer capsid protein VP4 into VP8* (28 kDa) and VP5* (60 kDa) domains by trypsin-like proteases of the host gastrointestinal tract (19–22). The VP8* lectin domain mediates RV attachment to different host cell receptors such as sialic acid-containing glycans, histo-blood group antigens, and integrins, depending on the virus strain (18, 23–25). Following endocytosis, low calcium levels in endosomes trigger the dissociation of VP7 and VP4, releasing the transcriptionally active DLP into the cytoplasm (3). Here, DLPs transcribe capped, nonpolyadenylated, positive-sense single-stranded RNA transcripts, which act as templates for viral protein translation (3). The 11 mRNAs share a conserved terminal 3′-UGUGACC sequence that contains *cis*-acting signals important for transcription by the RNA-dependent RNA polymerase (VP1) (26–30). The 3′-GACC sequence is bound by the N terminal of NSP3, while the C terminus of NSP3 interacts with the eukaryotic protein initiation factor 4GI (eIF4GI) (31–33). NSP3 displaces poly-A binding protein from ribosomal complexes to favor the translation of viral over host cell proteins (34–37).

The molecular characterization of RVs has historically proven challenging due to inefficiencies of the helper virus-dependent reverse genetics systems available for controlled mutagenesis studies (38–41). Recently, however, Kanai et al. (42) developed a plasmid only-based reverse genetics system for the simian RV strain SA11. The plasmid-only reverse genetics system has been optimized to study the functions of RV proteins, to generate RV reassortants and RV reporter expression systems, and conceptualize novel vaccine platforms (43–45). Using the plasmid only-based reverse genetics system, recombinant RVs harboring chimeric VP4 proteins that showed efficient replication in cell culture and neutralizing activity in mice have also been engineered (46–48). Protection from infection with RV is primarily mediated by heterotypic neutralizing antibodies that target VP4 and/or VP7 (47, 49). VP4 is therefore highly immunogenic and an important target for adaptive immunity (49, 50). Thus, the potential for VP4 to express heterologous epitopes from different RV strains may provide a delivery platform for expression of different vaccine antigens, although peptide insertions into VP4 has not previously been tested. Additionally, the plasmid only-based reverse genetics system has been utilized to generate a repertoire of recombinant RVs expressing fluorescent reporter proteins (48, 51). The C terminus of the SA11 NSP3 open reading frame (ORF) was fused to a porcine teschovirus translational 2A element followed by various reporters including UnaG, mKate, mRuby, or TagBFP to successfully yield two uncoupled proteins without compromising virus replication (51). A more recent study showed that the C terminus of SA11 NSP3 can express different peptides of the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) spike protein with minimal impact on endpoint titers (52). This has highlighted the potential to use RVs as expression vectors for development of polyvalent vaccines for enteric viruses.

In response to the recent emergence of SARS-CoV-2, the causative agent of coronavirus disease 2019 (COVID-19), several platforms have been utilized for vast global vaccine production (53, 54). Current licensed vaccines include examples that employ mRNA (Pfizer-BioNTech, Moderna) or viral vectors (AstraZeneca, Janssen) to deliver genetic material encoding the SARS-CoV-2 spike protein in order to stimulate the

production of neutralizing antibodies and T-cell-mediated immune responses that target this protein (55–58). High neutralizing antibody titers are strongly associated with specificity for the receptor binding domain (RBD) of the spike protein, making it the most immunogenic antigen (59–62). To assess the potential for generating chimeric vaccines using RV, we used SARS-CoV-2 as a timely and important example, introducing spike peptides into the RV genetic backbone and determining whether chimeric viruses showed cross-reactivity with spike antibodies. For this, the hypervariable region of SA11 VP4 (VP8* lectin domain) and the C terminus of the bovine RV RF strain NSP3 were modified to express SARS-CoV-2 spike epitopes.

We found that mutating the hypervariable region of SA11 VP4 reduced RV infectivity and mutants expressing spike peptides did not cross-react with SARS-CoV-2 spike antibodies, suggesting limitations of VP4 tagging as a strategy for live attenuated vaccine development. Using our newly established plasmid-only-based reverse genetics system for the bovine RV RF strain, we rescued infectious viruses expressing either the RBD or the RBM of the SARS-CoV-2 spike protein, with similar titers and replication kinetics to those of the wild-type (WT) virus. These viruses cross-reacted with RBD antibodies in enzyme-linked immunosorbent assay (ELISA) and were able to infect bovine gut enteroids, inferring the potential of the system for use in live attenuated vaccine development.

RESULTS

***In vitro* synthesis of RV proteins tagged with SARS-COV-2 spike peptides.** We engineered a panel of SA11 strain VP4 plasmids with SARS-CoV-2 spike peptide sequences inserted into the hypervariable region, and a panel of RF strain NSP3 plasmids with 3′ tags of SARS-CoV-2 RBM or RBD with or without a separating Thosea asigna virus 2A (T2A) peptide (Fig. 1A). VP4 comprises two major domains, VP5* and VP8*, which undergo conformational change upon tryptic cleavage that enhances viral entry (20, 63–65). For peptide insertion into VP4, the hypervariable region (residues L164 to N198) within the VP8* lectin domain was targeted due to the genome plasticity of this region and its virion surface expression (57, 63, 64). The insertion site has been mapped onto the crystal structure of VP4 (Fig. 1B).

To confirm the expression of the spike epitopes by the mutated VP4 and NSP3 segments, coupled *in vitro* transcription and translation (IVT) reactions were carried out using rabbit reticulocyte lysate system supplemented with [^{35}S]methionine (Fig. 1B and C).

Following sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and autoradiography, SA11 WT and mutated VP4 constructs produced polypeptide products of the expected protein size throughout (Fig. 1C). The WT RF NSP3 construct expressed a protein of expected size (Fig. 1D, lane 2). RF NSP3 segment incorporating the RBM motif produced a protein of a higher molecular weight consistent with its predicted size (Fig. 1D, lane 3). The T2A_RBM construct generated a product containing the NSP3 protein that ran slightly higher than WT NSP3 due to the residual C-terminal fusion of partial T2A sequence (66) (Fig. 1D, lane 4, red asterisk). The small size (~10 kDa) of the unconjugated RBM peptide made it difficult to visualize due to comigration with the dye front. Unseparated NSP3-T2A_RBM, seen as a minor product, was probably produced as a result of unsuccessful ribosome skipping at the T2A site (66) (Fig. 1D, lane 4, black asterisk). As expected, translation of NSP3 fused to the RBD peptide also produced a protein of a higher molecular weight than the WT (Fig. 1D, lane 5). For the T2A_RBD construct, untagged NSP3 was readily identifiable, and again a fainter band corresponding to the predicted molecular weight of fused NSP3-T2A_RBD was also seen (Fig. 1D, lane 6, red and black asterisks, respectively). The RBD product (~23 kDa) was reproducibly not detectable, possibly due to discontinued translation as a result of ribosome falloff at the T2A site, or degradation in the rabbit reticulocyte lysate system. These results show that the spike epitopes were successfully translated in a cell-free system and that the T2A peptide was functional.

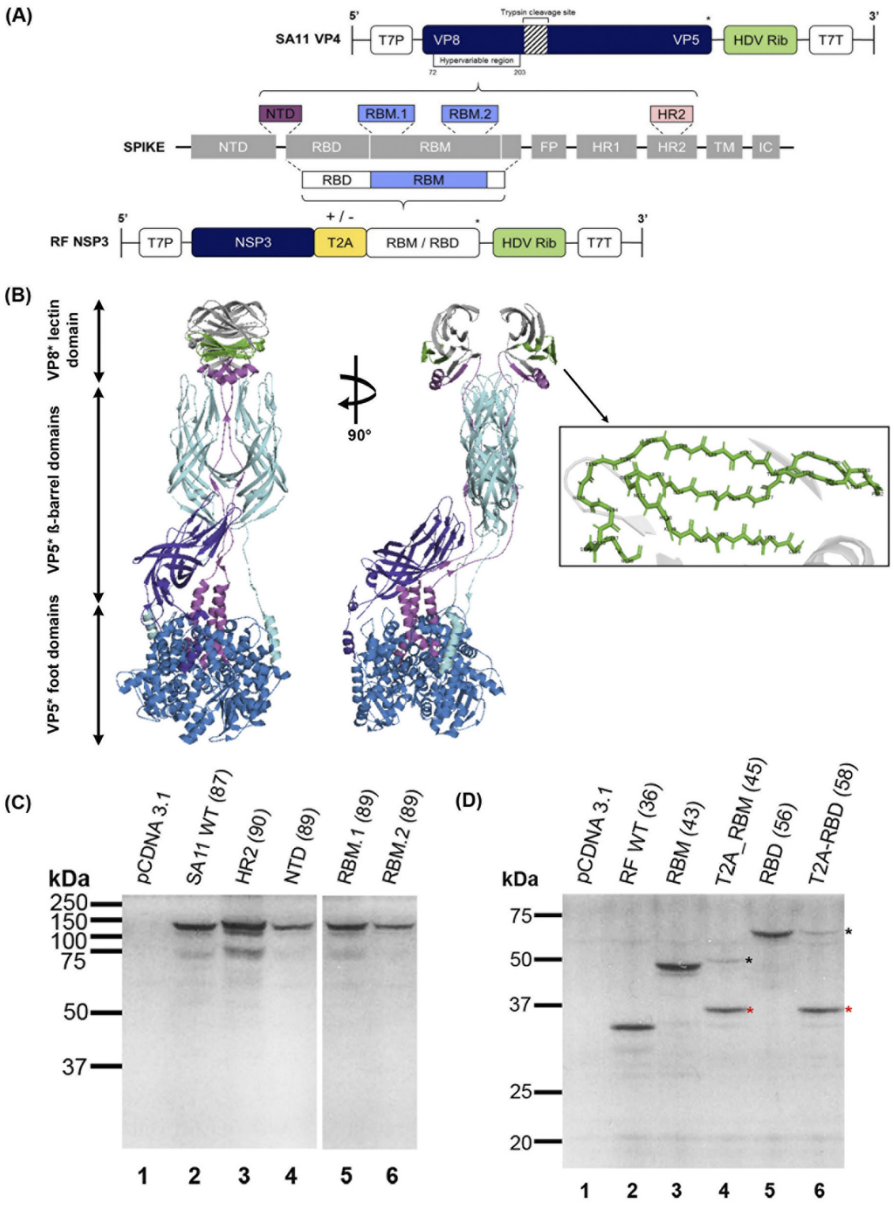


FIG 1 Design and validation of rotavirus (RV) VP4 and nonstructural protein 3 (NSP3) plasmid constructs used to generate mutant viruses. (A) Schematic showing overall topology of SARS-CoV-2 spike protein in gray boxes: N-terminal domain (NTD), receptor binding domain (RBD), which contains the receptor binding motif (RBM), fusion peptide (FP), heptad repeats 1 and 2 (HR1 and HR2), transmembrane region (TM), and the intracellular domain (IC) (adapted from Lan et al. [62]). Dashed lines represent selected regions of the spike protein (colored boxes) that were inserted into the hypervariable region of the SA11 viral protein 4 (VP4) gene (top) and the C terminus of the RF NSP3 gene (bottom). SA11 VP4 was edited to incorporate either NTD, RBM.1, RBM.2, or HR2 spike peptide sequences. RF NSP3 was fused with either the RBD or RBM sequence with or without Thosea asigna virus 2A (T2A) (yellow box),

(Continued on next page)

Establishment of a reverse genetics system for bovine RV RF strain and rescue of mutant viruses. To develop a reverse genetics system for the bovine RF strain of RV, BSR-T7 cells were transfected with plasmids encoding each of the 11 gene segments of the RF strain along with plasmids encoding the vaccinia virus capping enzyme (D1R and D12L) (Fig. 2A). Transfected cells were cocultured with MA104 cells and supernatant was then passaged on fresh MA104 cells. Following amplification in MA104 cells, supernatant was harvested as virus stock (Fig. 2B). To compare the replication kinetics of the engineered RF virus with the RF isolate, we first performed multicycle infections in MA104 cells at a low multiplicities of infection (MOI) (Fig. 2B). The two viruses showed similar overall kinetics, exhibiting exponential growth for around 16 h and then plateauing (Fig. 2B). From 24 hours postinfection (hpi), the titers of rescued RF were approximately a $\log_{10}$ lower than the RF isolate (24 and 48 hpi: *, $P \leq 0.05$; 32 hpi: ***, $P \leq 0.001$) (Fig. 2B). Although the RF isolate replicated better than the engineered RF, the endpoint titer of engineered RF showed that it grows efficiently in cultured cells. These data suggest that the bovine RV RF strain was successfully rescued using a plasmid only-based reverse genetics system.

Next, we used our established reverse genetics system for the bovine RV RF strain to generate the VP4 and NSP3 mutant viruses. With the exception of the N-terminal domain (NTD) mutant, all VP4 mutants displayed significantly lower titers compared to the WT, with an approximately two-$\log_{10}$ decrease in the titers observed across the panel (Fig. 2C). Nonetheless, all VP4 mutants were successfully rescued on all attempts. The VP4 mutants exhibited a similar but smaller speckled plaque phenotype to the WT (Fig. 2D). Statistical analyses of plaque sizes were not performed due to the ambiguities in determining the peripheries of individual plaques and their nonuniform diameters within a well.

In contrast, compared to the WT virus, there were no statistically significant differences observed in titers of any of the NSP3 mutants (Fig. 2E). However, out of three attempts, RBM was successfully rescued twice while T2A_RBM was rescued at all times, with one of the rescues showing a two-$\log_{10}$ lower titer (Fig. 2E). The RBD mutant was only rescued once with a titer that was a $\log_{10}$ lower than the WT (Fig. 2E). Although T2A_RBD rescued on all attempts, it delivered a five-$\log_{10}$ variation in titers between the rescues (Fig. 2E). The viruses with NSP3 tags displayed smaller plaques than the WT, except T2A_RBM, whose plaques had a similar morphology to that of WT (Fig. 2F). Notably, in the presence of T2A peptide, plaque sizes were larger than when RBM or RBD was fused directly to the C terminus of NSP3 without T2A (Fig. 2F).

Thus, small peptide insertions into the hypervariable region of SA11 VP4 significantly reduced the virus titer, while the efficiency of viral rescue was affected by the size of the peptides fused to the C terminus of RF NSP3. Furthermore, these data demonstrated that the RF strain of RV was successfully rescued from cloned cDNAs (confirmed by sequencing, data not shown).

Characteristics of VP4 and NSP3 mutant viruses. Following virus rescue, we first examined the effects of peptide insertion on RV replication in cell culture. Multistep growth curves were performed for the VP4 and NSP3 mutants after infection of MA104 cells at a low MOI. From 24 hpi, all VP4 mutants had significantly lower titers ($\geq$1-$\log_{10}$) than the WT virus for at least one time point during the infection (Fig. 3A). Next, we assessed the impact of the introduced mutations on the total RNA expression levels for

FIG 1 Legend (Continued)

represented by +/− sign. Both gene segments were flanked by the T7 promoter (T7P) and the antigenomic hepatitis delta virus (HDV) ribozyme ("HDV Rib," green boxes) followed by T7 terminator (T7T). *Stop codons. Schematic not to scale. (B) Ribbon representation of VP4 (adapted from Settembre et al. [63]). Two orthogonal views are shown. The VP8* fragment is in magenta extending into the VP5* foot domains (in blue), the hypervariable region of VP4 is in gray and the region where SARS-CoV-2 peptides (omitted for clarity) were inserted is in green. VP5* β-barrel domains are in cyan and purple. (C and D) Coupled in vitro transcription and translation reactions of mutated SA11 VP4 (C) and RF NSP3 (D) segments were carried out using the TnT rabbit reticulocyte lysate system supplemented with [^{35}S]methionine. Samples were analyzed using SDS-PAGE and autoradiography. The molecular weight marker and the expected product sizes of each segment (in brackets) are indicated (kDa). Empty pCDNA 3.1 vector was used as a negative control. In panel D, black asterisks indicate T2A read-through product and red asterisks identify separated products. WT, wild type.

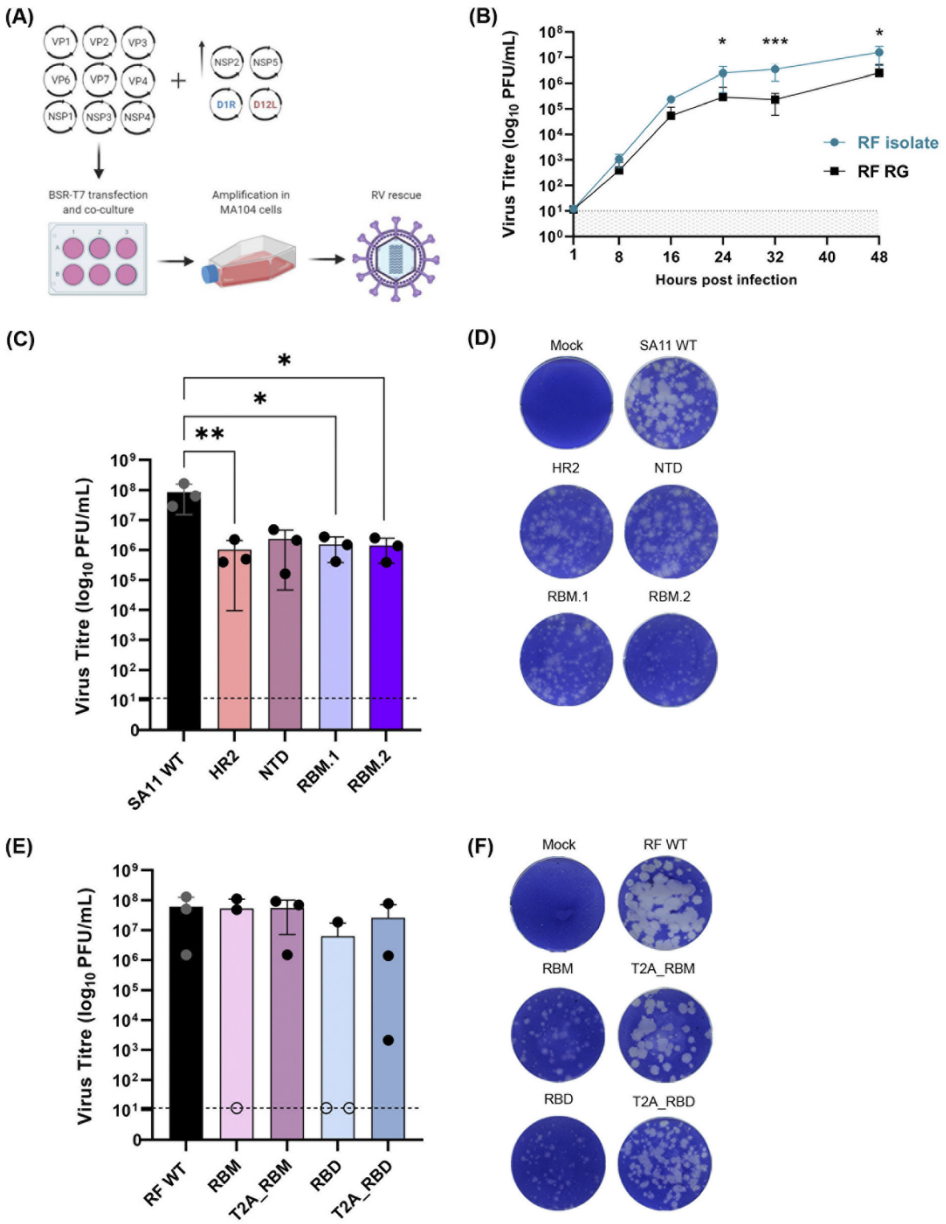


FIG 2 Generation of VP4 and NSP3 mutant viruses using plasmid-only based reverse genetics system. (A) Schematic representation of a 13-plasmid system for generation of mutant viruses expressing spike epitopes (adapted from Kanai et al. [42] and Komoto et al. [43]; created with BioRender.com). Full-length cDNAs representing each of the 11 gene segments were transfected into BSR-T7 cells with increasing amounts of two plasmids carrying NSP2 and NSP5 genes, along with two plasmids expressing vaccinia virus capping enzyme genes (D1R and D12L). Mutant viruses were rescued following amplification in MA104 cells. (B) Multistep growth curves comparing the bovine rotovirus (RV) RF isolate with reverse genetics (RG) virus. *, $P \leq 0.05$ (24 hours postinfection [hpi], $P = 0.0473$; 48 hpi, $P = 0.0407$). ***, $P \leq 0.001$ (32 hpi,

(Continued on next page)

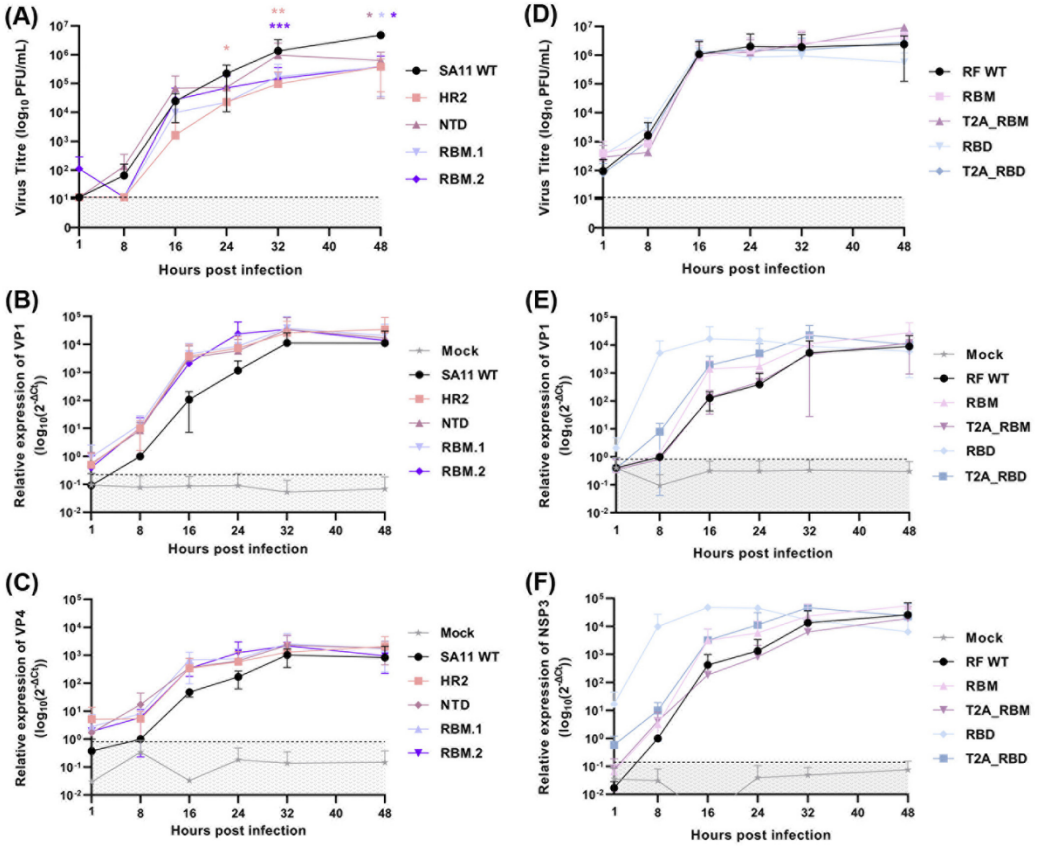


FIG 3 Replication kinetics and viral RNA content of WT and mutant viruses. (A and B) Multistep growth curves for VP4 and NSP3 mutant viruses respectively. *, $P \leq 0.05$ (24 hpi HR2, $P = 0.041$; 48 hpi NTD, $P = 0.047$; RBM.1, $P = 0.045$; RBM.2, $P = 0.031$). **, $P \leq 0.01$ (32 hpi HR2, $P = 0.009$). ***, $P \leq 0.001$ (32 hpi RBM.2, $P = 0.0006$). (C to F) Total cellular RNA levels of VP1, and VP4 (SA11 panel) or NSP3 (RF panel) transcripts. Data are from three independent experiments. Dashed line represents detection threshold.

VP1 and VP4 viral transcripts. No marked differences in transcript levels of either VP1 (Fig. 3B) or VP4 (Fig. 3C) were observed ($P \geq 0.05$, paired *t* test).

Conversely, the NSP3 mutants followed similar replication kinetics to RF WT with three log_{10} increases in titers between 8 and 16 hpi, with titers plateauing thereafter (Fig. 3D). Throughout, no major differences in titers between any of the viruses in the panel were observed. For the NSP3 virus panel, mutants produced higher levels of VP1 (Fig. 3E) and NSP3 (Fig. 3F) transcripts than the WT virus earlier in the time course. Higher RNA levels of VP1 and NSP3 in the RBD mutant are likely a result of higher RNA input (Fig. 3E and F, respectively). Possibly, a higher proportion of viral particles generated in the RBD virus stocks are nonviable, and therefore, more defective particles are present in the inoculum when equivalent PFU are used for the infections. Neither VP1

FIG 2 Legend (Continued)

$P = 0.0005$). Data are from three independent experiments. Dashed line represents detection threshold. Viral titers and plaque morphology versus WT for VP4 (C and D) and NSP3 (E and F) mutants. Dashed line represents detection threshold. (C) *, $P \leq 0.05$ (RBM.1, $P = 0.0205$; RBM.2, $P = 0.0217$). **, $P \leq 0.01$ (HR2, $P = 0.0018$). Representative results from three independent rescues are shown (except for RBD, which only rescued once). In panel E, open circles show failed rescues plotted at the limit of detection.

TABLE 1 Properties of VP4 and NSP3 mutant viruses[a]

Viruses				Protein products of mutated gene			
						VP4 cleaved products	
Name	Titers (PFU/mL)	Total size of VP4 gene (bp)	Rescue success	VP4 size (aa)	VP4 size (kDa)	VP5* size (kDa)	VP8* size (kDa)
SA11 WT	1.65E + 08	2,331	3/3	776	86.1	60	28
HR2	2.25E + 06	2,430		809	89.8		31
NTD	4.75E + 06	2,400		799	88.7		30
RBM.1	2.75E + 06	2,400		799	88.7		30
RBM.2	2.50E + 06	2,412		803	89.1		30
Name	**Titers (PFU/mL)**	**Total size of NSP3 gene (bp)**	**Rescue success**	**NSP3 size (aa)**	**NSP3 size (kDa)**	**T2A product/s (aa)**	**T2A product/s (kDa)**
RF WT	1.30E + 08	942	3/3	313	34.7	n/a	n/a
RBM	1.10E + 08	1,158	2/3	385	42.7		
T2A_RBM	9.20E + 07	1,212	3/3	403	44.7	330 + 73	36.6 + 8.1
RBD	1.88E + 07	1,521	1/3	506	56.1	n/a	n/a
T2A_RBD	7.70E + 07	1,575	3/3	524	58.1	330 + 194	36.6 + 21.5

[a]The hypervariable region of SA11 VP4 was tagged with SARS-CoV-2 spike peptides: heptad repeat 2 (HR2), N-terminal domain (NTD), and receptor binding motif (RBM) regions. The C terminus of RF nonstructural protein (NSP) was tagged with either the RBM or receptor binding domain (RBD) of SARS-CoV-2 spike with or without a Thosea asigna virus 2A (T2A) peptide; WT, wild type; n/a, not applicable.

nor NSP3 transcript levels differed significantly throughout the time course ($P \geq 0.05$, paired *t* test) (Fig. 3E and F).

Western blot analyses were then used to characterize the production of SARS-CoV-2 spike polypeptides and RV VP6 using whole-cell lysates from the infection time course. As different SARS-CoV-2 peptides were introduced in each mutant, antispike antibody affinity may vary between mutants and so expression levels cannot be compared. Due to lack of available antibodies, we were unable to measure RV VP4 and NSP3 protein levels directly.

For the detection of various SARS-CoV-2 spike peptides in cells infected with the VP4 mutants, a polyclonal SARS-CoV-2 spike antibody was used. As the spike peptides were introduced into the hypervariable region of VP4, we considered the possibility of detecting the full-length VP4 (sizes indicated in Table 1) and/or the VP8* cleaved product containing the spike peptides. At 32 hpi, full-length VP4 product was detected in heptad repeat 2 (HR2), NTD, and RBM.2 mutants (Fig. 4A, black asterisks). No SARS-CoV-2 spike peptide signal was detected for the RBM.1 mutant, likely reflecting selective clonality of the antibody used (Fig. 4A). Interestingly, at 32 hpi, the upper band representing the uncleaved VP4 product was brighter for RBM.2 (Fig. 4A, black asterisks), whereas for NTD, the lower band possibly representing the VP5* cleaved product was stronger than the upper band (Fig. 4A, green asterisks). These differences could be due to the distinct efficiencies of VP4 processing caused by the inserted peptides. The hypervariable region where the spike peptides were introduced is within the VP8* domain, and we would expect a product of around 30 to 31 kDa in the event of VP4 cleavage. Possibly, this is evident in NTD at 32 and 48 hpi (Fig. 4A, blue asterisks), although the NTD mutant showed several unexpected bands and so this may reflect nonspecific antibody binding; a faint product of similar molecular weight was also detected in the mock samples at 32 hpi (Fig. 4A).

We also examined RV VP6 production, which was detected in all VP4 viruses at 32 hpi (Fig. 4A). VP6 signal intensity did not correlate with spike signal intensity across viruses, again likely due to variable affinities of the spike antibody for the different spike peptides incorporated into VP4 (Fig. 4A).

In cells infected with NSP3 mutants, detection of both RBD and RBM peptides was possible using a polyclonal SARS-CoV-2 RBD antibody. Cross-reactivity for the RBM-expressing virus was first detected at 8 hpi, with increased levels present at 16 and 32 hpi (Fig. 4B). Antibody cross-reactivity for the T2A_RBM virus was only visible between 16 and 32 hpi (Fig. 4B). The band detected is consistent in size with an unseparated NSP3-T2A_RBM protein product (Table 1 and Fig. 4B, black asterisks). The RBM peptide of around 10 kDa was not detected, consistent with its *in vitro* translation efficiency

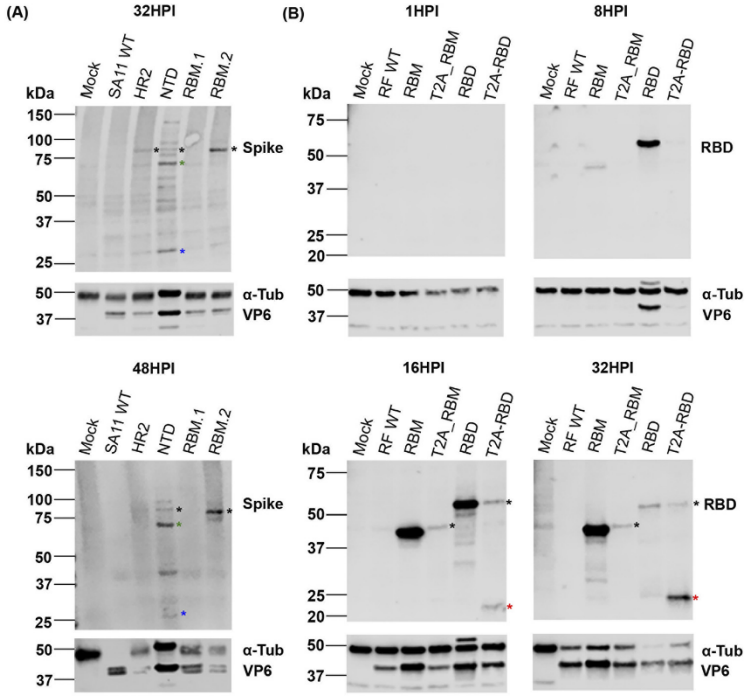


FIG 4 Expression of VP6 and SARS-CoV-2 spike peptides in infected cells. Cells infected at low MOI were harvested at 1, 8, 16, 24, 32, and 48 hpi. Whole-cell lysates were analyzed by SDS-PAGE and Western blot using polyclonal antibodies against RV VP6, and spike for SA11 mutants (A) or RBD for RF mutants (B). Alpha-tubulin (α-Tub) was used as a loading control. In panel A, black asterisks denote uncleaved VP4 product. Cleaved VP4 products VP5* and VP8* are marked by green and blue asterisks respectively. In panel B, black asterisks show T2A read-through product and red asterisks identify separated products. Representative results from three independent experiments are shown. Position of molecular weight markers are indicated (kDa).

(Fig. 4B and 1B). For the RBD virus, expression of the RBD peptide was first observed at 8 hpi but declined from 16 hpi paralleling the disappearance of tubulin (likely due to cell death) (Fig. 4B). In the T2A_RBD mutant, both NSP3-conjugated and unconjugated RBD products were detectable from 16 hpi (Fig. 4B, black and red asterisks, respectively). Over time, the RBD peptide (~23 kDa) became more apparent, as the NSP3-RBD signal diminished, confirming the functionality of the T2A element (Fig. 4B, red asterisks).

It is unclear why the T2A-induced ribosomal skipping appeared to improve in efficiency over the course of infection. It is possible that the stability of the fused peptides is lower than the separated peptides. Similarly, over the course of infection, RBM protein levels increased throughout, whereas RBD protein levels increased until 16 hpi after which they dropped dramatically (Fig. 4B).

The expression of VP6 was first observed at 8 hpi during infection with the RBD mutant, coinciding with the signal of the RBD peptide, but was only detected for the remaining viruses from 16 hpi (Fig. 4B). Since equal MOIs were used for time course infections, higher input of genomic RNA copies could explain earlier V6 detection in the RBD mutant. This is consistent with the higher transcript levels detected for the RBD mutant at 8 hpi (Fig. 3E and F). It is possible that in the event of a packaging defect of the RBD mutant more input genome copies would be required to deliver equal numbers of infectious particles.

Infection with both mutants resulted in a similar drop in tubulin levels (Fig. 4A and B), suggesting that the two mutants have different protein turnover rates.

In summary, introducing short SARS-CoV-2 peptides into the hypervariable region of VP4 impacted the virus yield, whereas fusing SARS-CoV-2 spike peptides to the C terminus of NSP3 with or without T2A did not affect the virus titer or replication kinetics. Nevertheless, SARS-CoV-2 spike peptides were detectable in the majority of mutants using polyclonal antibodies, encouraging further investigations into the potential of these strategies for heterologous peptide presentation. Use of monoclonal SARS-CoV-2 spike antibodies could be considered in the future studies as these may be more specific for the detection of spike epitopes expressed by mutant viruses.

Effect of gene mutation on viral genome packaging. For viruses of the *Reoviridae* family, genome packaging is a tightly orchestrated and controlled process, so increasing the RV genome size may affect packaging efficiency (29, 67–69). To examine whether our mutations may have affected genome packaging, RNA was extracted from equal volume of purified viruses and analyzed by urea-PAGE and silver staining.

The VP4 virus panel showed the expected constellation of genome segments, but as samples were resolved for an extended period of time to visualize the small changes in segment 4 (VP4) band sizes expected as a result of SARS-CoV-2 spike peptide insertion, segments 10 and 11 ran off the gel (Fig. 5A). When gels were run for a shorter time, no differences in band densities of segments 10 and 11 were seen (data not shown). Mutated VP4 segments migrated more slowly than the WT VP4 segment, reflecting the various sizes of inserted spike sequences (Fig. 5A, highlighted in red). Densitometry analysis showed no substantial differences in band density between mutated VP4 and WT VP4 segments (normalized to WT: HR2 = 0.96, NTD = 1.11, RBM.1 = 0.87, and RBM.2 = 0.96), suggesting no obvious packaging defects were introduced by small sequence insertions into the hypervariable region of VP4 (Fig. 5A).

The resolved segments for the NSP3 virus panel also showed the expected pattern of 11 RNA segments, as well as a prominent background band present in a mock infected sample that migrated between segments 1 and 2 (Fig. 5B). Segment 7 (encoding NSP3) from all mutants migrated notably slower than from WT virus, corresponding with its increased gene size (Fig. 5B, highlighted in red). Comigration of segments 7, 8, and 9 in the WT made it difficult to reliably separate segment 7, precluding direct quantitative analyses (Fig. 5B). Nevertheless, no obvious defect in packaging was observed through visualization of the complete genome of RF mutants.

To further evaluate whether virion infectivity may have been affected by genome mutagenesis, the genome copy number to PFU ratio was determined for the two panels of viruses, measuring VP1 and either VP4 (SA11 panel) or NSP3 (RF panel) segments (Fig. 5C to F). No significant differences were observed in the levels of VP1 and VP4 segments across the VP4 mutants relative to WT, although for all mutants a 1 to 2 $\log_{10}$ increase in RNA copies required to make an infectious virion was observed for both VP1 and VP4 ($P \geq 0.05$, paired *t* test) (Fig. 5C and D).

In contrast, all NSP3 mutants, with the exception of RBD, had equivalent numbers of VP1 and NSP3 segments (Fig. 5E and F). RBD had a significantly higher VP1 ($P = 0.005$, paired *t* test) and NSP3 ($P = 0.010$, paired *t* test) segment copy number:PFU ratio, with over 100-fold more copies of RNA required to make a fully infectious particle (Fig. 5C and D).

VP4 is important for viral attachment and entry as well as for the maturation of TLPs that constitute an infectious virus (3). Therefore, it was considered that mutation of VP4 may affect the assembly of the viral structural proteins. To test this, purified VP4 viruses were further analyzed by SDS-PAGE followed by Coomassie staining (NSP3 viruses were not included because the spike epitopes were fused to NSP3 which is not incorporated into virions) (Fig. 5G). Apart from VP3 and VP7 proteins, major structural proteins were detected throughout (Fig. 5G). VP4 was readily detectable for all viruses in the panel and so amounts were quantified by densitometry and normalized to that of WT virus. This showed no obvious VP4 incorporation defects, with all mutant viruses

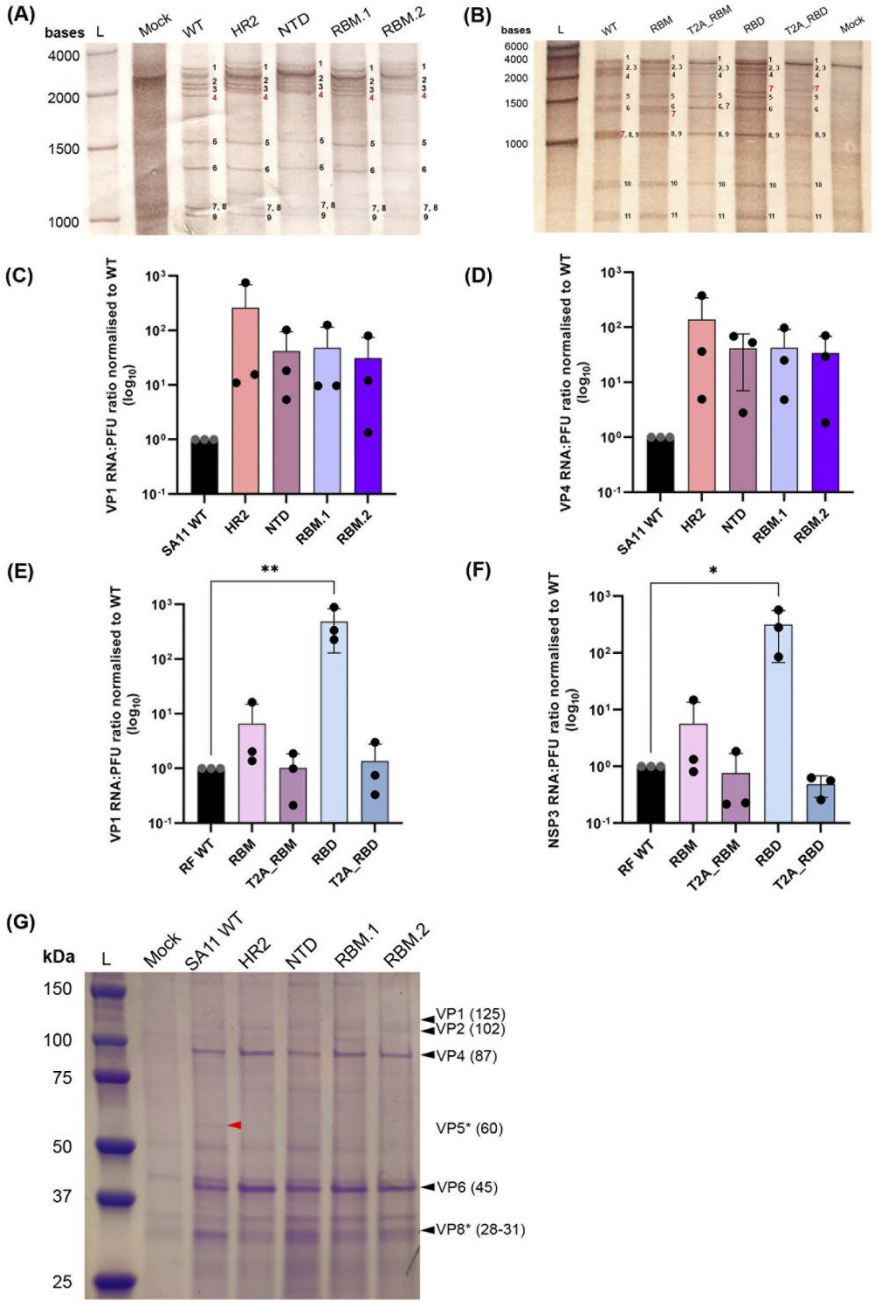


FIG 5 Impact of VP4 and NP3 mutation on packaging of RV RNA and proteins. Extracted RNA from virus stocks was analyzed by urea-PAGE and silver staining for SA11 mutants (A) and RF mutants (B). Individual RNA segments are labeled in black, and mutated

(Continued on next page)

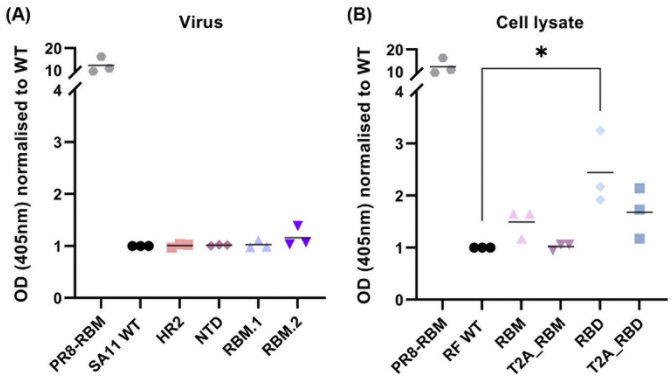


FIG 6 Detection of spike antigens expressed by mutant viruses. Presence of spike antigens were evaluated by indirect ELISA using viral supernatant of VP4 mutants (A) and cell lysates for NSP3 mutants (B). Colored symbols represent individual data points obtained from three independent experiments. Optical density (OD) (405 nm) signal for all mutants was normalized to WT (background signal). PR8-RBM represents positive control. (A) $P > 0.05$. (B) *, $P \leq 0.05$ (RBD, $P = 0.032$).

showing very similar VP4 amounts to the WT virus (relative to WT, values were as follows: HR2 = 1.07, NTD = 0.99, RBM.1 = 1.09, and RBM.2 = 1.06). Thus, the mutant VP4 proteins appear to be incorporated into virus particles with unaltered efficiency.

A notable difference between the WT virus and its mutants was the cleavage pattern of VP4 in the presence of trypsin used for propagating the panel of mutant viruses. A band of the predicted size for VP5* was visible in the WT (Fig. 5G, red arrow). However, no corresponding band was visible for any of the VP4 mutants expressing heterologous peptides, further supporting our earlier hypothesis that peptide insertion into the VP4 hypervariable region impairs its processing into VP8* and VP5* (Fig. 4A).

In contrast, VP8* was detected in all mutants, and it migrated slightly higher than WT VP8, corresponding to the additional peptide sequences present in the hypervariable domain of VP4 (Fig. 5G). Nevertheless, VP8* band density was stronger for WT than for the mutants, again consistent with inefficient tryptic cleavage of VP4.

Overall, no major effect on viral assembly was observed in either of the mutant viral panels, with the exception of inserting the large RBD tag into NSP3, which caused a defect that could be overcome by incorporation of a T2A element.

Detection of SARS-COV-2 spike antigens by indirect ELISA. To further investigate the viability of using RV as a delivery vector for SARS-CoV-2 spike antigens, we assessed whether the RV mutants expressing spike peptides would cross-react with RBD antibodies in an indirect ELISA. For the VP4 mutants, no antibody cross-reactivity was identified in this assay using the polyclonal SARS-CoV-2 RBD antibody (Fig. 6A). Additionally, no cross-reactivity was observed using the polyclonal SARS-CoV-2 spike antibodies to target the various spike peptides in the VP4 mutants, or with the PR8-RBM-positive control (data not shown), reflecting the lack of suitability of these antibodies for use in this assay.

Except the T2A_RBM mutant, all NSP3 mutants consistently showed a higher signal than the background signal from the WT virus, although only the RBD mutant showed a significantly higher signal than WT ($P = 0.032$, paired t test) (Fig. 6B).

FIG 5 Legend (Continued)

RNA segment notations are in red. Lane "L" represents High Range RNA Ladder showing band size of RNA transcripts. RNA:PFU ratios of VP1 and either SA11 VP4 (C and D) or RF NSP3 (E and F) genes were determined by RT-qPCR and a ratio of copy number to viral titer was calculated; values were then normalized to WT. Dots represent individual samples from three independent rescues with each performed in triplicate. (E) **, $P \leq 0.01$ (RBD, $P = 0.005$). (F) *, $P \leq 0.05$ (RBD, $P = 0.010$). (G) Electrophoretic profile of viral proteins from purified VP4 mutants visualized using Coomassie brilliant blue. Lane "L" represents protein ladder indicating the molecular weight markers (kDa). Black triangles show the positions of viral proteins and their expected protein sizes (in brackets) are indicated (kDa). VP3 and VP7 proteins were not detected. VP5* protein was detected in WT only (red triangle).

Based on these observations, RF NSP3 may be a better target for heterologous peptide conjugation than SA11 VP4 for expression of immunogenic antigens, which is at least partly attributable to the larger insertion site tolerated in the corresponding genome region.

Bovine enteroids as a species-specific model for viral infection. Cell culture-based assays identified the RF NSP3 gene as a strong candidate for expressing immunogenic heterologous peptides. Orally administered live attenuated RV vaccines replicate in the gastrointestinal tract, and so bovine intestinal organoids containing enterocytes, goblet, Paneth, enteroendocrine, and stem cells (70) were used to investigate if a more physiologically representative system was susceptible to infection with the NSP3 mutants. A bovine organoid system was chosen as the RF strain was first identified in diarrhoeic calves (71).

3D bovine enteroids were infected with NSP3 mutants at an approximate MOI of 10, stained with anti-VP6 and anti-RBD antibodies, and imaged by confocal microscopy (Fig. 7A and B, respectively). At 24 hpi, VP6 was predominantly detected in the epithelium comprised of mature enterocytes lining the apical surface of the organoid lumen (Fig. 7A), which is consistent with previous findings showing their preferential infection by RVs (17, 18). In contrast, VP6 distribution in RBM and T2A_RBM mutants was detected around the nuclei of cells located within the organoid lumen and in the punctate cytoplasmic inclusion bodies, most likely viroplasms, sites of RV replication and assembly (3) (Fig. 7A, VP6 panel). However, antibodies targeting NSP2 and NSP5 proteins, which together are necessary and sufficient for the formation of viroplasms (72), are not commercially available and so viroplasm identification could not be verified. VP6 signal in the RBD and T2A_RBD mutants was not apparent in the lumen and was mainly detected around the nuclei of cells in the organoid lining (Fig. 7A, VP6 panel).

Following staining with the SARS-CoV-2 RBD antibody, diffuse cytoplasmic signal was detected in a small number cells of organoids infected with the RBM mutant, while other mutants did not produce any visible anti-RBD cross-reactivity (Fig. 7B). However, the signal for RBM was weak relative to the background, indicating either that the expression of SARS-CoV-2 spike peptides was hindered in bovine enteroids or that the antibody does not work well in this context.

DISCUSSION

We have used simian SA11 and bovine RF RV strains as viral vectors to express various SARS-CoV-2 spike epitopes as a model system with which to test the potential for expression of multivalent antigens. Tagging of the VP4 protein found on the surface of the virion with smaller peptides consistently impaired viral growth and did not yield strong antibody cross-reactivity. Conversely, relatively large foreign sequences could be tagged to the C terminus of NSP3 protein, mostly without impairing viral titers and replication kinetics, and cross-reactivity with SARS-CoV-2 RBD antibodies was also demonstrable.

To investigate the feasibility of using RV as an expression vector, we analyzed the effect of introducing SARS-CoV-2 spike peptides into the "head" region of the VP4 (i.e., VP8* domain), outside the sialic acid binding domain. Since VP8* is used in several vaccine platforms such as protein subunit or nanoparticle vaccines to induce RV-specific neutralizing antibodies (73–78), we hypothesized that the expression of SARS-CoV-2 spike peptides by VP8* may similarly generate neutralizing antibodies. However, incorporation of various spike epitopes into the VP8* lectin domain between amino acid positions 164 to 198 consistently showed significant decrease in viral titers, despite having no obvious effect on the rescue efficiency (Fig. 2C).

It is possible that reduced infectivity and lack of uniformity of plaque sizes of VP4 mutants (Fig. 2C and D and Fig. 3A) could be due to inefficient conformational transition of VP4 following proteolytic cleavage (3, 20, 21). Following cell attachment, VP8* lectin domains at the tip of VP4 dissociate and expose the hydrophobic loops of the VP5* β-barrel domains (19, 63, 64). This enables interaction of the VP5* hydrophobic loops with the lipid bilayer and perforation of the target membrane by the VP5* foot

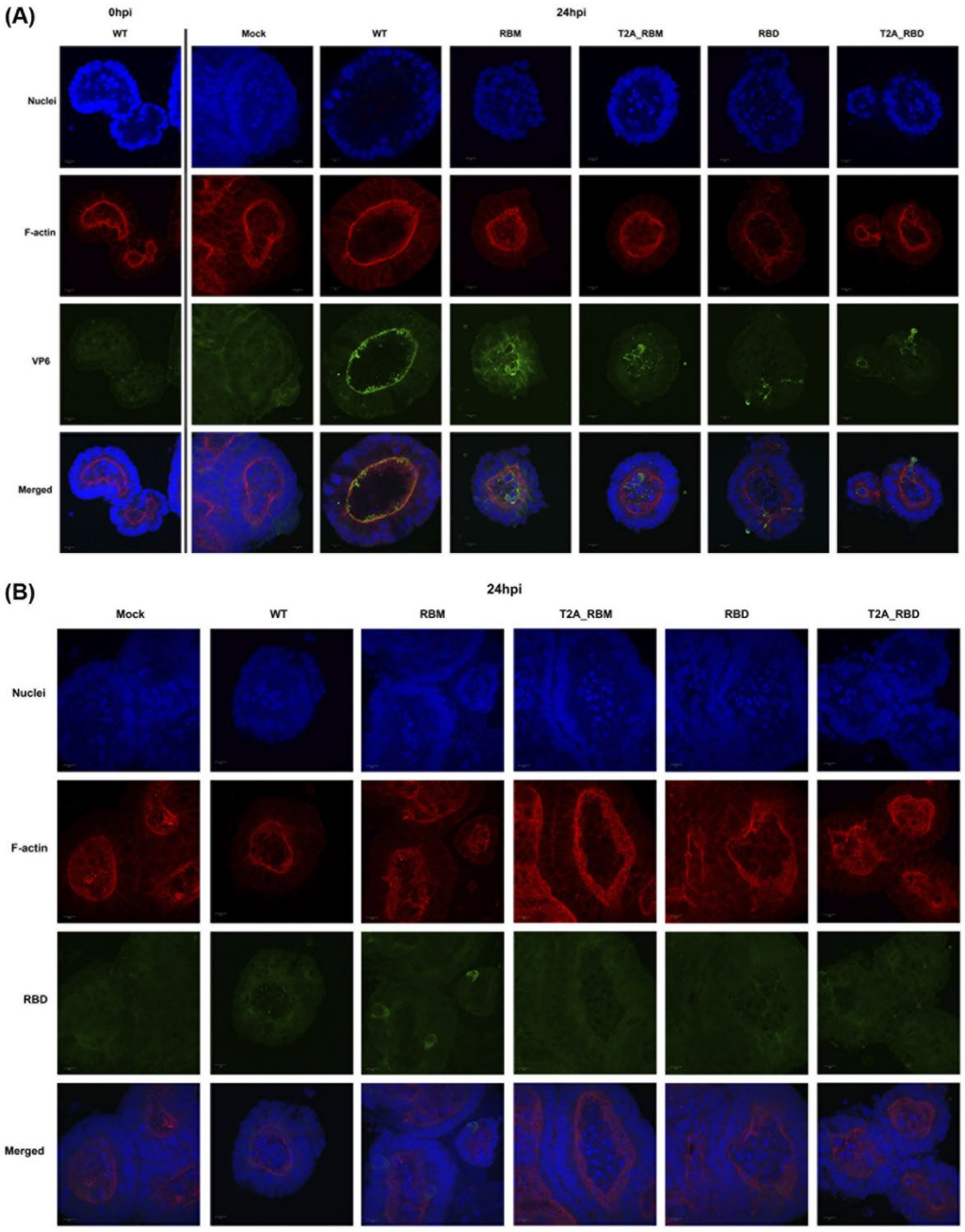


FIG 7 Bovine enteroids infected with NSP3 mutant viruses. Three-dimensional bovine enteroids were infected with the panel of NSP3 mutants, or mock at an approximate multiplicities of infection of 10 PFU per cell and fixed at 24 hpi, or at 0 hpi for WT as a control for background signal. Cells were stained for RV VP6 (A) and RBD of SARS-CoV-2 spike (B). Scale bar represents 10 μm.

domain, leading to RV entry, analogous to refolding of influenza virus hemagglutinin during membrane fusion (63, 64, 79–84). Therefore, our mutation of the VP8* lectin domain may have indirectly affected the association of the hydrophobic loops of the VP5* β-barrel domains with the target membrane.

Alternatively, while we saw no obvious effects of VP4 mutagenesis on genome packaging (Fig. 5A, C, and D), the absence of the VP5* detection in the mutants (where a band of the correct size was identified in the WT virus) could indicate inefficient cleavage of VP4 on TLPs (Fig. 5G). *In vitro* assembly of TLPs has shown that DLPs require addition of VP4 before VP7, and that the tryptic cleavage occurs following addition of VP7 (82, 85). Interaction of VP7 subunits with the VP5* foot domain stabilize and anchor the VP4 onto the virion, allowing the protease to cleave only the linker sequence that bridges VP8* and VP5*, which in our mutants is intact (20, 21, 63, 64). It is possible that our peptide insertions may have altered the VP4 conformation at this interaction site, disrupting VP4 cleavage and rendering VP5* undetectable in our assays (Fig. 4A and 5G). Finally, we were unable to detect any cross-reactivity of the SARS-CoV-2 spike peptides with spike (not shown) or RBD antibodies in ELISA (Fig. 6A), further suggesting limitations of using SA11 VP4 as an expression vector for foreign peptides, at least with the chosen peptides. It is possible, however, that these peptides may be immunogenic *in vivo*, which could be tested in a suitable animal model such as mice (86). Using a multiplexed approach, challenge with a combination of VP4 mutants could be applied to establish a more robust SARS-CoV-2 specific immune response.

Previous studies have shown that SA11 NSP3 is able to tolerate insertions of foreign sequences at its C terminus (45, 51, 52); we have demonstrated that this is also possible using the bovine RF strain NSP3. This is advantageous as the current live attenuated pentavalent RV vaccine RotaTeq utilizes a bovine RV backbone reassorted with different human strains, and the use of a nonprimate origin RV vector would lower susceptibility to preexisting immunity (9, 87, 88). In contrast, a monovalent, human live attenuated RV vaccine Rotarix contains only one RV strain of G1P[8] specificity (89, 90). The bovine backbone is amenable to multiplexing, and therefore will offer a broader protective immunity than those containing one specific RV strain.

The RF strain of RVs demonstrates a positive feedback mechanism for NSP3 mRNA translation where the 3′-GACC end of mRNA molecules and both domains (N and C terminal) of NSP3 are required (91). Accumulation of the first NSP3 molecules generated during infection triggers NSP3-dependent translation by specifically binding the terminal 3′ GACC of mRNA, and once established, NSP3 is available for translation of other viral mRNAs (91). Depending on the virus strain, the 3′-GACC canonical sequence differs in some segments of RV, and the same positive feedback mechanism is not observed in the SA11 strain (3, 91).

Rescue of the RF mutant with the RBD peptide (193 amino acids) fused directly to NSP3 was achieved only once in three attempts (Fig. 2E), possibly reflecting an impaired function of NSP3. However, introducing the RBM peptide produced a virus of similar titers to the WT and its rescue efficiency was not as affected (Fig. 2E). The only notable difference between the RBM and RBD was the size of the introduced peptide, suggesting that insertions larger than 72 amino acids may affect the function of NSP3. Increasing the size of the ORF by incorporating the T2A element improved the titer and the rescue efficiency of the T2A_RBD mutant (Fig. 2E), suggesting that the rescue defect imparted by peptide introduction is at the protein rather than the RNA level. Previous studies showed that inclusion of the porcine teschovirus 2A element (P2A) between SA11 NSP3 ORF and the coding sequence for UnaG resulted in NSP3 retaining its ability to induce nuclear localization of PABP (51). These recombinant viruses expressed both NSP3 and UnaG proteins, and grew to peak titers similar to those of SA11 WT (45, 51). Recombinant simian rhesus rotavirus (RRV) strain expressing green fluorescent protein from NSP3 C terminus, separated by P2A element, also had comparable growth kinetics to the WT virus (44). However, a more recent study showed that

the titers of recombinant SA11 RVs expressing SARS-CoV-2 peptides from the C terminus of NSP3 ORF with P2A were 0.5 to 1 log lower than the WT (52). Furthermore, the recombinant SA11 RV expressing the whole of SARS-CoV-2 spike from NSP3 with P2A showed genetic instability (by passage three) in which sequence deletions were apparent with continued passage (52). Our results confirm that introducing peptides to the C terminus of NSP3 separated by the T2A element did not affect the function or genetic stability of NSP3 (Fig. 2E, 4B, and 5B). In the absence of a T2A element, there may be a limit to the length of additional foreign sequences that the bovine RF NSP3 can accommodate. However, it is not possible to rule out that the activity of different 2A elements and the effects of different RG systems used for viral recue can vary (44, 45, 51, 52, 66).

We also found that the RNA:PFU ratios as well as the copy number of NSP3 and VP1 transcripts were affected for the RBD mutant only, suggesting that a higher proportion of viral particles generated are nonviable (Fig. 5E and F). This cannot be attributed to the longer length of this segment, as the T2A_RBD mutant did not have a raised RNA:PFU ratio (Fig. 5E and F). A higher number of defective virions in the viral progeny may explain the variable plaque sizes of NSP3 mutants (Fig. 2F). NSP3 exerts its role in virus replication by regulating viral mRNA translation. This is consistent with the observed increase in RNA:PFU ratios being nonspecific to a particular segment.

With the exception of T2A_RBM, NSP3 mutants expressing SARS-CoV-2 spike peptides cross-reacted with the SARS-CoV-2 RBD antibody (Fig. 6B), supporting the previously proposed viability of using NSP3 as a tagging system (52) and demonstrating its application to other RV strains. Follow up studies using sera from COVID-19 patients are needed to confirm antibody responses to the antigens produced by the NSP3 mutants.

Overall, we conclude that further investigation is required to evaluate the potential for peptide insertion into VP4 as a potential vaccine expression platform, using animal models to examine immunogenicity. Here, small peptide insertions into VP4 resulted in a log reduction in virus titer, which is tolerable when we consider other live attenuated vaccines currently in use, such as the influenza A vaccine FluMist (reviewed in reference 92). On the other hand, including the T2A element is beneficial for expressing foreign antigens as it allows coexpression of NSP3 and large peptides with little effect on the viral rescue efficiency, titer, and replication, all which are important traits for live attenuated vaccine development. In the absence of the T2A element, there may be a limit to the amount of additional foreign sequences that the RF NSP3 can accommodate. Nevertheless, there is potential to package additional heterologous peptides into the RV core, which may be several thousands of nucleotides in excess of the native dsRNA genome (93). Our results offer a possibility of utilizing a bovine RV RF strain as a backbone to facilitate the development of recombinant RV-based vaccines; here we used SARS-CoV-2 as an example, but this can be extrapolated to other gastrointestinal pathogens.

MATERIALS AND METHODS

Cell lines. African Green monkey kidney epithelial (MA104) cells were cultured in Dulbecco's modified Eagle's medium (DMEM) (Sigma-Aldrich) supplemented with heat inactivated 10% fetal bovine serum (FBS) (Gibco) and 1% penicillin- streptomycin (Gibco). Cells of the BSR-T7 clone of baby hamster kidney fibroblasts (BHK-21 cells) (94), constitutively expressing T7 RNA polymerase, were cultured in Glasgow's minimal essential medium (Gibco) supplemented with 1% tryptose phosphate broth (Gibco), heat inactivated 10% FBS (Gibco), and 1% penicillin-streptomycin (Gibco). Both cell lines were a kind gift from the laboratory of Prof. Massimo Palmarini (MRC-University of Glasgow Centre for Virus Research, UK). Cells were passaged twice weekly and maintained at 37°C, 5% CO_2. At every fifth passage, BSR-T7 cells were maintained in medium containing G-418 selection (1 mg/mL) (Invitrogen).

Viruses. RV strains used were simian RV SA11 strain (G3P[2]) and bovine RV RF strain (G6P[1]) and were rescued using reverse genetics (detailed below) (42, 71, 95, 96). The natural isolate of the RF strain was a kind gift from the laboratory of Dr. Ulrich Desselberger (University of Edinburgh, UK). Viruses were propagated in MA104 cells cultured in DMEM supplemented with 0.5 μg/mL porcine pancreatic trypsin type IX (Sigma-Aldrich), and viral titers were measured by plaque assay as described below.

Design of RV expression vectors containing spike epitopes. To engineer mutant RVs expressing SARS-CoV-2 spike peptides, we utilized two strategies involving both the simian RV SA11 strain and the

bovine RV RF strain. It was previously suggested that the surface protein VP4 of the SA11 strain tolerates short immunogenic peptides inserted at specific sites with minimal impact on viral replication or particle assembly (79). Guided by the available structure model of rhesus rotavirus (RRV) VP4, serotype G3 P5B[3] (strain RVA/Monkey/United States/RRV/1975/G3P5B[3]) (PDB: 4V7Q) (63), we chose an exposed loop outside the sialic acid-binding domain located within the "head" of the VP4 spike (Fig. 1A and B) to investigate the possibility of VP4 alteration. Moreover, several neutralization escape mutants were ascribed to the amino acid changes within this region (97), consistent with its accessibility to antibodies. A number of B-cell linear epitopes derived from either the heptad repeat 2 (HR2), N-terminal domain (NTD), or RBM regions of the SARS-CoV-2 spike protein were selected for insertion (61, 98, 99) (Fig. 1A). These epitopes were introduced into the hypervariable region of SA11 VP4 between amino acid position 164 and 198 with linker sequences (63, 64) to increase their accessibility. Plasmid constructs expressing HR2 and NTD of the SARS-CoV-2 spike peptides were generated using Q5 site-directed mutagenesis kit (New England BioLabs) with primers listed in Table 2 and pT7-VP4SA11 (Addgene no. 89165) expression vector according to the manufacturer's protocol. RBM.1 and RBM.2 constructs were ordered as gene fragments (gBlocks; Integrated DNA Technologies) and cloned into pT7-VP4SA11 expression vector using NheI and MfeI restriction enzymes. All plasmids were amplified by transformation into chemically competent *Escherichia coli* DH5α and purified using Qiagen Plasmid Midi Kit (Qiagen) according to the manufacturer's protocol. The presence and the size of the mutation in each plasmid was verified by Sanger sequencing (GATC Biotech or Genewiz, Germany). These epitopes were curated from the scientific literature by the Immune Epitope Database available at VIPR (https://www.viprbrc.org/brc/). Only unstructured epitopes of up to 15 residues were chosen (Table 3). The SA11 strain was selected for this mutagenesis due to the structural characterization of VP4 of its close relative RRV, its user-friendly plasmid only-based reverse genetics system and its rapid growth kinetics.

To test whether tagging a RV strain more closely related to the bovine virus backbone used in the pentavalent RotaTeq vaccine (87), we used the bovine RV RF strain. Here we modified the C terminus of the RF strain NSP3 ORF to express spike epitopes coding for the RBD or RBM of SARS-CoV-2, with or without an intervening Thosea asigna virus 2A (T2A) peptide (66, 100, 101) (Fig. 1A). The inclusion of the T2A sequence was employed in order to lower the risk of interfering with the function of the NSP3 gene. On the other hand, increasing the segment size by incorporating the T2A peptide could further affect the antigenic processing; hence, both approaches were trialed. The panel of constructs was assigned the notation RBM, T2A-RBM, RBD, and T2A-RBD.

Plasmid construction. pT7 plasmids used for reverse genetics of SA11 RV were kindly provided by Takeshi Kobayashi (42) through the Addgene plasmid repository under ID number 89162-72. To generate plasmids used for reverse genetics of the bovine RV RF strain, constructs were designed to encode each of the 11 RF gene segments, flanked at the 5′ end by a T7 promoter (T7P) and at the 3′ end by an antigenomic hepatitis delta virus ribozyme sequence, followed by the T7 terminator sequence as in Kanai et al. (42). The GenBank accession numbers of the bovine RV RF strain genome sequences are as follows: segment 1 (KF729687), segment 2 (KF729642), segment 3 (KF729645), segment 4 (KF729650), segment 5 (KF729656), segment 6 (KF729658), segment 7 (KF729664), segment 8 (KF729667), segment 9 (KF729675), segment 10 (KF729677), and segment 11 (KF729684) (96). The constructs were synthesized by Invitrogen GeneArt on either pMK-RQ (kanamycin resistance), pMA-RQ, or pMA-T (ampicillin resistance) vectors. RF strain NSP3 constructs RBM, T2A-RBM, RBD, and T2A-RBD were ordered as gene fragments from Invitrogen GeneArt and cloned into pT7-NSP2SA11 expression plasmid (Addgene no. 89169) after the NSP2 ORF was removed using SmaI and SalI restriction enzymes. All plasmids were amplified by transformation into chemically competent *E. coli* DH5α, except for the RF VP7 encoding plasmid for which DH10β cells were used, and purified using Qiagen Plasmid Midi Kit (Qiagen) according to the manufacturer's protocol. The presence and the size of the mutation in each plasmid were verified by Sanger sequencing (GATC Biotech or Genewiz, Germany) using primers listed in Table 2. Sequence results were analyzed in SSE v1.2 software (102).

***In vitro* transcription and translation assay.** Coupled *in vitro* transcription and translation reactions were carried out using the Promega TnT Coupled Reticulate Lysate System labeled with radioactive [^{35}S] methionine (PerkinElmer Inc.) according to the manufacturer's protocol. Briefly, TnT reactions were set up as follows: 8 μL TnT mix, 1 μCi [^{35}S]methionine, and 200 ng plasmid DNA, made up to 10 μL with H_2O. Mixes were prepared on ice before reactions were incubated at 30°C for 90 min. The reactions were denatured in 8.5 μL of 2× Laemmli buffer (65.8 mM Tris-HCl [pH 6.8], 100 mM DTT [pH 6.8], 2.1% sodium dodecyl sulfate [SDS], 26.3% [wt/vol] glycerol, 0.01% bromophenol blue) and boiled for 10 min at 95°C. Samples were analyzed using SDS-PAGE and autoradiography.

Autoradiography of dried polyacrylamide gels. Gels were fixed in gel fixing solution (50% [vol/vol] methanol in water with 10% [vol/vol] acetic acid) on a rocker for 45 min with the gel fixing solution being replaced every 15 min. Fixed gels were transferred onto 3MM Whatman filter paper (Scientific Laboratory Supplies), covered with cling film and dried in a gel dryer (model 543; Bio-Rad) by heating up to 80°C for 2 h under vacuum. Dried gels were placed in a sealed cassette with an X-ray film (Fisher Scientific) overnight. X-ray films were developed using a Konica SRX-101A X-ograph film processor following manufacturer's protocol.

Reverse genetics system. Viruses were rescued using a combination of the protocols described by Kanai et al. (42) and Komoto et al. (43), with slight modifications. At 70% confluence, monolayers of BSR-T7 cells in 6-well plates were cotransfected with 11 plasmids corresponding to each RV genome segment (2.5 μg for plasmids encoding NSP2 and NSP5; 0.8 μg for the remaining plasmids) and plasmids encoding two vaccinia virus capping enzyme subunits (pCAG-D1R and pCAG-D12L: 0.8 μg each) using 16 μL Lipofectamine 2000 (Invitrogen) per transfection reaction in a total volume of 200 μL of Opti-MEM

TABLE 2 Names and nucleotide sequences of primers used in this study[a]

Target gene	Primer name	Sequence (5′ to 3′)	Use	Expected size of amplicon (bp)
SA11 VP4 in HR2 mutant plasmid	T7_FW	TAATACGACTCACTATAGGG	Sequence pT7-VP4SA11 plasmid containing SARS-CoV-2 HR2 peptide	2,517
	VP4_559_F	GCTAGGACAGCACATTATTC		
	HDV_REV	TCGTCCACTCGGATGGCTA		
HR2 of SARS-CoV-2 spike	FW	GATTCATTTAAATTTTTAGCGACCTCATTGAGGCGGT	Site-directed mutagenesis of HR2 peptide	n/a
	REV	CAATTTCTCCGGAAGTTGTTTTTACGACATCAATA AATTTC ACTCATTGACCTTCAAGAATTGGGAAAATATGAGCA ATATGGAGCTAGCGCAAATGGAAGTATAGGACA ATATG		
NTD of SARS-CoV-2 spike	FW	TAAATTCCTTTGTCAATCTCAAAGCTCTTTCCGGAAGTT	Site-directed mutagenesis of NTD peptide	n/a
	REV	GTTTTTACGACATCAATAAATTTC CCAGACCTCTAATTTCAGGGTTGTTGGAGCTAGCG CAAATGGAAGTATAGGACAATATG		
SA11 VP4	SA11_VP4_1F	GGCTATAAAATGGCTTCGCTC	Sequence pT7-VP4SA11 plasmids containing SARS-CoV-2 spike peptides; PCR and Sanger sequencing of SA11 VP4 gene in viruses	1,192–1,275
	SA11_VP4_1193R	GTCTACCAGTTGGACAATGGC		
SA11 VP4	SA11_VP4 FWD	CATCAGCATGGACGGATGTAT	RT-qPCR of SA11 VP4 gene	83
	SA11_VP4 REV	CCTCAGTTTGTGTTGCCATTTC		
SA11 VP1	SA11_VP1_448F	GCAGTCATGTTCTGGTTGGA	RT-qPCR of SA11 VP1 gene	89
	SA11_VP1_576R	GTTATGCCTTGGTACGCCAT		
RF NSP3	FW_331	CGCGTTGGCCGATTCATTAAT	Sequence RF NSP3 plasmids	1,769–2,186
	HDV_REV	TCGTCCACTCGGATGGCTA		
	SARS_CoV2_FW	GCTTTTCAGTGGTTGATGCT	PCR and Sanger sequencing of RF NSP3 gene in viruses	1,178–1,595
	SARS_CoV2_REV	TCATAGAGGGTCATGTGAAG		
	FW_NSP3_qPCR	GATGCTCAAGATGGAGTCTAC	RT-qPCR of RF NSP3 gene	134
	REV_NSP3_qPCR	CACCTGAATCATCCATTACGT		
RF VP1	FW_RF_VP1_qPCR	CACAATCCGCAGTTCAAATTC	RT-qPCR of RF VP1 gene	119
	REV_RF_VP1_qPCR	CAGTGTGGCATTCTCCATAAC		

[a]VP, viral protein; n/a, not applicable.

TABLE 3 SARS-CoV-2 spike amino acid sequences selected for insertion (61, 98, 99) into the hypervariable region of SA11 VP4 and tagging of RF NSP3[a]

SARS-CoV-2 peptide insertions	Amino acid sequence (N to C terminal)
SA11 VP4 construct	
HR2	EIDRLNEVAKNLNESLIDLQELGKYEQY
NTD	KSFEIDKGIYQTSNFRVV
RBM.1	SNNLDSKVGGNYNYLYRL
RBM.2	FRKSNLKPFERDISTEIYQAGS
RF NSP3 construct	
RBM	NSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPY
T2A_RBM	EGRGSLLTCGDVEENPGPNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPY
RBD	LCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKL PDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVL SFELLHAPATVCGP
T2A_RBD	EGRGSLLTCGDVEENPGPLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFT NVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPC NGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGP

[a]Sequence of the T2A peptide (66) is underlined.

(Gibco). After 24-h incubation at 37°C 5% CO_2, MA104 cells (1×10^5 cells/well) were added to transfected BSR-T7 cells and cocultured for 4 days in FBS-free DMEM supplemented with 0.5 μg/mL porcine pancreatic trypsin type IX. Cocultured cells were then lysed three times by freeze/thaw and lysates were incubated with trypsin at a final concentration of 10 μg/mL for 30 min at 37°C 5% CO_2 to activate the virus. Lysates were then transferred to fresh MA104 cells in T25 flasks and incubated at 37°C 5% CO_2 for 1 h. After adsorption, MA104 cells were washed and cultured in FBS-free DMEM supplemented with 0.5 μg/mL porcine pancreatic trypsin type IX for up to 7 days or until complete cytopathic effect was observed. Cells were then lysed three times by freeze/thaw, pelleted at 3,220 $\times$ *g* for 30 min and virus-containing supernatants (P1 stocks) were aliquoted and stored at −80°C. To generate mutant SA11 or RF mutant viruses, plasmids encoding either the SA11 VP4 or RF NSP3 gene segment were replaced with corresponding plasmids encoding SARS-CoV-2 spike epitopes. Mock preparations with the mutated segment omitted were generated for use as negative controls throughout. All rescue experiments were performed three times for each virus panel. NSP3 mutants were passaged in cells up to five times maintaining stable transgene expression (confirmed by Sanger sequencing, data not shown). The panels of viruses were titred by plaque assays, and the presence of mutations in the target gene segments was confirmed by RT-PCR and Sanger sequencing (GATC Biotech or Genewiz, Germany) as described below. Properties of mutant viruses are summarized in Table 1.

RT-PCR. RNA was purified from viral stocks using the spin protocol of the QIAamp Viral RNA minikit (Qiagen) followed by RQ1 RNase-Free DNase treatment (Promega) to remove possible plasmid contamination according to the manufacturer's specifications. cDNA was synthesized with SuperScript III Reverse Transcriptase (RT) kit (Invitrogen) using 5 μL RNA and 1 μL random hexamer primers (10 μM) (Invitrogen) according to the manufacturer's protocol. PCR was then used to amplify across the regions containing the SARS-CoV-2 spike epitopes in SA11 VP4 or RF NSP3 with primers listed in Table 2 (amplicon product sizes: ~1.2 to 1.6 kbp). Using the Q5 High-Fidelity DNA polymerase PCR Kit (New England BioLabs) for SA11 VP4 and Platinum *Taq* DNA polymerase (Invitrogen) for RF NSP3, PCRs were set up according to the manufacturer's instructions in a T100 Thermal Cycler (Bio-Rad). SA11 VP4 PCR conditions were 1 cycle of 98°C for 30 sec, followed by 35 cycles of 20 sec at 98°C, 20 sec at 52°C, and 2 min at 72°C, finishing with a 2 min incubation at 72°C. RF NSP3 PCR conditions were 1 cycle of 5 min at 95°C, followed by 35 cycles of 30 sec at 95°C, 30 sec at 55°C, and 2 min at 72°C, finishing with a 5-min incubation at 72°C. PCR product length was confirmed by 0.8% agarose gel electrophoresis containing SYBR Safe DNA Gel Stain and gels were imaged using the Odyssey XF imaging system (LI-COR). Analyses were performed with Image Studio Lite software (LI-COR). The presence of sequence insertions was confirmed by sequencing at GATC Biotech or Genewiz, Germany using primers listed in Table 2. Sequence results were analyzed in SSE v1.2 software (102).

Plaque assay. Plaque assays for RV infectivity were performed using adapted methods (103, 104). In brief, confluent monolayers of MA104 cells in 6-well plates were washed with FBS-free DMEM and infected with 800 μL of 10-fold serially diluted virus for 1 h at 37°C 5% CO_2. Following virus adsorption, 2 mL/well overlay medium was added (1:1 ratio of 2.4% Avicel [FMC Biopolymer] and FBS-free DMEM supplemented with 0.5 μg/mL trypsin type IX) and incubated for 4 days. Cells were then fixed for 1 h with 1 mL/well of 10% neutral buffered formalin (CellPath) and stained for 1 h with 0.1% Toluidine blue (Sigma-Aldrich) dissolved in H_2O.

Multistep virus growth kinetics. To compare growth kinetics of WT and mutant viruses, monolayers of MA104 cells at 70% confluence in 24-well plates were infected in technical triplicate with viruses at a multiplicity of infection (MOI) of 0.03 PFU/cell. A low MOI of 0.03 was selected to permit multiple cycles of infection. After 1 h, cells were washed three times with FBS-free DMEM and cultured with FBS-free DMEM supplemented with 0.5 μg/mL trypsin type IX. Viral supernatant and whole cell lysates were harvested at 1, 8, 16, 24, 32, and 48 h postinfection (hpi) and frozen at −80°C. Virus titers were

TABLE 4 Primary and secondary antibodies used in this study[a]

Antibody	Catalog no. (supplier)	Clonality/host species or conjugate	Dilution	Used in
Primary antibody				
RV VP6	Ab181695 (Abcam)	Monoclonal/mouse	1 in 1,000	Western blotting Immunofluorescence
Alpha-tubulin	NB600-506 (Novus Biologicals)	Monoclonal/rat	1 in 5,000	Western blotting
SARS-CoV-2 Spike Protein	DA123 (MRC protein phosphorylation and Ubiquitylation unit, Dundee)	Polyclonal/sheep	1 in 100	
SARS-CoV-2 Spike RBD	PA5-114451 (Invitrogen)	Polyclonal/rabbit	1 in 100	ELISA
			1 in 1,000	Immunofluorescence
			1 in 1,000	Western blotting
Secondary antibody				
Donkey anti-mouse	A32766 (Invitrogen)	Alexa Fluor 488	1 in 1,000	Immunofluorescence
Donkey anti-mouse	926-32212 (LI-COR)	Alexa Fluor 800	1 in 10,000	Western blotting
Goat anti-rat	926-68076 (LI-COR)	Alexa Fluor 680	1 in 10,000	
Rabbit anti-sheep	SA5-10060 (Invitrogen)	Alexa Fluor 800	1 in 10,000	

[a]RV, rotovirus.

determined by plaque assay. Whole cell lysates harvested in 350 μL buffer RLT (Qiagen) were used to determine the RNA concentration by quantitative RT-PCR (RT-qPCR) (see below) and lysates harvested in 100 μL 2× Laemmli buffer were used to analyze protein expression by Western blotting as described below.

Viral RNA extraction and quantification by RT-qPCR. Primers for RT-qPCR targeting segments VP1 (SA11 and RF), VP4 (SA11), and NSP3 (RF) were designed using SSE v1.2 and OligoCalc software (102, 105) (Table 2).

RNA was extracted from infected cell lysates using RNeasy minikit (Qiagen) with an on-column DNase treatment (Qiagen) to remove possible DNA contamination according to manufacturer's specification. The extracted RNA was dissolved in 20 μL nuclease-free water (Qiagen) and stored at −80°C.

To quantify total RNA levels of each segment relative to WT from time course experiments, one-step RT-qPCR was performed using SensiFAST SYBR Lo-ROX One-Step Kit (Meridian Bioscience) with 2 μL of RNA samples according to the manufacturer's protocol on a Rotor-Gene Q apparatus (Qiagen) using primers listed in Table 2. RT-qPCR cycling conditions were: 45°C for 10 min for reverse transcription, 2 min at 95°C, followed by 40 cycles of 10 sec at 95°C and 30 sec at 60°C. The conditions then increased from 50°C to 99°C at 1° increments to generate a melting curve to confirm specific amplification of each gene. For relative expression of each gene, the results were analyzed according to the $2^{-\Delta Ct}$ method (106). Average threshold cycle (C_T) values of each gene were normalized to WT at 8 hpi (when RNA was first reliably detected for VP1, VP4, and NSP3), and the resulting ΔC_T values were adjusted for primer efficiency. Three independent experiments were performed in technical duplicates.

To quantify the RNA copies of corresponding gene segments in mutant viruses for RNA:PFU ratios, RNA from cDNA (approximately 100 ng) of each gene segment was synthesized using a MEGAscript T7 transcription kit (Invitrogen) followed by TURBO DNase treatment (Invitrogen) and a cleanup with MinElute PCR purification kit (Qiagen) according to manufacturer's instructions. The extracted RNA was dissolved in 10 μL nuclease-free water (Qiagen) and stored at −80°C. RNA was measured using Qubit RNA broad range assay kit (Invitrogen) and serial dilutions of RNA were used as standards in one-step RT-qPCR. A ratio of the transcript copy number of each gene segment in the mutant viruses to virus titer was calculated and normalized to WT. Three independent experiments were performed in technical triplicates with the standard done in duplicate.

The PCR efficiency (E) of each primer pair set was established by measuring serial dilutions of cDNA of each segment in triplicate and calculated based on the slope of the standard curve according to the formula $E = [10^{(-1/\text{slope})} - 1] \times 100$ (107). Threshold cycle (C_T) values equivalent to mock samples and non-template control were considered to be negative.

Western blotting. Whole cell lysates harvested in 2× Laemmli buffer from time course infections were boiled at 95°C for 10 min. Proteins were separated by SDS-PAGE on 10% or 4 to 20% Mini-PROTEAN TGX Precast Protein Gels (Bio-Rad) according to the manufacturer's instructions. Proteins were transferred onto a 0.2 μm Cytiva Amersham Protran Nitrocellulose membrane (Fisher Scientific) and blocked with 2% horse serum in 1× Tris-buffered saline (20 mM Tris, 150 mM NaCl, pH 7.6) containing 0.01% Tween 20 detergent (Sigma-Aldrich) (TBS-T) for 1 h at room temperature. Membranes were then probed with primary antibodies diluted in 2% horse serum/TBS-T and incubated overnight at 4°C with agitation (Table 4). After three 5-min washes with TBS-T, membranes were incubated for 1 h at room temperature with secondary antibodies diluted in TBS-T (Table 4). Following three 5 min washes with TBS-T, membranes were imaged using an Odyssey XF imaging system (LI-COR). Analyses were performed with Image Studio Lite software (LI-COR).

Virus purification and electrophoretic analysis of dsRNA and protein. Virions were semipurified from SA11 and RF RV stocks using ultracentrifugation as described previously (103). Briefly, for RNA visualization, 50 mL of clarified RV stock was pelleted through a buffered 25% (wt/vol) sucrose cushion

(100 mM NaCl, 10 mM Tris-HCl [pH 7], 1 mM EDTA) using a SW32Ti rotor in a Beckman Coulter Optima Max-E ultracentrifuge at 106,750 × *g* for 90 min at 4°C. Viral pellets were resuspended in 350 μL buffer RLT containing 3.5 μL β-mercaptoethanol (Sigma-Aldrich) and dsRNA was extracted using an RNeasy minikit as described above. Purified RNA was treated with RNase A (Thermo Scientific) for 30 min at 37°C to specifically degrade single-stranded RNA and retain viral genomic dsRNA.

Extracted dsRNA was separated using 5% urea polyacrylamide gels in 1× Tris-borate-EDTA (TBE) buffer (89 mM Tris-borate, 2 mM EDTA [pH 8.3]). Gels were fixed in gel fixing solution (30% [vol/vol] methanol in 1× TBE buffer with 10% [vol/vol] acetic acid) and RNA was visualized using the Silver Stain Plus Kit (Bio-Rad) according to the manufacturer's instructions. Gels were dried onto filter paper and imaged using a Samsung Xpress C480FW scanner.

To visualize viral proteins of VP4 mutants, 25 mL of clarified RV stock was used for ultracentrifugation as described above. Viral pellets were resuspended in 1× TNC buffer (20 mM Tris-HCl [pH 8], 100 mM NaCl, 1 mM $CaCl_2$) and mixed 1:1 with 2× Laemmli buffer followed by SDS-PAGE using 4 to 20% Mini-PROTEAN TGX Precast Protein Gels and staining with Coomassie brilliant blue R-250 (Bio-Rad) according to manufacturer's instructions. Gels were imaged using the Samsung Xpress C480FW scanner. Quantification was performed by densitometry of scanned gel images using Image J. Values were corrected for background noise and normalized to those of the WT virus.

Indirect-binding ELISA. To determine the antigenicity of viruses expressing SARS-CoV-2 peptides, 200 μl of either viral supernatant (VP4 mutants) or cell lysate (NSP3 mutants) grown in the presence of trypsin (which cleaves VP4 into VP5* and VP8*, resulting in enhanced infectivity) (20) diluted in 0.1 M carbonate-bicarbonate buffer (pH 9.6) (Sigma-Aldrich) was immobilized on clear 96-well ELISA plates (Greiner Bio-One) and incubated overnight at 4°C with agitation. Following two washes with PBS containing 0.05% Tween 20 detergent (PBS-T) using a Biochrom Asys Atlantis microplate washer (Scientific Laboratory Supplies), plates were blocked with 2% horse serum/PBS at room temperature for 2 h. After blocking, plates were incubated with 100 μL/well primary antibody (Table 4) diluted in 2% horse serum/PBS overnight at 4°C with agitation. Plates were then washed six times with PBS-T and incubated with horseradish peroxidase-conjugated goat anti-rabbit IgG (H+L) (1:2,000) (Bio-Rad) diluted in 2% horse serum/PBS for 1 h at room temperature with agitation. After 1 h, plates were washed three times with PBS-T and incubated with 2,2′-azino-bis(3-ethylbenzothiazoline-6-sulfonic acid) (Scientific Laboratory Supplies) substrate to allow the color to develop after which the reaction was terminated by adding 70 μL per/well of 1% SDS solution. The optical density (OD) was measured at 405 nm using Cytation 3 Cell Imaging Multi-Mode Reader (Agilent) and data was analyzed using BioTek Gen5 software (Agilent). Three independent experiments were performed with standards done in triplicate. For a positive control, influenza A virus (IAV) strain A/Puerto Rico/8/1934 (PR8) was tagged with the same RBM sequence of the SARS-CoV-2 spike protein in the hemagglutinin (HA) protein (generated and validated in-house, unpublished data). The RBM sequence was inserted (by gene synthesis) between the C-terminal end of the HA signal peptide and the start of the head domain (between codons 17 and 18), placing the RBM sequence on the amino-terminus of the HA_0 molecule. This strategy was adapted from Li et al. (108), who successfully introduced large polypeptides (up to 140 amino acids) from *Bacillus anthracis* into IAV Aichi strain at this position. The synthetic constructs were cloned into a modified pHW2000 plasmid containing bidirectional RNA polymerase I and II promoters (109) and used for virus rescue as previously described (110, 111).

Infection of bovine enteroids with NSP3 mutants. 3D bovine enteroids were prepared as described in Hamilton et al. (112). Infection of organoids was carried out as described by Derricott et al. (70) and Blake et al. (113) with modifications. 3D organoids were mechanically disrupted into multicellular fragments by pipetting to expose the apical surface of the cells in 80% of the enteroids. The sheared enteroids were counted using a bright-field microscope and diluted to the final concentration of 2,500 enteroids/well. Each enteroid was estimated to contain approximately 40 cells as described by Blake et al. (113). The sheared enteroids were then diluted to the appropriate concentration with IntestiCult Organoid Growth Medium (STEMCELL Technologies) supplemented with 10 μM ROCK pathway inhibitor (Cayman chemicals), 10 μM Galunisertib (Cayman Chemicals), and 10 μM p38 inhibitor (Enzo) and aliquoted into 15 mL falcon tubes for each condition. The enteroid fragment suspension was infected with NSP3 mutants at an approximate MOI of 10 and incubated for 1 h at 37°C. After 1 h, the enteroid/virus suspension was centrifuged at 48 × *g* for 2 min and washed five times by replacing the supernatant with equal volumes of fresh IntestiCult. Enteroid pellets were then resuspended in the appropriate volume of IntestiCult and 500 μL plated onto Corning Matrigel (Scientific Laboratory Supplies) coated coverslips in a prewarmed 24-well plate.

Immunofluorescent staining of bovine enteroids. Enteroids were fixed at 1 and 24 hpi with 4% paraformaldehyde for 1.5 h at 4°C with agitation. Following three washes with PBS, enteroids were permeabilized with 0.5% Triton X-100 in PBS for 15 min and then blocked with 2% horse serum in PBS for 1 h, all at room temperature. Enteroids were incubated with primary antibodies (Table 4) overnight at 4°C with agitation and then washed three times with PBS and incubated with secondary antibodies (Table 4) and phalloidin (F-actin detection) (1:100) (Invitrogen) for 1 h at room temperature. All antibodies were diluted in 2% horse serum/PBS. Enteroids were then washed with PBS three times with the addition of 4′,6-diamidino-2-phenylindole (DAPI) nuclear stain (1:5,000) (Invitrogen) for the final 10 min wash. Coverslips were rinsed in water, mounted onto microscope slides (Thermo Scientific) with ProLong Gold Antifade Mountant (Invitrogen) and imaged using a Zeiss LSM 710 confocal microscope at ×630 magnification. Images were analyzed using the Zen Black software and processed using Photoshop v23.1.1.

Statistical analysis. GraphPad Prism v9 was used for all statistical analyses. Data are presented as mean and standard deviation from three independent experiments with technical duplicates unless

otherwise stated. *P* values were determined by ratio paired *t* test and were considered statistically significant at <0.05, unless data were normalized in which case paired *t* tests were used.

ACKNOWLEDGMENTS

We want to thank lab members, central support unit, bioimaging, and technical staff for their support and assistance with this project. We thank the MRC-Protein Phosphorylation and Ubiquitylation Unit, University of Dundee and Christine Tait-Burkard for providing SARS-CoV-2 spike antibodies and Marius Diebold for various software support.

This study was supported by an ISSF3 Award from the Wellcome Trust. J.S., P.D., and E.G. are supported by BBSRC Institute Strategic Program Grant funding (BB/P013740/1) from the British Biotechnology and Biological Sciences Research Council. O.D. is supported by a Roslin Studentship Award. R.B. is supported by the University of Edinburgh scholarship. S.C. is supported by a Wellcome Trust Clinical Research Career Development Fellowship (211138/Z/18/Z). A.B. and E.G. are Sir Henry Dale Fellows supported by the Wellcome Trust (213437/Z/18/Z and 211222_Z_18_Z respectively). Funding for open access charge was from the Wellcome Trust.

We declare no conflicts of interest.

REFERENCES

1. Tate JE, Burton AH, Boschi-Pinto C, Parashar UD; World Health Organization–Coordinated Global Rotavirus Surveillance Network. 2016. Global, regional, and national estimates of rotavirus mortality in children <5 years of age, 2000-2013. Clin Infect Dis 62(Suppl 2):S96–S105. https://doi.org/10.1093/cid/civ1013.
2. Troeger C, Khalil IA, Rao PC, Cao S, Blacker BF, Ahmed T, Armah G, Bines JE, Brewer TG, Colombara DV, Kang G, Kirkpatrick BD, Kirkwood CD, Mwenda JM, Parashar UD, Petri WA, Riddle MS, Steele AD, Thompson RL, Walson JL, Sanders JW, Mokdad AH, Murray CJL, Hay SI, Reiner RC. 2018. Rotavirus vaccination and the global burden of rotavirus diarrhea among children younger than 5 years. JAMA Pediatr 172:958–965. https://doi.org/10.1001/jamapediatrics.2018.1960.
3. Estes MK, G HB. 2013. *Rotaviruses*, p 1347–1401. *In* Knipe DM (ed), Fields virology. Wolters Kluwer Health/Lippincott Williams & Wilkins: Philadelphia, PA.
4. Papp H, László B, Jakab F, Ganesh B, De Grazia S, Matthijnssens J, Ciarlet M, Martella V, Bányai K. 2013. Review of group A rotavirus strains reported in swine and cattle. Vet Microbiol 165:190–199. https://doi.org/10.1016/j.vetmic.2013.03.020.
5. Urie NJ, Lombard JE, Shivley CB, Kopral CA, Adams AE, Earleywine TJ, Olson JD, Garry FB. 2018. Preweaned heifer management on US dairy operations: part V. factors associated with morbidity and mortality in preweaned dairy heifer calves. J Dairy Sci 101:9229–9244. https://doi.org/10.3168/jds.2017-14019.
6. Windeyer MC, Leslie KE, Godden SM, Hodgins DC, Lissemore KD, LeBlanc SJ. 2014. Factors associated with morbidity, mortality, and growth of dairy heifer calves up to 3 months of age. Prev Vet Med 113:231–240. https://doi.org/10.1016/j.prevetmed.2013.10.019.
7. Kim H-J, Park J, Matthijnssens J, Lee JH, Bae YC, Alfajaro MM, Park SI, Kang MI, Cho KO. 2011. Intestinal and extra-intestinal pathogenicity of a bovine reassortant rotavirus in calves and piglets. Vet Microbiol 152:291–303. https://doi.org/10.1016/j.vetmic.2011.05.017.
8. Bergman H, Henschke N, Hungerford D, Pitan F, Ndwandwe D, Cunliffe N, Soares-Weiser K. 2021. Vaccines for preventing rotavirus diarrhoea: vaccines in use. Cochrane Database Syst Rev 11:CD008521. https://doi.org/10.1002/14651858.CD008521.pub6.
9. Ciarlet M, Schödel F. 2009. Development of a rotavirus vaccine: clinical safety, immunogenicity, and efficacy of the pentavalent rotavirus vaccine, RotaTeq. Vaccine 27(Suppl 6):G72–G81. https://doi.org/10.1016/j.vaccine.2009.09.107.
10. Burnett E, Jonesteller CL, Tate JE, Yen C, Parashar UD. 2017. Global impact of rotavirus vaccination on childhood hospitalizations and mortality from diarrhea. J Infect Dis 215:1666–1672. https://doi.org/10.1093/infdis/jix186.
11. Meganck V, Hoflack G, Piepers S, Opsomer G. 2015. Evaluation of a protocol to reduce the incidence of neonatal calf diarrhoea on dairy herds. Prev Vet Med 118:64–70. https://doi.org/10.1016/j.prevetmed.2014.11.007.
12. Viidu D-A, Mõtus K. 2022. Implementation of a pre-calving vaccination programme against rotavirus, coronavirus and enterotoxigenic Escherichia coli (F5) and association with dairy calf survival. BMC Vet Res 18:59. https://doi.org/10.1186/s12917-022-03154-2.
13. Saif LJ, Fernandez FM. 1996. Group A rotavirus veterinary vaccines. J Infect Dis 174:S98–S106. https://doi.org/10.1093/infdis/174.Supplement_1.S98.
14. Chattha KS, Roth JA, Saif LJ. 2015. Strategies for design and application of enteric viral vaccines. Annu Rev Anim Biosci 3:375–395. https://doi.org/10.1146/annurev-animal-022114-111038.
15. Rainsford EW, McCrae MA. 2007. Characterization of the NSP6 protein product of rotavirus gene 11. Virus Res 130:193–201. https://doi.org/10.1016/j.virusres.2007.06.011.
16. Mattion NM, Mitchell DB, Both GW, Estes MK. 1991. Expression of rotavirus proteins encoded by alternative open reading frames of genome segment 11. Virology 181:295–304. https://doi.org/10.1016/0042-6822(91)90495-W.
17. Crawford SE, Ramani S, Tate JE, Parashar UD, Svensson L, Hagbom M, Franco MA, Greenberg HB, O'Ryan M, Kang G, Desselberger U, Estes MK. 2017. Rotavirus infection. Nat Rev Dis Primers 3:17083. https://doi.org/10.1038/nrdp.2017.83.
18. Amimo JO, Raev SA, Chepngeno J, Mainga AO, Guo Y, Saif L, Vlasova AN. 2021. Rotavirus interactions with host intestinal epithelial cells. Front Immunol 12:793841. https://doi.org/10.3389/fimmu.2021.793841.
19. Dormitzer PR, Nason EB, Prasad BVV, Harrison SC. 2004. Structural rearrangements in the membrane penetration protein of a non-enveloped virus. Nature 430:1053–1058. https://doi.org/10.1038/nature02836.
20. Crawford SE, Mukherjee SK, Estes MK, Lawton JA, Shaw AL, Ramig RF, Prasad BV. 2001. Trypsin cleavage stabilizes the rotavirus VP4 spike. J Virol 75:6052–6061. https://doi.org/10.1128/JVI.75.13.6052-6061.2001.
21. Arias CF, Romero P, Alvarez V, López S. 1996. Trypsin activation pathway of rotavirus infectivity. J Virol 70:5832–5839. https://doi.org/10.1128/jvi.70.9.5832-5839.1996.
22. Estes MK, Graham DY, Mason BB. 1981. Proteolytic enhancement of rotavirus infectivity: molecular mechanisms. J Virol 39:879–888. https://doi.org/10.1128/jvi.39.3.879-888.1981.
23. Hu L, Crawford SE, Czako R, Cortes-Penfield NW, Smith DF, Le Pendu J, Estes MK, Prasad BVV. 2012. Cell attachment protein VP8* of a human rotavirus specifically interacts with A-type histo-blood group antigen. Nature 485:256–259. https://doi.org/10.1038/nature10996.
24. Sun X, Li D, Duan Z. 2021. Structural basis of glycan recognition of rotavirus. Front Mol Biosci 8:658029. https://doi.org/10.3389/fmolb.2021.658029.
25. Fiore L, Greenberg HB, Mackow ER. 1991. The VP8 fragment of VP4 is the rhesus rotavirus hemagglutinin. Virology 181:553–563. https://doi.org/10.1016/0042-6822(91)90888-I.
26. Patton JT. 1996. Rotavirus VP1 alone specifically binds to the 3′ end of viral mRNA, but the interaction is not sufficient to initiate minus-strand synthesis. J Virol 70:7940–7947. https://doi.org/10.1128/jvi.70.11.7940-7947.1996.

27. Trask SD, McDonald SM, Patton JT. 2012. Structural insights into the coupling of virion assembly and rotavirus replication. Nat Rev Microbiol 10:165–177. https://doi.org/10.1038/nrmicro2673.
28. Lu X, McDonald SM, Tortorici MA, Tao YJ, Vasquez-Del Carpio R, Nibert ML, Patton JT, Harrison SC. 2008. Mechanism for coordinated RNA packaging and genome replication by rotavirus polymerase VP1. Structure 16:1678–1688. https://doi.org/10.1016/j.str.2008.09.006.
29. Borodavka A, Desselberger U, Patton JT. 2018. Genome packaging in multi-segmented dsRNA viruses: distinct mechanisms with similar outcomes. Curr Opin Virol 33:106–112. https://doi.org/10.1016/j.coviro.2018.08.001.
30. Ding K, Celma CC, Zhang X, Chang T, Shen W, Atanasov I, Roy P, Zhou ZH. 2019. In situ structures of rotavirus polymerase in action and mechanism of mRNA transcription and release. Nat Commun 10:2216. https://doi.org/10.1038/s41467-019-10236-7.
31. Chizhikov V, Patton JT. 2000. A four-nucleotide translation enhancer in the 3'-terminal consensus sequence of the nonpolyadenylated mRNAs of rotavirus. RNA 6:814–825. https://doi.org/10.1017/S1355838200992264.
32. Piron M, Delaunay T, Grosclaude J, Poncet D. 1999. Identification of the RNA-binding, dimerization, and eIF4GI-binding domains of rotavirus nonstructural protein NSP3. J Virol 73:5411–5421. https://doi.org/10.1128/JVI.73.7.5411-5421.1999.
33. Piron M, Vende P, Cohen J, Poncet D. 1998. Rotavirus RNA-binding protein NSP3 interacts with eIF4GI and evicts the poly(A) binding protein from eIF4F. EMBO J 17:5811–5821. https://doi.org/10.1093/emboj/17.19.5811.
34. Harb M, Becker MM, Vitour D, Baron CH, Vende P, Brown SC, Bolte S, Arold ST, Poncet D. 2008. Nuclear localization of cytoplasmic poly(A)-binding protein upon rotavirus infection involves the interaction of NSP3 with eIF4G and RoXaN. J Virol 82:11283–11293. https://doi.org/10.1128/JVI.00872-08.
35. Rubio RM, Mora SI, Romero P, Arias CF, López S. 2013. Rotavirus prevents the expression of host responses by blocking the nucleocytoplasmic transport of polyadenylated mRNAs. J Virol 87:6336–6345. https://doi.org/10.1128/JVI.00361-13.
36. Padilla-Noriega L, Paniagua O, Guzmán-León S. 2002. Rotavirus protein NSP3 shuts off host cell protein synthesis. Virology 298:1–7. https://doi.org/10.1006/viro.2002.1477.
37. Montero H, Arias CF, Lopez S. 2006. Rotavirus nonstructural protein NSP3 is not required for viral protein synthesis. J Virol 80:9031–9038. https://doi.org/10.1128/JVI.00437-06.
38. Komoto S, Sasaki J, Taniguchi K. 2006. Reverse genetics system for introduction of site-specific mutations into the double-stranded RNA genome of infectious rotavirus. Proc Natl Acad Sci U S A 103:4646–4651. https://doi.org/10.1073/pnas.0509385103.
39. Trask SD, Taraporewala ZF, Boehme KW, Dermody TS, Patton JT. 2010. Dual selection mechanisms drive efficient single-gene reverse genetics for rotavirus. Proc Natl Acad Sci U S A 107:18652–18657. https://doi.org/10.1073/pnas.1011948107.
40. Troupin C, Dehée A, Schnuriger A, Vende P, Poncet D, Garbarg-Chenon A. 2010. Rearranged genomic RNA segments offer a new approach to the reverse genetics of rotaviruses. J Virol 84:6711–6719. https://doi.org/10.1128/JVI.00547-10.
41. Johne R, Reetz J, Kaufer BB, Trojnar E. 2016. Generation of an avian-mammalian rotavirus reassortant by using a helper virus-dependent reverse genetics system. J Virol 90:1439–1443. https://doi.org/10.1128/JVI.02730-15.
42. Kanai Y, Komoto S, Kawagishi T, Nouda R, Nagasawa N, Onishi M, Matsuura Y, Taniguchi K, Kobayashi T. 2017. Entirely plasmid-based reverse genetics system for rotaviruses. Proc Natl Acad Sci U S A 114:2349–2354. https://doi.org/10.1073/pnas.1618424114.
43. Komoto S, Fukuda S, Ide T, Ito N, Sugiyama M, Yoshikawa T, Murata T, Taniguchi K. 2018. Generation of recombinant rotaviruses expressing fluorescent proteins by using an optimized reverse genetics system. J Virol 92:e00588-18. https://doi.org/10.1128/JVI.00588-18.
44. Sánchez-Tacuba L, Feng N, Meade NJ, Mellits KH, Jaïs PH, Yasukawa LL, Resch TK, Jiang B, López S, Ding S, Greenberg HB. 2020. An optimized reverse genetics system suitable for efficient recovery of simian, human, and murine-like rotaviruses. J Virol 94:e01294-20. https://doi.org/10.1128/JVI.01294-20.
45. Philip AA, Perry JL, Eaton HE, Shmulevitz M, Hyser JM, Patton JT. 2019. Generation of recombinant rotavirus expressing NSP3-UnaG fusion protein by a simplified reverse genetics system. J Virol 93:e01616-19. https://doi.org/10.1128/JVI.01616-19.
46. Kanai Y, Onishi M, Kawagishi T, Pannacha P, Nurdin JA, Nouda R, Yamasaki M, Lusiany T, Khamrin P, Okitsu S, Hayakawa S, Ebina H, Ushijima H, Kobayashi T. 2020. Reverse genetics approach for developing rotavirus vaccine candidates carrying VP4 and VP7 genes cloned from clinical isolates of human rotavirus. J Virol 95:e01374-20. https://doi.org/10.1128/JVI.01374-20.
47. Falkenhagen A, Huyzers M, van Dijk AA, Johne R. 2021. Rescue of infectious rotavirus reassortants by a reverse genetics system is restricted by the receptor-binding region of VP4. Viruses 13:363. https://doi.org/10.3390/v13030363.
48. Uprety T, Wang D, Li F. 2021. Recent advances in rotavirus reverse genetics and its utilization in basic research and vaccine development. Arch Virol 166:2369–2386. https://doi.org/10.1007/s00705-021-05142-7.
49. Nair N, Feng N, Blum LK, Sanyal M, Ding S, Jiang B, Sen A, Morton JM, He X-S, Robinson WH, Greenberg HB. 2017. VP4- and VP7-specific antibodies mediate heterotypic immunity to rotavirus in humans. Sci Transl Med 9:eaam5434. https://doi.org/10.1126/scitranslmed.aam5434.
50. Feng N, Hu L, Ding S, Sanyal M, Zhao B, Sankaran B, Ramani S, McNeal M, Yasukawa LL, Song Y, Prasad BVV, Greenberg HB. 2019. Human VP8* mAbs neutralize rotavirus selectively in human intestinal epithelial cells. J Clin Invest 129:3839–3851. https://doi.org/10.1172/JCI128382.
51. Philip AA, Patton JT. 2020. Expression of separate heterologous proteins from the rotavirus nsp3 genome segment using a translational 2A stop-restart element. J Virol 94:e00959-20. https://doi.org/10.1128/JVI.00959-20.
52. Philip AA, Patton JT. 2021. Rotavirus as an expression platform of domains of the SARS-CoV-2 spike protein. Vaccines (Basel) 9:449. https://doi.org/10.3390/vaccines9050449.
53. Krammer F. 2020. SARS-CoV-2 vaccines in development. Nature 586:516–527. https://doi.org/10.1038/s41586-020-2798-3.
54. Kyriakidis NC, López-Cortés A, González EV, Grimaldos AB, Prado EO. 2021. SARS-CoV-2 vaccines strategies: a comprehensive review of phase 3 candidates. NPJ Vaccines 6:28. https://doi.org/10.1038/s41541-021-00292-w.
55. Hassan AO, Kafai NM, Dmitriev IP, Fox JM, Smith BK, Harvey IB, Chen RE, Winkler ES, Wessel AW, Case JB, Kashentseva E, McCune BT, Bailey AL, Zhao H, VanBlargan LA, Dai Y-N, Ma M, Adams LJ, Shrihari S, Danis JE, Gralinski LE, Hou YJ, Schäfer A, Kim AS, Keeler SP, Weiskopf D, Baric RS, Holtzman MJ, Fremont DH, Curiel DT, Diamond MS. 2020. A single-dose intranasal ChAd vaccine protects upper and lower respiratory tracts against SARS-CoV-2. Cell 183:169–184.e13. https://doi.org/10.1016/j.cell.2020.08.026.
56. Bos R, Rutten L, van der Lubbe JEM, Bakkers MJG, Hardenberg G, Wegmann F, Zuijdgeest D, de Wilde AH, Koornneef A, Verwilligen A, van Manen D, Kwaks T, Vogels R, Dalebout TJ, Myeni SK, Kikkert M, Snijder EJ, Li Z, Barouch DH, Vellinga J, Langedijk JPM, Zahn RC, Custers J, Schuitemaker H. 2020. Ad26 vector-based COVID-19 vaccine encoding a prefusion-stabilized SARS-CoV-2 Spike immunogen induces potent humoral and cellular immune responses. NPJ Vaccines 5:91–91. https://doi.org/10.1038/s41541-020-00243-x.
57. Sahin U, Muik A, Derhovanessian E, Vogler I, Kranz LM, Vormehr M, Baum A, Pascal K, Quandt J, Maurus D, Brachtendorf S, Lörks V, Sikorski J, Hilker R, Becker D, Eller A-K, Grützner J, Boesler C, Rosenbaum C, Kühnle M-C, Luxemburger U, Kemmer-Brück A, Langer D, Bexon M, Bolte S, Karikó K, Palanche T, Fischer B, Schultz A, Shi P-Y, Fontes-Garfias C, Perez JL, Swanson KA, Loschko J, Scully IL, Cutler M, Kalina W, Kyratsous CA, Cooper D, Dormitzer PR, Jansen KU, Türeci Ö. 2020. COVID-19 vaccine BNT162b1 elicits human antibody and TH1 T cell responses. Nature 586:594–599. https://doi.org/10.1038/s41586-020-2814-7.
58. Wang Z, Schmidt F, Weisblum Y, Muecksch F, Barnes CO, Finkin S, Schaefer-Babajew D, Cipolla M, Gaebler C, Lieberman JA, Oliveira TY, Yang Z, Abernathy ME, Huey-Tubman KE, Hurley A, Turroja M, West KA, Gordon K, Millard KG, Ramos V, Da Silva J, Xu J, Colbert RA, Patel R, Dizon J, Unson-O'Brien C, Shimeliovich I, Gazumyan A, Caskey M, Bjorkman PJ, Casellas R, Hatziioannou T, Bieniasz PD, Nussenzweig MC. 2021. mRNA vaccine-elicited antibodies to SARS-CoV-2 and circulating variants. Nature 592:616–622. https://doi.org/10.1038/s41586-021-03324-6.
59. Yang J, Wang W, Chen Z, Lu S, Yang F, Bi Z, Bao L, Mo F, Li X, Huang Y, Hong W, Yang Y, Zhao Y, Ye F, Lin S, Deng W, Chen H, Lei H, Zhang Z, Luo M, Gao H, Zheng Y, Gong Y, Jiang X, Xu Y, Lv Q, Li D, Wang M, Li F, Wang S, Wang G, Yu P, Qu Y, Yang L, Deng H, Tong A, Li J, Wang Z, Yang J, Shen G, Zhao Z, Li Y, Luo J, Liu H, Yu W, Yang M, Xu J, Wang J, Li H, Wang H, et al. 2020. A vaccine targeting the RBD of the S protein of SARS-CoV-2 induces protective immunity. Nature 586:572–577. https://doi.org/10.1038/s41586-020-2599-8.

60. Liu Z, Xu W, Xia S, Gu C, Wang X, Wang Q, Zhou J, Wu Y, Cai X, Qu D, Ying T, Xie Y, Lu L, Yuan Z, Jiang S. 2020. RBD-Fc-based COVID-19 vaccine candidate induces highly potent SARS-CoV-2 neutralizing antibody response. Signal Transduct Target Ther 5:282. https://doi.org/10.1038/s41392-020-00402-5.
61. Piccoli L, Park Y-J, Tortorici MA, Czudnochowski N, Walls AC, Beltramello M, Silacci-Fregni C, Pinto D, Rosen LE, Bowen JE, Acton OJ, Jaconi S, Guarino B, Minola A, Zatta F, Sprugasci N, Bassi J, Peter A, De Marco A, Nix JC, Mele F, Jovic S, Rodriguez BF, Gupta SV, Jin F, Piumatti G, Lo Presti G, Pellanda AF, Biggiogero M, Tarkowski M, Pizzuto MS, Cameroni E, Havenar-Daughton C, Smithey M, Hong D, Lepori V, Albanese E, Ceschi A, Bernasconi E, Elzi L, Ferrari P, Garzoni C, Riva A, Snell G, Sallusto F, Fink K, Virgin HW, Lanzavecchia A, Corti D, Veesler D. 2020. Mapping neutralizing and immunodominant sites on the SARS-CoV-2 spike receptor-binding domain by structure-guided high-resolution serology. Cell 183:1024–1042.e21. https://doi.org/10.1016/j.cell.2020.09.037.
62. Lan J, Ge J, Yu J, Shan S, Zhou H, Fan S, Zhang Q, Shi X, Wang Q, Zhang L, Wang X. 2020. Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. Nature 581:215–220. https://doi.org/10.1038/s41586-020-2180-5.
63. Settembre EC, Chen JZ, Dormitzer PR, Grigorieff N, Harrison SC. 2011. Atomic model of an infectious rotavirus particle. EMBO J 30:408–416. https://doi.org/10.1038/emboj.2010.322.
64. Herrmann T, Torres R, Salgado EN, Berciu C, Stoddard D, Nicastro D, Jenni S, Harrison SC. 2021. Functional refolding of the penetration protein on a non-enveloped virus. Nature 590:666–670. https://doi.org/10.1038/s41586-020-03124-4.
65. Rodríguez JM, Chichón FJ, Martín-Forero E, González-Camacho F, Carrascosa JL, Castón JR, Luque D. 2014. New insights into rotavirus entry machinery: stabilization of rotavirus spike conformation is independent of trypsin cleavage. PLoS Pathog 10:e1004157. https://doi.org/10.1371/journal.ppat.1004157.
66. Liu Z, Chen O, Wall JBJ, Zheng M, Zhou Y, Wang L, Vaseghi HR, Qian L, Liu J. 2017. Systematic comparison of 2A peptides for cloning multi-genes in a polycistronic vector. Sci Rep 7:2193. https://doi.org/10.1038/s41598-017-02460-2.
67. Borodavka A, Dykeman EC, Schrimpf W, Lamb DC. 2017. Protein-mediated RNA folding governs sequence-specific interactions between rotavirus genome segments. Elife 6:e27453. https://doi.org/10.7554/eLife.27453.
68. McDonald SM, Patton JT. 2011. Assortment and packaging of the segmented rotavirus genome. Trends Microbiol 19:136–144. https://doi.org/10.1016/j.tim.2010.12.002.
69. Bravo JPK, Bartnik K, Venditti L, Acker J, Gail EH, Colyer A, Davidovich C, Lamb DC, Tuma R, Calabrese AN, Borodavka A. 2021. Structural basis of rotavirus RNA chaperone displacement and RNA annealing. Proc Natl Acad Sci U S A 118:e2100198118. https://doi.org/10.1073/pnas.2100198118.
70. Derricott H, Luu L, Fong WY, Hartley CS, Johnston LJ, Armstrong SD, Randle N, Duckworth CA, Campbell BJ, Wastling JM, Coombes JL. 2019. Developing a 3D intestinal epithelium model for livestock species. Cell Tissue Res 375:409–424. https://doi.org/10.1007/s00441-018-2924-9.
71. L'Haridon R, Scherrer R. 1976. [In vitro culture of a rotavirus associated with neonatal calf diarrhea] Ann Rech Vet 7:373–381.
72. Papa G, Borodavka A, Desselberger U. 2021. Viroplasms: assembly and functions of rotavirus replication factories. Viruses 13:1349. https://doi.org/10.3390/v13071349.
73. Wen X, Cao D, Jones RW, Li J, Szu S, Hoshino Y. 2012. Construction and characterization of human rotavirus recombinant VP8* subunit parenteral vaccine candidates. Vaccine 30:6121–6126. https://doi.org/10.1016/j.vaccine.2012.07.078.
74. Park W-J, Yoon Y-K, Park J-S, Pansuriya R, Seok Y-J, Ganapathy R. 2021. Rotavirus spike protein ΔVP8* as a novel carrier protein for conjugate vaccine platform with demonstrated antigenic potential for use as bivalent vaccine. Sci Rep 11:22037. https://doi.org/10.1038/s41598-021-01549-z.
75. Xia M, Huang P, Jiang X, Tan M. 2021. A nanoparticle-based trivalent vaccine targeting the glycan binding VP8* domains of rotaviruses. Viruses 1:72. https://doi.org/10.3390/v13010072.
76. Tan M, Huang P, Xia M, Fang P-A, Zhong W, McNeal M, Wei C, Jiang W, Jiang X. 2011. Norovirus P particle, a novel platform for vaccine development and antibody production. J Virol 85:753–764. https://doi.org/10.1128/JVI.01835-10.
77. Wen X, Cao D, Jones RW, Hoshino Y, Yuan L. 2015. Tandem truncated rotavirus VP8* subunit protein with T cell epitope as non-replicating parenteral vaccine is highly immunogenic. Hum Vaccin Immunother 11:2483–2489. https://doi.org/10.1080/21645515.2015.1054583.
78. Li Y, Xue M, Yu L, Luo G, Yang H, Jia L, Zeng Y, Li T, Ge S, Xia N. 2018. Expression and characterization of a novel truncated rotavirus VP4 for the development of a recombinant rotavirus vaccine. Vaccine 36:2086–2092. https://doi.org/10.1016/j.vaccine.2018.03.011.
79. Komoto S, Kugita M, Sasaki J, Taniguchi K. 2008. Generation of recombinant rotavirus with an antigenic mosaic of cross-reactive neutralization epitopes on VP4. J Virol 82:6753–6757. https://doi.org/10.1128/JVI.00601-08.
80. Harrison SC. 2008. Viral membrane fusion. Nat Struct Mol Biol 15:690–698. https://doi.org/10.1038/nsmb.1456.
81. Trask SD, Kim IS, Harrison SC, Dormitzer PR. 2010. A rotavirus spike protein conformational intermediate binds lipid bilayers. J Virol 84:1764–1770. https://doi.org/10.1128/JVI.01682-09.
82. Dormitzer PR, Greenberg HB, Harrison SC. 2001. Proteolysis of monomeric recombinant rotavirus VP4 yields an oligomeric VP5* core. J Virol 75:7339–7350. https://doi.org/10.1128/JVI.75.16.7339-7350.2001.
83. Kim IS, Trask SD, Babyonyshev M, Dormitzer PR, Harrison SC. 2010. Effect of mutations in VP5 hydrophobic loops on rotavirus cell entry. J Virol 84:6200–6207. https://doi.org/10.1128/JVI.02461-09.
84. Yoder JD, Trask SD, Vo TP, Binka M, Feng N, Harrison SC, Greenberg HB, Dormitzer PR. 2009. VP5* rearranges when rotavirus uncoats. J Virol 83:11372–11377. https://doi.org/10.1128/JVI.01228-09.
85. Trask SD, Dormitzer PR. 2006. Assembly of highly infectious rotavirus particles recoated with recombinant outer capsid proteins. J Virol 80:11293–11304. https://doi.org/10.1128/JVI.01346-06.
86. Caddy SL, Vaysburd M, Wing M, Foss S, Andersen JT, O'Connell K, Mayes K, Higginson K, Iturriza-Gómara M, Desselberger U, James LC. 2020. Intracellular neutralisation of rotavirus by VP6-specific IgG. PLoS Pathog 16:e1008732. https://doi.org/10.1371/journal.ppat.1008732.
87. Vesikari T, Matson DO, Dennehy P, Van Damme P, Santosham M, Rodriguez Z, Dallas MJ, Heyse JF, Goveia MG, Black SB, Shinefield HR, Christie CDC, Ylitalo S, Itzler RF, Coia ML, Onorato MT, Adeyi BA, Marshall GS, Gothefors L, Campens D, Karvonen A, Watt JP, O'Brien KL, DiNubile MJ, Clark HF, Boslego JW, Offit PA, Heaton PM, Rotavirus Efficacy and Safety Trial (REST) Study Team. 2006. Safety and efficacy of a pentavalent human-bovine (WC3) reassortant rotavirus vaccine. N Engl J Med 354:23–33. https://doi.org/10.1056/NEJMoa052664.
88. Angel J, Franco MA, Greenberg HB. 2007. Rotavirus vaccines: recent developments and future considerations. Nat Rev Microbiol 5:529–539. https://doi.org/10.1038/nrmicro1692.
89. Bernstein DI, Smith VE, Sherwood JR, Schiff GM, Sander DS, DeFeudis D, Spriggs DR, Ward RL. 1998. Safety and immunogenicity of live, attenuated human rotavirus vaccine 89-12. Vaccine 16:381–387. https://doi.org/10.1016/S0264-410X(97)00210-7.
90. Bernstein DI, Sack DA, Rothstein E, Reisinger K, Smith VE, O'Sullivan D, Spriggs DR, Ward RL. 1999. Efficacy of live, attenuated, human rotavirus vaccine 89–12 in infants: a randomised placebo-controlled trial. Lancet 354:287–290. https://doi.org/10.1016/S0140-6736(98)12106-2.
91. Gratia M, Vende P, Charpilienne A, Baron HC, Laroche C, Sarot E, Pyronnet S, Duarte M, Poncet D. 2016. Challenging the roles of NSP3 and untranslated regions in rotavirus mRNA translation. PLoS One 11:e0145998. https://doi.org/10.1371/journal.pone.0145998.
92. Jang YH, Seong BL. 2012. Principles underlying rational design of live attenuated influenza vaccines. Clin Exp Vaccine Res 1:35–49. https://doi.org/10.7774/cevr.2012.1.1.35.
93. Desselberger U. 2020. What are the limits of the packaging capacity for genomic RNA in the cores of rotaviruses and of other members of the Reoviridae? Virus Res 276:197822. https://doi.org/10.1016/j.virusres.2019.197822.
94. Buchholz UJ, Finke S, Conzelmann K-K. 1999. Generation of bovine respiratory syncytial virus (BRSV) from cDNA: BRSV NS2 is not essential for virus replication in tissue culture, and the human RSV leader region acts as a functional BRSV genome promoter. J Virol 73:251–259. https://doi.org/10.1128/JVI.73.1.251-259.1999.
95. Taniguchi K, Nishikawa K, Kobayashi N, Urasawa T, Wu H, Gorziglia M, Urasawa S. 1994. Differences in plaque size and VP4 sequence found in SA11 virus clones having simian authentic VP4. Virology 198:325–330. https://doi.org/10.1006/viro.1994.1035.
96. Richards JE, Desselberger U, Lever AM. 2013. Experimental pathways towards developing a rotavirus reverse genetics system: synthetic full length rotavirus ssRNAs are neither infectious nor translated in permissive cells. PLoS One 8:e74328. https://doi.org/10.1371/journal.pone.0074328.
97. Dormitzer PR, Sun Z-YJ, Wagner G, Harrison SC. 2002. The rhesus rotavirus VP4 sialic acid binding domain has a galectin fold with a novel carbohydrate binding site. EMBO J 21:885–897. https://doi.org/10.1093/emboj/21.5.885.

98. Yi Z, Ling Y, Zhang X, Chen J, Hu K, Wang Y, Song W, Ying T, Zhang R, Lu H, Yuan Z. 2020. Functional mapping of B-cell linear epitopes of SARS-CoV-2 in COVID-19 convalescent population. Emerg Microbes Infect 9:1988–1996. https://doi.org/10.1080/22221751.2020.1815591.
99. Polyiam K, Phoolcharoen W, Butkhot N, Srisaowakarn C, Thitithanyanont A, Auewarakul P, Hoonsuwan T, Ruengjitchatchawalya M, Mekvichitsaeng P, Roshorm YM. 2021. Immunodominant linear B cell epitopes in the spike and membrane proteins of SARS-CoV-2 identified by immunoinformatics prediction and immunoassay. Sci Rep 11:20383. https://doi.org/10.1038/s41598-021-99642-w.
100. Daniels RW, Rossano AJ, Macleod GT, Ganetzky B. 2014. Expression of multiple transgenes from a single construct using viral 2A peptides in Drosophila. PLoS One 9:e100637. https://doi.org/10.1371/journal.pone.0100637.
101. Jeong YS, Ku H-K, Jung Y-J, Kim JK, Lee KB, Kim J-K, Lim S-H, Lee D, Ha S-H. 2021. 2A-linked bi-, tri-, and quad-cistrons for the stepwise biosynthesis of β-carotene, zeaxanthin, and ketocarotenoids in rice endosperm. Metab Eng Commun 12:e00166. https://doi.org/10.1016/j.mec.2021.e00166.
102. Simmonds P. 2012. SSE: a nucleotide and amino acid sequence analysis platform. BMC Res Notes 5:50. https://doi.org/10.1186/1756-0500-5-50.
103. Arnold M, Patton JT, McDonald SM. 2009. Culturing, storage, and quantification of rotaviruses. Curr Protoc Microbiol 15:Unit-15C.3.https://doi.org/10.1002/9780471729259.mc15c03s15.
104. Matrosovich M, Matrosovich T, Garten W, Klenk H-D. 2006. New low-viscosity overlay medium for viral plaque assays. Virol J 3:63. https://doi.org/10.1186/1743-422X-3-63.
105. Kibbe WA. 2007. OligoCalc: an online oligonucleotide properties calculator. Nucleic Acids Res 35:W43–W46. https://doi.org/10.1093/nar/gkm234.
106. Schmittgen TD, Livak KJ. 2008. Analyzing real-time PCR data by the comparative CT method. Nat Protoc 3:1101–1108. https://doi.org/10.1038/nprot.2008.73.
107. Mygind T, Birkelund S, Birkebaek NH, Østergaard L, Jensen JS, Christiansen G. 2002. Determination of PCR efficiency in chelex-100 purified clinical samples and comparison of real-time quantitative PCR and conventional PCR for detection of Chlamydia pneumoniae. BMC Microbiol 2:17. https://doi.org/10.1186/1471-2180-2-17.
108. Li Z-N, Mueller SN, Ye L, Bu Z, Yang C, Ahmed R, Steinhauer DA. 2005. Chimeric influenza virus hemagglutinin proteins containing large domains of the Bacillus anthracis protective antigen: protein characterization, incorporation into infectious influenza viruses, and antigenicity. J Virol 79:10003–10012. https://doi.org/10.1128/JVI.79.15.10003-10012.2005.
109. de Wit E, Spronken MIJ, Bestebroer TM, Rimmelzwaan GF, Osterhaus ADME, Fouchier RAM. 2004. Efficient generation and growth of influenza virus A/PR/8/34 from eight cDNA fragments. Virus Res 103:155–161. https://doi.org/10.1016/j.virusres.2004.02.028.
110. Hutchinson EC, Curran MD, Read EK, Gog JR, Digard P. 2008. Mutational analysis of cis-acting RNA signals in segment 7 of influenza A virus. J Virol 82:11869–11879. https://doi.org/10.1128/JVI.01634-08.
111. Wise HM, Foeglein A, Sun J, Dalton RM, Patel S, Howard W, Anderson EC, Barclay WS, Digard P. 2009. A complicated message: Identification of a novel PB1-related protein translated from influenza A virus segment 2 mRNA. J Virol 83:8021–8031. https://doi.org/10.1128/JVI.00826-09.
112. Hamilton CA, Young R, Jayaraman S, Sehgal A, Paxton E, Thomson S, Katzer F, Hope J, Innes E, Morrison LJ, Mabbott NA. 2018. Development of in vitro enteroids derived from bovine small intestinal crypts. Vet Res 49:54. https://doi.org/10.1186/s13567-018-0547-5.
113. Blake R, Jensen K, Mabbott N, Hope J, Stevens J. 2022. Development of 3D bovine intestinal organoid derived models to investigate Mycobacterium avium ssp paratuberculosis pathogenesis. bioRxiv. https://www.biorxiv.org/content/10.1101/2022.05.13.491821v1.

4.5 Conclusion

In this study, two RV strains, SA11 and RF, were used as viral vaccine vectors to express different SARS-CoV-2 spike peptides. Recombinant viruses were generated using the established reverse genetics system for the bovine strain RF (*see Chapter 3*). The main findings from this study were: i) peptide expression by SA11 VP4 resulted in reduced viral titres and RF NSP3 demonstrated a limit to the length of additional sequences it can accommodate and ii) SA11 VP4 did not show any cross-reactivity with SARS-CoV-2 antibodies in ELISA whereas cross-reactivity was observed with RF NSP3, highlighting the potential of this system for the expression of foreign peptides.

There were some limitations to this study. Firstly, the small SARS-CoV-2 spike peptides expressed by SA11 VP4 did not include other potential neutralising antibody targets present on the full-length spike. On the other hand, tagging of RF NSP3 with large SARS-CoV-2 spike peptides affected the viral rescue even in the presence of the helper plasmids expressing VV capping enzyme. To further optimise the reverse genetics efficiency, introducing the ASFV capping enzyme, NP868R, could improve the viral protein expression and enhance the viral rescue as demonstrated by other studies [290, 291, 299, 396]. Sánchez-Tacuba *et al*., (2020) also showed that using a genetically modified MA104 cell line in which some components of the innate immunity were inhibited, enhanced the replication of different RV strains [292]. Future experiments could include a more efficient capping enzyme as well as a genetically modified MA104 cell line for virus propagation to facilitate expression of larger heterologous peptides with less impact on the viral rescue efficiency. Furthermore, RV constructs containing heterologous inserts above a certain size were previously shown to be genetically unstable upon five serial passages in cell culture [299]. In this study, although RF NSP3 was genetically stable up to three passages in cell culture, there was a limit to the length of foreign sequence it could accommodate. Thus, the size of the foreign peptides should be carefully considered in the future.

While the bovine enteroids are more physiologically representative of the intestinal epithelium than the commonly used cell-lines, this model still remains a reductionist approach for the study of host-pathogen interactions. Unlike the avian apical-out enteroids that contain the components of the immune system [397], the 3D basal-out bovine enteroids used in this study lacked immune cells. Previous studies showed that infection of human intestinal organoids with RV and SARS-CoV-2 predominantly induced a type I and III IFN responses [398, 399]. Although, the presence of antiviral cytokines, such as IL-6, IL-8, TNF-α, TGF and IFN-γ, has been characterised [400], the expression of the immune cell repertoire would need

further investigation if the immune response (innate or adaptive) to SARS-CoV-2 peptides was to be assessed in this system. Several protocols have been established to co-culture organoids with microbiome [401] and different types of immune cells, such as T lymphocytes [402] and macrophages [403], as well as to replicate the peristaltic-like motility *in vitro [404]*, all of which better represent the gut microenvironment. Therefore, future studies could investigate the presence of enteric immune cells in the bovine enteroids to elucidate the effect they may have on the outcome of RV and SARS-CoV-2 infections. Notably, the bovine enteroids could recapitulate *in vivo* observations of vaccine attenuation as seen by attenuated replication of the Rotarix vaccine strain in human intestinal organoids that was not observed in MA104 cells, suggesting a useful system for testing the attenuation of other candidate RV vaccines [89]. Additionally, monoclonal antibodies targeting the VP8* domain of RV VP4, that were previously found to be non-neutralising in the traditional MA104-based assays efficiently neutralised human RV infection in human intestinal organoids [42]. Thus, this system could also be utilised for detection of anti-RV neutralisation responses that may have been missed in the conventional MA104 cell lines.

Despite the above limitations, this study's findings present many advantages and potential downstream applications of RV as an expression platform for next-generation vaccines against SARS-CoV-2. Table 13 summarises several characteristics of selected currently licenced SARS-CoV-2 vaccines compared with the most commonly used RV vaccine, RotaTeq.

Although current licenced SARS-CoV-2 vaccines offer effective protection against severe disease and mortality, several new variants of concern with acquired mutations within the RBD have led to increased transmissibility and immune escape [405]. Therefore, the duration of protective immunity from first-generation vaccines against the new emerging SARS-CoV-2 variants of concern could be diminished [406]. Different strategies have been employed to increase the efficacy of current vaccines such as vaccination with a booster dose and developing variant-specific or multivalent vaccines [407-409]. Thus, the RV vector could be utilised to express a combination of RBD epitopes from multiple strains to induce a broad SARS-CoV-2 immune response against the emerging variants of concern. However, the high mutation rate of the SARS-CoV-2 genome would require an update of the vaccine analogous to the seasonal influenza vaccine strategy [410].

Breakthrough infections in fully vaccinated individuals have been associated with poor or absent mucosal immunity, which is critical for blocking SARS-CoV-2 infection and disease progression [405]. Several studies have demonstrated that the first-generation SARS-CoV-2

vaccines were unable to elicit effective mucosal immunity in the upper respiratory tract [405, 411, 412]. Recent reports have shown that many mucosal vaccine candidates have entered clinical trials in response to the weakening of SARS-CoV-2 vaccine-induced immunity [413]. Therefore, generating a combined RV-SARS-CoV-2 live-attenuated oral vaccine could induce strong cellular and humoral immune responses [320]. In particular, orally administered vaccines would be more suited for use in rural areas, especially in low- and middle-income countries, due to their capability for long-term storage at refrigeration temperature compared to the ultralow-freezer temperatures required for mRNA vaccines. In contrast to the intramuscular injection of the current vaccines, oral vaccines could improve accessibility by eliminating the need for sterile needles thereby reducing manufacturing costs. RV vaccines could also be administered to children under the age of 6 months and without the need for adjuvants, which could reduce the infectious disease burden in early childhood in countries with high SARS-CoV-2 related deaths [388].

The RV vaccine expression platform has the potential for development of a safe and effective vaccine that could elicit a protective mucosal immune response, as well as help to overcome some challenges of vaccine distribution to low- and middle-income countries. Additionally, the RV-based combination vaccine could be amenable to other enteric pathogens and deployed in countries where child mortality is associated with comorbidities.

Table 13 Characteristics of selected licenced SARS-CoV-2 vaccines versus RV RotaTeq vaccine

Manufacturer	Vaccine name	Vaccine type	Immune response	Species	Vaccine efficacy*	Dose (route)	Storage	Population	Booster required
Merck	RotaTeq [224]	Pentavalent live-attenuated	IgG and IgA, T cell immunity (Th1)	Human-bovine reassortant	98%	3 doses at 2, 4 and 6 months (oral)	24 months at 2-8°C	≤ 6 months	No
AstraZeneca	Vaxzevria [414]	Non-replicating adenovirus-vector	Neutralising antibodies (IgG) and T cell immunity (Th1-skewed specific CD4+ and CD8+)	Rhesus macaques	74%	2 doses 4 - 12 weeks apart (intramuscular)	9 months at 2-8°C	≥ 18 years old	Yes
Janssen	Jcovden [415]				67%	1 dose (intramuscular)	6-9 months at 2-8°C		
Moderna	Spikevax [416]	mRNA		N/A	94.1%	2 doses 28 days apart (intramuscular)	6 months at -25°C; 30 days at 4°C	≥ 6 months	
BioNTech	Comirnaty [417]				95%	2 doses 21 days apart (intramuscular)			
Novavax	Nuvaxovid [418]	Spike protein subunit with adjuvant			89.7%	2 doses 3 weeks apart (intramuscular)	9 months at 2 - 8°C	≥ 12 years old	

*Based on overall vaccine efficacy from Phase III clinical trials.

Chapter 5: Analysis of potential accessory gene products expressed by the bovine RV strain RF

5.1 Background and Aims

RNA viruses have very compact genomes and rely on the host translational machinery for viral protein synthesis. To maximise their coding capacity, RNA viruses utilise various non-canonical translational mechanisms that facilitate access to multiple, often overlapping ORFs. Examples of such mechanisms include the use of IRESs, ribosomal leaky scanning, non-AUG translation initiation (e.g. CUG, GUG and UUG) and ribosome re-initiation (Figure 5.1) [59, 419, 420]. Thus, by exploiting the host's translational regulatory mechanisms, RNA viruses can express several proteins, including accessory proteins, from a single mRNA transcript during infection. Although non-essential, viral accessory proteins can play important roles in immune evasion (e.g. ORF7a of SARS-CoV-2 [421]), virulence (e.g. PR8 PA-X of IAV [57]) and/or transmission (e.g. protein UP of enteroviruses [422]) [423].

Based on the conventional cap-dependent model of translation initiation, the 40S ribosomal subunit is recruited to the 5'-end of mRNA and scans along in a 5' to 3' direction until encountering a start codon (AUG) [424]. The efficiency of initiation at a potential AUG codon is modulated by the context of its flanking Kozak motif [328, 425, 426]. According to Kozak, the strength of the AUG codon is particularly dependent on the nucleotides at positions -3 and +4 relative to the AUG codon (A of the AUG is designated as +1) [425]. If an AUG is flanked by the nucleotide sequences – GCC*R*CC**AUG***G* **–** (R represents purine) it is considered to be strong, diminishing to moderate when either A/G is present at -3 or when G is present at +4, and AUG sites with neither the A/G at -3 nor the G at +4 are considered weak [425].

A transcript with a weak or moderate AUG start site can induce ribosomal leaky scanning [59, 427]. For example, one of the accessory proteins encoded by IAV segment 2, PB1-F2, is accessed by leaky scanning and modulates the host immune response, contributing to viral pathogenesis in mice [428-430]. For Andes virus (a hantavirus), leaky scanning produces both the nucleocapsid protein and an additional non-structural protein from a single small mRNA whose relevance during viral replication has not yet been described [431].

For some viruses within the *Sedoreoviridae* family, overlapping ORFs have been identified. The bicistronic S1 gene segment of mammalian reovirus encodes the σ1 protein and also a second protein, σ1s, which is translated from the first-out of frame start codon and plays a role in pathogenesis and virulence [432-434]. Similarly, the M3 gene segment of reovirus encodes a second non-structural protein, μNS, translated from a downstream in-

frame AUG codon that is important for viral inclusion body formation in which RNA replication, assortment, and packaging converge for particle assembly [435-437].

In most RV species A, segment 11 contains an overlapping ORF that leads to expression of two non-structural proteins, the main product NSP5 and an additional NSP6 polypeptide probably accessed by leaky scanning [45, 60]. NSP6 has been proposed to localise to viroplasms (sites of RV replication) through its interactions with NSP5 and NSP2 [62]. Mice infected with NSP6-deficient virus still exhibited diarrhoea but viral replication and pathological changes were impaired [66]. Gene segment 7 of some RV strains including SA11 possesses two in-frame initiation codons where the canonical AUG has a suboptimal Kozak context, thus facilitating expression of two polypeptides in a cell-free system [72, 73]. In species B RVs, gene segment 5 contains two overlapping ORFs which were predicted to encode putative proteins NSP1-1 and NSP1-2, where NSP1-1 mediates fusion and syncytium formation, while the role of NSP1-2 has not been characterised [75].

Besides the Kozak context surrounding the first AUG codon, the distance of the AUG to the 5' end or to a preceding AUG codon can also promote the mechanism of leaky scanning [438, 439]. In murine norovirus, translation of a virulence factor 1 from an overlapping ORF initiates at the third AUG codon, 13 bases downstream of first AUG and 7 from the second, despite the second AUG occurring in an optimal Kozak context [440]. The bicistronic segment 6 of influenza B virus encodes two proteins from overlapping ORFs, canonical NB and an additional protein NA, from a downstream AUG codon located only 4 nucleotides away from canonical AUG which is 5 nucleotides from the 5' end [441]. Additionally, ORFX in genogroup I of astroviruses, overlaps the 5' region of the capsid-encoding ORF2 and is predicted to be translated via ribosomal leaky scanning, possibly enhanced by the very short leader sequence [58]. The 5' and 3' UTRs of RV genes vary in length, with the shortest 5' UTR found to be only 9 nucleotides in the gene segment encoding VP4 (*see Introduction, Figure 1.3*). Thus, the close proximity of an AUG codon to the 5' end of the transcript can promote leaky scanning as it is less efficiently recognised by the ribosome with efficiency diminishing as the 5' UTR length decreases below 30 nucleotides [438].

The presence of a short upstream ORF may stimulate a mechanism of ribosome re-initiation where the 40S subunit can resume scanning and re-initiate translation at a downstream ORF, thus bypassing intervening AUG codons [442]. Indeed, Wise *et al.*, (2011) showed that translation from segment 2 mRNA of IAV to produce multiple polypeptide species was regulated by a combination of ribosomal leaky scanning and translation re-initiation [443]. The initiation codon for ORF3 in the feline calicivirus genome is positioned close to the

termination codon of ORF2 and is translated via re-initiation after translation of ORF2 [444, 445]. This is thought to occur through interaction of the 40S ribosomal subunit with the termination upstream ribosome-binding site (TURBS) sequence in the viral mRNA that tethers the translation initiation factor eIF3 to the mRNA [446]. A similar mechanism is used by influenza B virus for the translation of the bicistronic segment 7 where the coding sequences for M1 and BM2 products overlap and translation initiation of BM2 depends on the presence of TURBS [447].

Although AUG codons are preferred, several examples have been identified where multiple proteins are translated from a single transcript that involve initiation at non-AUG codons [442, 448]. Examples of viruses that employ this approach include the Sendai virus that utilises an in-frame 'ACG' to initiate translation of the C' protein (N-terminal extension of C), and human parainfluenza virus 1 translates its C' protein from a 'GUG' codon [449, 450].

Other mechanisms involving *cis*-acting RNA structural elements, called IRESs, presented to the ribosome can support translation of viral mRNAs in cap-independent manner [420]. There are four different classes of IRESs each differing in their mechanism of internal translation initiation [451]. The picornavirus type I IRES found in poliovirus spans the 5' UTR of 743 nucleotides and promotes efficient translation initiation in the absence of cap but require all other initiating factors [452]. Whereas class I IRES of human rhinovirus 2, under some circumstances, can initiate translation of downstream AUGs without scanning where the ribosome is loaded onto mRNA by direct base pairing [453, 454]. Translation of two in-frame AUG codons, separated by 84 nucleotides, in the mRNA of the foot and mouth disease virus (FMDV) (aphthovirus) is driven by class II IRES present in the long 5'-UTR [455]. It does not bind the 40S subunit directly and requires almost all canonical translation initiation factors (except eIF4E) as well as *trans*-acting factors [420]. Some members of the *Flaviviridae* and *Picornaviridae* families express class III IRESs that are found only at the 5' end of viral RNAs that directly bind the 40S ribosomal subunit and eIF3 using multiple contact points in their domains [456, 457]. The compact class IV IRES (< 200 nucleotides) can initiate translation on a non-AUG start codon as it binds the 40S ribosomal subunit directly onto the mRNAs without the need for scanning or any canonical initiating factors [451]. This IRES is exclusively found in the intergenic region of the viral genome of the *Dicistoviridae* family that separates two ORFs and drives the translation of ORF2 [458].

In silico genomic analysis of species A RVs identified conserved RNA structures at the 5' and 3' terminal regions of all segments that sometimes extended into the coding region, suggesting presence of *cis*-acting functions [50]. Structurally conserved elements such as

long-range interactions were found in certain RV transcripts encoding proteins VP4, VP6, NSP2, NSP3 and NSP5 and were further biochemically validated for segment 11 only [50]. This raises the possibility that the translation initiation of some RV proteins could potentially be mediated through presence of IRES elements.

In addition to the translational regulation at initiation phase, non-canonical translation mechanisms that act during elongation and termination have also been identified in RNA viruses and include ribosomal frameshifting (e.g. pp1ab of mouse hepatitis virus [459]), stop codon read-through (e.g. non-structural protein of Sindbis virus [460, 461]) and stop-carry on mechanisms (e.g. segment 6 of species C RV [77]) [59].

Many RNA viruses have evolved various non-canonical translation strategies to express multiple protein isoforms from a single viral transcript that can enhance viral replication and pathogenicity. Based on the existing evidence for the expression of additional polypeptides in other viruses and in RVs (described in more detail in *Introduction*), this chapter aimed to identify potential accessory proteins in the genome of bovine RV strain RF. To test this, *in vitro* methods using RF plasmids and bioinformatic analyses were employed. Subsequently, the developed reverse genetics system for the RF strain was used to manipulate the viral genome and examine the potential effects of putative accessory proteins on viral replication.

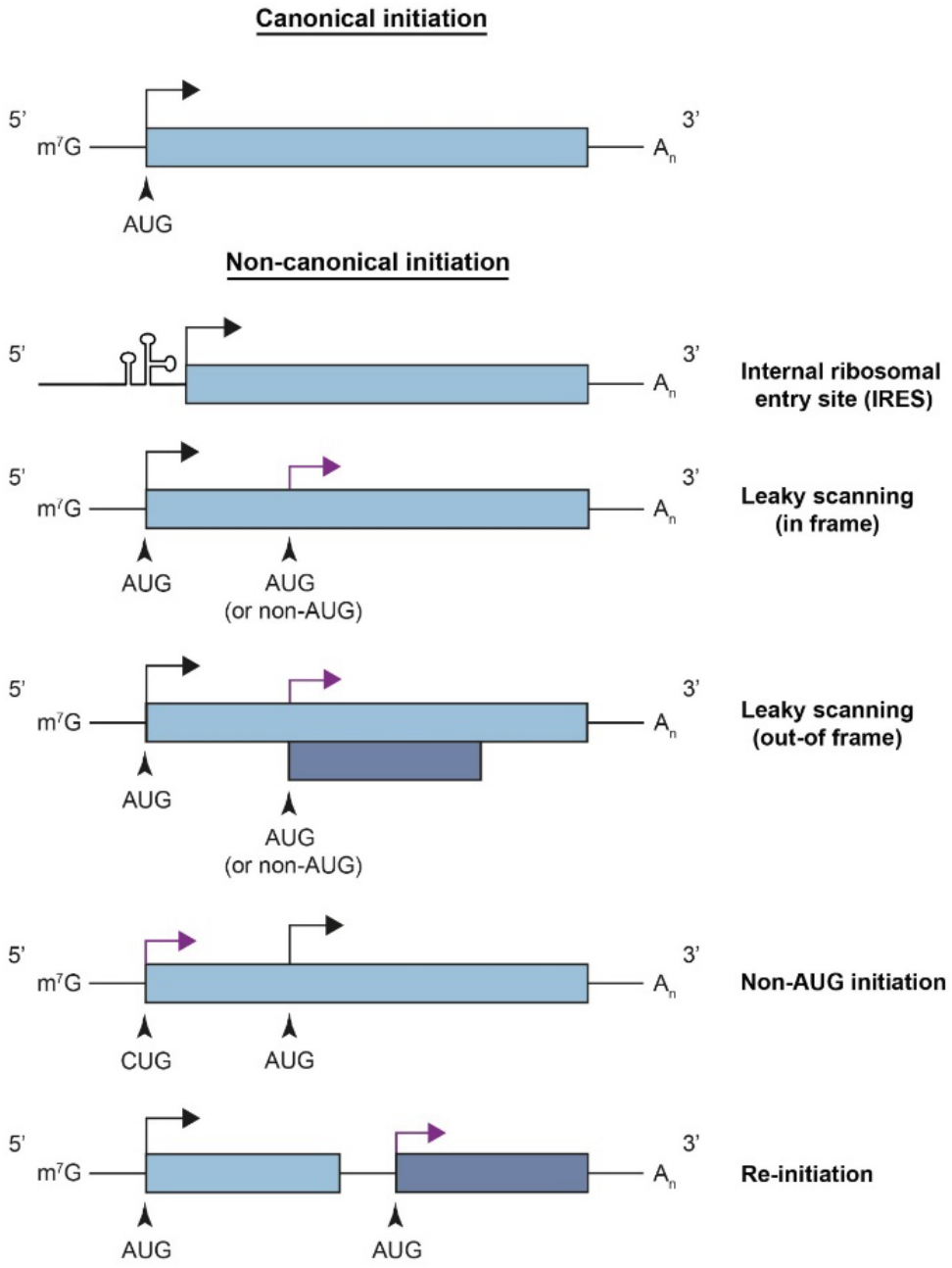


Figure 5.1 Canonical and non-canonical translation initiation mechanisms.
Top panel shows canonical mRNA translation and examples of non-canonical translation initiation are shown below. The main ORF is shown in blue in 5' to 3' orientation. Black arrows indicate initiation of protein synthesis at the start of an ORF and purple arrows show synthesis of an alternative product in a non-canonical manner. Expression of an additional polypeptide is coloured in dark blue. (*Adapted from Firth et al., 2012 [59]*).

5.2 Results

5.2.1 Detecting potential accessory proteins encoded by the RV genome

To test the hypothesis that RV segments express additional polypeptides other than those encoded by the primary ORF, coupled *in vitro* transcription and translation (IVT) reactions were carried out with the rabbit reticulocyte lysate system supplemented with ^{35}S protein labelling mix and programmed with cDNAs representing RV gene segments (Figure 5.2A). Empty pCDNA3.1 vector was used as a negative control to ascertain a background signal (Figure 5.2B and C, lane 1). Following detection by SDS-PAGE and autoradiography, the majority of RF and SA11 gene segments produced polypeptides of the expected product sizes for the main ORF (Figure 5.2A and B respectively). Translation of RF and SA11 gene segment 7 (VP7) produced a band of lower than the expected product size of 37 kDa (Figure 5.2B and C respectively, lanes 8). However, the RF NSP5 band migrated more slowly than expected (Figure 5.2B, lane 12). Additionally, no obvious protein products were detected for SA11 NSP1 and NSP5 (Figure 5.2C, lanes 6 and 12 respectively). In the case of SA11 NSP4, a slightly darker background product migrating at around 20 kDa was observed, but no unique product that could be unequivocally labelled as NSP4 was observed (Figure 5.2C, lane 11).

In addition to the products corresponding with canonical RV proteins, unknown polypeptide species were reproducibly detected in the RF and SA11 IVT reactions, which could be indicative of bona fide alternative gene products (Figure 5.2B and C, pink arrows). However, some unknown bands present in RF VP1 and VP4 translations could be strain-specific as these were not detected in SA11 reactions (Figure 5.2B, green arrows). Examination of replicate gels as well as non-radioactive IVT reactions supplemented with a tRNA charged with biotinylated lysine (data not shown) demonstrated consistent translation of an additional band in VP1 of ~100 kDa (Figure 5.2B and C, black arrow). Accordingly, downstream alternative translation initiation was further assessed for VP1.

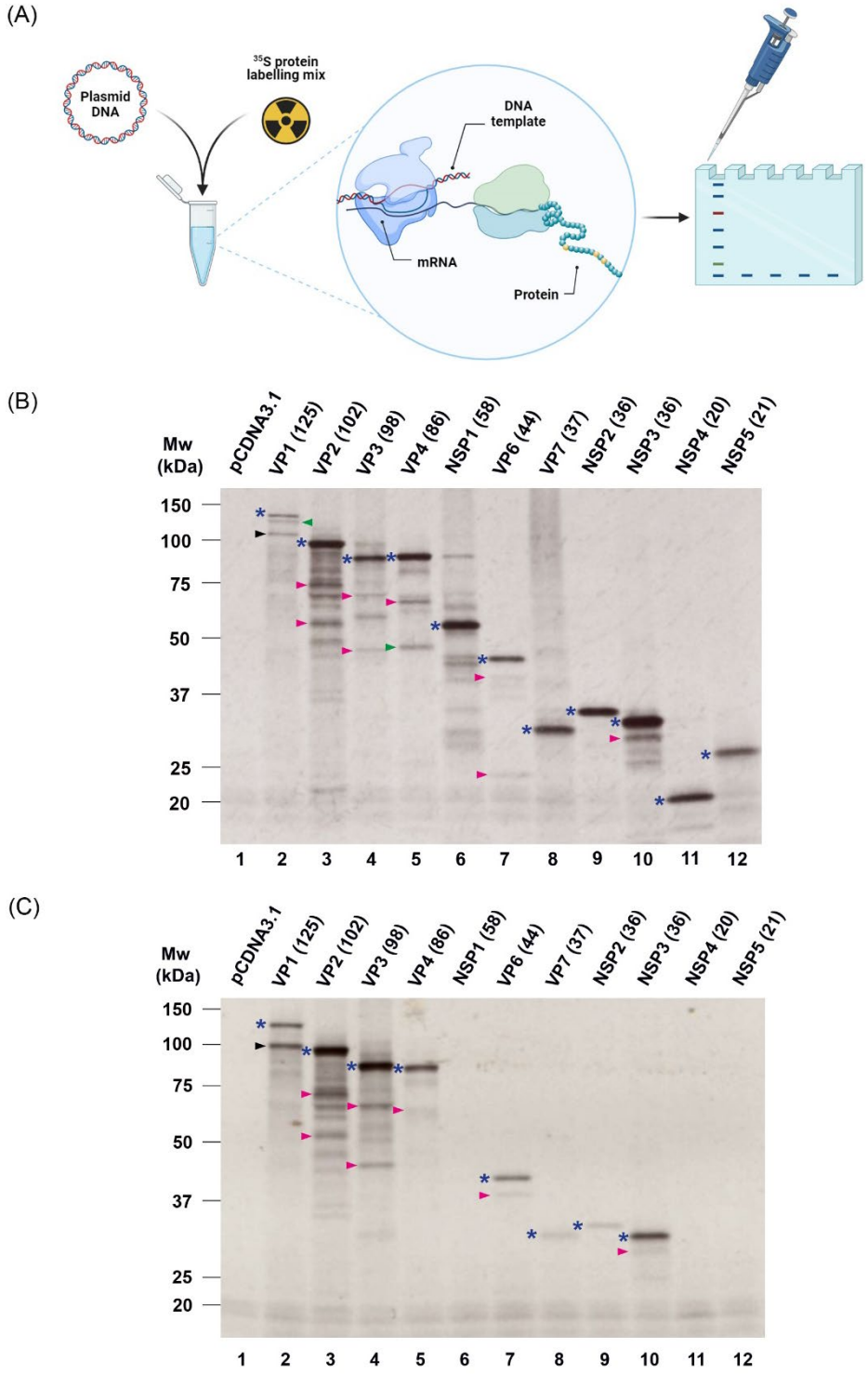


Figure 5.2 *In vitro* expression of RV gene segments.
(A) Schematic of coupled *in vitro* transcription and translation reactions carried out using rabbit reticulocyte lysate system supplemented with ^{35}S protein labelling mix. Newly synthesised proteins were separated by SDS-PAGE and detected by autoradiography. *(Created with BioRender.com).* **(B)** Analysis of 11 plasmids representing RF and **(C)** SA11 gene segments. Empty pCDNA3.1 vector was used as negative control. The molecular weight markers and the expected product sizes of each segment (in brackets) are indicated (kDa). Main protein products are highlighted by blue asterisks and pink arrowheads indicate additional polypeptides. Green arrowheads show products identified only in RF translations and black arrowheads highlight the VP1 band that was analysed further. Lane numbers are shown.

5.2.2 Bioinformatic analyses of AUG codons in segment 1

One hypothesis to explain the 100 kDa product identified in the IVT reactions of segment 1 (VP1) was the result of internal translation initiation from downstream AUGs. To identify potential candidate AUGs that could be responsible, out of 3,302 nucleotides the first 1,400 nucleotides of segment 1 were analysed, as the scanning efficiency of ribosomes reduces with increasing length of the mRNA to encounter potential downstream AUG codons [442]. Based on the observed size of the novel VP1 product and the predicted size of polypeptides hypothetically derived from internal translation initiation, five AUG codon candidates: AUG6, AUG12, AUG14, AUG20 and AUG22, were selected from frame 1 for downstream analysis (Figure 5.3A). No ORF capable of producing a protein product of suitable length was predicted from frame 2 or 3.

To evaluate the conservation of the identified AUG candidates, all available nucleotide sequences for VP1 were downloaded from the NCBI viral genome resource and aligned using SSE software [325, 327]. The dataset was refined by including only the field isolates, thus the laboratory and vaccine strains were not included in the analysis. Following elimination of poor quality and fragmented sequences as well as sequences with ambiguous nucleotides, 1,689 sequences were further assessed using R Scripts written by Dr Samantha Lycett and Mr Marius Diebold (*Appendix A, section 8.1*). Initially, the presence of AUG codons was recorded for all sequences, after which the duplicate positions were removed generating a list of all positions containing an AUG codon. Next, the prevalence of specific AUG codons in all reading frames was plotted (Figure 5.3B). As expected, the main canonical product expressed by AUG1 at nucleotide position 19 was 100% conserved. In contrast, the AUG6 at nucleotide position 214 was poorly conserved and was only detected in 11 sequences (Figure 5.3B, dashed line). However, the remaining candidate AUGs were 100% prevalent in all sequences analysed (Figure 5.3B).

The Kozak motif surrounding the AUG codon influences protein translation efficiency [328, 426]. Thus, all nucleotide sequences from frame 1 were further assessed using R Scripts to identify the strength of the Kozak signalling. Sequences were categorised into strong (A/G at -3 and G at +4), medium (either A/G at -3 or G at +4) and weak (neither A/G at -3 nor G at +4) [328]. For determining the Kozak strength of the canonical AUG, the majority of the sequences were unassigned due to the absence of the 5' UTR in the published sequences (Figure 5.4). For those sequences for which it was possible to assign Kozak strength (36.9%), all sequences had a strong signal except for one that displayed a moderate strength. In the 11 sequences in which AUG6 was present, it was in a strong Kozak context. Moderate Kozak

signalling was predominant for the remaining candidate AUGs and was over 98% conserved: AUG12 at 99%, AUG14 at 98.9%, AUG20 at 99% and AUG22 at 98.9%. Weak Kozak signalling for candidate AUGs was detected in less than 0.5% of the sequences and only AUG22 had a single sequence with a strong Kozak.

In summary, bioinformatic analyses of segment 1 sequences revealed conserved AUG codons that may be responsible for the expression of the novel polypeptide present in VP1 translation reactions. Most AUG candidates were almost 100% prevalent in clinical isolates and had over 98% conserved moderate Kozak signalling context. Despite AUG6 being the least conserved, it was incorporated into downstream analyses as it was present in the bovine strain RF genome which was used throughout this part of the study.

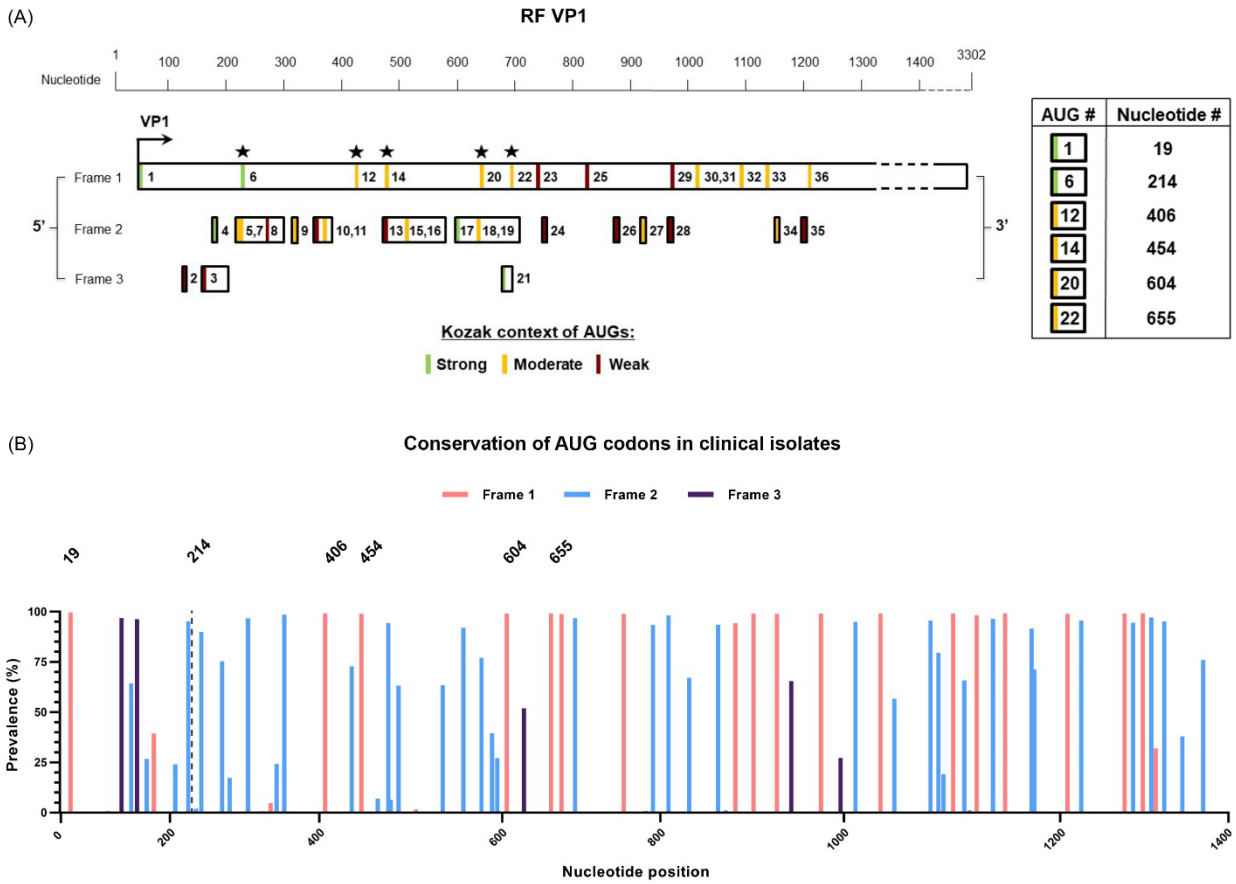


Figure 5.3 Conservation analysis of AUG codons in segment 1 (VP1).
(A) Schematic of ORFs for the first 1400 nucleotides of the 5' end of VP1 mRNA. The AUG codons in all reading frames are numbered according to their position and coloured according to the strength of their Kozak context. Green – strong with A/G at -3 and G at +4; yellow – moderate with either A/G at -3 or G at +4; red – weak neither A/G at -3 nor G at +4. Length of the boxes correspond to the ORF length. Stars indicate candidate AUGs that may serve as translation initiation sites for the unknown polypeptide identified in VP1 of ~100kDa. **(B)** Conservation of AUG codons in three frames from all available VP1 nucleotide sequences (1,689 sequences analysed) downloaded from NCBI viral genome resource shown as (%). Nucleotide positions of candidate AUGs are marked and dashed line represents a poorly conserved AUG6 at position 214.

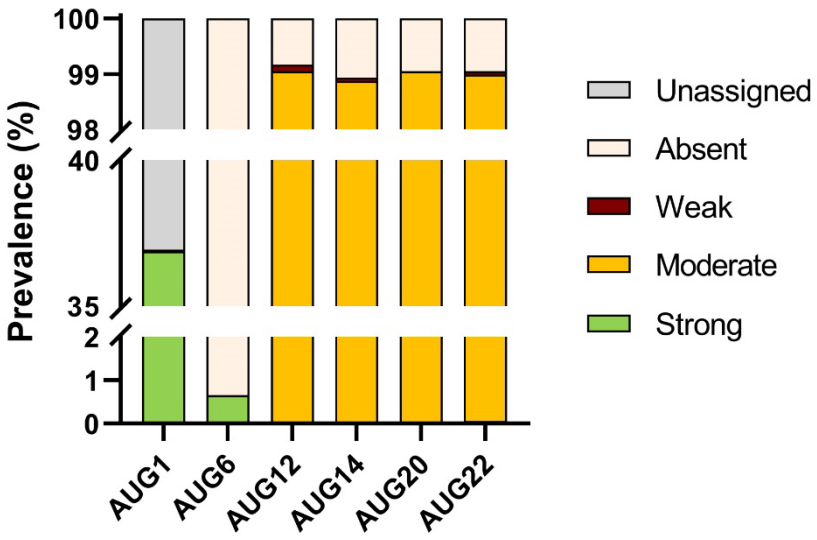


Figure 5.4 Analysis of the Kozak context of candidate AUG codons in segment 1. The strength of the Kozak motif for each AUG codon in the dataset was categorised as strong (green – A/G at -3 and G at +4), medium (yellow – either A/G at -3 or G at +4) and weak (red – neither A/G at -3 nor G at +4). 'Unassigned' (in grey) refers to sequences in which the Kozak context was not determined due to missing sequence information and 'absent' (beige) highlights sequences lacking the AUG codon.

5.2.3 Mutations of candidate AUGs in RF segment 1

Following identification of candidate AUGs in segment 1 (VP1), site-directed mutagenesis was performed to determine whether an AUG codon was responsible for internal translation initiation of the unknown polypeptide of ~100 kDa. Point mutations were introduced to change the methionine codon at AUGs 6, 12, 14, 20 and 22 to leucine or valine (Table 14). Leucine and valine were selected from the same hydrophobicity group to minimise the risk of any detrimental effect on protein folding of VP1 [462]. Subsequently, AUG candidates will be referred to by their corresponding position within the VP1 protein (Table 14).

IVT reactions were carried out with the VP1 mutant panel following the same method as described in Figure 5.2. In addition, segment 1 of the bovine RV isolate ZAF strain (RVA/Cow-wt/ZAF/1603/2007/G6P[5]) was included in the IVT reactions to test if the presence of the identified unknown polypeptide was specific to the laboratory strains of RF and SA11 [463]. SA11, ZAF and WT RF all produced the canonical VP1 product and the unknown polypeptide of ~100 kDa, suggesting its conservation across clinical and laboratory strains (Figure 5.5, lanes 2, 3 and 4 respectively). Furthermore, a stop codon was introduced at nucleotide position 30, downstream of AUG1, to check that the annotated band corresponding to the canonical VP1 product was assigned correctly (lane 5). As expected, in the presence of a stop codon, the top band was no longer detected, confirming that it corresponded to the canonical VP1 protein product, while the additional lower bands were still present (lane 5). Mutation of AUG6 (M66L) suggested that it is required for the expression of the polypeptide running just below the main VP1 protein product (lane 6). However, this polypeptide was not identified in SA11 nor in ZAF as they do not encode AUG6, complementing its low conservation observed in Figure 5.3 (Figure 5.5, compare lanes 2 and 3 to 4). Individual mutations of AUG12 (M130L) and AUG14 (M146L) showed no effect on the translation of the unknown polypeptide, although a slight size difference in the band migration in the M146L translation reaction was observed (lanes 7 and 8). However, mutation of AUGs 12 and 14 together (M130L+M146L) presented a decrease in translation of the unknown polypeptide, consistent with the unknown band arising from co-migration of two unresolvable products expressed from the AUG12 and AUG14 codons (lane 9, pink arrowhead).

It is plausible that when the AUG codons 12 (M130) and 14 (M146) were mutated to CUG, residual translation initiation still occurred in the IVT reactions, explaining why a faint band remained for these double mutants (lanes 7 and 8). Thus, the AUG codon at these positions was mutated to GUG (valine) to test whether a clearer effect on the unknown band of ~100 kDa would be observed. However, mutation of M130V and M146V had no effect on

the unknown polypeptide and similar to M146L (lane 8), a slight change in the band migration was seen in M146V (lanes 10 and 11). Additionally, the double mutation of M130V+M146V displayed similar phenotype as M130L+M146L with a faint band still present (lane 12, pink arrowhead).

To test whether the remaining AUG candidates, AUG20 (M196) and AUG22 (M213), affected the translation of the unknown band, leucine mutations were also introduced at these sites. The expression of the unknown polypeptide remained unchanged by single or double mutation of M196 and M213, suggesting that the observed band was a result of co-migrating polypeptides translated from M130 and M146 residues (lanes 13, 14 and 15).

Overall, the data in this section showed that the simultaneous mutation of both AUGs 12 and 14 (M130+M146), to ‘CUG’ (leucine) or to ‘GUG’ (valine), resulted in a decreased expression of the identified polypeptide. This suggests that two isoforms of VP1 products were being translated from M130 and M146 as single mutations at these sites had no effect on the expression of the identified band. Thus, it is likely that the co-migration of the two products represent the band of ~100 kDa.

Table 14 Summary of AUG candidates and their respective nucleotide and codon changes
Resulting mutation in the VP1 protein and the corresponding product sizes are shown (nucleotides (nt), amino acids (aa) and kDa). Introduction of a stop codon is denoted by N/A under the 'AUG #' column.

AUG #	Nucleotide change	Codon change	VP1 mutation	ORF size (nt)	Product size (aa)	Product size (kDa)
N/A	T30G	UAU → UAG	Y4*	3255	1084	120.4
6	A214C	AUG → CUG	M66L	3072	1023	113.6
12	A406C		M130L	2880	959	106.6
14	A454C		M146L	2832	943	104.8
12	A406G	AUG → GUG	M130V	2880	959	106.6
14	A454G		M146V	2832	943	104.8
20	A604C	AUG → CUG	M196L	2682	893	99.2
22	A655C		M213L	2631	876	97.3

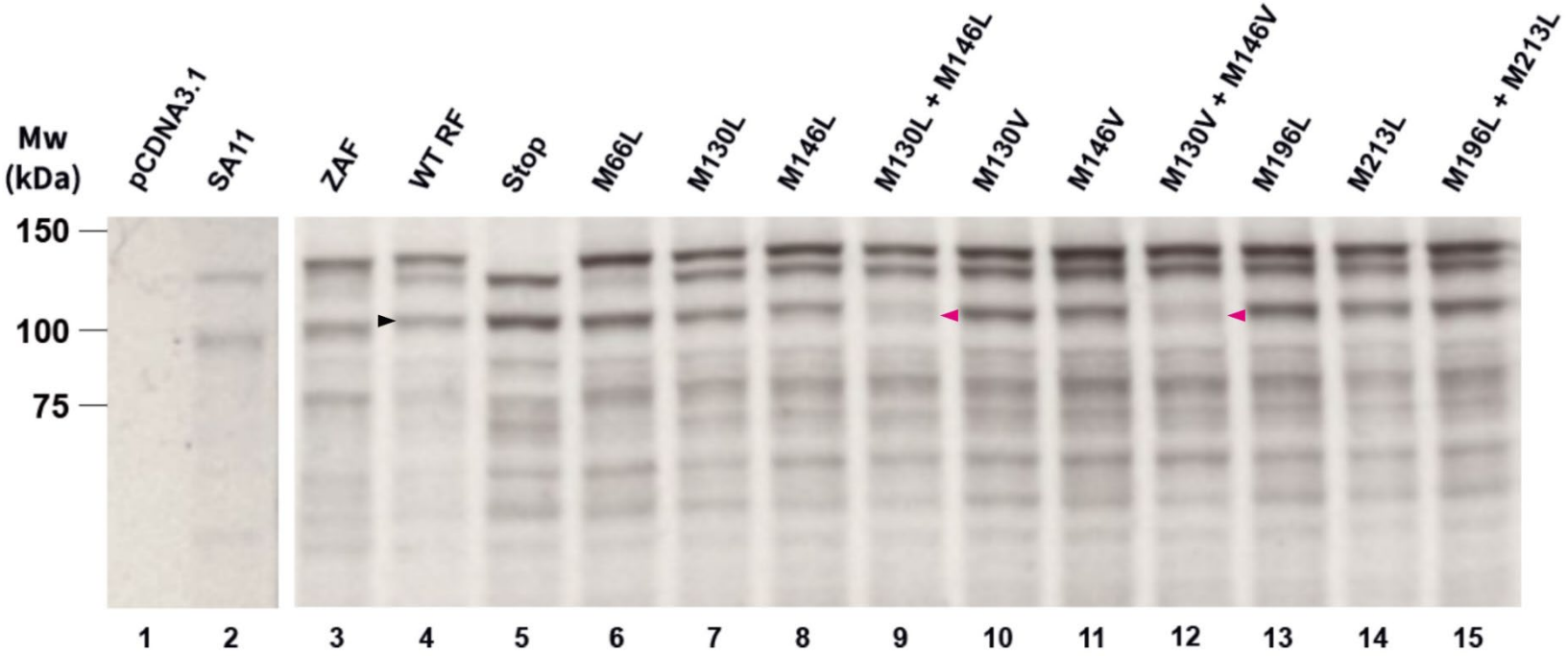


Figure 5.5 Site-directed mutagenesis of candidate AUGs in segment 1.
IVT reactions were carried out as described in **Figure 5.2** and samples were analysed using SDS-PAGE and autoradiography. Empty pCDNA 3.1 vector was used as a negative control. Segment 1 of the bovine RV isolate ZAF strain (RVA/Cow-wt/ZAF/1603/2007/G6P[5]) was included to test if the identified unknown polypeptide was specific to the laboratory strains of SA11 and RF. 'Stop' represents a construct in which a stop codon was introduced between AUGs 1 and 6 of segment 1 ORF. The positions of molecular weight markers are shown (kDa) and lane numbers are indicated. The black arrowhead marks the unknown polypeptide of interest and pink arrowheads highlight instances of a decrease in its translation.

5.2.4 Expression of VP1 truncated products in mammalian cells

To test if the novel VP1 isoforms were an artefact of the IVT reactions or were a result of alternative translation initiation, the expression of mutated VP1 products was also tested in cells. For this, the first 670 nucleotides of segment 1 (VP1) containing all five AUG candidates were cloned into an expression plasmid upstream of, but in frame with the EGFP gene (Figure 5.6). The EGFP fusion constructs were made due to the unavailability of antibodies against VP1. This construct, termed WT RF, was then used to introduce frame 1 mutations of the candidate AUGs by site-directed mutagenesis as described in Table 14. For AUG candidates that could initiate downstream translation the ORF structures with their predicted protein products and product sizes are shown in Figure 5.6.

Following confirmation by DNA sequencing, the panel of VP1 constructs was transfected into HEK293T cells and after 48 hr analysed by western blot. Early attempts at expressing the VP1 constructs in cells repeatedly gave poor results, suggesting potential protein degradation (data not shown). Therefore, after the initial 24 hr, cells were incubated with either DMSO or 20 μM of MG132 drug (proteasome inhibitor) for an additional 24 hr (Figure 5.7A and B respectively). The β-actin loading control was consistent across all lanes in both panels except for pEGFP and ORF9B, which were underloaded to avoid signal saturation and interference (panel A and B, lanes 1 and 2 respectively). The empty plasmid expressing EGFP (pEGFP) was used as a transfection control and produced a protein product that migrated at the predicted molecular weight of 27 kDa (panel A and B, lane 1). A plasmid expressing SARS-CoV-2 ORF9B tagged with GFP (36 kDa) was used as an additional positive control; this also produced a major polypeptide species that migrated as expected and whose abundance was not drastically affected by the MG132 treatment (panels A and B, lane 2). The negative control, termed ‘mock’, included the transfection reagent only (panel A and B, lane 3). In DMSO only, with the exposure settings required to visualise the VP1-EGFP fusion proteins in other samples, faint background bands were visible in ORF9B and mock samples, which were absent at the lower exposure settings required in the presence of MG132 (panel A and B, lanes 2 and 3 respectively).

In the absence of MG132 treatment, all VP1 fusions gave readily detectable amounts of an anti-EGFP reactive species migrating similarly to the untagged EGFP itself (Figure 5.7A, red arrowhead, lanes 4-14). This product probably resulted from accumulation of a stably folded EGFP domain, potentially avoiding the proteolytic degradation of a longer fusion protein. The majority of the VP1 constructs also produced lower amounts of a peptide

migrating slightly slower than the 50 kDa marker, likely representing a full length VP1-EGFP (52.6 kDa) protein (panel A, black arrowhead, lanes 4-14). In confirmation of this, the ~52 kDa band was absent from the Stop construct sample (panel A, lane 5). A further novel species seen above background levels was also observed in most VP1 samples, running slightly faster than the 50 kDa marker (panel A, blue arrowhead, lanes 4-14) which could possibly correspond to the product initiated from M66 in the IVTs (Figure 5.5, lane 6). However, the intensity of this band was not affected by leucine mutation (M66L) (Figure 5.7A, lane 6, blue arrowhead), but it was reduced by the Stop and M130L mutations (lanes 5 and 7), but not the M130V mutation (lane 10).

In the presence of MG132, most but not all VP1 constructs still produced an EGFP signal (Figure 5.7B, red arrowhead, lanes 4-14) and a more predominant 52 kDa product corresponding to the full-length VP1-EGFP fusion protein (black arrowhead, lanes 4-14), consistent with this polypeptide being unstable and subject to proteasomal degradation. Also prominent in most samples was the slightly faster migrating product of just under 50 kDa (blue arrowhead), as well as a family of peptides migrating at around 37 kDa (yellow arrowhead); these latter bands potentially matching the predicted molecular weights of proteins initiated at M130 and M146 (Figure 5.6).

The band of ~45 kDa was not drastically affected by the M66L mutation, but was reduced by the Stop mutation (panel B, lanes 5 and 6, blue arrowhead). Its accumulation was however reduced by M130L and M130V but not by changes at M146 (compare lanes 7 and 10 with 8 and 11 respectively). However, the ~45 kDa polypeptide was no longer detected in mutations of both M130L and M146L, whilst M130V and M146V together had no effect (compare lane 9 with 12, blue arrowhead), further supporting the conclusion that unlike in the IVT reactions, this peptide does not result from translation initiation at M66.

Contrasting the IVT result, expression of the 37 kDa peptide was reduced with the M130L and M130V mutation, while alterations of M146 had much less of an effect (panel B, yellow arrowhead, compare lanes 7 and 10 with 8 and 11 respectively). However, in agreement with the IVT data, accumulation of the 37 kDa species was reduced by double mutation of M130L and M146L (lane 9) whereas only a minimal effect was observed with both M130V and M146V mutations (lane 12). Unlike in the IVT translations, mutation of M196L reduced the expression of the 37 kDa band as well as the EGFP signal (lane 13, yellow and red arrowhead respectively). It is possible that the M196 was utilised for the translation of the observed band of ~25 kDa (lane 13, red arrowhead) based on its predicted protein size (Figure

5.6). On the other hand, similar to the IVT reactions, no effect was observed with M213 mutation (lane 14).

Quantification of individual peptides by densitometry from three independent experiments containing MG132 showed that the above pattern was consistent (Figure 5.7C). For each construct, the observed polypeptides were normalised to the total density of VP1 and presented as its fraction. The polypeptide species within each construct were then compared to the WT RF. As expected, introducing a Stop significantly suppressed the expression of the VP1-EGFP fusion protein as well as the 45.1 kDa polypeptide although not significantly, whereas the 37 kDa peptide was upregulated in contrast to WT RF. It further confirmed that the mutation of M130L with or without M146L reduced the expression of both 45.1 kDa and 37 kDa peptides (significantly with M146L), while significantly increasing the accumulation of the VP1-EGFP fusion protein. The remaining constructs were only subtly different to WT RF.

In summary, the “in cell” test of alternative translations initiation events of RF segment 1 mRNA confirmed that a novel VP1 isoform, named VP1-N129, was expressed from M130. This result did not support the hypothesis that M66 and M146 were used, with the caveat that these outcomes may have been obscured by post-translational modifications (in particular degradation) of the VP1 fusion proteins.

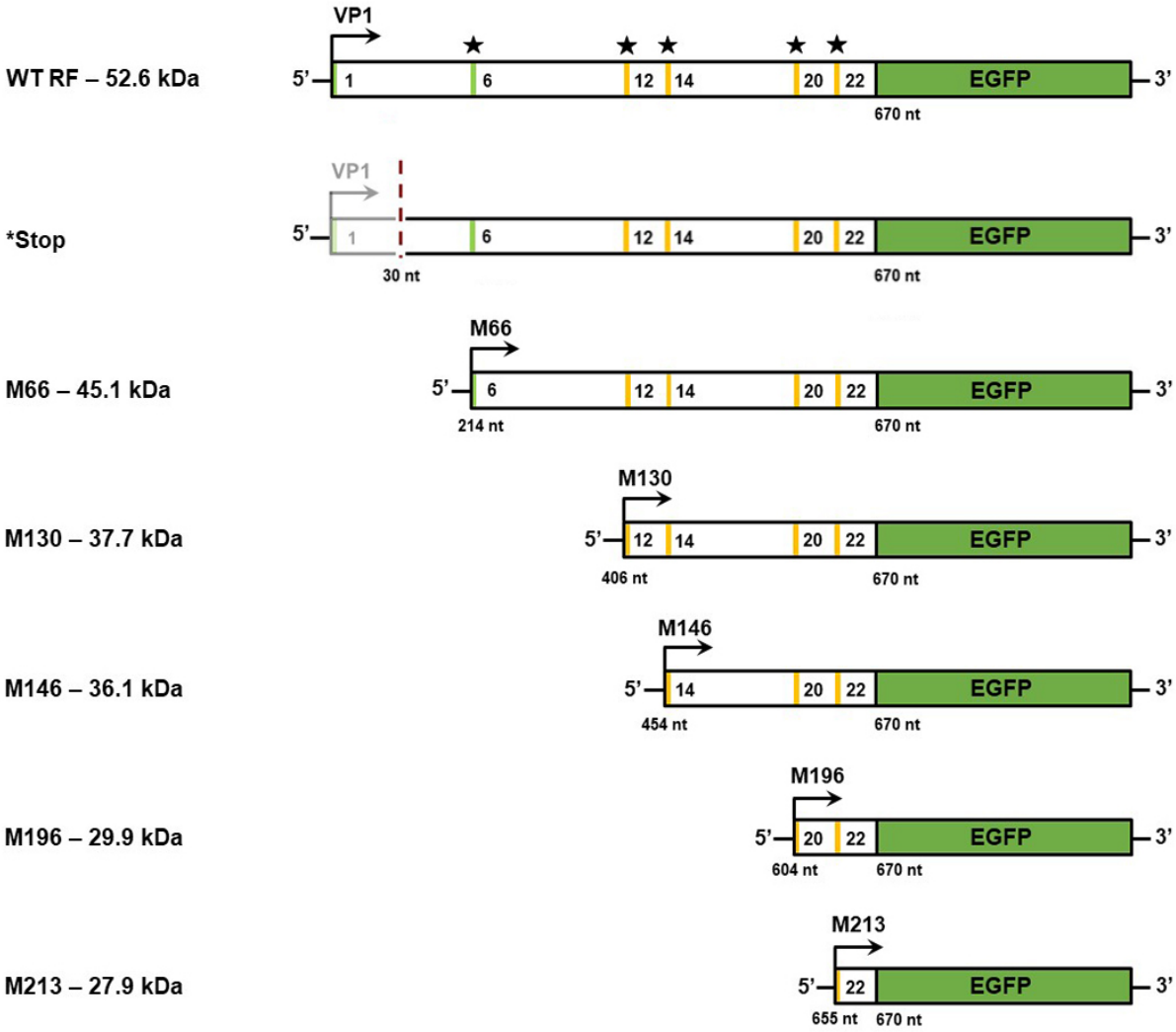


Figure 5.6 Schematic of the first 670 nucleotides of segment 1 tagged with the EGFP gene.
ORF structures with their predicted protein products and molecular weight (kDa) are shown. AUG codons are numbered according to their position and coloured based on the strength of their Kozak context (green – strong consensus with A/G at -3 and G at +4; yellow – medium with either A/G at -3 or G at +4; red – weak with neither A/G at -3 nor G at +4). Stars highlight candidate AUGs that may serve as translation initiation sites to yield the identified peptide of ~100kDa. Red dashed line represents the location of a stop codon, 30 nucleotides downstream of the first AUG codon. The nucleotide position of each candidate AUG is also shown.

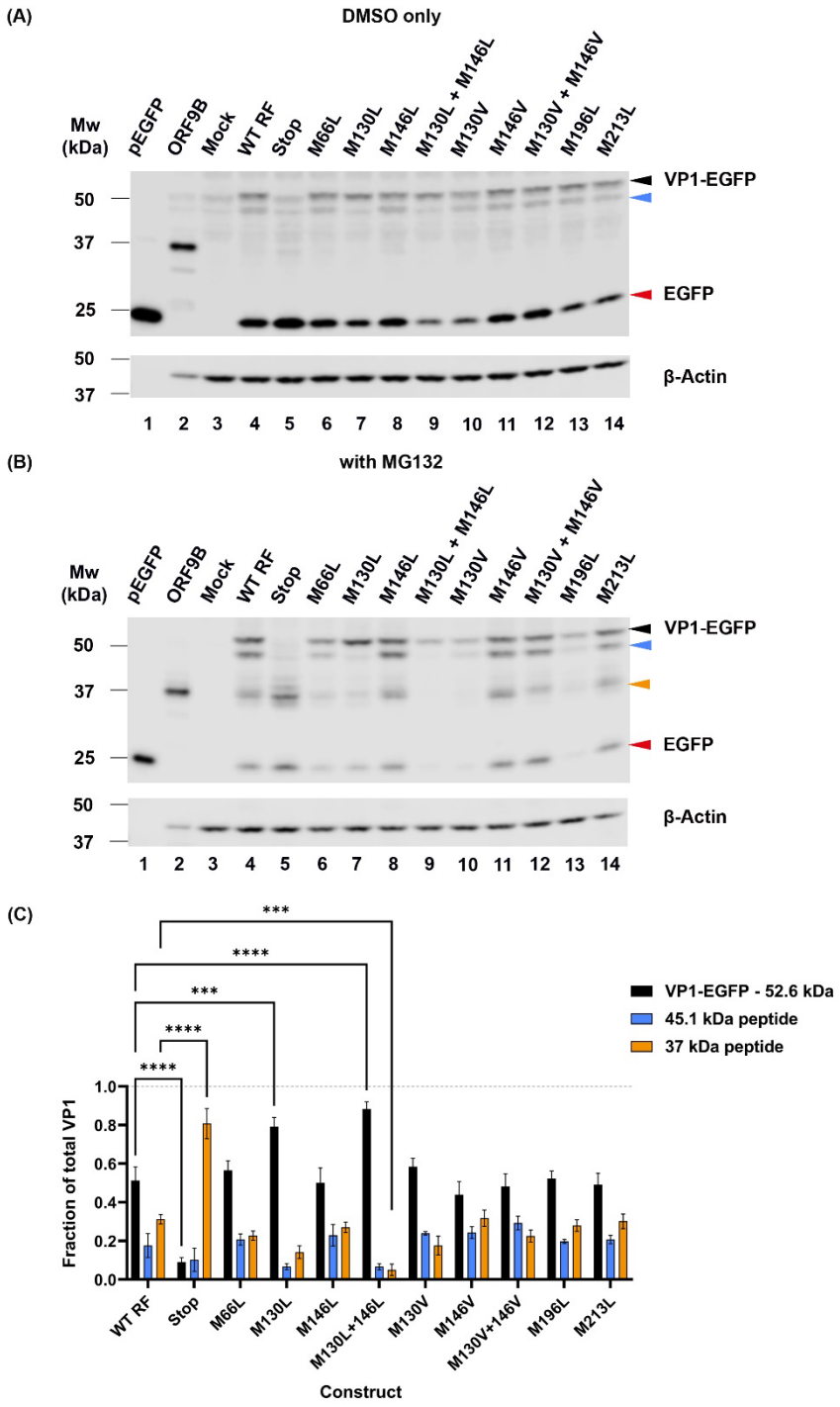


Figure 5.7 Expression of VP1 products in transfected cells.
HEK293T cells were transfected with the indicated VP1 and control constructs tagged with EGFP and incubated in the absence **(A)** or presence of MG132 drug **(B)**. After 48 hr, lysates were collected and analysed by western blot. β-actin was used as a loading control. The positions of molecular weight makers (kDa) are indicated. pEGFP was used as a transfection control and SARS-CoV-2 ORF9B was used as a positive control. Note that the pEGFP and ORF9B samples were intentionally underloaded to avoid signal interference in adjacent lanes. **(C)** Quantification of the relative expression of VP1-pEGFP – 52.6 kDa, 45.1 kDa and 37 kDa peptides in each construct was performed by densitometry in the presence of MG132 drug. Data are mean ± SEM from three independent experiments. Statistic annotations are the result of a two-way ANOVA test with Dunnett's multiple comparisons test performed for each construct for each peptide. ****p-value ≤ 0.001, *****p-value ≤ 0.0001.*

5.2.5 Generation of VP1 mutant viruses

Data so far has suggested that the 100 kDa N-terminally truncated isoform of VP1, VP1-N129, could be a functionally important but unidentified accessory protein. Mutation of specific AUGs in segment 1 (VP1) identified likely sites of downstream translation initiation resulting in production of this peptide. Therefore, to examine the effect of the mutations introduced into VP1 on viral fitness, the mutated segment 1 was used to rescue viruses using reverse genetics for the bovine strain RF following the same protocol as described in *Chapter 3*. Following the viral rescue (from three independent experiments), MA104 cells were infected with the panel of viruses and titrated by plaque assay.

No virus was detected in the mock-infected controls, where the segment 1 plasmid was omitted (data not shown). As expected, no viable virus was generated when a stop codon downstream of the canonical VP1 AUG was incorporated (Figure 5.8A). The M66L mutant viruses produced a comparable titre (>10^7 PFU/mL) to WT RF, suggesting no major effect on viral fitness. In contrast, despite being only 16 amino acids apart, the M130L change was detrimental, whereas M146L had no effect on viral rescue. Double mutation of M130L and M146L also failed to produce any virus. In contrast with the mutation to leucine, a valine mutation at the M130 position did not impact viral rescue and the M146V virus showed a similar phenotype to WT RF (Figure 5.8A). Similar to double mutation of M130L+M146L, mutation of M130V+M146V, resulted in failure to rescue virus. M196L and M213L mutants had comparable titres to that of WT RF, whereas the double mutation of M196L+M213L did not produce any virus. Sequencing of VP1 mutants from P2 stocks confirmed the presence of the desired mutations (data not shown).

To further characterise the fitness of the VP1 mutants, plaque morphology was assessed and compared to WT RF. MA104 cells were infected with P1 stocks from three different rescues and incubated under a semi-solid overlay medium. Following a four-day incubation, plaques were fixed and stained with toluidine blue. Plates were then scanned and cell-absent areas were quantified using ImageJ. A heterogeneous population of large and small plaques was observed from WT RF and all mutants, although the M130V and M196L viruses had an obvious small plaque phenotype (Figure 5.8B). Quantification of individual plaques showed that, with the exception of M196L, no significant difference in plaque phenotype was observed between WT RF and the VP1 mutants, all of which displayed plaques ranging between ~0.25 – 3.5 mm in diameter (Figure 5.8C). M196L had significantly smaller plaques than the WT RF, suggesting a loss of fitness in cultured cells (Figure 5.8C).

To compare the replication kinetics of the VP1 mutants with the WT RF, multicycle infections were performed in MA104 cells infected at a low MOI for 48 hr. Samples were collected at several times post infection and titred by plaque assay. No virus was detected in mock infected controls (data not shown). Overall, all mutants displayed broadly similar growth characteristics to WT RF, increasing steadily in titre across the 48 hr timecourse from $<10^3$ PFU/mL at 0 hr (residual inoculum) to peak titres of around 10^6 PFU/mL (Figure 5.8D). The M196L mutant was the only exception, reaching the peak titre by 32 hr and thereafter plateauing with significantly lower titre than WT RF at 40 and 48 hpi, complementing its small plaque phenotype.

Overall, data from the viral rescue experiments demonstrated that only the M130L mutation was lethal to the virus; all other changes were tolerated (including M130V alone), though the M196L mutation led to a reduction in virus fitness. This further supports the hypothesis that the polypeptide expressed from M130 (AUG12) is a potential accessory protein correlating with its synthesis in cells. However, the complexities of M130L versus M130V effects on the viral rescue could potentially be attributed to their different effect on the polymerase activity of VP1 where M130L mutation affects the polymerase and M130V does not.

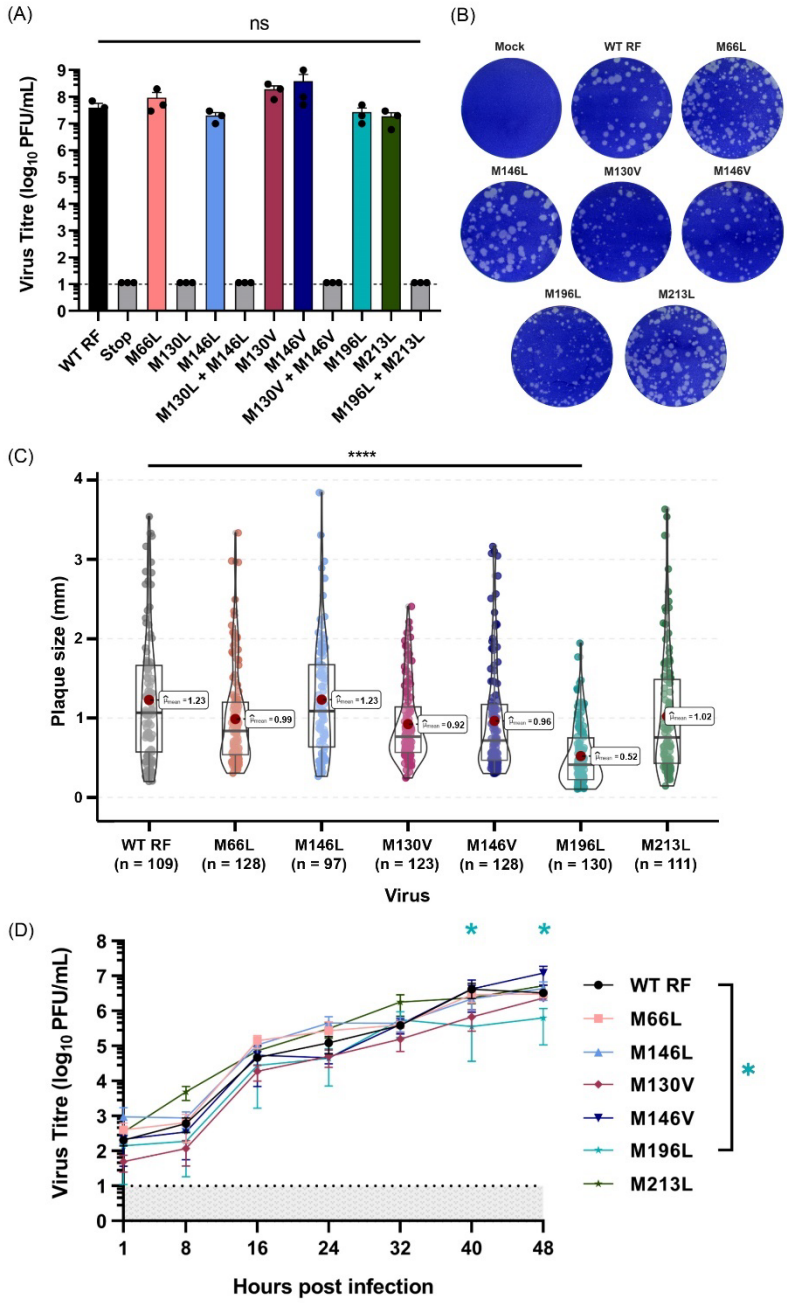


Figure 5.8 Rescue of VP1 mutant viruses and growth kinetic analysis.
(A) Viruses were rescued using the established reverse genetics system (see **Chapter 3**) and viral titres of P1 stocks were titrated by plaque assay on MA104 cells. Data are mean ± SD from three independent experiments plaqued in duplicate. Statistical annotations are the result of an ordinary one-way ANOVA test. Multiple comparison tests were performed between each mutant. NS: Non-significant. Dashed line represents detection threshold. **(B)** Plaque phenotype of VP1 mutants versus WT RF. **(C)** Violin plot of plaque size distribution for VP1 mutants versus WT RF. A minimum of 40 plaques were measured for each virus from three independent experiments plaqued at the same time. Each dot represents a plaque. Statistical annotations are the result of a Welch ANOVA test with Holm's adjustment. *****p-value ≤ 0.0001*. The mean values are shown. **(D)** Multi-step growth curves comparing VP1 mutants with WT RF. MA104 cells were infected at MOI of 0.03 and samples were collected at the indicated hours post infection. Dashed line represents detection threshold. Data are mean ± SEM from three independent experiments. Statistical annotations are the result of an ordinary one-way ANOVA test (performed individually for each time-point). Only significant values are annotated. **p-value ≤ 0.05.*

5.2.6 Establishing a mini-genome assay to test the polymerase function of VP1

To test whether the polymerase activity was affected by VP1 mutations, a system was needed to assay the RdRp function outside of the context of virus replication. Therefore, we sought to establish a mini-genome assay to reconstitute the polymerase activity.

To set up a plasmid based mini-genome assay for RV, a similar approach to the reverse genetics system was utilised. cDNA constructs were designed to express the firefly luciferase (Fluc) gene in a positive or negative orientation flanked by the T7 RNA polymerase promoter and HDV ribozyme followed by the T7 transcription terminator (Figure 5.9A). The 5' and 3' UTRs from the NSP1 gene were incorporated into the construct to retain the *cis*-acting signals for polymerase recognition, similar to the reporter design for IAV mini-genome assay [464]. Thus, if the synthetic RV segment was recognised and transcribed and/or replicated by the viral polymerase, the luciferase levels in transfected cells would represent a measure of the transcriptional activity of the polymerase.
These custom constructs were synthesised by Invitrogen GeneArt in a pMA plasmid backbone. Positive and negative sense reporters were named 5'-reporter and 3'-reporter respectively. Following the same method as described in *Chapter 3*, 11 plasmids corresponding to each RV genome segment were co-transfected with either the positive or the negative sense reporter and luminescence was measured after 48 hr (Figure 5.9A). The plasmids carrying the NSP2 and NSP5 genes were transfected at a 3.125 times higher amount (312.5 ng each) than the remaining plasmids as described previously (100 ng each) [288].

Before assessing the effects of VP1 mutations on the polymerase activity, initially, increasing amounts of the reporter plasmids were transfected into BSR-T7 cells to determine the quantity required for an optimal signal without saturating the system. A plasmid expressing the Fluc gene driven by the cytomegalovirus immediate early promoter was used as a positive control (named PVR1255) and this produced a strong luciferase signal that increased with plasmid dose (Figure 5.9B). A negative control "mock" sample included the transfection reagent only; this produced constant low levels of background signal. When increasing amounts of the two synthetic RV reporter plasmids were transfected, only the baseline signal of luminescence was observed until 200 ng. At 200 ng both reporters showed a significant increase above background signal with the 5'-reporter having a higher luminescence value compared to the 3'-reporter (Figure 5.9B). Results indicated that the maximum luciferase signal detected was at 200 ng of reporter plasmid; an amount which was then used in further analysis.

Next, a dose-dependent titration was performed for all 11 plasmids in the presence of 200 ng of the 5'- or the 3'-reporter to determine the minimal amount of RV plasmids needed to generate the best luciferase signal (Figure 5.9C and D respectively). Transfected 5'-reporter or 3'-reporter plasmids only were used to establish a baseline signal that was then used for normalisation (data not shown). The PVR1255 plasmid was included as a positive control and produced a strong signal as described above. Co-transfection of the 11 RV plasmids with 200 ng of the 5'-reporter showed a significant increase in the luciferase expression compared to the 11 plasmids only signal (Figure 5.9C). The maximum luminescence reading was reached at a 100 ng dose of 11 plasmids (312.5 ng of NSP2 and NSP5) suggesting that the system was saturated, after which the luciferase activity started to decline with decreasing amounts of RV plasmids added (Figure 5.9C). In the absence of the 3'-reporter plasmid, 11 plasmids alone showed a background signal slightly above the mock sample that contained the transfection reagent only (Figure 5.9D). However, the luciferase activity did not reach significant levels when the 3'-reporter was co-transfected with 11 plasmids at varying quantities (Figure 5.9D). Thus, the 3'-reporter construct was discarded from subsequent experiments due to the low levels of luminescence readings despite addition of RV plasmids. Overall, these experiments suggested that an RV mini-genome system based on transfecting 200 ng of the 5'-reporter along with 100 ng of the RV plasmids (312.5 ng for NSP2 and NSP5 plasmids) might be viable, as it showed the highest luminescence signal.

The RV polymerase domain resembles a 'right-handed' architecture made up of "palm, finger and thumb" subdomains [108]. Ogden *et al.*, (2012) showed that the conserved aspartate residues within the 'GDD' motif in the palm subdomain of RV RdRp were critical for RNA synthesis (Figure 5.10A) [159]. To test if the Fluc signal in the RV mini-genome system was dependent on the VP1 activity, these conserved aspartate residues, D631 and D632, were mutated to alanine, creating mutants D631A and D632A respectively (Figure 5.10B). Following the introduction of these mutations into the VP1 plasmid by site-directed mutagenesis, viruses were rescued using the reverse genetics to confirm RdRp inactivity. As expected, no virus was recovered in three independent experiments in the presence of the mutated polymerase (Figure 5.10C). This confirmed that the mutations rendered the virus incompetent of replication, most likely due to polymerase inactivation.

Subsequently, to test whether VP1 mutations had an effect on the polymerase activity, 11 RV gene segments (100 ng each except 312.5 ng for NSP2 and NSP5), were co-transfected into BSR-T7 with the 5'-reporter at 200 ng. The WT VP1 plasmid was substituted with mutated VP1 plasmid and transfected at 100 ng (Figure 5.10D). All luminescence readings were normalised to the background signal of the 5'-reporter only (data not shown)

and compared to D631A and D632A samples to assess whether RdRp activity was affected. As before, a PVR1255 positive control produced a strong luciferase signal. The mock and 11 plasmids only samples showed similar background luminescence which were below levels of detection (Figure 5.10D). Despite the presence of D631A and D632A mutations within the catalytic site of VP1, these samples still produced luminescence readings higher than the 5'-reporter alone, suggesting that some residual transcriptional activity of the polymerase remained (Figure 5.10D). Nonetheless, the luciferase activity was still significantly higher for WT RF and M66L samples compared to D632A but not to D631A (Figure 5.10D). The M130L and M146L mutants showed an increase in luminescence in contrast to D632A but not as significant when compared to the D631A mutant. Although, the M130L+M146L and M130V+M146V mutants affected the viral rescue, in this assay their luciferase activity was similar to that of WT RF, with the M130L+M146L mutant displaying a significantly higher signal than D631A and D632A (Figure 5.10D). Contrasting to the slow replication and small plaque phenotype of M196L, its luciferase signal was not as affected as the D631A or D632A. The M213L mutant was the only sample that displayed comparable luciferase activity to D631A, indicating that the mutation introduced at this site had a similar effect as the mutation made in the catalytic site that intended to disrupt the polymerase activity (Figure 5.10D).

In summary, preliminary results from the mini-genome assay showed that the VP1 mutations had a minimal impact on the polymerase activity, though further optimisation and characterisation of this assay is required. Thus, within the limits of the data, the results are consistent with the hypothesis that a potential accessory protein is expressed from M130 (AUG12). Although, it is plausible that the mutation at M130 affected the overall VP1 protein structure or it's inter-molecular interactions.

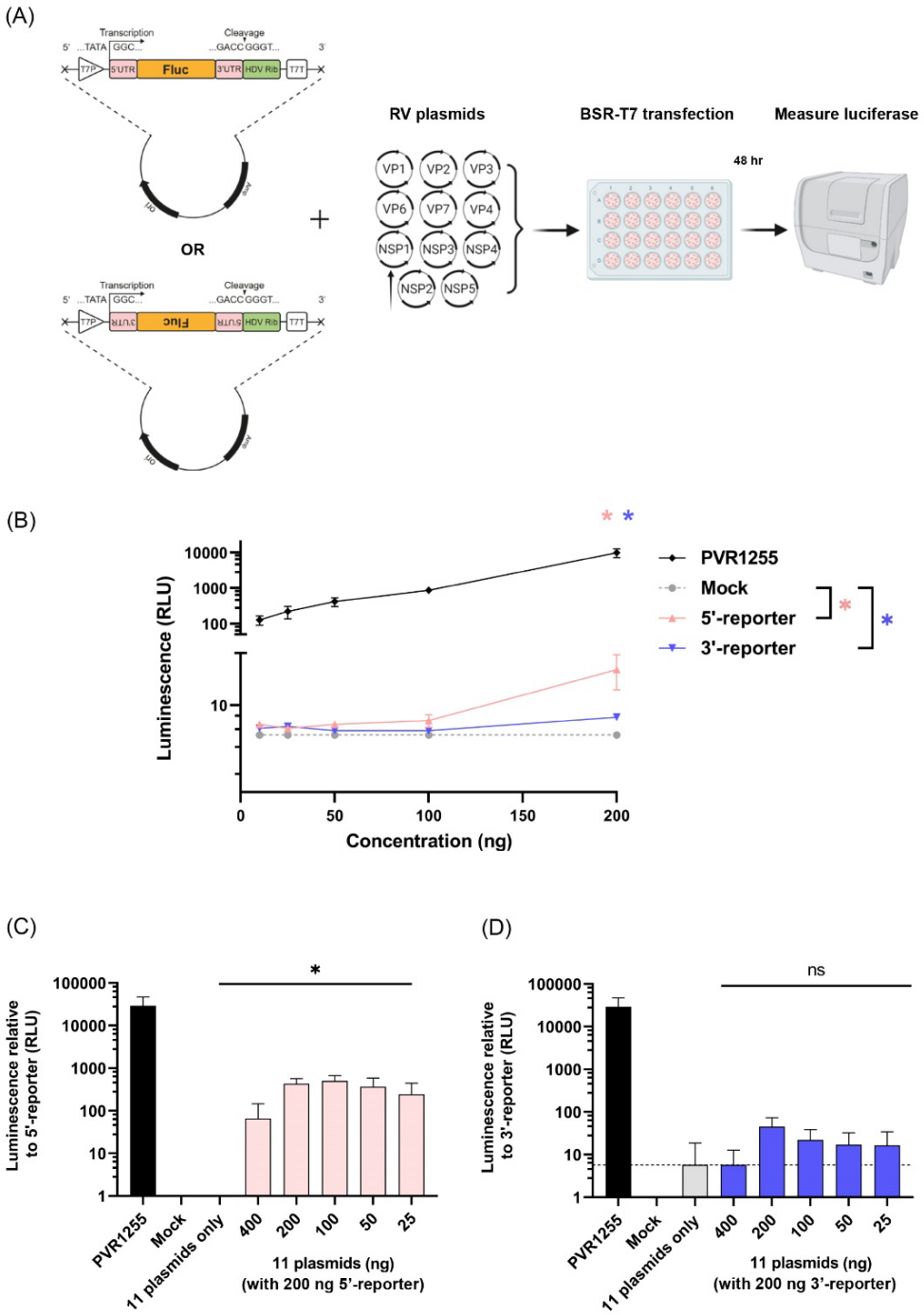


Figure 5.9 Establishment of a mini-genome assay.
(A) Schematic of proposed mini-genome assay. RV plasmids encoding each bovine RF stain gene were co-transfected with T7 reporter plasmids expressing the Fluc gene in either positive (5'-reporter) or negative sense (3'-reporter). Luciferase activity was measured after 48 hr post transfection. *(created with BioRender.com)*. **(B)** Dose-dependent titration of the 5'- and 3'-reporter plasmids. PVR1255 plasmid expressing FLuc gene was used as a positive control. Mock sample contained transfection reagent only. Data are mean ± SEM from four independent experiments each performed in duplicate. Statistical annotations are the results of ratio paired t-test. **p-value ≤ 0.05*. Dose-dependent titration of 11 RV plasmids with 200 ng of the 5'-reporter **(C)** and 3'-reporter **(D)**. Normalised values to 5'-reporter and 3'-reporter plasmids only are shown (**C** and **D** respectively). Data are mean ± SEM from four independent experiments each performed in duplicate. Statistical annotations are the result of ratio paired t-test. **p-value ≤ 0.05*, NS: Non-significant.

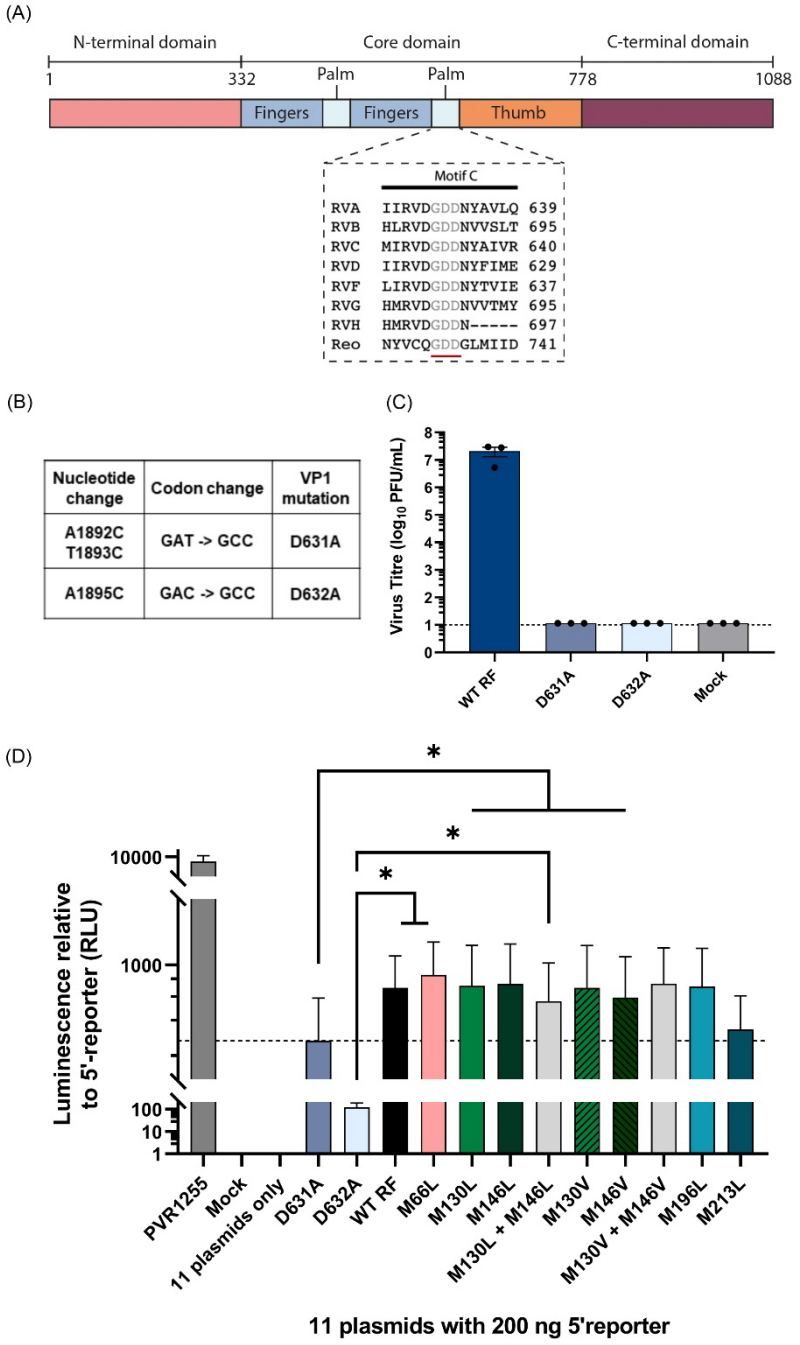


Figure 5.10 Measuring polymerase activity of VP1 mutants using the mini-genome assay. **(A)** Generation of inactive VP1 RdRp through mutation of the catalytic site. VP1 shown as a linear schematic and coloured according to the domain organisation with amino acid numbers labelled above. *(Adapted from Steger et al., 2019 [157] and Oden et al., 2012 [159])*. The N-terminal domain (pink) and C-terminal domain (burgundy) flank the core domain containing fingers (blue), palm (light blue) and thumb (orange). Sequence-based alignment of RdRps from representative species of RV and reovirus showing the conserved catalytic 'GDD' site (underlined in red) within motif C. 'GDD' mutations are shown in **(B)**. **(C)** Viral titres of mutants and WT RF were titrated by plaque assay on MA104 cells. Data are mean ± SD from three independent experiments each plaqued in triplicate. **(D)** Luciferase activity following transfection of VP1 mutants normalised to the 5'-reporter only sample. PVR1255 plasmid expressing FLuc gene was used as a positive control. Mock sample contained the transfection reagent only. D631A and D632A samples were used as negative controls. Data are mean ± SEM from four independent experiments each performed in duplicate. Statistical annotations are the result of ratio paired t-test. *$p< 0.05$. Only significant values are annotated.

5.2.7 Predicted crystal structure of VP1 M130 and M146 mutants

To further test the hypothesis of whether the VP1 mutations were affecting the function of the protein, potential structural differences between the VP1 mutants and the WT RF were evaluated. Crystal structures of VP1 for the SA11 and RRV, but not RF, strains have been previously resolved revealing a cage-like globular enzyme that shares structural similarities with other polymerases [107, 108]. Therefore, AlphaFold was used to predict the structures of WT RF and VP1 mutants. Only the M130 and M146 mutants were included in this analysis due to their varied effect on the virus rescue despite being only 16 amino acids apart.

Using the available resolved structure of VP1 for the RRV strain in a transcript-elongated state (TES) and in a duplex-open state (DOS), the VP1 molecule was coloured according to its polymerase domains: N-terminal domain (blue), core (active site) (grey) and C-terminal domain (purple) (Figure 5.11A). In the TES, the enzyme splits the dsRNA genome, releasing the nascent RNA transcript into the cytoplasm whilst retaining the template (-)RNA within the core to use for further (+)RNA synthesis that occurs in the DOS state [107]. The M130 and M146 residues were then mapped onto the structure and were shown to lie within the N-terminal domain of RdRp (red and magenta spheres respectively). Specifically, the M130 residue (red spheres) was located closer to the core of the polymerase (grey) containing the template RNA (black arrows). On the other hand, the M146 residue (magenta spheres) was near the surface of the protein and closer to the coding RNA strand (black arrow) (Figure 5.11A).

When the predicted structure of WT RF (pink) was superimposed with the RRV VP1 molecule in TES no major differences in the overall structure were seen, despite some areas of mismatch (Figure 5.11B, black arrowheads). Due to their similarities in structure, the VP1 domains in the WT RF molecule were coloured to match those of RRV VP1. The mapped locations of M130 and M146 residues in WT RF were similar to those in RRV VP1 (Figure 5.11C, red and magenta spheres respectively). Thus, it is plausible that mutating the M130 residue had influenced the interaction of the N-terminus with the template RNA within the core, while the M146 location could explain its non-detrimental effect on the virus rescue.

Subsequently, predicted structures of M130 and M146 mutants (Figure 5.12A and B respectively, coloured) were superimposed with the WT RF (in grey) and showed no conformational differences in the overall VP1 structure. Additionally, the spatial arrangement of the mutated residues and their interactions with the neighbouring amino acids was not affected (compare WT RF top magnified windows (grey) with coloured magnified windows).

However, an additional interaction was observed between the M146V residue and the neighbouring leucine at position 149 (L149) represented by an additional yellow bond, but this mutant did not affect the viral rescue (Figure 5.12B, light blue magnified window, black arrowhead).

Data in this section shows that no changes were made to the overall structure of the VP1 protein when M130 or M146 were mutated and their interactions with the neighbouring residues were not disrupted. Thus, this strengthens the hypothesis that the identified novel VP1 isoform, VP1-N129, could be a potential accessory protein being expressed from the M130 residue.

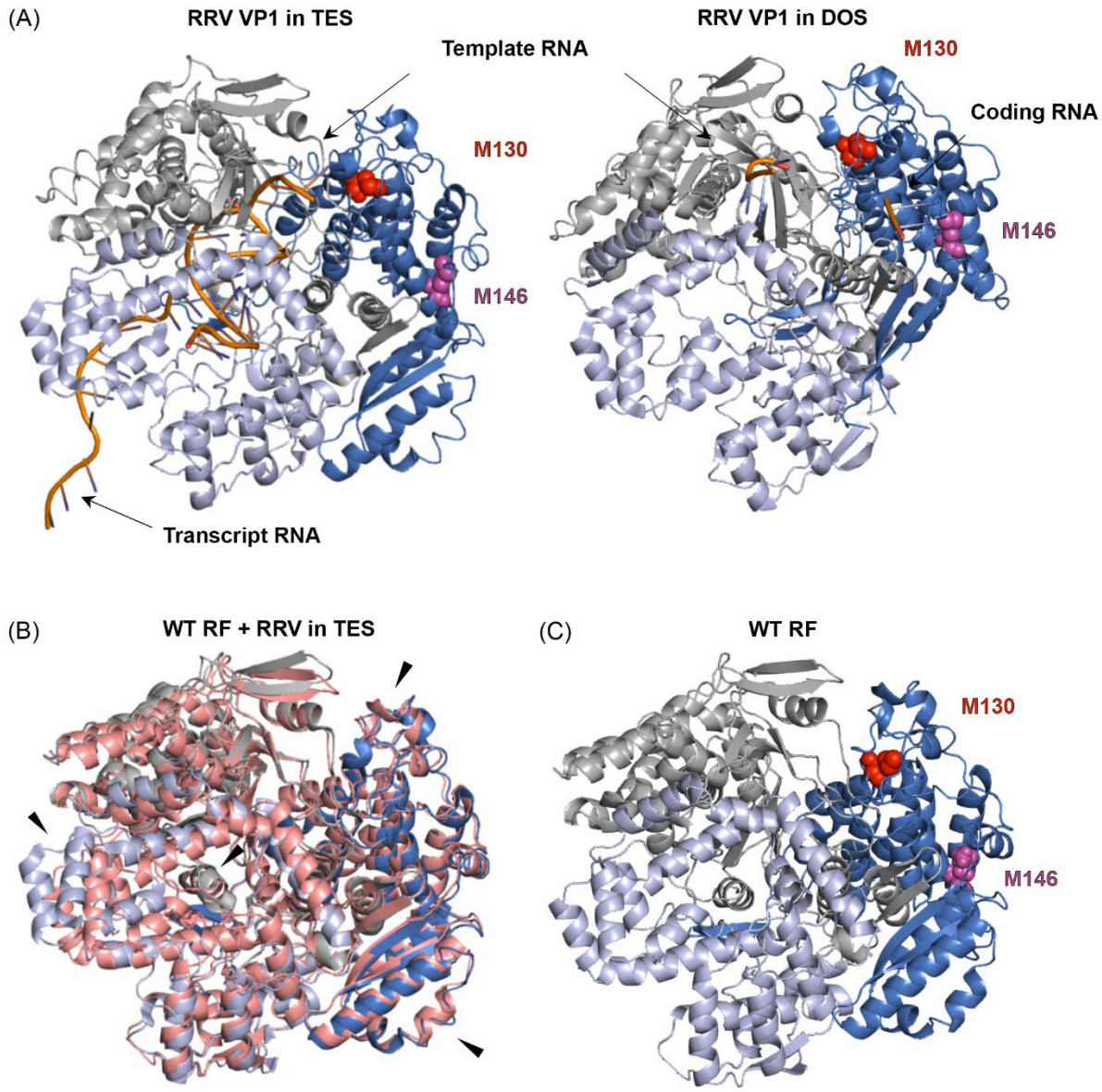


Figure 5.11 Crystal structure of VP1 RdRp showing M130 and M146 residues.
(A) Ribbon model of VP1 for RRV strain in transcript-elongated state (TES) and in a duplex-open state (DOS) showing N-terminal domain (blue), core (grey) and C-terminal domain (purple). In TES, the core is occupied by dsRNA genome and incoming NTPs (not shown) but contains no RNA in DOS. Template and transcript RNA are coloured by element and indicated by black arrows. A-helices are shown as spirals and β-sheets as arrows. Coloured spheres represent mapped M130 and M146 residues annotated in red and magenta spheres respectively. (*Adapter from Ding et al., 2019 [107] and annotated in PyMOL [330]*). **(B)** Predicted structure of WT RF using AlphaFold (pink) superimposed with RRV VP1 molecule in TES with its domains coloured as in **(A)**. Black arrowheads indicate areas of mismatch. **(C)** Predicted structure of WT RF with its domains coloured like those of RRV in **(A)** with mapped M130 (red spheres) and M146 (magenta spheres) residues.

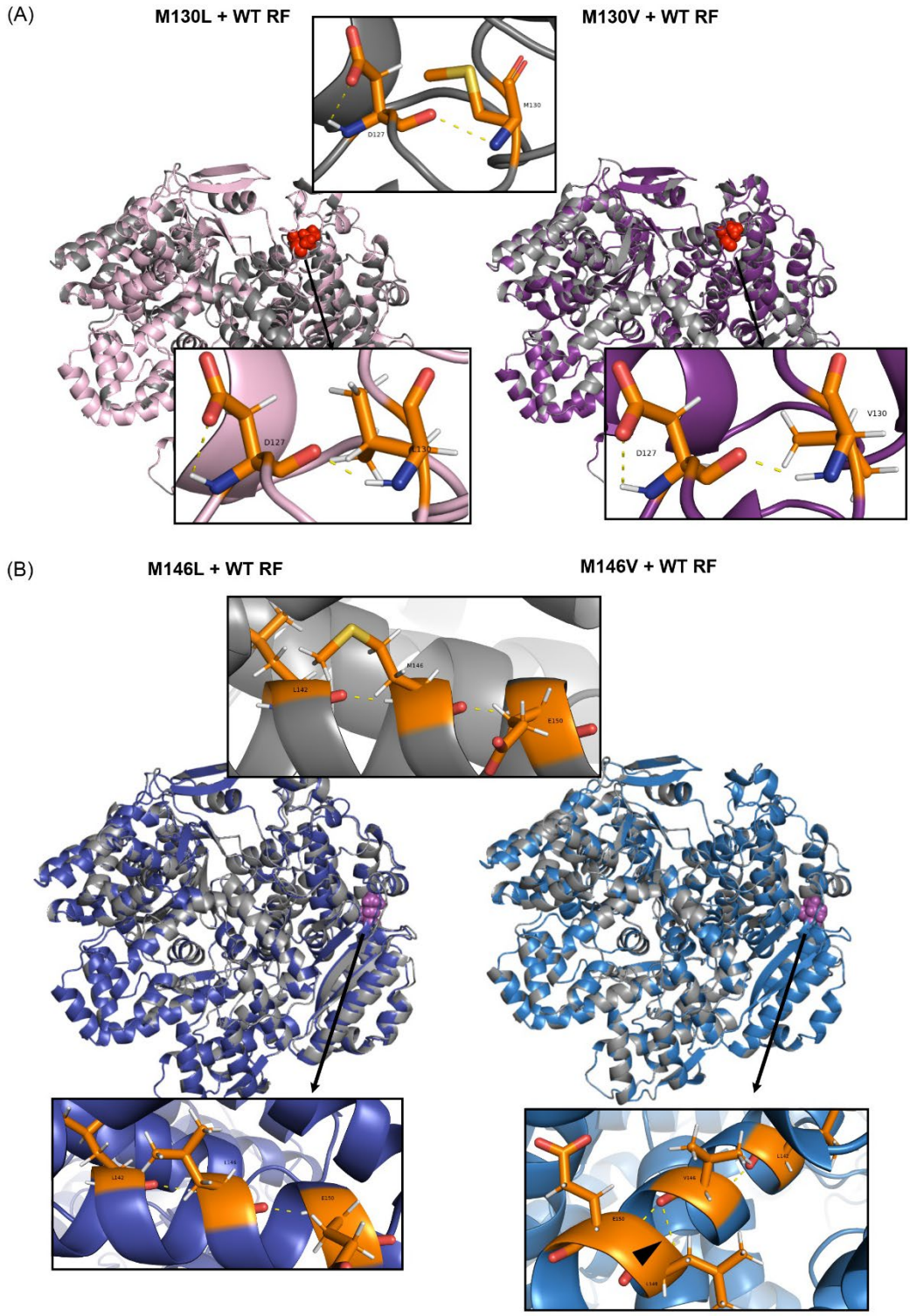


Figure 5.12 Predicted structures of M130 and M146 mutants.
Predicted structures of M130 **(A)** and M146 **(B)** mutants superimposed with WT RF (grey) in ribbon representation with coloured spheres showing the location of mutated residues. In **(A)** the structure of M130L and M130V is in pink and purple respectively, while in **(B)** M146L and M146V are in dark and light blue respectively. Top magnified window in each panel shows the respective residues in WT RF (grey) and their interactions. Bottom magnified windows show interactions between mutated and nearby residues. All residues are coloured by element: carbon = orange, hydrogen = grey, oxygen = red, nitrogen = blue and sulphur = yellow. One letter code of the amino acid residue with their corresponding location is shown. Dashed yellow lines represent hydrogen bonds. Black arrowhead shows additional hydrogen bond between M146V and the neighbouring residue L149.

5.3 Discussion

Expression of alternative ORFs through non-canonical translational mechanisms has been identified in many RNA viruses. The presence of a polycistronic segment 11 and alternative ORFs in multiple RV species led to the hypothesis that other segments of RV may also possess multiple ORFs within their genome, possibly accessed through similar mechanisms.

Initially, IVT reactions were set up to test the hypothesis of whether gene segments of RV RF and SA11 strains could express additional polypeptides. Results demonstrated that the majority of segments in both strains produced bands corresponding to the canonical products (Figure 5.2B and C). However, in some segments the main products were absent or were not of predicted size, possibly due to the absence of microsomal membranes in the IVT reactions, which facilitate the study of co-translational and initial post-translational processing of proteins *in vitro* [465].

In the IVT reactions performed, the observed VP7 band was smaller (~33 kDa) than the predicted canonical product of 37 kDa (Figure 5.2B and C, lane 8). In cells, the VP7 glycoprotein is synthesised by the ER membrane-bound ribosomes, which co-translationally translocate the nascent polypeptide into the ER as a result of signal sequences present at the N-termini [466]. Cleavage of the signal peptide occurs co-translationally and thus it is plausible that the detected VP7 band is a secondary product first described by Ericson *et al.*, (1983) and Chan *et al.*, (1986) which was shown to lack the N-terminal cleavage signal sequence that retained the VP7 within the ER [72, 73]. Indeed, multiple products may be produced from the VP7-encoding segment by downstream initiation. In the RF strain, three in frame AUGs exist proximal to the 5' end of the mRNA sequence, but only the first and the third are conserved amongst all RVs [47]. The canonical AUG corresponding with the main VP7 product of 37 kDa has a weak Kozak context but the conserved downstream in-frame AUG is located in a favourable context which if utilised would generate a product of ~33 kDa (*Appendix C in section 8.3*). This could provide an alternative explanation for the unexpected size of the observed protein product resulting from translation of the VP7-segment in the IVT reactions (Figure 5.2B and C, lane 8).

Despite NSP4 being a transmembrane protein with the same post-translational modification as VP7 [127], the band corresponding to its main product size was present in RF IVTs but not in the SA11 translations, even though the canonical AUG in both strains has a strong Kozak context (Figure 5.2B and C, lane 11).

Similarly, the protein product of segment 11 (NSP5) was only detected in RF and not in SA11 IVT reactions, although the peptide that was identified migrated higher (over 25 kDa) than the predicted 21 kDa (Figure 5.2B and C, lane 12). NSP5 undergoes two post-translational modifications: phosphorylation and *O*-linked glycosylation [467, 468]. Depending on the degree of phosphorylation, NSP5 can exist as several isoforms with different molecular weights of 26, 28 and 30-34 kDa [141, 468, 469]. Hyper-phosphorylation of NSP5 is required for its interaction with NSP2 to form a lattice structure for viroplasm assembly [135, 301]. Previous studies showed that the 28 kDa isoform was present *in vitro* and in the lysates of RV infected cells [61, 141, 469]. Thus, it is plausible that the identified band of ~28 kDa could be the phosphorylated isoform of NSP5 resulting from autocatalytic kinase activity *in vitro* as described previously [61, 470, 471].

The IVT assay was not sufficiently sensitive to detect the 12 kDa accessory protein, NSP6, from an overlapping ORF in segment 11 (Figure 5.2B and C, lane 12) [45]. This could be explained by the presence of abundant endogenous globin (16 kDa) which distorts the gel, making detection of proteins of less than 16 kDa challenging [465]. Small proteins can also be degraded by proteases or by an ubiquitin-dependent pathway within the rabbit reticulocyte lysate component of the IVT [465, 472]. It is also plausible that the variation in the number of methionine residues affected the expression of the NSP6 protein, which contains only five thereof, however translation of the NSP5 product containing only four was still detectable in RF IVTs (Figure 5.2B, lane 12). The IVT reaction conditions, such as magnesium and potassium chloride concentrations, have been shown to affect the selection of AUG start codons, which may have played a role in failed recognition of the downstream AUG corresponding to NSP6 [442, 473].

Additional treatment with λ-phosphatase could validate the phosphorylated isoform of NSP5. Translation of smaller proteins such as NSP6 may be more efficient using the Wheat Germ Extract System in the future [465]. Human cell-derived IVT systems (HeLa and hybridoma cell extracts) may also be utilised in the future for synthesis of proteins with post and co-translational modifications [474, 475].

In addition to the detection of canonical proteins, unknown polypeptides were also identified in most of RF and SA11 segment translations, consistent with the hypothesis that RVs possess multiple ORFs. Some of these peptides were conserved across strains while others appeared only in RF translations (Figure 5.2B and C, pink and green arrowheads respectively). For example, a product of 50 kDa produced in the reactions with the VP4-encoding plasmid for RF but not SA11 could be explained by the presence of an in-frame

downstream AUG codon in a suboptimal Kozak context (see *Appendix C in section 8.3* and Figure 5.2B, lane 5, green arrowhead). Similarly, RF segment 1 (VP1) that codes for the RdRp contains an in-frame AUG codon located 192 nucleotides downstream of the canonical AUG in a strong Kozak context but is absent in segment 1 of SA11 (*Appendix C in section 8.3*). Thus, the detected band in RF translations of segment 1 but not in SA11, could be the protein product with a predicted size of ~117.5 kDa translated from this downstream AUG (Figure 5.2B, lane 2, green arrowhead,). Additionally, a polypeptide of ~100 kDa of unknown origin was consistently translated in the IVT reactions of RF and SA11 VP1 segments (Figure 5.2B and C, lane 2, black arrowhead), including in non-radioactive IVT reactions containing a pre-charged ε-labelled biotinylated lysine-tRNA complex (data not shown). Due to the consistency of the results, segment 1 was further analysed to investigate whether the unknown polypeptide of ~100 kDa could be a putative accessory protein translated from a downstream AUG.

Following identification of candidate AUGs in RF segment 1 that could explain the production of an unknown band of ~100 kDa, all available RV isolate sequences were analysed to evaluate the conservation of these selected AUG codons (Figure 5.3A and B). Bioinformatic analyses revealed high prevalence of AUGs 12, 14, 20 and 22 (M130, M146, M196 and M213 respectively) as well as presence of a moderate Kozak context that was more than 98% conserved in RV isolates (Figure 5.3B and Figure 5.4). The candidate AUGs were then mutated to determine which AUG codon was responsible for the unknown polypeptide. The mutation of candidate AUGs showed that a combination of two similarly sized polypeptides of ~100 kDa were synthesised from AUG12 and AUG14, likely representing N-terminally truncated isoforms of VP1, suggesting possible internal translation initiation (Figure 5.5).

By what mechanisms the ribosomes accessed the downstream AUG codons to initiate translation at M130 and M146 was not explored in this chapter. However, previously Rainsford *et al.*, (2007) suggested that the expression of NSP6 from segment 11 in species A of RV likely occurred via the mechanism of leaky scanning [62]. Despite the optimal Kozak context around the canonical AUG, instances of leaky scanning mechanisms have been observed when the first AUG codon was close to the 5' end or when a G in position +4 was combined with a U in position +5 [476]. Studies from orthoreoviruses showed that the translation initiation complexes can bypass over 600 nucleotides to access downstream ORFs, demonstrating that the distance scanned by the ribosomes can be considerable [477]. Additionally, isoforms of the phosphoprotein in the rabies virus and the M protein in the vesicular stomatitis virus are expressed via sub-optimal Kozak signalling preceded by a canonical AUG in an optimal context [478, 479]. Indeed, the canonical AUG for the main VP1 protein has strong Kozak

signalling and is located at nucleotide position 19, whereas the downstream AUGs at positions 406 and 454 (M130 and M146 respectively) are in a sub-optimal context and do not possess a U at position +5 (Figure 5.3A). Thus, it is plausible that the expression of the identified band from the M130 and/or the M146 occurred by leaky scanning, which could be tested by ribosomal profiling in the future studies [58].

In contrast to the IVT result, when segment 1 amplicon fused to the EGFP gene was transfected in cells, M130 mutation resulted in the loss of a 37 kDa EGFP-tagged product, whereas M146 was not utilised (Figure 5.7B). In support of this, quantification of the peptides showed that mutation of M130, with or without M146, reduced the expression of the identified polypeptide (Figure 5.7C). This result however contradicted the hypothesis of downstream translation initiation as mutation of an AUG codon to GUG (M130V) had much less of an effect on the expression of identified band in contrast to a CUG mutation (M130L). A comparison of the findings between the cell-free and the cell-based assays is summarised in Figure 5.13.

One possibility to explain this is that the non-canonical initiation was occurring by leaky scanning mechanism resulting in initiation at a GUG but not a CUG codon. Non-AUG initiation by leaky scanning has been described for the expression of the C protein in the Sendai respiratory virus and for RNA viruses infecting plants [449, 480-484]. For example, in the panicum mosaic virus up to four ORFs are translated using a combination of leaky scanning initiation at GUG codons [481]. Interestingly, expression of the C protein in human parainfluenza virus 1 is initiated at a GUG codon with high efficiency [450]. However, downstream non-AUG initiation by leaky scanning utilising a GUG codon in a sub-optimal context has rarely been observed in RNA viruses [59].

Another plausible mechanism of expression is the potential presence of an IRES, a *cis*-acting RNA element, within the RNA structure of segment 1. Evidence from FMDV studies showed presence of two in-frame AUG codons separated by IRES where both AUGs have been utilised for translation initiation [485-488]. Mutation of AUG to CUG at AUG12 diminished protein production, whereas mutation to GUG did not, but under the IRES scenario it is plausible that translation initiation could occur also at a GUG codon. Class IV IRESs, generally around ~150-200 nucleotides long, directly recruit the translation machinery to the start codon and proceed without the need for scanning or an AUG codon [420, 451]. Additionally, translation of two ORFs for the expression of L protein products in Theiler's virus was also dependent on the presence of an IRES where one of the ORFs utilised a non-AUG codon [489]. The mutation of the AUG codon to GUG represents a transition mutation within the nucleotide sequence, which is biochemically a less dramatic change than the AUG to CUG

substitution. The GUG mutation is therefore less likely to impact the RNA structure, and so it follows, the structure of an IRES – providing a possible explanation as to why translation initiation may occur at GUG but not CUG. Furthermore, RV studies showed the presence of conserved *cis*-acting signals within the highly structured RNAs produced from this segment [50]. Additionally, the distance between the canonical AUG and the nucleotide position 406 for M130 could support the formation of an IRES, although initial prediction using online tools could not identify presence of an IRES (data not shown).

It is also possible that the N-truncated VP1 isoform was expressed by a combination of mechanisms such as leaky scanning *and* through presence of an IRES. Future work to investigate the mechanism of expression could involve utilising bicistronic reporter assays to help identify the presence of an IRES [490], whilst leaky scanning could be further explored through modifications of the Kozak motifs surrounding AUG12. Nonetheless, a novel VP1 isoform, named VP1-N129, expressed from M130 was identified although it is unclear how this ORF was accessed by the ribosomes or what its biological relevance is.

To evaluate the effect of mutation of candidate AUGs in the context of viral infection, reverse genetics was used to generate mutant viruses. Interestingly, leucine substitution at position M130 but not valine affected the virus rescue (Figure 5.8). In contrast to leucine and valine, methionine is an unbranched amino acid that allows considerable conformational flexibility, which may play an important role in RdRp stabilisation during its interaction with the RNA template and/or transcript [462]. All three amino acids are hydrophobic, non-polar in nature and are almost always found in the interior of protein molecules [462]. However, leucine differs from valine by the presence of an additional methylene group and a branch point at the γ-carbon [491]. Therefore, the leucine side-chain might have limited the internal flexibility of the HLH subdomain, creating steric hindrance and/or disrupting the stoichiometry of the VP1 protein. It is unclear why mutation to valine but not leucine at this position was tolerated.

The effects of VP1 mutations on the polymerase activity were also examined by establishing a plasmid-based mini-genome assay. Previous studies showed that absence of the aspartate residues in the ‘GDD’ motif abolished or significantly altered the RNA-dependent polymerase activity *in vitro* and *in vivo* [159, 492-494]. Indeed, the D631A and D632A mutations affected viral rescue and reduced the luciferase signal in the mini-genome assay, although the values were not at baseline (Figure 5.10C and D respectively). In contrast to the D631A mutation, D632A showed significantly lower luciferase activity compared to WT RF suggesting that changes to the first aspartic acid residue within the ‘GDD’ motif appeared more tolerable. In other viral RdRps, single or multiple substitutions of the equivalent residues in the

‘GDD’ motif showed that the first aspartate was a strict requirement for the catalytic activity of the polymerase but this depended on the assay conditions used [492, 495, 496]. The preliminary data from the RV polymerase assay developed here does not support the notion that the VP1 mutants affected the polymerase function as their luciferase signal was comparable to that of the WT RF (Figure 5.10D). Further work is needed to establish a more reliable and sensitive assay that can be used to investigate the impact of VP1 mutations on the polymerase activity. The plasmid based mini-genome assay could be modified in the future to determine the minimum number of plasmids required to reconstitute the polymerase activity *in vitro* as demonstrated for influenza virus [464, 497]. Future studies could explore varying the molar ratios of VP1 and VP2 in the mini-genome assay as the molar ratio of 1/11 (VP1 to VP2) produced the highest level of dsRNA synthesis in the cell-free system similar to that found in virion cores [498]. Different or synchronous mutations of the conserved residues in the ‘GDD’ motif, for example GDD → GAA/NN or GDD → A/YDD, could help improve the baseline signal of inactive polymerase in the future [492, 494, 495, 499].

For the mutants generated, it is possible that the packaging of the (+)RNA strand and/or dsRNA synthesis was affected, although this was not explored in this chapter (though this can be analysed by the RNA:PFU ratio calculations and RNA gel methods described in Diebold *et al.*, (2022) [300]. Assembly of the progeny cores is initiated by the interaction of VP1 and VP3 with (+)RNA that forms replication intermediates lacking polymerase activity [49, 108, 498, 500]. VP1 specifically binds the 3’ consensus sequence of the (+)RNA containing *cis*-acting signals but requires VP2 to become transcriptionally active [49, 498, 501, 502]. The *cis*-acting elements important for packaging and translation initiation have been identified in the 5’ UTR leader regions and coding sequences of influenza A and B viruses [441, 503, 504]. In RVs, highly conserved *cis*-acting elements that enhance (-)RNA synthesis were also shown to be present at the 5’-end of (+)RNA which sometimes extended into the coding region [49, 295, 501, 505, 506]. Base-pairing between the 5’ and 3’ in *cis* enhancement signals of (+)RNAs lead to circularisation and formation of panhandle structures [49, 501]. Thus, introduced mutations that result in a failure of the 3’ consensus sequence to base-pair with the 5’ terminus inhibit synthesis of dsRNA [49, 507]. It is possible that the leucine mutation at M130 (M130L) disrupted the secondary RNA structure and blocked VP1 from recognising the *cis* elements that may be present near the 5’ end leading to inhibition of dsRNA synthesis (whereas this was tolerated with valine mutation). Indeed, this could also explain why attempts to create a recombinant RV in which a FLAG epitope was fused to the C-terminus of VP1 was unsuccessful (data not shown).

The M130L may have disrupted the molecular interaction or the flexibility with which VP1 was able to interact with VP2 and/or VP3. Earlier studies with RV temperature sensitive mutants showed that assembly of VP2 into cores was required for RNA replication [508]. Furthermore, the N-terminus of VP2 was required for encapsidation of VP1 and VP3 although its function in RNA transcription is not known [158]. Interestingly, Kumar *et al.*, (2020) showed that VP3 was not detected in the DLP reconstructions *in situ* and it is still unclear how VP3 coordinates capping inside the constrained particle [110]. Steger *et al.*, (2019) and Estrozi *et al.*, (2013) mapped amino acids 264 – 267 within the N-terminal domain of VP1 as important contact sites for interaction with VP2 for *in vitro* dsRNA synthesis but these were not the residues mutated here (Table 14) [28, 157]. Other studies also raised the possibility that the VP1-VP2 interactions might differ during transcription and replication as well as being strain dependent [509]. Thus, future studies involving characterisation of protein interactions through co-immunoprecipitation or pull-down assays could help provide insight into the effects of mutation in VP1 on its interaction with VP2 and/or VP3.

To assess whether the protein folding may have been impacted by the introduced amino acid substitutions at M130 residue, AlphaFold-based protein folding predictions were made with the assistance of Dr Barbara Shih, Lancaster University. Within the crystal structure of VP1, the M130 residue was found within the cap-binding site of the N-terminal domain that forms an α-helix motif (Figure 5.11A and C) [107]. *In situ* structural studies showed that this cap-binding site is present in RdRps of other members of the *Sedoreoviridae* family [160, 510, 511]. The N-terminal domain splits the genomic dsRNA through its interaction with the conserved m7GpppGGC residue of the (+)RNA strand at the cap-binding site [105, 107, 108]. This results in the HLH subdomain (residues 39 – 69) further separating the dsRNA to allow RNA transcription while the template RNA is directed back into the core interior [105, 107]. The isolated (-)RNA is then transcribed within the core domain of RdRp with complementary NTPs that form a backbone with the 5’ end of the nascent RNA [105, 107, 108]. The cap-binding site was also shown to accommodate the NTP molecules during transcription [107]. The neighbouring HLH subdomain near the cap-binding site undergoes conformational changes to accommodate the interaction between the polymerase core and the RNA template and/or transcript [107]. However, the structure predictions of the M130 mutant showed no major differences in the overall molecule of RdRp compared to the WT and no intra molecular interactions with neighbouring residues were predicted to be disturbed (Figure 5.12A). *In vitro* dsRNA synthesis assays using open cores or baculovirus-expressed recombinant proteins could address the effects of VP1 mutation on RNA synthesis and/or replication in the future [157].

Overall, this chapter demonstrated that in segment 1 (VP1), an in-frame downstream AUG at nucleotide position 406 was responsible for the expression of a novel N-truncated VP1 isoform – VP1-N129. In cells, its expression was decreased by M130L mutation alone or in combination with M146L, but M130V mutation did not exert this effect. Similarly, in the context of a viral infection, only leucine substitution at M130 affected the viral rescue, whereas valine did not. The effects of leucine replacement at M130 on VP1 could be attributed to non-exclusive reasons below:

1) Translation initiation at a downstream AUG was disrupted, either by affecting propensity of ribosomal leaky scanning, the utility of an IRES, or a combination thereof.
2) Viral polymerase activity was impacted.
3) Genome packaging was disrupted.
4) Core particle assembly was impaired through VP1 interactions with VP2, VP3 and/or viral RNAs.

Given the findings presented in this chapter, the priorities for future investigations are (1) and (2).

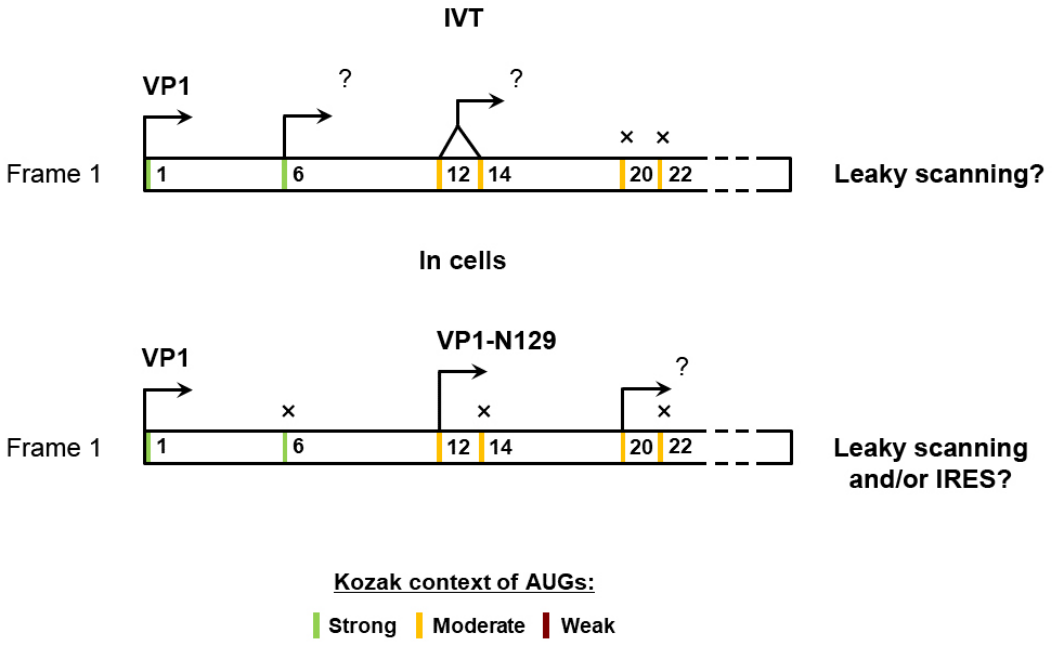


Figure 5.13 Possible mechanisms for expression of VP1 products.
Top panel – in the IVTs evidence of expression exists for the main product VP1 and for the product expressed by AUG6 (black arrows). Mutation of AUG12 and AUG14 together decreased the expression of the identified band of ~100 kDa (joined arrow with question mark). Both AUGs were utilised for its expression possibly via the mechanism of leaky scanning. No effect from mutating AUG20 and/or AUG22 was observed (black crosses).
Bottom panel – in cells, expression of the main product VP1 was detected as well as the novel VP1 isoform VP1-N129 expressed by AUG12 (black arrows). No evidence of expression of peptides from AUG6, AUG14 and AUG22 was observed (black crosses). Evidence for peptide expression from AUG20 was observed but not confirmed (black arrow with question mark). The mechanism of expression of VP1-N129 was not examined but could involve presence of an IRES and/or leaky scanning.

Chapter 6: Concluding remarks

6.1 General conclusion

In this project, a plasmid-based reverse genetics system was established for the bovine RV strain RF with the aim to identify expression of potential accessory proteins encoded by its genome.

The data presented in this study describe identification of an unknown polypeptide of ~100 kDa in the RF segment 1 (VP1). Based on the predicted size of the peptide, bioinformatic analyses revealed candidate AUGs with high prevalence and conservation that were selected for further investigation. In the cell free system, mutation of both AUG12 (M130) and AUG14 (M146) decreased the expression of the unknown polypeptide, suggesting that both AUGs were utilised for downstream translation initiation.

Production of additional polypeptides was also tested in cells transfected with VP1-EGFP tagged constructs, which revealed an alternative hypothesis. In contrast to the IVTs, only M130 (AUG12) was utilised for the expression of the unknown polypeptide of ~100 kDa. Thus, a novel VP1 isoform named VP1-N129 was identified. Expression of additional polypeptides by either M66, M146 or M213 was not detected in cells.

The effects of mutation were examined in the context of RV infection by generating a panel of mutant viruses using reverse genetics. Interestingly, only the leucine substitution at M130 residue (M130L) resulted in no virus production, whereas the valine mutation (M130V) had a similar phenotype to the WT virus. Mutation of M146 had no effect on the viral rescue nor on replication. To examine the function of the VP1 mutants, their effect on the viral polymerase activity was investigated by attempting to establish a mini-genome assay. Preliminary results suggest that the VP1 mutations had no effect on the polymerase function, although further experiments are needed to improve the sensitivity of the assay. Additionally, predicted protein structures of M130 and M146 mutants did not show any major differences in the protein folding compared to the WT RF.

6.2 Future work and directions

6.2.1 Identification of potential function and mechanism of expression

Data presented in *Chapter 5* showed that additional polypeptides were being expressed *in vitro* but whether these products provide any functional benefit for the virus still requires further investigation. It is plausible that during viral infection, downstream alternative translation initiation that results in the shorter VP1 isoforms acts as a decoy mechanism to either antagonise the innate immune response or control viral gene expression. In future studies, their potential effect on the antiviral response could be examined through a type I interferon assay. Rescuing the VP1 mutants in cell lines expressing WT VP1 could help elucidate whether the absent signals in the truncated VP1 isoforms are important for virus replication. Additionally, peptides of interest could be gel extracted and further characterised by mass spectrometry.

Ribosome profiling could help identify whether the downstream AUGs are being utilised for translation initiation [459]. Furthermore, generation of recombinant baculoviruses expressing the VP1 mutants could provide insight into their effect on dsRNA synthesis *in vitro* [157]. The mini-genome assay can be further modified to identify the minimum number of plasmids required for transfection in order to improve the sensitivity of the assay. Additionally, different mutations of the conserved residues in the ‘GDD’ motif may be more effective at inactivating the polymerase activity in this assay, thus generating a more reliable negative control. Potential mechanisms of downstream translation initiation could be assessed by modifying the Kozak strength around the initiating AUG codons.

The presence of an IRES can be investigated by utilising a bicistronic reporter assay [512-514]. For example, to demonstrate IRES activity, the sequence upstream of VP1-N129 containing putative IRES could be cloned in-frame between two cistrons, which can be any fluorescent reporters such as Renilla and firefly luciferases (Figure 6.1) [513, 514]. The test sequence can either include or exclude the upstream AUG6 (M66), although both options can be trialled (shaded in grey). Translation of the first cistron will be 5’ cap-dependent whereas the expression of the second cistron may be IRES-dependent, providing proof for internal ribosome loading onto the intercistronic sequence (Figure 6.1) [512, 514]. Thus, a sequence would qualify as an IRES if there is a significant increase in the Cistron-2/Cistron-1 expression ratio. Potential read-through could be inhibited by introducing a stop codon after the first

cistron and an empty bicistronic reporter (control) could be used for comparison and establishing baselines [512, 514].

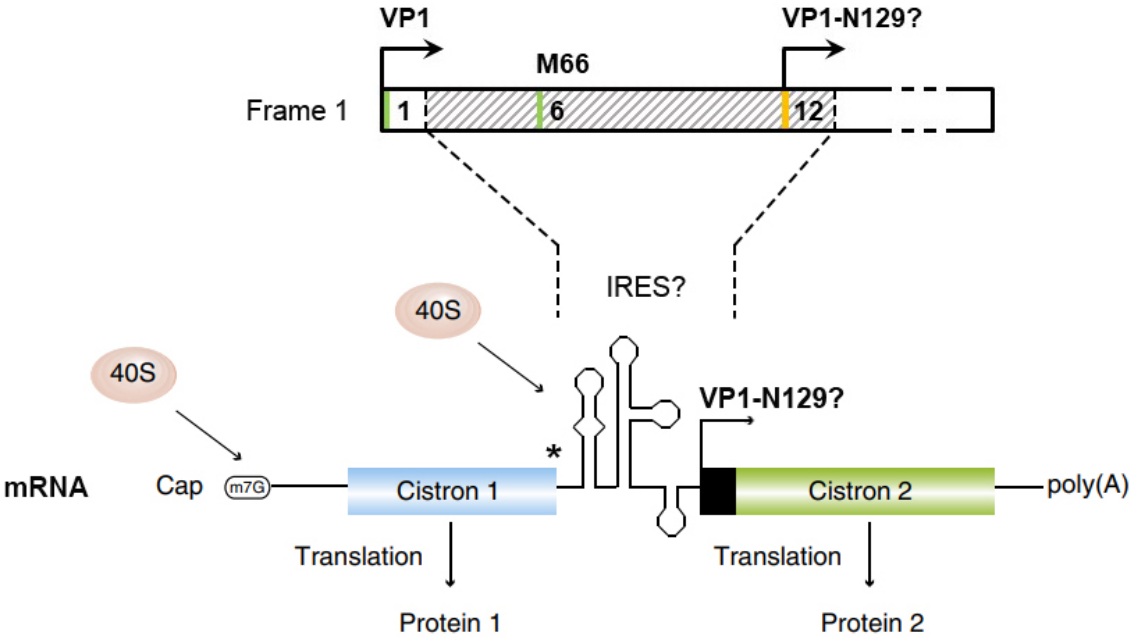


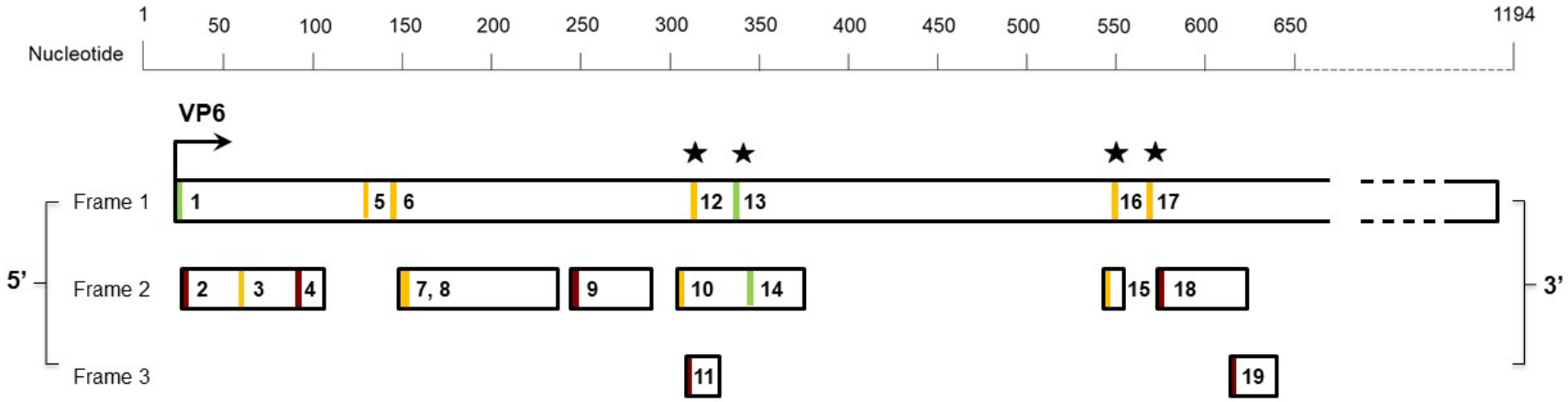
Figure 6.1 Schematic representation of the bicistronic reporter construct.
Simplified schematic of VP1 ORF from **Figure 5.3A** with predicted protein products. Positions of AUG codons are indicated. Sequence containing putative IRES upstream of VP1-N129 is placed in-frame between two cistrons. Shaded area represents test sequence that can either include or exclude AUG6 (M66). In mRNA transcribed from the bicistronic construct, the first cistron (blue) is translated by 5' cap-dependent scanning mechanism while translation of the second cistron (green) depends on the presence of an IRES in the intercistronic sequence. For diagram simplicity, only the 40S ribosomal subunit is shown, other proteins known to participate in these processes have been omitted. Asterisk denotes a stop codon. *(Adapted from Sonenberg et al., 2016 [513]).*

6.2.2 Analysis of accessory protein expression by other segments

The presence of alternative translation initiation in the RV genome was tested using the IVT reactions that enabled detection of synthesised polypeptides some of which contained radiolabelled ^{35}S-methionine residues. The majority of RV segments showed bands migrating at sizes corresponding to the main protein products, with some segments producing additional polypeptides conserved in translations of RF and SA11 segments (Figure 5.2B and C).

Following the IVT reactions, preliminary data was generated from further analysis of segment 6 (VP6) of the bovine RV strain RF that identified potential AUGs that could be responsible for the detected unknown polypeptides (Chapter 5, Figure 5.2B, lane 7, pink arrowheads). Bioinformatic analysis of the first 600 nucleotides of segment 6 showed presence of downstream AUG codons in all three frames, some of which had a strong Kozak context (Figure 6.2A). Based on predicted sizes of additional polypeptides hypothetically derived from internal translation initiation, four AUG candidates: AUG12, AUG13, AUG16 and AUG17, were selected from frame 1 for downstream experiments (Figure 6.2A, black stars and table). Subsequently, site-directed mutagenesis of candidate AUGs followed by the IVT reactions revealed that mutation of AUG13 (M100) and AUG17 (M180) reduced the expression of additional unknown polypeptides (Figure 6.2B, pink arrowheads). Based on this initial result, further experiments could involve expressing VP6 truncated products in cells due to the availability of a commercial antibody targeting this protein. Additionally, future work could evaluate the conservation of identified AUG codons and their Kozak context across clinical isolates to determine their importance in viral pathogenicity.

Similarly, potential downstream translation initiation in other segments could be further investigated through site-directed mutagenesis as well as applying the R script presented in this work. Furthermore, expression of candidate accessory proteins in other segments could be investigated through introduction of N- or C-terminal tags. The effects of accessory proteins on viral pathogenesis could reveal innovative control mechanisms allowing improvements to be made to the current vaccine designs as well as increase our understanding of RV biology.



Kozak context of AUGs:

Strong | Moderate | Weak

AUG #	Nucleotide #	Nucleotide change	VP6 mutation
1	24	-	-
12	312	A312C	M97L
13	321	A321C	M100L
16	552	A552C	M177L
17	561	A561C	M180L

Figure 6.2 Analysis of AUG codons in RV segment 6 (VP6).
Schematic of the first 650 nucleotides of VP6 mRNA ORF showing AUG codons in all reading frames. AUG codons are numbered according to their position and coloured according to the strength of their Kozak context. Green – strong with A/G at -3 and G at +4; yellow – medium with either A/G at -3 or G at +4; red – weak neither A/G at -3 nor G at +4. Length of the boxes correspond to ORF length. Stars indicate candidate AUGs that may serve as translation initiation sites for the unknown polypeptides identified in IVT reactions of VP6 (**see Chapter 5, Figure 5.2B, lane 7**). Box shows a summary of AUG candidates with their nucleotide positions and corresponding mutations in the VP6 protein.

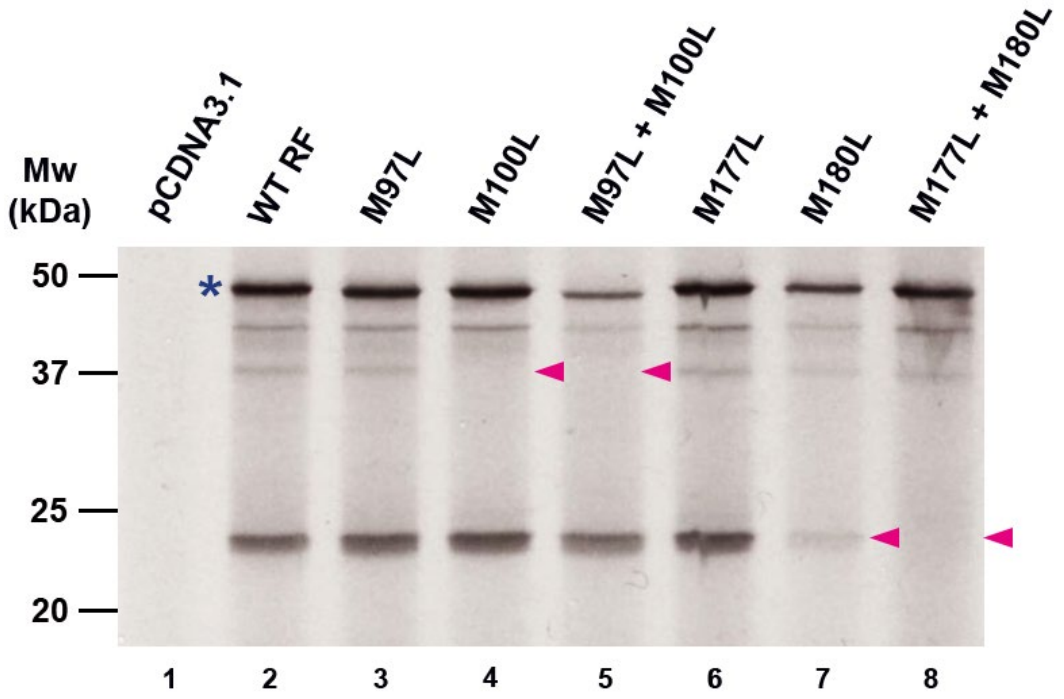


Figure 6.3 Site-directed mutagenesis of candidate AUGs in segment 6 (VP6).
IVT reactions were carried out as described in **section 5.2.1** and proteins were separated by SDS-PAGE and detected by autoradiography. Empty pCDNA 3.1 vector was used as a negative control. The positions of molecular weight markers are indicated (kDa). The blue asterisk shows the canonical peptide (44 kDa) and pink arrowheads highlight instances of a decrease in translation of unknown polypeptide species.

Chapter 7: References

1. Troeger, C., et al., *Rotavirus Vaccination and the Global Burden of Rotavirus Diarrhea Among Children Younger Than 5 Years.* JAMA pediatrics, 2018. **172**(10): p. 958-965.
2. Bishop, R.F., et al., *Virus particles in epithelial cells of duodenal mucosa from children with acute non-bacterial gastroenteritis.* Lancet, 1973. **2**(7841): p. 1281-3.
3. Adams, W.R. and L.M. Kraft, *Epizootic diarrhea of infant mice: indentification of the etiologic agent.* Science, 1963. **141**(3578): p. 359-60.
4. Malherbe, H.H. and M. Strickland-Cholmley, *Simian virus SA11 and the related O agent.* Arch Gesamte Virusforsch, 1967. **22**(1): p. 235-45.
5. Mebus, C.A., Underdahl, N. R., Rhodes, M. B. and Twiehaus, M. J., *"Calf Diarrhea (Scours): Reproduced with a Virus from a Field Outbreak".* Historical Research Bulletins of the Nebraska Agricultural Experiment Station (1913-1993). 1969(69).
6. Miyabe, F.M., et al., *Porcine rotavirus B as primary causative agent of diarrhea outbreaks in newborn piglets.* Scientific Reports, 2020. **10**(1): p. 22002.
7. Uddin Ahmed, N., et al., *Risk factors for bovine rotavirus infection and genotyping of bovine rotavirus in diarrheic calves in Bangladesh.* PLOS ONE, 2022. **17**(2): p. e0264577.
8. Papp, H., et al., *Review of group A rotavirus strains reported in swine and cattle.* Vet Microbiol, 2013. **165**(3-4): p. 190-9.
9. Urie, N.J., et al., *Preweaned heifer management on US dairy operations: Part V. Factors associated with morbidity and mortality in preweaned dairy heifer calves.* J Dairy Sci, 2018. **101**(10): p. 9229-9244.
10. Crawford, S.E., Ding, S.D., Greenberg, H.B, Estes, M.K., *Rotaviruses*, in *Fields Virology: RNA Viruses*, D.M.K. Peter M. Howley, Editor. 2023.
11. Matthijnssens, J., et al., *VP6-sequence-based cutoff values as a criterion for rotavirus species demarcation.* Archives of Virology, 2012. **157**(6): p. 1177-1182.
12. Mihalov-Kovács, E., et al., *Candidate new rotavirus species in sheltered dogs, Hungary.* Emerg Infect Dis, 2015. **21**(4): p. 660-3.
13. Bányai, K., et al., *Candidate new rotavirus species in Schreiber's bats, Serbia.* Infection, Genetics and Evolution, 2017. **48**: p. 19-26.
14. Johne, R., et al., *Genome analysis of the novel putative rotavirus species K.* Virus Research, 2023. **334**: p. 199171.
15. Johne, R., et al., *Whole Genome Sequence Analysis of a Prototype Strain of the Novel Putative Rotavirus Species L.* Viruses, 2022. **14**(3).
16. Matthijnssens, J., et al., *Recommendations for the classification of group A rotaviruses using all 11 genomic RNA segments.* Arch Virol, 2008. **153**(8): p. 1621-9.
17. Matthijnssens, J. *Rotavirus Classification Working Group: RCWG*. April 2023; Available from: https://rega.kuleuven.be/cev/viralmetagenomics/virus-classification/rcwg.
18. Matthijnssens, J., et al., *Full genome-based classification of rotaviruses reveals a common origin between human Wa-Like and porcine rotavirus strains and human DS-1-like and bovine rotavirus strains.* J Virol, 2008. **82**(7): p. 3204-19.
19. Matthijnssens, J., et al., *Uniformity of rotavirus strain nomenclature proposed by the Rotavirus Classification Working Group (RCWG).* Arch Virol, 2011. **156**(8): p. 1397-413.
20. Scherrer, R., et al., *Reovirus-like agent (rotavirus) associated with neonatal calf gastroenteritis in France.* Ann Rech Vet, 1976. **7**(1): p. 25-31.
21. Crawford, S.E., et al., *Rotavirus infection.* Nat Rev Dis Primers, 2017. **3**: p. 17083.
22. Li, Z., et al., *Rotavirus architecture at subnanometer resolution.* J Virol, 2009. **83**(4): p. 1754-66.

23. Settembre, E.C., et al., *Atomic model of an infectious rotavirus particle.* EMBO J, 2011. **30**(2): p. 408-16.
24. Flewett, T.H., et al., *Relation between viruses from acute gastroenteritis of children and newborn calves.* Lancet, 1974. **2**(7872): p. 61-3.
25. Lawton, J.A., M.K. Estes, and B.V. Prasad, *Three-dimensional visualization of mRNA release from actively transcribing rotavirus particles.* Nat Struct Biol, 1997. **4**(2): p. 118-21.
26. Prasad, B.V.V., et al., *Three-dimensional structure of rotavirus.* Journal of Molecular Biology, 1988. **199**(2): p. 269-275.
27. McClain, B., et al., *X-ray crystal structure of the rotavirus inner capsid particle at 3.8 A resolution.* J Mol Biol, 2010. **397**(2): p. 587-99.
28. Estrozi, L.F., et al., *Location of the dsRNA-dependent polymerase, VP1, in rotavirus particles.* J Mol Biol, 2013. **425**(1): p. 124-32.
29. Prasad, B.V.V., et al., *Visualization of ordered genomic RNA and localization of transcriptional complexes in rotavirus.* Nature, 1996. **382**(6590): p. 471-473.
30. Guglielmi, K.M., S.M. McDonald, and J.T. Patton, *Mechanism of Intraparticle Synthesis of the Rotavirus Double-stranded RNA Genome*.* Journal of Biological Chemistry, 2010. **285**(24): p. 18123-18128.
31. Grimes, J.M., et al., *The atomic structure of the bluetongue virus core.* Nature, 1998. **395**(6701): p. 470-478.
32. Reinisch, K.M., M.L. Nibert, and S.C. Harrison, *Structure of the reovirus core at 3.6?Å resolution.* Nature, 2000. **404**(6781): p. 960-967.
33. Shaw, A.L., et al., *The structure of aquareovirus shows how the different geometries of the two layers of the capsid are reconciled to provide symmetrical interactions and stabilization.* Structure, 1996. **4**(8): p. 957-967.
34. Hill, C.L., et al., *The structure of a cypovirus and the functional organization of dsRNA viruses.* Nature Structural Biology, 1999. **6**(6): p. 565-568.
35. Mathieu, M., et al., *Atomic structure of the major capsid protein of rotavirus: implications for the architecture of the virion.* Embo j, 2001. **20**(7): p. 1485-97.
36. Trask, S.D., S.M. McDonald, and J.T. Patton, *Structural insights into the coupling of virion assembly and rotavirus replication.* Nat Rev Microbiol, 2012. **10**(3): p. 165-77.
37. Dormitzer, P.R., et al., *Structural rearrangements in the membrane penetration protein of a non-enveloped virus.* Nature, 2004. **430**(7003): p. 1053-1058.
38. Herrmann, T., et al., *Functional refolding of the penetration protein on a non-enveloped virus.* Nature, 2021. **590**(7847): p. 666-670.
39. Estes, M.K., D.Y. Graham, and B.B. Mason, *Proteolytic enhancement of rotavirus infectivity: molecular mechanisms.* J Virol, 1981. **39**(3): p. 879-88.
40. Crawford, S.E., et al., *Trypsin cleavage stabilizes the rotavirus VP4 spike.* Journal of virology, 2001. **75**(13): p. 6052-6061.
41. Ruggeri, F.M. and H.B. Greenberg, *Antibodies to the trypsin cleavage peptide VP8 neutralize rotavirus by inhibiting binding of virions to target cells in culture.* J Virol, 1991. **65**(5): p. 2211-9.
42. Feng, N., et al., *Human VP8* mAbs neutralize rotavirus selectively in human intestinal epithelial cells.* J Clin Invest, 2019. **130**: p. 3839-3851.
43. Park, W.-J., et al., *Rotavirus spike protein ΔVP8* as a novel carrier protein for conjugate vaccine platform with demonstrated antigenic potential for use as bivalent vaccine.* Scientific Reports, 2021. **11**(1): p. 22037.
44. Tan, M., et al., *Norovirus P particle, a novel platform for vaccine development and antibody production.* Journal of virology, 2011. **85**(2): p. 753-764.
45. Mattion, N.M., et al., *Expression of rotavirus proteins encoded by alternative open reading frames of genome segment 11.* Virology, 1991. **181**(1): p. 295-304.
46. Desselberger, U., *Rotaviruses.* Virus Res, 2014. **190**: p. 75-96.
47. Poncet, D., *II, 5. Translation of rotavirus mRNAs in the infected cell*, in *Perspectives in Medical Virology*. 2003, Elsevier. p. 185-205.

48. Patton, J.T. and E. Spencer, *Genome replication and packaging of segmented double-stranded RNA viruses.* Virology, 2000. **277**(2): p. 217-25.
49. Tortorici, M.A., B.A. Shapiro, and J.T. Patton, *A base-specific recognition signal in the 5' consensus sequence of rotavirus plus-strand RNAs promotes replication of the double-stranded RNA genome segments.* Rna, 2006. **12**(1): p. 133-46.
50. Li, W., et al., *Genomic analysis of codon, sequence and structural conservation with selective biochemical-structure mapping reveals highly conserved and dynamic structures in rotavirus RNAs with potential cis-acting functions.* Nucleic Acids Res, 2010. **38**(21): p. 7718-35.
51. Chen, D. and J.T. Patton, *Rotavirus RNA replication requires a single-stranded 3' end for efficient minus-strand synthesis.* J Virol, 1998. **72**(9): p. 7387-96.
52. Dermody, T.S., Parker, J. S. L., Sherry, B., *Orthoreoviruses*, in *Fields Virology: RNA Viruses*, P. Howley, Knipe, D., Whelan, S. and Freed, E. , Editor. 2022.
53. Roy, P., *Orbiviruses*, in *Fields Virology: RNA Viruses*, P. Howley, Knipe, D., Whelan, S. and Freed, E. , Editor. 2022.
54. Bae, S.H., et al., *Structural features of an influenza virus promoter and their implications for viral RNA synthesis.* Proc Natl Acad Sci U S A, 2001. **98**(19): p. 10602-7.
55. Barr, J.N. and G.W. Wertz, *Bunyamwera bunyavirus RNA synthesis requires cooperation of 3'- and 5'-terminal sequences.* J Virol, 2004. **78**(3): p. 1129-38.
56. Coffin, J.M., S.H. Hughes, and H.E. Varmus, *Retroviruses*, in *Retroviruses*. 1997, Cold Spring Harbor Laboratory Press
Copyright © 1997, Cold Spring Harbor Laboratory Press.: Cold Spring Harbor (NY).
57. Hussain, S., et al., *Mutation of Influenza A Virus PA-X Decreases Pathogenicity in Chicken Embryos and Can Increase the Yield of Reassortant Candidate Vaccine Viruses.* J Virol, 2019. **93**(2).
58. Lulla, V. and A.E. Firth, *A hidden gene in astroviruses encodes a viroporin.* Nature Communications, 2020. **11**(1): p. 4070.
59. Firth, A.E. and I. Brierley, *Non-canonical translation in RNA viruses.* J Gen Virol, 2012. **93**(Pt 7): p. 1385-409.
60. Rainsford, E., *Functional Studies on the Rotavirus Non-Structural Proteins NSP5 and NSP6*, in *Department Biological Sciences*. 2005, University of Warwick: Coventry, United Kingdom. p. 305.
61. Torres-Vega, M.A., et al., *The C-terminal domain of rotavirus NSP5 is essential for its multimerization, hyperphosphorylation and interaction with NSP6.* J Gen Virol, 2000. **81**(Pt 3): p. 821-30.
62. Rainsford, E.W. and M.A. McCrae, *Characterization of the NSP6 protein product of rotavirus gene 11.* Virus Res, 2007. **130**(1-2): p. 193-201.
63. Holloway, G., et al., *Rotavirus NSP6 localizes to mitochondria via a predicted N-terminal a-helix.* J Gen Virol, 2015. **96**(12): p. 3519-24.
64. Gonzalez, R.A., et al., *In vivo interactions among rotavirus nonstructural proteins.* Arch Virol, 1998. **143**(5): p. 981-96.
65. Komoto, S., et al., *Reverse Genetics System Demonstrates that Rotavirus Nonstructural Protein NSP6 Is Not Essential for Viral Replication in Cell Culture.* J Virol, 2017. **91**(21).
66. Fukuda, S., et al., *Rotavirus incapable of NSP6 expression can cause diarrhea in suckling mice.* 2022. **103**(5).
67. Kojima, K., et al., *Sequence analysis of normal and rearranged NSP5 genes from human rotavirus strains isolated in nature: implications for the occurrence of the rearrangement at the step of plus strand synthesis.* Virology, 1996. **224**(2): p. 446-52.
68. Wu, H., et al., *Serological and genomic characterization of human rotaviruses detected in China.* J Med Virol, 1998. **55**(2): p. 168-76.
69. Gorziglia, M., K. Nishikawa, and N. Fukuhara, *Evidence of duplication and deletion in super short segment 11 of rabbit rotavirus Alabama strain.* Virology, 1989. **170**(2): p. 587-90.

70. González, S.A. and O.R. Burrone, *Porcine OSU rotavirus segment II sequence shows common features with the viral gene of human origin.* Nucleic Acids Res, 1989. **17**(15): p. 6402.
71. Trojnar, E., et al., *The genome segments of a group D rotavirus possess group A-like conserved termini but encode group-specific proteins.* J Virol, 2010. **84**(19): p. 10254-65.
72. Chan, W.K., et al., *Two glycoproteins are produced from the rotavirus neutralization gene.* Virology, 1986. **151**(2): p. 243-52.
73. Ericson, B.L., et al., *Two types of glycoprotein precursors are produced by the simian rotavirus SA11.* Virology, 1983. **127**(2): p. 320-32.
74. Eiden, J.J., *Expression and sequence analysis of gene 7 of the IDIR agent (group B rotavirus): similarity with NS53 of group A rotavirus.* Virology, 1994. **199**(1): p. 212-8.
75. Diller, J.R., et al., *Rotavirus Species B Encodes a Functional Fusion-Associated Small Transmembrane Protein.* J Virol, 2019. **93**(20).
76. Langland, J.O., et al., *Products of the porcine group C rotavirus NSP3 gene bind specifically to double-stranded RNA and inhibit activation of the interferon-induced protein kinase PKR.* J Virol, 1994. **68**(6): p. 3821-9.
77. Luke, G.A., et al., *Occurrence, function and evolutionary origins of '2A-like' sequences in virus genomes.* J Gen Virol, 2008. **89**(Pt 4): p. 1036-42.
78. Piron, M., et al., *Identification of the RNA-binding, dimerization, and eIF4GI-binding domains of rotavirus nonstructural protein NSP3.* J Virol, 1999. **73**(7): p. 5411-21.
79. Ito, H., et al., *Complete nucleotide sequence of a group A avian rotavirus genome and a comparison with its counterparts of mammalian rotaviruses.* Virus Research, 2001. **75**(2): p. 123-138.
80. Gault, E., et al., *A human rotavirus with rearranged genes 7 and 11 encodes a modified NSP3 protein and suggests an additional mechanism for gene rearrangement.* J Virol, 2001. **75**(16): p. 7305-14.
81. Poncet, D., S. Laurent, and J. Cohen, *Four nucleotides are the minimal requirement for RNA recognition by rotavirus non-structural protein NSP3.* Embo j, 1994. **13**(17): p. 4165-73.
82. Arnold, M., J.T. Patton, and S.M. McDonald, *Culturing, storage, and quantification of rotaviruses.* Current protocols in microbiology, 2009. **Chapter 15**: p. Unit-15C.3.
83. Isa, P., C.F. Arias, and S. López, *Role of sialic acids in rotavirus infection.* Glycoconjugate Journal, 2006. **23**(1): p. 27-37.
84. Amimo, J.O., et al., *Rotavirus Interactions With Host Intestinal Epithelial Cells.* Frontiers in immunology, 2021. **12**: p. 793841-793841.
85. López, S. and C.F. Arias, *Multistep entry of rotavirus into cells: a Versaillesque dance.* Trends Microbiol, 2004. **12**(6): p. 271-8.
86. Dormitzer, P.R., et al., *The rhesus rotavirus VP4 sialic acid binding domain has a galectin fold with a novel carbohydrate binding site.* The EMBO journal, 2002. **21**(5): p. 885-897.
87. Martínez, M.A., et al., *Gangliosides Have a Functional Role during Rotavirus Cell Entry.* 2013. **87**(2): p. 1115-1122.
88. Imbert-Marcille, B.M., et al., *A FUT2 gene common polymorphism determines resistance to rotavirus A of the P[8] genotype.* J Infect Dis, 2014. **209**(8): p. 1227-30.
89. Saxena, K., et al., *Human Intestinal Enteroids: a New Model To Study Human Rotavirus Infection, Host Restriction, and Pathophysiology.* 2016. **90**(1): p. 43-56.
90. Payne, D.C., et al., *Epidemiologic Association Between FUT2 Secretor Status and Severe Rotavirus Gastroenteritis in Children in the United States.* JAMA Pediatr, 2015. **169**(11): p. 1040-5.
91. Tan, M. and X. Jiang, *Histo-blood group antigens: a common niche for norovirus and rotavirus.* Expert Reviews in Molecular Medicine, 2014. **16**: p. e5.
92. Le Pendu, J., K. Nyström, and N. Ruvoën-Clouet, *Host-pathogen co-evolution and glycan interactions.* Curr Opin Virol, 2014. **7**: p. 88-94.

93. Guerrero, C.A., et al., *Heat Shock Cognate Protein 70 Is Involved in Rotavirus Cell Entry.* 2002. **76**(8): p. 4096-4102.
94. Isa, P., et al., *Rotavirus RRV associates with lipid membrane microdomains during cell entry.* Virology, 2004. **322**(2): p. 370-81.
95. Hewish, M.J., Y. Takada, and B.S. Coulson, *Integrins α2β1 and α4β1 Can Mediate SA11 Rotavirus Attachment and Entry into Cells.* 2000. **74**(1): p. 228-236.
96. Zárate, S., et al., *Interaction of rotaviruses with Hsc70 during cell entry is mediated by VP5.* J Virol, 2003. **77**(13): p. 7254-60.
97. Gualtero, D.F., et al., *Amino acid domains 280–297 of VP6 and 531–554 of VP4 are implicated in heat shock cognate protein hsc70-mediated rotavirus infection.* Archives of Virology, 2007. **152**(12): p. 2183-2196.
98. Abdelhakim, A.H., et al., *Structural Correlates of Rotavirus Cell Entry.* PLOS Pathogens, 2014. **10**(9): p. e1004355.
99. Kim, I.S., et al., *Effect of mutations in VP5 hydrophobic loops on rotavirus cell entry.* J Virol, 2010. **84**(12): p. 6200-7.
100. Yoder, J.D., et al., *VP5* rearranges when rotavirus uncoats.* J Virol, 2009. **83**(21): p. 11372-7.
101. Harrison, S.C., *Viral membrane fusion.* Nature Structural & Molecular Biology, 2008. **15**(7): p. 690-698.
102. Gutiérrez, M., et al., *Different rotavirus strains enter MA104 cells through different endocytic pathways: the role of clathrin-mediated endocytosis.* J Virol, 2010. **84**(18): p. 9161-9.
103. Salgado, E.N., et al., *Visualization of Calcium Ion Loss from Rotavirus during Cell Entry.* J Virol, 2018. **92**(24).
104. Periz, J., et al., *Rotavirus mRNAS are released by transcript-specific channels in the double-layered viral capsid.* Proc Natl Acad Sci U S A, 2013. **110**(29): p. 12042-7.
105. Jenni, S., et al., *In situ Structure of Rotavirus VP1 RNA-Dependent RNA Polymerase.* J Mol Biol, 2019. **431**(17): p. 3124-3138.
106. Zhang, X., et al., *In situ structures of the segmented genome and RNA polymerase complex inside a dsRNA virus.* Nature, 2015. **527**(7579): p. 531-534.
107. Ding, K., et al., *In situ structures of rotavirus polymerase in action and mechanism of mRNA transcription and release.* Nature Communications, 2019. **10**(1): p. 2216.
108. Lu, X., et al., *Mechanism for coordinated RNA packaging and genome replication by rotavirus polymerase VP1.* Structure, 2008. **16**(11): p. 1678-88.
109. Chen, D., et al., *Rotavirus Open Cores Catalyze 5'-Capping and Methylation of Exogenous RNA: Evidence That VP3 Is a Methyltransferase.* Virology, 1999. **265**(1): p. 120-130.
110. Kumar, D., et al., *2.7 Å cryo-EM structure of rotavirus core protein VP3, a unique capping machine with a helicase activity.* Sci Adv, 2020. **6**(16): p. eaay6410.
111. Lawton, J.A., M.K. Estes, and B.V. Prasad, *Mechanism of genome transcription in segmented dsRNA viruses.* Adv Virus Res, 2000. **55**: p. 185-229.
112. Lawton, J.A., M.K. Estes, and B.V. Prasad, *Identification and characterization of a transcription pause site in rotavirus.* J Virol, 2001. **75**(4): p. 1632-42.
113. Kohli, E., et al., *Inhibition of in vitro reconstitution of rotavirus transcriptionally active particles by anti-VP6 monoclonal antibodies.* Arch Virol, 1994. **135**(1-2): p. 193-200.
114. Feng, N., et al., *Inhibition of rotavirus replication by a non-neutralizing, rotavirus VP6-specific IgA mAb.* J Clin Invest, 2002. **109**(9): p. 1203-13.
115. Vende, P., et al., *Efficient translation of rotavirus mRNA requires simultaneous interaction of NSP3 with the eukaryotic translation initiation factor eIF4G and the mRNA 3′ end.* J Virol, 2000. **74**(15): p. 7064-71.
116. Chung, K.T. and M.A. McCrae, *Regulation of gene expression by the NSP1 and NSP3 non-structural proteins of rotavirus.* Arch Virol, 2011. **156**(12): p. 2197-203.
117. Chizhikov, V. and J.T. Patton, *A four-nucleotide translation enhancer in the 3'-terminal consensus sequence of the nonpolyadenylated mRNAs of rotavirus.* Rna, 2000. **6**(6): p. 814-25.

118. Groft, C.M. and S.K. Burley, *Recognition of eIF4G by Rotavirus NSP3 Reveals a Basis for mRNA Circularization.* Molecular Cell, 2002. **9**(6): p. 1273-1283.
119. Deo, R.C., et al., *Recognition of the rotavirus mRNA 3' consensus by an asymmetric NSP3 homodimer.* Cell, 2002. **108**(1): p. 71-81.
120. Piron, M., et al., *Rotavirus RNA-binding protein NSP3 interacts with eIF4GI and evicts the poly(A) binding protein from eIF4F.* The EMBO journal, 1998. **17**(19): p. 5811-5821.
121. Gratia, M., et al., *Rotavirus NSP3 Is a Translational Surrogate of the Poly(A) Binding Protein-Poly(A) Complex.* J Virol, 2015. **89**(17): p. 8773-82.
122. Padilla-Noriega, L., O. Paniagua, and S. Guzmán-León, *Rotavirus protein NSP3 shuts off host cell protein synthesis.* Virology, 2002. **298**(1): p. 1-7.
123. Arnold, M.M., et al., *Rotavirus variant replicates efficiently although encoding an aberrant NSP3 that fails to induce nuclear localization of poly(A)-binding protein.* J Gen Virol, 2012. **93**(Pt 7): p. 1483-1494.
124. Rubio, R.M., et al., *Rotavirus prevents the expression of host responses by blocking the nucleocytoplasmic transport of polyadenylated mRNAs.* J Virol, 2013. **87**(11): p. 6336-45.
125. Montero, H., et al., *Rotavirus infection induces the phosphorylation of eIF2alpha but prevents the formation of stress granules.* J Virol, 2008. **82**(3): p. 1496-504.
126. Rojas, M., C.F. Arias, and S. López, *Protein kinase R is responsible for the phosphorylation of eIF2alpha in rotavirus infection.* J Virol, 2010. **84**(20): p. 10457-66.
127. Zhang, M., et al., *A functional NSP4 enterotoxin peptide secreted from rotavirus-infected cells.* J Virol, 2000. **74**(24): p. 11663-70.
128. González, R.A., et al., *Relative localization of viroplasmic and endoplasmic reticulum-resident rotavirus proteins in infected cells.* Archives of Virology, 2000. **145**(9): p. 1963-1973.
129. Papa, G., A. Borodavka, and U. Desselberger, *Viroplasms: Assembly and Functions of Rotavirus Replication Factories.* Viruses, 2021. **13**(7): p. 1349.
130. Fabbretti, E., et al., *Two non-structural rotavirus proteins, NSP2 and NSP5, form viroplasm-like structures in vivo.* J Gen Virol, 1999. **80 (Pt 2)**: p. 333-339.
131. Campagna, M., et al., *RNA interference of rotavirus segment 11 mRNA reveals the essential role of NSP5 in the virus replicative cycle.* J Gen Virol, 2005. **86**(Pt 5): p. 1481-1487.
132. Lopez, T., et al., *Reduced expression of the rotavirus NSP5 gene has a pleiotropic effect on virus replication.* J Gen Virol, 2005. **86**(Pt 6): p. 1609-17.
133. Taraporewala, Z.F., et al., *Analysis of a temperature-sensitive mutant rotavirus indicates that NSP2 octamers are the functional form of the protein.* J Virol, 2002. **76**(14): p. 7082-93.
134. Eichwald, C., J.F. Rodriguez, and O.R. Burrone, *Characterization of rotavirus NSP2/NSP5 interactions and the dynamics of viroplasm formation.* J Gen Virol, 2004. **85**(Pt 3): p. 625-634.
135. Cheung, W., et al., *Rotaviruses associate with cellular lipid droplet components to replicate in viroplasms, and compounds disrupting or blocking lipid droplets inhibit viroplasm formation and viral replication.* J Virol, 2010. **84**(13): p. 6782-98.
136. Gaunt, E.R., et al., *Inhibition of rotavirus replication by downregulation of fatty acid synthesis.* J Gen Virol, 2013. **94**(Pt 6): p. 1310-7.
137. Gaunt, E.R., et al., *Lipidome analysis of rotavirus-infected cells confirms the close interaction of lipid droplets with viroplasms.* J Gen Virol, 2013. **94**(Pt 7): p. 1576-86.
138. Contin, R., et al., *Rotavirus replication requires a functional proteasome for effective assembly of viroplasms.* J Virol, 2011. **85**(6): p. 2781-92.
139. López, T., et al., *Replication of the rotavirus genome requires an active ubiquitin-proteasome system.* J Virol, 2011. **85**(22): p. 11964-71.
140. Arnoldi, F., et al., *Rotavirus Increases Levels of Lipidated LC3 Supporting Accumulation of Infectious Progeny Virus without Inducing Autophagosome Formation.* PLOS ONE, 2014. **9**(4): p. e95197.

141. Afrikanova, I., et al., *Phosphorylation generates different forms of rotavirus NSP5.* J Gen Virol, 1996. **77 (Pt 9)**: p. 2059-65.
142. Criglar, J.M., et al., *A novel form of rotavirus NSP2 and phosphorylation-dependent NSP2-NSP5 interactions are associated with viroplasm assembly.* J Virol, 2014. **88**(2): p. 786-98.
143. Criglar, J.M., et al., *Phosphorylation cascade regulates the formation and maturation of rotaviral replication factories.* 2018. **115**(51): p. E12015-E12023.
144. Martin, D., et al., *Sequestration of Free Tubulin Molecules by the Viral Protein NSP2 Induces Microtubule Depolymerization during Rotavirus Infection.* 2010. **84**(5): p. 2522-2532.
145. Jiang, X., et al., *Cryoelectron microscopy structures of rotavirus NSP2-NSP5 and NSP2-RNA complexes: implications for genome replication.* J Virol, 2006. **80**(21): p. 10829-35.
146. Geiger, F., et al., *Liquid-liquid phase separation underpins the formation of replication factories in rotaviruses.* Embo j, 2021. **40**(21): p. e107711.
147. Shah, P.N.M., et al., *Characterization of the rotavirus assembly pathway in situ using cryoelectron tomography.* Cell Host Microbe, 2023. **31**(4): p. 604-615.e4.
148. Borodavka, A., U. Desselberger, and J.T. Patton, *Genome packaging in multi-segmented dsRNA viruses: distinct mechanisms with similar outcomes.* Current Opinion in Virology, 2018. **33**: p. 106-112.
149. McDonald, S.M., et al., *Reassortment in segmented RNA viruses: mechanisms and outcomes.* Nature Reviews Microbiology, 2016. **14**(7): p. 448-460.
150. Silvestri, L.S., Z.F. Taraporewala, and J.T. Patton, *Rotavirus replication: plus-sense templates for double-stranded RNA synthesis are made in viroplasms.* Journal of virology, 2004. **78**(14): p. 7763-7774.
151. Vasquez-Del Carpio, R., et al., *Histidine triad-like motif of the rotavirus NSP2 octamer mediates both RTPase and NTPase activities.* J Mol Biol, 2006. **362**(3): p. 539-54.
152. Hu, L., et al., *Crystallographic Analysis of Rotavirus NSP2-RNA Complex Reveals Specific Recognition of 5' GG Sequence for RTPase Activity.* J Virol, 2012. **86**(19): p. 10547-57.
153. Taraporewala, Z., D. Chen, and J.T. Patton, *Multimers Formed by the Rotavirus Nonstructural Protein NSP2 Bind to RNA and Have Nucleoside Triphosphatase Activity.* 1999. **73**(12): p. 9934-9943.
154. Taraporewala, Z.F., et al., *Structure-Function Analysis of Rotavirus NSP2 Octamer by Using a Novel Complementation System.* 2006. **80**(16): p. 7984-7994.
155. Jayaram, H., et al., *Rotavirus protein involved in genome replication and packaging exhibits a HIT-like fold.* Nature, 2002. **417**(6886): p. 311-5.
156. Tortorici, M.A., et al., *Template recognition and formation of initiation complexes by the replicase of a segmented double-stranded RNA virus.* J Biol Chem, 2003. **278**(35): p. 32673-82.
157. Steger, C.L., et al., *In Vitro Double-Stranded RNA Synthesis by Rotavirus Polymerase Mutants with Lesions at Core Shell Contact Sites.* J Virol, 2019. **93**(20).
158. Zeng, C.Q., et al., *The N terminus of rotavirus VP2 is necessary for encapsidation of VP1 and VP3.* J Virol, 1998. **72**(1): p. 201-8.
159. Ogden, K.M., H.N. Ramanathan, and J.T. Patton, *Mutational analysis of residues involved in nucleotide and divalent cation stabilization in the rotavirus RNA-dependent RNA polymerase catalytic pocket.* Virology, 2012. **431**(1): p. 12-20.
160. Tao, Y., et al., *RNA Synthesis in a Cage - Structural Studies of Reovirus Polymerase λ3.* Cell, 2002. **111**(5): p. 733-745.
161. Gridley, C.L. and J.T. Patton, *Regulation of rotavirus polymerase activity by inner capsid proteins.* Curr Opin Virol, 2014. **9**: p. 31-8.
162. Borodavka, A., et al., *Protein-mediated RNA folding governs sequence-specific interactions between rotavirus genome segments.* Elife, 2017. **6**: p. e27453.
163. Strauss, S., et al., *Principles of RNA recruitment to viral ribonucleoprotein condensates in a segmented dsRNA virus.* eLife, 2023. **12**: p. e68670.

164. Bravo, J.P.K., et al., *Structural basis of rotavirus RNA chaperone displacement and RNA annealing.* 2021. **118**(41): p. e2100198118.
165. Vende, P., et al., *Rotavirus NSP2 interferes with the core lattice protein VP2 in initiation of minus-strand synthesis.* Virology, 2003. **313**(1): p. 261-273.
166. Ayala-Breton, C., et al., *Analysis of the kinetics of transcription and replication of the rotavirus genome by RNA interference.* J Virol, 2009. **83**(17): p. 8819-31.
167. Papa, G., et al., *CRISPR-Csy4-Mediated Editing of Rotavirus Double-Stranded RNA Genome.* Cell Reports, 2020. **32**(13).
168. Crawford, S.E., et al., *Autophagy hijacked through viroporin-activated calcium/calmodulin-dependent kinase kinase-β signaling is required for rotavirus replication.* Proc Natl Acad Sci U S A, 2012. **109**(50): p. E3405-13.
169. Berkova, Z., et al., *Rotavirus NSP4 induces a novel vesicular compartment regulated by calcium and associated with viroplasms.* J Virol, 2006. **80**(12): p. 6061-71.
170. Hyser, J.M., et al., *Rotavirus disrupts calcium homeostasis by NSP4 viroporin activity.* mBio, 2010. **1**(5).
171. Hyser, J.M., et al., *Activation of the endoplasmic reticulum calcium sensor STIM1 and store-operated calcium entry by rotavirus requires NSP4 viroporin activity.* J Virol, 2013. **87**(24): p. 13579-88.
172. Pham, T., et al., *The Rotavirus NSP4 Viroporin Domain is a Calcium-conducting Ion Channel.* Sci Rep, 2017. **7**: p. 43487.
173. Crawford, S.E., et al., *COPII Vesicle Transport Is Required for Rotavirus NSP4 Interaction with the Autophagy Protein LC3 II and Trafficking to Viroplasms.* J Virol, 2019. **94**(1).
174. Taylor, J.A., et al., *The RER-Localized Rotavirus Intracellular Receptor: A Truncated Purified Soluble Form Is Multivalent and Binds Virus Particles.* Virology, 1993. **194**(2): p. 807-814.
175. O'Brien, J.A., J.A. Taylor, and A.R. Bellamy, *Probing the Structure of Rotavirus NSP4: a Short Sequence at the Extreme C Terminus Mediates Binding to the Inner Capsid Particle.* 2000. **74**(11): p. 5388-5394.
176. Michelangeli, F., et al., *Selective depletion of stored calcium by thapsigargin blocks rotavirus maturation but not the cytopathic effect.* 1995. **69**(6): p. 3838-3847.
177. Ahmadian, S. and M.S. Shahrabadi, *Morphological study of the role of calcium in the assembly of the rotavirus outer capsid protein VP7.* Biotech Histochem, 1999. **74**(5): p. 266-73.
178. Poruchynsky, M.S., D.R. Maass, and P.H. Atkinson, *Calcium depletion blocks the maturation of rotavirus by altering the oligomerization of virus-encoded proteins in the ER.* J Cell Biol, 1991. **114**(4): p. 651-6.
179. Hyser, J.M., et al., *Epitope mapping and use of epitope-specific antisera to characterize the VP5* binding site in rotavirus SA11 NSP4.* Virology, 2008. **373**(1): p. 211-28.
180. Cuadras, M.A., et al., *Dissecting rotavirus particle-raft interaction with small interfering RNAs: insights into rotavirus transit through the secretory pathway.* J Virol, 2006. **80**(8): p. 3935-46.
181. Trask, S.D. and P.R. Dormitzer, *Assembly of highly infectious rotavirus particles recoated with recombinant outer capsid proteins.* J Virol, 2006. **80**(22): p. 11293-304.
182. Shaw, A.L., et al., *Three-dimensional visualization of the rotavirus hemagglutinin structure.* Cell, 1993. **74**(4): p. 693-701.
183. Jourdan, N., et al., *Rotavirus is released from the apical surface of cultured human intestinal cells through nonconventional vesicular transport that bypasses the Golgi apparatus.* J Virol, 1997. **71**(11): p. 8268-78.
184. Nejmeddine, M., et al., *Rotavirus spike protein VP4 is present at the plasma membrane and is associated with microtubules in infected cells.* J Virol, 2000. **74**(7): p. 3313-20.
185. Trejo-Cerro, Ó., et al., *Actin-Dependent Nonlytic Rotavirus Exit and Infectious Virus Morphogenetic Pathway in Nonpolarized Cells.* J Virol, 2018. **92**(6).

186. Gardet, A., et al., *Rotavirus spike protein VP4 binds to and remodels actin bundles of the epithelial brush border into actin bodies.* J Virol, 2006. **80**(8): p. 3947-56.
187. Gentsch, J.R., et al., *Serotype diversity and reassortment between human and animal rotavirus strains: implications for rotavirus vaccine programs.* J Infect Dis, 2005. **192 Suppl 1**: p. S146-59.
188. Bányai, K., et al., *Systematic review of regional and temporal trends in global rotavirus strain diversity in the pre rotavirus vaccine era: insights for understanding the impact of rotavirus vaccination programs.* Vaccine, 2012. **30 Suppl 1**: p. A122-30.
189. Matthijnssens, J. and M. Van Ranst, *Genotype constellation and evolution of group A rotaviruses infecting humans.* Curr Opin Virol, 2012. **2**(4): p. 426-33.
190. Velázquez, F.R., et al., *Rotavirus Infection in Infants as Protection against Subsequent Infections.* 1996. **335**(14): p. 1022-1028.
191. Johansen, K. and L. Svensson, *Neutralization of rotavirus and recognition of immunologically important epitopes on VP4 and VP7 by human IgA.* Arch Virol, 1997. **142**(7): p. 1491-8.
192. Franco, M.A., J. Angel, and H.B. Greenberg, *Immunity and correlates of protection for rotavirus vaccines.* Vaccine, 2006. **24**(15): p. 2718-31.
193. Desselberger, U. and H.I. Huppertz, *Immune responses to rotavirus infection and vaccination and associated correlates of protection.* J Infect Dis, 2011. **203**(2): p. 188-95.
194. Glass, R.I., et al., *Rotavirus vaccines: successes and challenges.* J Infect, 2014. **68 Suppl 1**: p. S9-18.
195. Desselberger, U., *Differences of Rotavirus Vaccine Effectiveness by Country: Likely Causes and Contributing Factors.* 2017. **6**(4): p. 65.
196. Parker, E.P., et al., *Causes of impaired oral vaccine efficacy in developing countries.* Future Microbiol, 2018. **13**(1): p. 97-118.
197. Vesikari, T., et al., *Clinical efficacy of the RIT 4237 live attenuated bovine rotavirus vaccine in infants vaccinated before a rotavirus epidemic.* J Pediatr, 1985. **107**(2): p. 189-94.
198. Vesikari, T., et al., *Protection of infants against rotavirus diarrhoea by RIT 4237 attenuated bovine rotavirus strain vaccine.* Lancet, 1984. **1**(8384): p. 977-81.
199. De Mol, P., et al., *Failure of live, attenuated oral rotavirus vaccine.* Lancet, 1986. **2**(8498): p. 108.
200. Senturia, Y.D., et al., *Live attenuated oral rotavirus vaccine.* The Lancet, 1987. **330**(8567): p. 1091-1092.
201. Ukae, S., et al., *Efficacy of rhesus rotavirus vaccine MMU-18006 against gastroenteritis due to serotype 1 rotavirus.* Vaccine, 1994. **12**(10): p. 933-939.
202. Vesikari, T., et al., *A comparative trial of rhesus monkey (RRV-1) and bovine (RIT 4237) oral rotavirus vaccines in young children.* J Infect Dis, 1986. **153**(5): p. 832-9.
203. Vesikari, T., et al., *Rhesus Rotavirus candidate vaccine. Clinical trial in children vaccinated between 2 and 5 months of age.* Am J Dis Child, 1990. **144**(3): p. 285-9.
204. Rennels, M.B., et al., *An efficacy trial of the rhesus rotavirus vaccine in Maryland. The Clinical Study Group.* Am J Dis Child, 1990. **144**(5): p. 601-4.
205. Losonsky, G.A., et al., *Systemic and mucosal immune responses to rhesus rotavirus vaccine MMU 18006.* Pediatr Infect Dis J, 1988. **7**(6): p. 388-93.
206. Christy, C., et al., *Field trial of rhesus rotavirus vaccine in infants.* Pediatr Infect Dis J, 1988. **7**(9): p. 645-50.
207. Kapikian, A.Z., et al., *Efficacy of a quadrivalent rhesus rotavirus-based human rotavirus vaccine aimed at preventing severe rotavirus diarrhea in infants and young children.* J Infect Dis, 1996. **174 Suppl 1**: p. S65-72.
208. Midthun, K., et al., *Reassortant rotaviruses as potential live rotavirus vaccine candidates.* J Virol, 1985. **53**(3): p. 949-54.
209. Midthun, K., et al., *Single gene substitution rotavirus reassortants containing the major neutralization protein (VP7) of human rotavirus serotype 4.* J Clin Microbiol, 1986. **24**(5): p. 822-6.

210. Santosham, M., et al., *Efficacy and safety of high-dose rhesus-human reassortant rotavirus vaccine in Native American populations.* J Pediatr, 1997. **131**(4): p. 632-8.
211. Pérez-Schael, I., et al., *Efficacy of the rhesus rotavirus-based quadrivalent vaccine in infants and young children in Venezuela.* N Engl J Med, 1997. **337**(17): p. 1181-7.
212. Murphy, T.V., et al., *Intussusception among infants given an oral rotavirus vaccine.* N Engl J Med, 2001. **344**(8): p. 564-72.
213. Simonsen, L., et al., *More on RotaShield and intussusception: the role of age at the time of vaccination.* J Infect Dis, 2005. **192 Suppl 1**: p. S36-43.
214. Murphy, B.R., et al., *Reappraisal of the Association of Intussusception with the Licensed Live Rotavirus Vaccine Challenges Initial Conclusions.* The Journal of Infectious Diseases, 2003. **187**(8): p. 1301-1308.
215. Kapikian, A.Z., et al., *A hexavalent human rotavirus-bovine rotavirus (UK) reassortant vaccine designed for use in developing countries and delivered in a schedule with the potential to eliminate the risk of intussusception.* J Infect Dis, 2005. **192 Suppl 1**: p. S22-9.
216. Bernstein, D.I., et al., *Safety and immunogenicity of live, attenuated human rotavirus vaccine 89-12.* Vaccine, 1998. **16**(4): p. 381-7.
217. Ruiz-Palacios, G.M., et al., *Safety and efficacy of an attenuated vaccine against severe rotavirus gastroenteritis.* N Engl J Med, 2006. **354**(1): p. 11-22.
218. Vesikari, T., et al., *Efficacy of human rotavirus vaccine against rotavirus gastroenteritis during the first 2 years of life in European infants: randomised, double-blind controlled study.* Lancet, 2007. **370**(9601): p. 1757-63.
219. Heaton, P.M. and M. Ciarlet, *The Pentavalent Rotavirus Vaccine: Discovery to Licensure and Beyond.* Clinical Infectious Diseases, 2007. **45**(12): p. 1618-1624.
220. Angel, J., M.A. Franco, and H.B. Greenberg, *Rotavirus vaccines: recent developments and future considerations.* Nat Rev Microbiol, 2007. **5**(7): p. 529-39.
221. Bernstein, D.I., et al., *Evaluation of WC3 rotavirus vaccine and correlates of protection in healthy infants.* The Journal of infectious diseases, 1990. **162**(5): p. 1055-1062.
222. Georges-Courbot, M.C., et al., *Evaluation of the efficacy of a low-passage bovine rotavirus (strain WC3) vaccine in children in Central Africa.* Res Virol, 1991. **142**(5): p. 405-11.
223. Heaton, P.M., et al., *Development of a pentavalent rotavirus vaccine against prevalent serotypes of rotavirus gastroenteritis.* J Infect Dis, 2005. **192 Suppl 1**: p. S17-21.
224. Vesikari, T., et al., *Safety and efficacy of a pentavalent human-bovine (WC3) reassortant rotavirus vaccine.* N Engl J Med, 2006. **354**(1): p. 23-33.
225. Groome, M.J., et al., *Safety and immunogenicity of a parenteral trivalent P2-VP8 subunit rotavirus vaccine: a multisite, randomised, double-blind, placebo-controlled trial.* Lancet Infect Dis, 2020. **20**(7): p. 851-863.
226. Resch, T.K., et al., *Inactivated rotavirus vaccine by parenteral administration induces mucosal immunity in mice.* Scientific Reports, 2018. **8**(1): p. 561.
227. Li, J., et al., *Effectiveness of Lanzhou lamb rotavirus vaccine in preventing gastroenteritis among children younger than 5 years of age.* Vaccine, 2019. **37**(27): p. 3611-3616.
228. Skansberg, A., et al., *Product review of the rotavirus vaccines ROTASIIL, ROTAVAC, and Rotavin-M1.* Hum Vaccin Immunother, 2021. **17**(4): p. 1223-1234.
229. Kawagishi, T., et al., *Mucosal and systemic neutralizing antibodies to norovirus induced in infant mice orally inoculated with recombinant rotaviruses.* 2023. **120**(9): p. e2214421120.
230. Burnett, E., et al., *Major Changes in Spatiotemporal Trends of US Rotavirus Laboratory Detections After Rotavirus Vaccine Introduction-2009-2021.* Pediatr Infect Dis J, 2022. **41**(9): p. 759-763.
231. Bergman, H., et al., *Vaccines for preventing rotavirus diarrhoea: vaccines in use.* The Cochrane database of systematic reviews, 2021. **11**(11): p. CD008521-CD008521.
232. Ella, R., et al., *A randomized, open-labelled, non-inferiority phase 4 clinical trial to evaluate the immunogenicity and safety of the live, attenuated, oral rotavirus vaccine,*

ROTAVAC® in comparison with a licensed rotavirus vaccine in healthy infants. Vaccine, 2019. **37**(31): p. 4407-4413.
233. Patel, M.M. and U.D. Parashar, *Assessing the Effectiveness and Public Health Impact of Rotavirus Vaccines after Introduction in Immunization Programs.* The Journal of Infectious Diseases, 2009. **200**(Supplement_1): p. S291-S299.
234. Boom, J.A., et al., *Sustained protection from pentavalent rotavirus vaccination during the second year of life at a large, urban United States pediatric hospital.* 2010. **29**(12): p. 1133-1135.
235. Vesikari, T., et al., *Sustained efficacy of the pentavalent rotavirus vaccine, RV5, up to 3.1 years following the last dose of vaccine.* Pediatr Infect Dis J, 2010. **29**(10): p. 957-63.
236. Tate, J.E., et al., *Uptake, impact, and effectiveness of rotavirus vaccination in the United States: review of the first 3 years of postlicensure data.* Pediatr Infect Dis J, 2011. **30**(1 Suppl): p. S56-60.
237. Paulke-Korinek, M., et al., *Herd immunity after two years of the universal mass vaccination program against rotavirus gastroenteritis in Austria.* Vaccine, 2011. **29**(15): p. 2791-6.
238. Velasquez-Portocarrero, D.E., et al., *Head-to-head comparison of the immunogenicity of RotaTeq and Rotarix rotavirus vaccines and factors associated with seroresponse in infants in Bangladesh: a randomised, controlled, open-label, parallel, phase 4 trial.* Lancet Infect Dis, 2022. **22**(11): p. 1606-1616.
239. Armah, G.E., et al., *Efficacy of pentavalent rotavirus vaccine against severe rotavirus gastroenteritis in infants in developing countries in sub-Saharan Africa: a randomised, double-blind, placebo-controlled trial.* Lancet, 2010. **376**(9741): p. 606-14.
240. Lopman, B.A., et al., *Understanding Reduced Rotavirus Vaccine Efficacy in Low Socio-Economic Settings.* PLOS ONE, 2012. **7**(8): p. e41720.
241. Parker, E.P., et al., *Causes of impaired oral vaccine efficacy in developing countries.* 2018. **13**(1): p. 97-118.
242. Patel, M.M., et al., *Intussusception risk and health benefits of rotavirus vaccination in Mexico and Brazil.* N Engl J Med, 2011. **364**(24): p. 2283-92.
243. Buttery, J.P., et al., *Intussusception following rotavirus vaccine administration: post-marketing surveillance in the National Immunization Program in Australia.* Vaccine, 2011. **29**(16): p. 3061-6.
244. Glass, R.I. and U.D.J.T.N.E.j.o.m. Parashar, *Rotavirus vaccines—balancing intussusception risks and health benefits.* 2014. **370**(6): p. 568.
245. Patel, N.C., et al., *Vaccine-acquired rotavirus in infants with severe combined immunodeficiency.* N Engl J Med, 2010. **362**(4): p. 314-9.
246. Boom, J.A., et al., *Symptomatic Infection and Detection of Vaccine and Vaccine-Reassortant Rotavirus Strains in 5 Children: A Case Series.* The Journal of Infectious Diseases, 2012. **206**(8): p. 1275-1279.
247. Hemming, M. and T.J.T.P.i.d.j. Vesikari, *Detection of rotateq vaccine-derived, double-reassortant rotavirus in a 7-year-old child with acute gastroenteritis.* 2014. **33**(6): p. 655-656.
248. Bucardo, F., et al., *Vaccine-derived NSP2 segment in rotaviruses from vaccinated children with gastroenteritis in Nicaragua.* 2012. **12**(6): p. 1282-1294.
249. Kanai, Y., et al., *Reverse Genetics Approach for Developing Rotavirus Vaccine Candidates Carrying VP4 and VP7 Genes Cloned from Clinical Isolates of Human Rotavirus.* J Virol, 2020. **95**(2).
250. hebdomadaire, W.H.O.J.W.E.R.R.é., *Rotavirus vaccines: WHO position paper—January 2013.* 2013. **88**(05): p. 49-64.
251. Alaoui Amine, S., et al., *Evidence for zoonotic transmission of species A rotavirus from goat and cattle in nomadic herds in Morocco, 2012–2014.* Virus Genes, 2020. **56**(5): p. 582-593.

252. Alaoui Amine, S., et al., *Full-length genome analysis of the first human G8P[14] rotavirus strain from Morocco suggests evidence of zoonotic transmission.* Virus Genes, 2019. **55**(4): p. 465-478.
253. Tacharoenmuang, R., et al., *Characterization of a G10P[14] rotavirus strain from a diarrheic child in Thailand: Evidence for bovine-to-human zoonotic transmission.* Infect Genet Evol, 2018. **63**: p. 43-57.
254. Martella, V., et al., *Zoonotic aspects of rotaviruses.* Vet Microbiol, 2010. **140**(3-4): p. 246-55.
255. Mascarenhas, J.D.P., et al., *Detection of a neonatal human rotavirus strain with VP4 and NSP4 genes of porcine origin.* J Med Microbiol, 2007. **56**(Pt 4): p. 524-532.
256. Meganck, V., et al., *Evaluation of a protocol to reduce the incidence of neonatal calf diarrhoea on dairy herds.* Prev Vet Med, 2015. **118**(1): p. 64-70.
257. Vlasova, A.N., J.O. Amimo, and L.J. Saif, *Porcine Rotaviruses: Epidemiology, Immune Responses and Control Strategies.* Viruses, 2017. **9**(3).
258. Martella, V., et al., *Prevalence of group C rotaviruses in weaning and post-weaning pigs with enteritis.* Vet Microbiol, 2007. **123**(1-3): p. 26-33.
259. Park, J.-G., et al., *Development of a live attenuated trivalent porcine rotavirus A vaccine against disease caused by recent strains most prevalent in South Korea.* Veterinary Research, 2019. **50**(1): p. 2.
260. Dhama, K., et al., *Rotavirus diarrhea in bovines and other domestic animals.* Vet Res Commun, 2009. **33**(1): p. 1-23.
261. Uprety, T., D. Wang, and F. Li, *Recent advances in rotavirus reverse genetics and its utilization in basic research and vaccine development.* Archives of virology, 2021. **166**(9): p. 2369-2386.
262. Pollard, A.J. and E.M. Bijker, *A guide to vaccinology: from basic principles to new developments.* Nature Reviews Immunology, 2021. **21**(2): p. 83-100.
263. Richards, J.E., U. Desselberger, and A.M. Lever, *Experimental pathways towards developing a rotavirus reverse genetics system: synthetic full length rotavirus ssRNAs are neither infectious nor translated in permissive cells.* PLoS One, 2013. **8**(9): p. e74328.
264. Gombold, J.L. and R.F. Ramig, *Analysis of reassortment of genome segments in mice mixedly infected with rotaviruses SA11 and RRV.* 1986. **57**(1): p. 110-116.
265. Greenberg, H.B., et al., *Rescue of noncultivatable human rotavirus by gene reassortment during mixed infection with ts mutants of a cultivatable bovine rotavirus.* 1981. **78**(1): p. 420-424.
266. Fukuda, S., et al., *Rapid generation of rotavirus single-gene reassortants by means of eleven plasmid-only based reverse genetics.* 2020. **101**(8): p. 806-815.
267. Kobayashi, T., et al., *A plasmid-based reverse genetics system for animal double-stranded RNA viruses.* Cell Host Microbe, 2007. **1**(2): p. 147-57.
268. Kobayashi, T., et al., *An improved reverse genetics system for mammalian orthoreoviruses.* Virology, 2010. **398**(2): p. 194-200.
269. Kawagishi, T., et al., *Reverse Genetics for Fusogenic Bat-Borne Orthoreovirus Associated with Acute Respiratory Tract Infections in Humans: Role of Outer Capsid Protein σC in Viral Replication and Pathogenesis.* PLoS Pathog, 2016. **12**(2): p. e1005455.
270. Boyce, M., C.C.P. Celma, and P. Roy, *Development of Reverse Genetics Systems for Bluetongue Virus: Recovery of Infectious Virus from Synthetic RNA Transcripts.* 2008. **82**(17): p. 8339-8348.
271. Matsuo, E., C.C.P. Celma, and P. Roy, *A reverse genetics system of African horse sickness virus reveals existence of primary replication.* FEBS Letters, 2010. **584**(15): p. 3386-3391.
272. Pretorius, J.M., H. Huismans, and J. Theron, *Establishment of an entirely plasmid-based reverse genetics system for Bluetongue virus.* Virology, 2015. **486**: p. 71-77.
273. Enami, M., et al., *Introduction of site-specific mutations into the genome of influenza virus.* Proc Natl Acad Sci U S A, 1990. **87**(10): p. 3802-5.

274. Komoto, S., J. Sasaki, and K. Taniguchi, *Reverse genetics system for introduction of site-specific mutations into the double-stranded RNA genome of infectious rotavirus.* Proc Natl Acad Sci U S A, 2006. **103**(12): p. 4646-51.
275. Komoto, S., et al., *Generation of recombinant rotavirus with an antigenic mosaic of cross-reactive neutralization epitopes on VP4.* J Virol, 2008. **82**(13): p. 6753-7.
276. Kobayashi, N., et al., *Selection of rotavirus VP7 gene in the genetic background of simian rotavirus SA11: implications for rotavirus reassortant vaccine development.* Antiviral Research, 1996. **31**(3): p. 185-190.
277. Troupin, C., et al., *Rearranged genomic RNA segments offer a new approach to the reverse genetics of rotaviruses.* J Virol, 2010. **84**(13): p. 6711-9.
278. Pedley, S., et al., *The Genomes of Rotaviruses Isolated from Chronically Infected Immunodeficient Children.* 1984. **65**(7): p. 1141-1150.
279. Hundley, F., et al., *Genome rearrangements of bovine rotavirus after serial passage at high multiplicity of infection.* Virology, 1985. **143**(1): p. 88-103.
280. Desselberger, U., *Genome rearrangements of rotaviruses.* Adv Virus Res, 1996. **46**: p. 69-95.
281. Schnepf, N., et al., *Rearrangements of rotavirus genomic segment 11 are generated during acute infection of immunocompetent children and do not occur at random.* J Virol, 2008. **82**(7): p. 3689-96.
282. Trask, S.D., et al., *Dual selection mechanisms drive efficient single-gene reverse genetics for rotavirus.* Proc Natl Acad Sci U S A, 2010. **107**(43): p. 18652-7.
283. Johne, R., et al., *Generation of an Avian-Mammalian Rotavirus Reassortant by Using a Helper Virus-Dependent Reverse Genetics System.* J Virol, 2016. **90**(3): p. 1439-43.
284. Kanai, Y., et al., *Entirely plasmid-based reverse genetics system for rotaviruses.* Proc Natl Acad Sci U S A, 2017. **114**(9): p. 2349-2354.
285. Brown, C.W., et al., *The p14 FAST protein of reptilian reovirus increases vesicular stomatitis virus neuropathogenesis.* J Virol, 2009. **83**(2): p. 552-61.
286. Ciechonska, M. and R. Duncan, *Reovirus FAST proteins: virus-encoded cellular fusogens.* Trends Microbiol, 2014. **22**(12): p. 715-24.
287. Salsman, J., et al., *Extensive syncytium formation mediated by the reovirus FAST proteins triggers apoptosis-induced membrane instability.* J Virol, 2005. **79**(13): p. 8090-100.
288. Komoto, S., et al., *Generation of Recombinant Rotaviruses Expressing Fluorescent Proteins by Using an Optimized Reverse Genetics System.* J Virol, 2018. **92**(13): p. e00588-18.
289. Komoto, S., et al., *Generation of Infectious Recombinant Human Rotaviruses from Just 11 Cloned cDNAs Encoding the Rotavirus Genome.* J Virol, 2019. **93**(8).
290. Philip, A.A., et al., *Generation of Recombinant Rotavirus Expressing NSP3-UnaG Fusion Protein by a Simplified Reverse Genetics System.* J Virol, 2019. **93**(24): p. e01616-19.
291. Eaton, H.E., et al., *African Swine Fever Virus NP868R Capping Enzyme Promotes Reovirus Rescue during Reverse Genetics by Promoting Reovirus Protein Expression, Virion Assembly, and RNA Incorporation into Infectious Virions.* J Virol, 2017. **91**(11).
292. Sánchez-Tacuba, L., et al., *An Optimized Reverse Genetics System Suitable for Efficient Recovery of Simian, Human, and Murine-Like Rotaviruses.* J Virol, 2020. **94**(18): p. e01294-20.
293. Komoto, S., et al., *Modification of the trypsin cleavage site of rotavirus VP4 to a furin-sensitive form does not enhance replication efficiency.* J Gen Virol, 2011. **92**(Pt 12): p. 2914-2921.
294. Schrauwen, E.J., et al., *The multibasic cleavage site in H5N1 virus is critical for systemic spread along the olfactory and hematogenous routes in ferrets.* J Virol, 2012. **86**(7): p. 3975-84.
295. Navarro, A., S.D. Trask, and J.T. Patton, *Generation of genetically stable recombinant rotaviruses containing novel genome rearrangements and heterologous sequences by reverse genetics.* J Virol, 2013. **87**(11): p. 6211-20.

296. Kanai, Y., et al., *Development of Stable Rotavirus Reporter Expression Systems.* J Virol, 2019. **93**(4).
297. Liu, Z., et al., *Systematic comparison of 2A peptides for cloning multi-genes in a polycistronic vector.* Sci Rep, 2017. **7**(1): p. 2193.
298. Philip, A.A. and J.T. Patton, *Expression of Separate Heterologous Proteins from the Rotavirus NSP3 Genome Segment Using a Translational 2A Stop-Restart Element.* J Virol, 2020. **94**(18): p. e00959-20.
299. Philip, A.A. and J.T. Patton, *Rotavirus as an Expression Platform of Domains of the SARS-CoV-2 Spike Protein.* Vaccines (Basel), 2021. **9**(5): p. 449.
300. Diebold, O., et al., *Using Species a Rotavirus Reverse Genetics to Engineer Chimeric Viruses Expressing SARS-CoV-2 Spike Epitopes.* 2022. **96**(14): p. e00488-22.
301. Papa, G., et al., *Recombinant Rotaviruses Rescued by Reverse Genetics Reveal the Role of NSP5 Hyperphosphorylation in the Assembly of Viral Factories.* 2019. **94**(1): p. 10.1128/jvi.01110-19.
302. Campagna, M., et al., *Impaired hyperphosphorylation of rotavirus NSP5 in cells depleted of casein kinase 1α is associated with the formation of viroplasms with altered morphology and a moderate decrease in virus replication.* 2007. **88**(10): p. 2800-2810.
303. Eichwald, C., et al., *Uncoupling substrate and activation functions of rotavirus NSP5: Phosphorylation of Ser-67 by casein kinase 1 is essential for hyperphosphorylation.* 2004. **101**(46): p. 16304-16309.
304. Haurwitz, R.E., S.H. Sternberg, and J.A. Doudna, *Csy4 relies on an unusual catalytic dyad to position and cleave CRISPR RNA.* 2012. **31**(12): p. 2824-2832.
305. Song, Y., et al., *Reverse Genetics Reveals a Role of Rotavirus VP3 Phosphodiesterase Activity in Inhibiting RNase L Signaling and Contributing to Intestinal Viral Replication In Vivo.* J Virol, 2020. **94**(9).
306. Mohanty, S.K., et al., *The SRL peptide of rhesus rotavirus VP4 protein governs cholangiocyte infection and the murine model of biliary atresia.* Hepatology, 2017. **65**(4).
307. Mohanty, S.K., et al., *A Point Mutation in the Rhesus Rotavirus VP4 Protein Generated through a Rotavirus Reverse Genetics System Attenuates Biliary Atresia in the Murine Model.* 2017. **91**(15): p. 10.1128/jvi.00510-17.
308. Zhu, N., et al., *A Novel Coronavirus from Patients with Pneumonia in China, 2019.* 2020. **382**(8): p. 727-733.
309. Worobey, M., et al., *The Huanan Seafood Wholesale Market in Wuhan was the early epicenter of the COVID-19 pandemic.* 2022. **377**(6609): p. 951-959.
310. Goldstein, S.A., Hogue, B. G., Leibowitz, J. L., Weiss, S. R., *Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2)*, in *Fields Virology: RNA Viruses*, P. Howley, Knipe, D., Whelan, S. and Freed, E. , Editor. 2023, Wolters Kluwer Health.: Philadelphia. p. 706 - 740.
311. Shang, J., et al., *Cell entry mechanisms of SARS-CoV-2.* 2020. **117**(21): p. 11727-11734.
312. Cui, J., F. Li, and Z.-L. Shi, *Origin and evolution of pathogenic coronaviruses.* Nature Reviews Microbiology, 2019. **17**(3): p. 181-192.
313. Zhong, N.S., et al., *Epidemiology and cause of severe acute respiratory syndrome (SARS) in Guangdong, People's Republic of China, in February, 2003.* Lancet, 2003. **362**(9393): p. 1353-8.
314. Zaki, A.M., et al., *Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia.* N Engl J Med, 2012. **367**(19): p. 1814-20.
315. Martin, J.E., et al., *A SARS DNA vaccine induces neutralizing antibody and cellular immune responses in healthy adults in a Phase I clinical trial.* Vaccine, 2008. **26**(50): p. 6338-43.
316. Graham, R.L., E.F. Donaldson, and R.S. Baric, *A decade after SARS: strategies for controlling emerging coronaviruses.* Nat Rev Microbiol, 2013. **11**(12): p. 836-48.
317. Lin, J.T., et al., *Safety and immunogenicity from a phase I trial of inactivated severe acute respiratory syndrome coronavirus vaccine.* Antivir Ther, 2007. **12**(7): p. 1107-13.

318. Hu, B., et al., *Characteristics of SARS-CoV-2 and COVID-19.* Nature Reviews Microbiology, 2021. **19**(3): p. 141-154.
319. Lan, J., et al., *Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor.* Nature, 2020. **581**(7807): p. 215-220.
320. Krammer, F., *SARS-CoV-2 vaccines in development.* Nature, 2020. **586**(7830): p. 516-527.
321. Dai, L. and G.F. Gao, *Viral targets for vaccines against COVID-19.* Nature Reviews Immunology, 2021. **21**(2): p. 73-82.
322. Buchholz, U.J., S. Finke, and K.-K. Conzelmann, *Generation of Bovine Respiratory Syncytial Virus (BRSV) from cDNA: BRSV NS2 Is Not Essential for Virus Replication in Tissue Culture, and the Human RSV Leader Region Acts as a Functional BRSV Genome Promoter.* 1999. **73**(1): p. 251-259.
323. DuBridge, R.B., et al., *Analysis of mutation in human cells by using an Epstein-Barr virus shuttle system.* Mol Cell Biol, 1987. **7**(1): p. 379-87.
324. Kibbe, W.A., *OligoCalc: an online oligonucleotide properties calculator.* Nucleic Acids Res, 2007. **35**(Web Server issue): p. W43-6.
325. Simmonds, P., *SSE: a nucleotide and amino acid sequence analysis platform.* BMC Res Notes, 2012. **5**: p. 50.
326. Derricott, H., et al., *Developing a 3D intestinal epithelium model for livestock species.* Cell Tissue Res, 2019. **375**(2): p. 409-424.
327. Brister, J.R., et al., *NCBI viral genomes resource.* Nucleic Acids Res, 2015. **43**(Database issue): p. D571-7.
328. Kozak, M., *Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes.* Cell, 1986. **44**(2): p. 283-92.
329. Jumper, J., et al., *Highly accurate protein structure prediction with AlphaFold.* Nature, 2021. **596**(7873): p. 583-589.
330. Schrodinger, LLC, *The PyMOL Molecular Graphics System, Version 1.8*. 2015.
331. Xie, X., et al., *An Infectious cDNA Clone of SARS-CoV-2.* Cell Host & Microbe, 2020. **27**(5): p. 841-848.e3.
332. Heaton, N.S., et al., *Targeting Viral Proteostasis Limits Influenza Virus, HIV, and Dengue Virus Infection.* Immunity, 2016. **44**(1): p. 46-58.
333. Kawagishi, T., et al., *Reverse Genetics System for a Human Group A Rotavirus.* J Virol, 2020. **94**(2).
334. Ciarlet, M. and F. Schödel, *Development of a rotavirus vaccine: clinical safety, immunogenicity, and efficacy of the pentavalent rotavirus vaccine, RotaTeq.* Vaccine, 2009. **27 Suppl 6**: p. G72-81.
335. Madeira, F., et al., *Search and sequence analysis tools services from EMBL-EBI in 2022.* Nucleic acids research, 2022. **50**(W1): p. W276-W279.
336. Brown, N.P., C. Leroy, and C. Sander, *MView: a web-compatible database search or multiple alignment viewer.* Bioinformatics, 1998. **14**(4): p. 380-381.
337. Stothard, P., *The sequence manipulation suite: JavaScript programs for analyzing and formatting protein and DNA sequences.* Biotechniques, 2000. **28**(6): p. 1102, 1104.
338. Roner, M.R. and W.K.J.P.o.t.N.A.o.S. Joklik, *Reovirus reverse genetics: Incorporation of the CAT gene into the reovirus genome.* 2001. **98**(14): p. 8036-8041.
339. Morgan, J.R., L.K. Cohen, and B.E.J.J.o.v. Roberts, *Identification of the DNA sequences encoding the large subunit of the mRNA-capping enzyme of vaccinia virus.* 1984. **52**(1): p. 206-214.
340. Niles, E.G., et al., *Vaccinia virus gene D12L encodes the small subunit of the viral mRNA capping enzyme.* 1989. **172**(2): p. 513-522.
341. Arias, C.F., et al., *Trypsin activation pathway of rotavirus infectivity.* Journal of virology, 1996. **70**(9): p. 5832-5839.
342. Scholz, J., et al., *Establishment of a Plasmid-Based Reverse Genetics System for the Cell Culture-Adapted Hepatitis E Virus Genotype 3c Strain 47832c.* Pathogens, 2020. **9**(3).

343. Huang, A.S. and D. Baltimore, *Defective Viral Particles and Viral Disease Processes.* Nature, 1970. **226**(5243): p. 325-327.
344. Smith, S.C., et al., *Reovirus RNA Recombination Is Sequence Directed and Generates Internally Deleted Defective Genome Segments during Passage.* 2021. **95**(8): p. e02181-20.
345. Odagiri, T. and K. Tobita, *Mutation in NS2, a nonstructural protein of influenza A virus, extragenically causes aberrant replication and expression of the PA gene and leads to generation of defective interfering particles.* 1990. **87**(15): p. 5988-5992.
346. Hutchinson, E.C., et al., *Genome packaging in influenza A virus.* 2010. **91**(2): p. 313-328.
347. Alam, M.M., et al., *Identical rearrangement of NSP3 genes found in three independently isolated virus clones derived from mixed infection and multiple passages of Rotaviruses.* Arch Virol, 2008. **153**(3): p. 555-9.
348. Genoyer, E. and C.B. López, *The Impact of Defective Viruses on Infection and Immunity.* 2019. **6**(1): p. 547-566.
349. Sánchez-Aparicio, M.T., et al., *Loss of Sendai virus C protein leads to accumulation of RIG-I immunostimulatory defective interfering RNA.* J Gen Virol, 2017. **98**(6): p. 1282-1293.
350. Xu, J., et al., *Replication defective viral genomes exploit a cellular pro-survival mechanism to establish paramyxovirus persistence.* Nat Commun, 2017. **8**(1): p. 799.
351. Xu, J., et al., *Identification of a Natural Viral RNA Motif That Optimizes Sensing of Viral RNA by RIG-I.* mBio, 2015. **6**(5): p. e01265-15.
352. Russell, A.B., et al., *Single-Cell Virus Sequencing of Influenza Infections That Trigger Innate Immunity.* 2019. **93**(14): p. e00500-19.
353. James, V.L.A., et al., *Molecular characterization of human group C rotavirus genes 6, 7 and 9.* J Gen Virol, 1999. **80 (Pt 12)**: p. 3181-3187.
354. Harb, M., et al., *Nuclear localization of cytoplasmic poly(A)-binding protein upon rotavirus infection involves the interaction of NSP3 with eIF4G and RoXaN.* J Virol, 2008. **82**(22): p. 11283-93.
355. Montero, H., C.F. Arias, and S. Lopez, *Rotavirus Nonstructural Protein NSP3 is not required for viral protein synthesis.* J Virol, 2006. **80**(18): p. 9031-8.
356. Gratia, M., et al., *Challenging the Roles of NSP3 and Untranslated Regions in Rotavirus mRNA Translation.* PLoS One, 2016. **11**(1): p. e0145998.
357. Berglund, P., et al., *Viral alteration of cellular translational machinery increases defective ribosomal products.* J Virol, 2007. **81**(13): p. 7220-9.
358. Jenkins, M.R., et al., *Addition of a prominent epitope affects influenza A virus-specific CD8+ T cell immunodominance hierarchies when antigen is limiting.* J Immunol, 2006. **177**(5): p. 2917-25.
359. Sigal, L.J., et al., *Cytotoxic T-cell immunity to virus-infected non-haematopoietic cells requires presentation of exogenous antigen.* Nature, 1999. **398**(6722): p. 77-80.
360. Lu, X., et al., *Endogenous viral antigen processing generates peptide-specific MHC class I cell-surface clusters.* Proc Natl Acad Sci U S A, 2012. **109**(38): p. 15407-12.
361. Dolan, B.P., et al., *MHC class I antigen processing distinguishes endogenous antigens based on their translation from cellular vs. viral mRNA.* Proc Natl Acad Sci U S A, 2012. **109**(18): p. 7025-30.
362. Yang, N., et al., *Defining Viral Defective Ribosomal Products: Standard and Alternative Translation Initiation Events Generate a Common Peptide from Influenza A Virus M2 and M1 mRNAs.* The Journal of Immunology, 2016. **196**(9): p. 3608-3617.
363. Hickman, H.D., et al., *Correction: Influenza A Virus Negative Strand RNA Is Translated for CD8+ T Cell Immunosurveillance.* The Journal of Immunology, 2018. **201**(7): p. 2187-2187.
364. In, J.G., et al., *Human mini-guts: new insights into intestinal physiology and host-pathogen interactions.* Nat Rev Gastroenterol Hepatol, 2016. **13**(11): p. 633-642.
365. Finkbeiner, S.R., et al., *Stem cell-derived human intestinal organoids as an infection model for rotaviruses.* mBio, 2012. **3**(4): p. e00159-12.

366. Yin, Y., et al., *Modeling rotavirus infection and antiviral therapy using primary intestinal organoids.* Antiviral Res, 2015. **123**: p. 120-31.
367. Guo, Y., et al., *Infection of porcine small intestinal enteroids with human and pig rotavirus A strains reveals contrasting roles for histo-blood group antigens and terminal sialic acids.* PLoS Pathog, 2021. **17**(1): p. e1009237.
368. Hamilton, C.A., et al., *Development of in vitro enteroids derived from bovine small intestinal crypts.* Veterinary Research, 2018. **49**(1): p. 54.
369. Blake, R., et al., *The Development of 3D Bovine Intestinal Organoid Derived Models to Investigate Mycobacterium Avium ssp Paratuberculosis Pathogenesis.* 2022. **9**.
370. Chang-Graham, A.L., et al., *Rotavirus induces intercellular calcium waves through ADP signaling.* Science, 2020. **370**(6519).
371. Foulke-Abel, J., et al., *Human enteroids as an ex-vivo model of host–pathogen interactions in the gastrointestinal tract.* 2014. **239**(9): p. 1124-1134.
372. Chen, S., et al., *Rotavirus Infection and Cytopathogenesis in Human Biliary Organoids Potentially Recapitulate Biliary Atresia Development.* mBio, 2020. **11**(4).
373. McCracken, K.W., et al., *Generating human intestinal tissue from pluripotent stem cells in vitro.* Nature Protocols, 2011. **6**(12): p. 1920-1928.
374. Holmes, E.C., et al., *The origins of SARS-CoV-2: A critical review.* Cell, 2021. **184**(19): p. 4848-4856.
375. Boni, M.F., et al., *Evolutionary origins of the SARS-CoV-2 sarbecovirus lineage responsible for the COVID-19 pandemic.* Nature Microbiology, 2020. **5**(11): p. 1408-1417.
376. Kyriakidis, N.C., et al., *SARS-CoV-2 vaccines strategies: a comprehensive review of phase 3 candidates.* npj Vaccines, 2021. **6**(1): p. 28.
377. Du, L., et al., *Identification of a receptor-binding domain in the S protein of the novel human coronavirus Middle East respiratory syndrome coronavirus as an essential target for vaccine development.* J Virol, 2013. **87**(17): p. 9939-42.
378. Kapadia, S.U., et al., *Long-term protection from SARS coronavirus infection conferred by a single immunization with an attenuated VSV-based vaccine.* Virology, 2005. **340**(2): p. 174-182.
379. Yang, Z.-y., et al., *A DNA vaccine induces SARS coronavirus neutralization and protective immunity in mice.* Nature, 2004. **428**(6982): p. 561-564.
380. Piccoli, L., et al., *Mapping Neutralizing and Immunodominant Sites on the SARS-CoV-2 Spike Receptor-Binding Domain by Structure-Guided High-Resolution Serology.* Cell, 2020. **183**(4): p. 1024-1042.e21.
381. Götzinger, F., et al., *COVID-19 in children and adolescents in Europe: a multinational, multicentre cohort study.* The Lancet Child & Adolescent Health, 2020. **4**(9): p. 653-661.
382. Buitrago-Garcia, D., et al., *Occurrence and transmission potential of asymptomatic and presymptomatic SARS-CoV-2 infections: A living systematic review and meta-analysis.* PLOS Medicine, 2020. **17**(9): p. e1003346.
383. Heald-Sargent, T., et al., *Age-Related Differences in Nasopharyngeal Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) Levels in Patients With Mild to Moderate Coronavirus Disease 2019 (COVID-19).* JAMA Pediatrics, 2020. **174**(9): p. 902-903.
384. Madera, S., et al., *Nasopharyngeal SARS-CoV-2 viral loads in young children do not differ significantly from those in older children and adults.* Scientific Reports, 2021. **11**(1): p. 3044.
385. Jones, T.C., et al., *Estimating infectiousness throughout SARS-CoV-2 infection course.* 2021. **373**(6551): p. eabi5273.
386. Johansson, M.A., et al., *SARS-CoV-2 Transmission From People Without COVID-19 Symptoms.* JAMA Netw Open, 2021. **4**(1): p. e2035057.
387. Oliveira, E.A., et al., *Clinical characteristics and risk factors for death among hospitalised children and adolescents with COVID-19 in Brazil: an analysis of a nationwide database.* The Lancet Child & Adolescent Health, 2021. **5**(8): p. 559-568.

388. Kitano, T., et al., *The differential impact of pediatric COVID-19 between high-income countries and low- and middle-income countries: A systematic review of fatality and ICU admission in children worldwide.* PLOS ONE, 2021. **16**(1): p. e0246326.
389. Lamers, M.M., et al., *SARS-CoV-2 productively infects human gut enterocytes.* 2020. **369**(6499): p. 50-54.
390. Cholankeril, G., et al., *High Prevalence of Concurrent Gastrointestinal Manifestations in Patients With Severe Acute Respiratory Syndrome Coronavirus 2: Early Experience From California.* Gastroenterology, 2020. **159**(2): p. 775-777.
391. Guo, M., et al., *Potential intestinal infection and faecal–oral transmission of SARS-CoV-2.* Nature Reviews Gastroenterology & Hepatology, 2021. **18**(4): p. 269-283.
392. Zhou, Z., et al., *Effect of Gastrointestinal Symptoms in Patients With COVID-19.* Gastroenterology, 2020. **158**(8): p. 2294-2297.
393. Redd, W.D., et al., *Prevalence and Characteristics of Gastrointestinal Symptoms in Patients With Severe Acute Respiratory Syndrome Coronavirus 2 Infection in the United States: A Multicenter Cohort Study.* Gastroenterology, 2020. **159**(2): p. 765-767.e2.
394. Periwal, S.B. and J.J. Cebra, *Respiratory Mucosal Immunization with Reovirus Serotype 1/L Stimulates Virus-Specific Humoral and Cellular Immune Responses, Including Double-Positive (CD4(+)/CD8(+) T Cells.* 1999. **73**(9): p. 7633-7640.
395. van Splunter, M., et al., *Oral cholera vaccination promotes homing of IgA(+) memory B cells to the large intestine and the respiratory tract.* Mucosal Immunol, 2018. **11**(4): p. 1254-1264.
396. Jaïs, P.H., et al., *C3P3-G1: first generation of a eukaryotic artificial cytoplasmic expression system.* Nucleic Acids Res, 2019. **47**(5): p. 2681-2698.
397. Nash, T.J., et al., *Inside-out chicken enteroids with leukocyte component as a model to study host–pathogen interactions.* Communications Biology, 2021. **4**(1): p. 377.
398. Saxena, K., et al., *A paradox of transcriptional and functional innate interferon responses of human intestinal enteroids to enteric virus infection.* 2017. **114**(4): p. E570-E579.
399. Stanifer, M.L., et al., *Critical Role of Type III Interferon in Controlling SARS-CoV-2 Infection in Human Intestinal Epithelial Cells.* Cell Reports, 2020. **32**(1).
400. Blake, R., *Investigating the early interaction between Mycobacterium avium ssp paratuberculosis and the host using a bovine enteroid system*, in *Infection and Immunity*. 2023, University of Edinburgh.
401. Puschhof, J., et al., *Intestinal organoid cocultures with microbes.* Nature Protocols, 2021. **16**(10): p. 4633-4649.
402. Schreurs, R.R.C.E., et al., *In vitro co-culture of human intestinal organoids and lamina propria-derived CD4+ T cells.* STAR Protocols, 2021. **2**(2): p. 100519.
403. Noel, G., et al., *A primary human macrophage-enteroid co-culture model to investigate mucosal gut physiology and host-pathogen interactions.* Scientific Reports, 2017. **7**(1): p. 45270.
404. Lee, K.K., et al., *Human stomach-on-a-chip with luminal flow and peristaltic-like motility.* Lab Chip, 2018. **18**(20): p. 3079-3085.
405. Fraser, R., et al., *Upper respiratory tract mucosal immunity for SARS-CoV-2 vaccines.* Trends Mol Med, 2023. **29**(4): p. 255-267.
406. Harvey, W.T., et al., *SARS-CoV-2 variants, spike mutations and immune escape.* Nature Reviews Microbiology, 2021. **19**(7): p. 409-424.
407. Shinde, V., et al., *Efficacy of NVX-CoV2373 Covid-19 Vaccine against the B.1.351 Variant.* N Engl J Med, 2021. **384**(20): p. 1899-1909.
408. Choi, A., et al., *Safety and immunogenicity of SARS-CoV-2 variant mRNA vaccine boosters in healthy adults: an interim analysis.* Nature Medicine, 2021. **27**(11): p. 2025-2031.
409. Halfmann, P.J., et al., *Multivalent S2-based vaccines provide broad protection against SARS-CoV-2 variants of concern and pangolin coronaviruses.* eBioMedicine, 2022. **86**.

410. Hay, A.J. and J.W. McCauley, *The WHO global influenza surveillance and response system (GISRS)-A future perspective.* Influenza Other Respir Viruses, 2018. **12**(5): p. 551-557.
411. Levin, E.G., et al., *Waning Immune Humoral Response to BNT162b2 Covid-19 Vaccine over 6 Months.* N Engl J Med, 2021. **385**(24): p. e84.
412. Hoffmann, M., et al., *SARS-CoV-2 variants B.1.351 and P.1 escape from neutralizing antibodies.* Cell, 2021. **184**(9): p. 2384-2393.e12.
413. Sunagar, R., A. Singh, and S. Kumar, *SARS-CoV-2: Immunity, Challenges with Current Vaccines, and a Novel Perspective on Mucosal Vaccines.* Vaccines (Basel), 2023. **11**(4).
414. Falsey, A.R., et al., *Phase 3 Safety and Efficacy of AZD1222 (ChAdOx1 nCoV-19) Covid-19 Vaccine.* 2021. **385**(25): p. 2348-2360.
415. Sadoff, J., et al., *Safety and Efficacy of Single-Dose Ad26.COV2.S Vaccine against Covid-19.* 2021. **384**(23): p. 2187-2201.
416. Baden, L.R., et al., *Efficacy and Safety of the mRNA-1273 SARS-CoV-2 Vaccine.* N Engl J Med, 2021. **384**(5): p. 403-416.
417. Polack, F.P., et al., *Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine.* N Engl J Med, 2020. **383**(27): p. 2603-2615.
418. Heath, P.T., et al., *Safety and Efficacy of NVX-CoV2373 Covid-19 Vaccine.* 2021. **385**(13): p. 1172-1183.
419. Jan, E., I. Mohr, and D. Walsh, *A Cap-to-Tail Guide to mRNA Translation Strategies in Virus-Infected Cells.* 2016. **3**(1): p. 283-307.
420. Jaafar, Z.A. and J.S. Kieft, *Viral RNA structure-based strategies to manipulate translation.* Nature Reviews Microbiology, 2019. **17**(2): p. 110-123.
421. Xia, H., et al., *Evasion of Type I Interferon by SARS-CoV-2.* Cell Reports, 2020. **33**(1).
422. Lulla, V., et al., *An upstream protein-coding region in enteroviruses modulates virus infection in gut epithelial cells.* Nature Microbiology, 2019. **4**(2): p. 280-292.
423. Rancurel, C., et al., *Overlapping genes produce proteins with unusual sequence properties and offer insight into de novo protein creation.* J Virol, 2009. **83**(20): p. 10719-36.
424. Hinnebusch, A.G., *Molecular mechanism of scanning and start codon selection in eukaryotes.* Microbiol Mol Biol Rev, 2011. **75**(3): p. 434-67, first page of table of contents.
425. Kozak, M., *Structural features in eukaryotic mRNAs that modulate the initiation of translation.* J Biol Chem, 1991. **266**(30): p. 19867-70.
426. Kozak, M., *Point mutations close to the AUG initiator codon affect the efficiency of translation of rat preproinsulin in vivo.* Nature, 1984. **308**(5956): p. 241-246.
427. Xu, H., et al., *Length of the ORF, position of the first AUG and the Kozak motif are important factors in potential dual-coding transcripts.* Cell Res, 2010. **20**(4): p. 445-57.
428. Chen, W., et al., *A novel influenza A virus mitochondrial protein that induces cell death.* Nature Medicine, 2001. **7**(12): p. 1306-1312.
429. Zamarin, D., M.B. Ortigoza, and P. Palese, *Influenza A virus PB1-F2 protein contributes to viral pathogenesis in mice.* J Virol, 2006. **80**(16): p. 7976-83.
430. Pinto, R.M., et al., *Accessory Gene Products of Influenza A Virus.* Cold Spring Harb Perspect Med, 2021. **11**(12).
431. Vera-Otarola, J., et al., *The Andes hantavirus NSs protein is expressed from the viral small mRNA by a leaky scanning mechanism.* J Virol, 2012. **86**(4): p. 2176-87.
432. Ernst, H. and A.J. Shatkin, *Reovirus hemagglutinin mRNA codes for two polypeptides in overlapping reading frames.* Proc Natl Acad Sci U S A, 1985. **82**(1): p. 48-52.
433. Hoyt, C.C., et al., *Nonstructural protein sigma1s is a determinant of reovirus virulence and influences the kinetics and severity of apoptosis induction in the heart and central nervous system.* J Virol, 2005. **79**(5): p. 2743-53.
434. Sarkar, G., et al., *Identification of a new polypeptide coded by reovirus gene S1.* 1985. **54**(3): p. 720-725.

435. Mustoe, T.A., et al., *Genetics of reovirus: Identification of the ds RNA segments encoding the polypeptides of the μ and σ size classes.* Virology, 1978. **89**(2): p. 594-604.
436. Kobayashi, T., et al., *Identification of Functional Domains in Reovirus Replication Proteins μNS and μ2.* 2009. **83**(7): p. 2892-2906.
437. Miller, C.L., et al., *Reovirus σNS Protein Localizes to Inclusions through an Association Requiring the μNS Amino Terminus.* 2003. **77**(8): p. 4566-4576.
438. Kozak, M., *A short leader sequence impairs the fidelity of initiation by eukaryotic ribosomes.* Gene Expr, 1991. **1**(2): p. 111-5.
439. Sedman, S.A., G.W. Gelembiuk, and J.E. Mertz, *Translation initiation at a downstream AUG occurs with increased efficiency when the upstream AUG is located very close to the 5' cap.* J Virol, 1990. **64**(1): p. 453-7.
440. McFadden, N., et al., *Norovirus Regulation of the Innate Immune Response and Apoptosis Occurs via the Product of the Alternative Open Reading Frame 4.* PLOS Pathogens, 2011. **7**(12): p. e1002413.
441. Williams, M.A. and R.A. Lamb, *Effect of mutations and deletions in a bicistronic mRNA on the synthesis of influenza B virus NB and NA glycoproteins.* J Virol, 1989. **63**(1): p. 28-35.
442. Kozak, M., *Pushing the limits of the scanning mechanism for initiation of translation.* Gene, 2002. **299**(1): p. 1-34.
443. Wise, H.M., et al., *Overlapping signals for translational regulation and packaging of influenza A virus segment 2.* Nucleic Acids Research, 2011. **39**(17): p. 7775-7790.
444. Herbert, T.P., I. Brierley, and T.D. Brown, *Detection of the ORF3 polypeptide of feline calicivirus in infected cells and evidence for its expression from a single, functionally bicistronic, subgenomic mRNA.* J Gen Virol, 1996. **77 (Pt 1)**: p. 123-7.
445. Luttermann, C. and G. Meyers, *A bipartite sequence motif induces translation reinitiation in feline calicivirus RNA.* J Biol Chem, 2007. **282**(10): p. 7056-65.
446. Powell, M.L., *Translational termination-reinitiation in RNA viruses.* Biochem Soc Trans, 2010. **38**(6): p. 1558-64.
447. Powell, M.L., et al., *Further characterisation of the translational termination-reinitiation signal of the influenza B virus segment 7 RNA.* PLoS One, 2011. **6**(2): p. e16822.
448. Touriol, C., et al., *Generation of protein isoform diversity by alternative initiation of translation at non-AUG codons.* Biol Cell, 2003. **95**(3-4): p. 169-78.
449. Gupta, K.C. and S. Patwardhan, *ACG, the initiator codon for a Sendai virus protein.* J Biol Chem, 1988. **263**(18): p. 8553-6.
450. Boeck, R., et al., *The parainfluenza virus type 1 P/C gene uses a very efficient GUG codon to start its C' protein.* J Virol, 1992. **66**(3): p. 1765-8.
451. Mailliot, J. and F. Martin, *Viral internal ribosomal entry sites: four classes for one goal.* 2018. **9**(2): p. e1458.
452. Pelletier, J. and N. Sonenberg, *Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA.* Nature, 1988. **334**(6180): p. 320-325.
453. Kaminski, A., et al., *Mechanism of initiation site selection promoted by the human rhinovirus 2 internal ribosome entry site.* J Virol, 2010. **84**(13): p. 6578-89.
454. Anderson, E.C., S.L. Hunt, and R.J. Jackson, *Internal initiation of translation from the human rhinovirus-2 internal ribosome entry site requires the binding of Unr to two distinct sites on the 5' untranslated region.* J Gen Virol, 2007. **88**(Pt 11): p. 3043-3052.
455. López de Quinto, S. and E. Martínez-Salas, *Involvement of the aphthovirus RNA region located between the two functional AUGs in start codon selection.* Virology, 1999. **255**(2): p. 324-36.
456. Belsham, G.J., *Divergent picornavirus IRES elements.* Virus Res, 2009. **139**(2): p. 183-92.
457. Lukavsky, P.J., et al., *Structures of two RNA domains essential for hepatitis C virus internal ribosome entry site function.* Nature Structural Biology, 2000. **7**(12): p. 1105-1110.

458. Hertz, M.I. and S.R. Thompson, *Mechanism of translation initiation by Dicistroviridae IGR IRESs.* Virology, 2011. **411**(2): p. 355-61.
459. Irigoyen, N., et al., *High-Resolution Analysis of Coronavirus Gene Expression by RNA Sequencing and Ribosome Profiling.* PLoS Pathog, 2016. **12**(2): p. e1005473.
460. Strauss, E.G., C.M. Rice, and J.H. Strauss, *Sequence coding for the alphavirus nonstructural proteins is interrupted by an opal termination codon.* Proc Natl Acad Sci U S A, 1983. **80**(17): p. 5271-5.
461. Firth, A.E., et al., *Stimulation of stop codon readthrough: frequent presence of an extended 3' RNA structural element.* Nucleic Acids Res, 2011. **39**(15): p. 6679-91.
462. Aledo, J.C., *Methionine in proteins: The Cinderella of the proteinogenic amino acids.* Protein Sci, 2019. **28**(10): p. 1785-1796.
463. Jere, K.C., et al., *Whole genome sequence analyses of three African bovine rotaviruses reveal that they emerged through multiple reassortment events between rotaviruses from different mammalian species.* Vet Microbiol, 2012. **159**(1-2): p. 245-50.
464. Lutz, A., et al., *Virus-inducible reporter genes as a tool for detecting and quantifying influenza A virus replication.* J Virol Methods, 2005. **126**(1-2): p. 13-20.
465. Sundquist, T., *Optimize Your TNT® Reticulocyte Lysate Systems Reaction.* Promega Corporation Web site.
466. Stirzaker, S.C., D. Poncet, and G.W. Both, *Sequences in rotavirus glycoprotein VP7 that mediate delayed translocation and retention of the protein in the endoplasmic reticulum.* J Cell Biol, 1990. **111**(4): p. 1343-50.
467. Welch, S.K., S.E. Crawford, and M.K. Estes, *Rotavirus SA11 genome segment 11 protein is a nonstructural phosphoprotein.* J Virol, 1989. **63**(9): p. 3974-82.
468. González, S.A. and O.R. Burrone, *Rotavirus NS26 is modified by addition of single O-linked residues of N-acetylglucosamine.* Virology, 1991. **182**(1): p. 8-16.
469. Poncet, D., et al., *In vivo and in vitro phosphorylation of rotavirus NSP5 correlates with its localization in viroplasms.* J Virol, 1997. **71**(1): p. 34-41.
470. Blackhall, J., et al., *Analysis of rotavirus nonstructural protein NSP5 phosphorylation.* J Virol, 1998. **72**(8): p. 6398-405.
471. Blackhall, J., et al., *Serine protein kinase activity associated with rotavirus phosphoprotein NSP5.* J Virol, 1997. **71**(1): p. 138-44.
472. Mumford, R.A., et al., *Protease activities present in wheat germ and rabbit reticulocyte lysates.* Biochemical and Biophysical Research Communications, 1981. **103**(2): p. 565-572.
473. Kozak, M., *Evaluation of the fidelity of initiation of translation in reticulocyte lysates from commercial sources.* Nucleic Acids Res, 1990. **18**(9): p. 2828.
474. Kobayashi, T., et al., *An improved cell-free system for picornavirus synthesis.* J Virol Methods, 2007. **142**(1-2): p. 182-8.
475. Mikami, S., et al., *A hybridoma-based in vitro translation system that efficiently synthesizes glycoproteins.* J Biotechnol, 2006. **127**(1): p. 65-78.
476. Kozak, M., *Recognition of AUG and alternative initiator codons is augmented by G in position +4 but is not generally affected by the nucleotides in positions +5 and +6.* Embo j, 1997. **16**(9): p. 2482-92.
477. Shmulevitz, M., et al., *Sequential partially overlapping gene arrangement in the tricistronic S1 genome segments of avian reovirus and Nelson Bay reovirus: implications for translation initiation.* J Virol, 2002. **76**(2): p. 609-18.
478. Chenik, M., K. Chebli, and D. Blondel, *Translation initiation at alternate in-frame AUG codons in the rabies virus phosphoprotein mRNA is mediated by a ribosomal leaky scanning mechanism.* 1995. **69**(2): p. 707-712.
479. Jayakar, H.R. and M.A. Whitt, *Identification of two additional translation products from the matrix (M) gene that contribute to vesicular stomatitis virus cytopathology.* J Virol, 2002. **76**(16): p. 8011-8.

480. de Breyne, S., R. Stalder, and J. Curran, *Intracellular processing of the Sendai virus C' protein leads to the generation of a Y protein module: Structure–functional implications.* FEBS Letters, 2005. **579**(25): p. 5685-5690.
481. Turina, M., B. Desvoyes, and K.B. Scholthof, *A gene cluster encoded by panicum mosaic virus is associated with virus movement.* Virology, 2000. **266**(1): p. 120-8.
482. Castaño, A., L. Ruiz, and C. Hernández, *Insights into the translational regulation of biologically active open reading frames of Pelargonium line pattern virus.* Virology, 2009. **386**(2): p. 417-426.
483. Kanyuka, K.V., et al., *Nucleotide sequence of shallot virus X RNA reveals a 5'-proximal cistron closely related to those of potexviruses and a unique arrangement of the 3'-proximal cistrons.* J Gen Virol, 1992. **73 (Pt 10)**: p. 2553-60.
484. Ling, R., et al., *An essential fifth coding ORF in the sobemoviruses.* Virology, 2013. **446**(1-2): p. 397-408.
485. Gao, Y., S.Q. Sun, and H.C. Guo, *Biological function of Foot-and-mouth disease virus non-structural proteins and non-coding elements.* Virology Journal, 2016. **13**: p. 107.
486. Luz, N. and E. Beck, *Interaction of a cellular 57-kilodalton protein with the internal translation initiation site of foot-and-mouth disease virus.* J Virol, 1991. **65**(12): p. 6486-94.
487. Andreev, D.E., et al., *Differential factor requirement to assemble translation initiation complexes at the alternative start codons of foot-and-mouth disease virus RNA.* Rna, 2007. **13**(8): p. 1366-74.
488. López de Quinto, S. and E. Martínez-Salas, *Involvement of the Aphthovirus RNA Region Located between the Two Functional AUGs in Start Codon Selection.* Virology, 1999. **255**(2): p. 324-336.
489. van Eyll, O. and T. Michiels, *Non-AUG-initiated internal translation of the L* protein of Theiler's virus and importance of this protein for viral persistence.* J Virol, 2002. **76**(21): p. 10665-73.
490. Loughran, G., et al., *Avoidance of reporter assay distortions from fused dual reporters.* Rna, 2017. **23**(8): p. 1285-1289.
491. Reitzer, L., *Amino Acid Synthesis*, in *Reference Module in Biomedical Sciences*. 2014, Elsevier.
492. Jablonski, S.A. and C.D. Morrow, *Mutation of the aspartic acid residues of the GDD sequence motif of poliovirus RNA-dependent RNA polymerase results in enzymes with altered metal ion requirements for activity.* 1995. **69**(3): p. 1532-1539.
493. Valverde-Garduño, V., P. Gariglio, and L. Gutiérrez, *Functional analysis of HIV-1 reverse transcriptase motif C: site-directed mutagenesis and metal cation interaction.* J Mol Evol, 1998. **47**(1): p. 73-80.
494. Vreede, F.T., T.E. Jung, and G.G. Brownlee, *Model Suggesting that Replication of Influenza Virus Is Regulated by Stabilization of Replicative Intermediates.* 2004. **78**(17): p. 9568-9572.
495. Wang, X. and S. Gillam, *Mutations in the GDD Motif of Rubella Virus Putative RNA-Dependent RNA Polymerase Affect Virus Replication.* Virology, 2001. **285**(2): p. 322-331.
496. Vázquez, A.L., J.M. Alonso, and F. Parra, *Mutation analysis of the GDD sequence motif of a calicivirus RNA-dependent RNA polymerase.* J Virol, 2000. **74**(8): p. 3888-91.
497. Huang, T.S., P. Palese, and M. Krystal, *Determination of influenza virus proteins required for genome replication.* J Virol, 1990. **64**(11): p. 5669-73.
498. Patton, J.T., et al., *Rotavirus RNA polymerase requires the core shell protein to synthesize the double-stranded RNA genome.* 1997. **71**(12): p. 9618-9626.
499. Tomar, S., et al., *Catalytic core of alphavirus nonstructural protein nsP4 possesses terminal adenylyltransferase activity.* J Virol, 2006. **80**(20): p. 9962-9.
500. Patton, J.T., *Rotavirus VP1 alone specifically binds to the 3' end of viral mRNA, but the interaction is not sufficient to initiate minus-strand synthesis.* J Virol, 1996. **70**(11): p. 7940-7.

501. Chen, D., et al., *Features of the 3'-consensus sequence of rotavirus mRNAs critical to minus strand synthesis.* Virology, 2001. **282**(2): p. 221-9.
502. McDonald, S.M. and J.T. Patton, *Rotavirus VP2 core shell regions critical for viral polymerase activation.* J Virol, 2011. **85**(7): p. 3095-105.
503. Gog, J.R., et al., *Codon conservation in the influenza A virus genome defines RNA packaging signals.* Nucleic Acids Res, 2007. **35**(6): p. 1897-907.
504. Hutchinson, E.C., et al., *Mutational analysis of cis-acting RNA signals in segment 7 of influenza A virus.* J Virol, 2008. **82**(23): p. 11869-79.
505. Patton, J.T., J. Chnaiderman, and E. Spencer, *Open Reading Frame in Rotavirus mRNA Specifically Promotes Synthesis of Double-Stranded RNA: Template Size Also Affects Replication Efficiency.* Virology, 1999. **264**(1): p. 167-180.
506. Barro, M., et al., *Identification of Sequences in Rotavirus mRNAs Important for Minus Strand Synthesis Using Antisense Oligonucleotides.* Virology, 2001. **288**(1): p. 71-80.
507. Chen, D. and J.T. Patton, *Rotavirus RNA Replication Requires a Single-Stranded 3' End for Efficient Minus-Strand Synthesis.* 1998. **72**(9): p. 7387-7396.
508. Mansell, E.A. and J.T. Patton, *Rotavirus RNA replication: VP2, but not VP6, is necessary for viral replicase activity.* J Virol, 1990. **64**(10): p. 4988-96.
509. Steger, C.L., et al., *Group A Rotavirus VP1 Polymerase and VP2 Core Shell Proteins: Intergenotypic Sequence Variation and In Vitro Functional Compatibility.* J Virol, 2019. **93**(2).
510. Ding, K., L. Nguyen, and Z.H. Zhou, *In Situ Structures of the Polymerase Complex and RNA Genome Show How Aquareovirus Transcription Machineries Respond to Uncoating.* J Virol, 2018. **92**(21).
511. Matsuo, E. and P. Roy, *Bluetongue Virus VP1 Polymerase Activity In Vitro: Template Dependency, Dinucleotide Priming and Cap Dependency.* PLOS ONE, 2011. **6**(11): p. e27702.
512. Balvay, L., et al., *Structural and functional diversity of viral IRESes.* Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms, 2009. **1789**(9): p. 542-557.
513. Sonenberg, N., Y. Svitkin, and N. Siddiqui, *Internal Ribosome Entry Site-Mediated Translation*, in *Encyclopedia of Cell Biology*, R.A. Bradshaw and P.D. Stahl, Editors. 2016, Academic Press: Waltham. p. 307-316.
514. Thompson, S.R., *So you want to know if your message has an IRES?* 2012. **3**(5): p. 697-705.

Chapter 8: Appendices

8.1 Appendix A: R script for finding AUGs and Kozak context (courtesy of Sam Lycett and Marius Diebold)

```
# R Scripts written by Sam Lycett and Marius Diebold


# Set working directory so R can access the files where they are saved
setwd("C:/Users/s1412131/Desktop/R")

# Loading packages / making them available
library("seqinr")
library("ape")
library("stringr")
library("dplyr")
library("tibble")

#############################################################################
# Part 1: Read and process data - Marius
#############################################################################

#1. To create data frame, read and assign the ‘File name’.FST file

Viral_proteins <- read.fasta(file = " File name’.FST ", seqtype = c("DNA"),
                as.string = TRUE, forceDNAtolower = FALSE,set.attributes = TRUE,
                legacy.mode = TRUE, seqonly = FALSE,strip.desc = FALSE,
                whole.header = TRUE,bfa = FALSE,
                sizeof.longlong = .Machine$sizeof.longlong,endian = "big",apply.mask = TRUE)

# 2. To read DNA and asign it to Data frame
DNA_Dataframe <- read.dna(file = " File name’.FST ",format = "fasta",skip = 0,nlines =
0,comment.char = 0,as.character = TRUE, as.matrix = FALSE)

# 3. To extract names from fasta file that reads DNA and assign it to a dataframe
DNA_name <- getName.SeqFastadna(DNA_Dataframe)
print(DNA_name)

# 4. To separate the title of the files by | and create a matrix
Name_split <- strsplit(DNA_name,split = "[|]")
print(Name_split)

# 5. To make a matrix of names for each sequence
sequence_matrix <- data.frame(matrix(unlist(Name_split), nrow=1689,
byrow=TRUE),stringsAsFactors=FALSE)
names(sequence_matrix)[1:8] <- list("Accession Number","Length", "Segment", "Genotype",
"Host", "Country", "Isolation", "Date")

# 6. To put Viral_proteins into a data frame - all in one function
Viral_proteins_sequence <-
rownames_to_column(data.frame((unlist(Viral_proteins,use.names = TRUE,))))
```

```
attributes(Viral_proteins_sequence)
# 7. Truncating column name
Truncate_1 <- data.frame(str_trunc(string=Viral_proteins_sequence$rowname,width=9,side
= "right",ellipsis = ""),Viral_proteins_sequence$X.unlist.Viral_proteins..use.names...TRUE....)

# 8. To get and rename column names
# get names of columns for use in next part
colnames(Truncate_1)
names(Truncate_1)[names(Truncate_1) ==
"str_trunc.string...Viral_proteins_sequence.rowname..width...9.."] <- "Accession Number"
names(Truncate_1)[names(Truncate_1) ==
"Viral_proteins_sequence.X.unlist.Viral_proteins..use.names...TRUE...."] <- "Sequence"

# 9. To match the accession number to sequence - correct way
Final_Table <- inner_join(sequence_matrix,Truncate_1,by=c("Accession Number" =
"Accession Number"))

#############################################################################
Use above in combination with set functions from Sam from Part 2 onwards or use this full
code from Sam below
#############################################################################

#####################################################################
# Import other utility code
#####################################################################

# 26 May 2022 - utility code
# setwd to the data path
# on my computer this is
setwd("Documents/Ola/")
# compile custom function from source (put getEl.R into this folder as well)
# else manually open file getEl.R and do run all
Rpath <- ""
source(paste0(Rpath,"getEl.R"))

#####################################################################
# Define custom functions - DO NOT CHANGE THESE
#####################################################################

# Custom functions
# this works on read.fasta sequence string
find_motif <- function(seq, motif="atg",return.NULL=TRUE) {
  pos <- gregexpr(tolower(motif), tolower(seq))[[1]]
  return(pos)
}

# use the output of find_motif in this function along with the list of sequence names
make_motif_list_to_matrix <- function(aug_list,taxa) {
  nseqs       <- length(aug_list)
  #npos_per_seq <- unlist(lapply(aug_list,length))
  all_pos     <- setdiff(sort(unique(unlist(aug_list))),-1) # 26 May remove -1 because this
means not present
  all_pos_matrix <- matrix(0, nseqs, length(all_pos))
  rownames(all_pos_matrix) <- taxa
  colnames(all_pos_matrix) <- paste0("Site",all_pos)
```

```
  for (j in 1:nseqs) {
    kk <- match(aug_list[[j]],all_pos)
    all_pos_matrix[j,kk] <- 1
  }
  return(all_pos_matrix)
}

# cn = column names
get_coding_frame_from_site_number <- function(cn, txt="Site") {
  cn <- gsub(txt,"",cn)
  cn <- as.integer(cn)
  codingFrame <- ((cn-1) %% 3)+1
  return(codingFrame)
}

# 26 May 2022
plot_aug_matrix <- function(aug_pos_matrix,cols=brewer.pal(8,"Blues")[8:1]) {
        # sname=sname, ttxt=gsub(".fas","",sname),
        #image(t(aug_pos_matrix), col=cols, axes=FALSE, xlab="Position",
ylab="Sequences")
        #nVarSites <- length(aug_pos_matrix[1,])
        #xpos <- (0:(nVarSites-1))/nVarSites
        #xlab <- gsub("Site","",colnames(aug_pos_matrix))
        #axis(1,at=xpos,lab=xlab)
        #title(ttxt)
        temp <- aug_pos_matrix
        rownames(temp) <- NULL
        heatmap(temp, Rowv=NA, Colv=NA, col=cols, xlab="Site", margins=c(5,0))
}

#############################################################################
# Pre-amble - ONLY DO THIS ONCE TO CONVERT TO FASTA - MIGHT NOT BE
NECESSARY
#############################################################################

#dataPath <- ""
#setwd(dataPath)
#sname <- "RVA VP1(dfs).fasta"

doConvert <- FALSE
if (doConvert) {
        # change directory to data directory
        # in this case Ola

        # convert to new file name
        dataPath <- "Ola/Alignments from Virus Variation_NCBI/"
        sname    <- paste0(dataPath,"Exported_NSP5_partially cleaned.txt")
        seqs  <- read.dna(sname,format="fasta",as.matrix=FALSE)

        # new folder and name
        dataPath <- "Ola/Alignments_VV/"
        sname    <- "NSP5.fas"
        dir.create(dataPath)
        write.dna(seqs, file=paste0(dataPath,sname), format="fasta", nbcol=-1, colsep="")
}
```

```
#########################################################################
# MAIN SCRIPT
#########################################################################

#########################################################################
# Set up paths and data - you may CHANGE THIS for different segments / data files
#########################################################################
# define path where the data is
dataPath <- "Alignments_VV/"

# CHANGE THIS FILENAME to be the fasta file you want to process
# choose segment file to process (in this case NSP5)
sname    <- "NSP5.fas"

# define output path
outPath  <- "Alignments_VV_results/"

# create path - only really need to do this once but it wont hurt to do it again
if (!dir.exists(outPath)) {
  dir.create(outPath)
}

#########################################################################
# Part 1: Read and process data - AUG - you DO NOT need to change this
#########################################################################

print("** START **")

# read fasta format sequences
#seqs  <- read.fasta(sname,seqtype="DNA", whole.header = TRUE, as.string=TRUE)
seqs  <- read.dna(paste0(dataPath,sname),format="fasta",as.matrix=FALSE)
nseqs <- length(seqs)
taxa  <- attributes(seqs)$name
print(paste("Number of sequences:",length(taxa)))

# find all aug - will come out a list
aug_list <- vector("list",nseqs)
for (j in 1:nseqs) {
  seqtxt       <- paste(unlist(as.character(seqs[j])),collapse="")
  aug_list[[j]] <- find_motif(seqtxt, motif="atg")
}

# use custom function to make to matrix
aug_pos_matrix <- make_motif_list_to_matrix(aug_list,taxa)

# write data to file
write.csv(aug_pos_matrix,
file=paste0(outPath,gsub(".fas","",sname),"_aug_pos_matrix.csv"))

# display
imageName <-
paste0(outPath,gsub(".fas","",sname),"_aug_pos_matrix_simple_image_plot.png")
png(file=imageName, height=6*300, width=6*300, res=300)
plot_aug_matrix(aug_pos_matrix)
```

```
dev.off()
print(imageName)

#fract_aug_per_site     <- apply(aug_pos_matrix,2,sum)/length(aug_pos_matrix[,1])

#############################################################################
# Part 2: Read and process data Strong & Moderate Kosak sequences - you DO NOT need
to change this
#############################################################################

npos        <- length(unlist(as.character(seqs[j])))
nseqs       <- length(seqs)

# Strong and Moderate individually - same method as above
# define motifs
strong_kosak <- "[ag]..atgg"
mod_kosak    <- "[ag]..atga"

# make lists
strong_list  <- vector("list",nseqs)
mod_list     <- vector("list",nseqs)

# populate lists
for (j in 1:nseqs) {
  seqtxt          <- paste(unlist(as.character(seqs[j])),collapse="")
  mod_list[[j]]    <- find_motif(seqtxt, motif=mod_kosak)
  strong_list[[j]] <- find_motif(seqtxt, motif=strong_kosak)
}

# use custom function to make to matrix
strong_pos_matrix    <- make_motif_list_to_matrix(strong_list,taxa)
mod_pos_matrix       <- make_motif_list_to_matrix(mod_list,taxa)

# write data to file
write.csv(strong_pos_matrix,
file=paste0(outPath,gsub(".fas","",sname),"_strong_pos_matrix.csv"))
write.csv(mod_pos_matrix,
file=paste0(outPath,gsub(".fas","",sname),"_mod_pos_matrix.csv"))

# display
imageName                                                                    <-
paste0(outPath,gsub(".fas","",sname),"_strong_pos_matrix_simple_image_plot.png")
png(file=imageName, height=6*300, width=6*300, res=300)
plot_aug_matrix(strong_pos_matrix)
dev.off()
print(imageName)

imageName                                                                    <-
paste0(outPath,gsub(".fas","",sname),"_mod_pos_matrix_simple_image_plot.png")
png(file=imageName, height=6*300, width=6*300, res=300)
plot_aug_matrix(mod_pos_matrix)
dev.off()
print(imageName)

# split to coding frames
```

```
strong_pos_coding_frame <- get_coding_frame_from_site_number(
colnames(strong_pos_matrix) )
mod_pos_coding_frame <- get_coding_frame_from_site_number(
colnames(mod_pos_matrix) )

# write separate coding frames to files
for (cf in 1:3) {
  strong_in_frame_matrix  <- strong_pos_matrix[,which(strong_pos_coding_frame==cf)]
  mod_in_frame_matrix     <- mod_pos_matrix[,which(mod_pos_coding_frame==cf)]

  write.csv(strong_in_frame_matrix,
file=paste0(outPath,gsub(".fas","",sname),"_strong_in_frame_matrix_codingpos",cf,".csv"))
  write.csv(mod_in_frame_matrix,
file=paste0(outPath,gsub(".fas","",sname),"_mod_in_frame_matrix_codingpos",cf,".csv"))
}

#############################################################################
# Part 3: Read and process data - Kosak sequences all in one - you DO NOT need to change
this
# !! except please check the weak def !!
#############################################################################

print("** Part 3 - AUG, weak, moderate, strong **")

# all in one
strong_kosak <- "[ag]..atgg"
mod_kosak    <- "[ag]..atga"
weak_kosak   <- "...atg[ag]" # what is a weak kosak ? just guessing here - this needs checking
motif_list   <- c("atg",weak_kosak,mod_kosak,strong_kosak)
motif_name   <- c("AUG","Weak","Moderate","Strong")
# 1 = aug, 2 = weak, 3 = mod, 4 = strong

# position normally reports as the first matching position in the motif
# but kosak sequences start 3 back from AUG so use offsets
# set all these to 0 if not want to use
motif_offset <- c(0,    3,      3,         3)

# use whole position matrix
# put 1 if AUG only; 2 if weak, 3 if moderate; 4 if strong
# note that the strongs will over-ride the moderate, weak and AUGs etc
# so the AUGs in here are the ones which are only AUG, and not also weak, moderate or
strong
kosak_matrix          <- matrix(0, nseqs, npos)
rownames(kosak_matrix) <- taxa
for (j in 1:nseqs) {
  seqtxt          <- paste(unlist(as.character(seqs[j])),collapse="")

  for (k in 1:length(motif_list)) {
    found_pos <- find_motif(seqtxt, motif=motif_list[k])
    if (length(found_pos)>0) {
      if (found_pos[1]!=-1) {
        kosak_matrix[j,found_pos-motif_offset[k]] <- k
      }
    }
  }
```

```
}

kcols <- rainbow(4,s=0.7,v=0.7)[4:1]
plot_aug_matrix(kosak_matrix, cols=c("grey80",kcols))


# write data to file
write.csv(kosak_matrix,
file=paste0(outPath,gsub(".fas","",sname),"_kosak_matrix_1234.csv"))

# display
imageName                                                                    <-
paste0(outPath,gsub(".fas","",sname),"_kosak_matrix_1234_image_plot.png")
png(file=imageName, height=6*300, width=6*300, res=300)
plot_aug_matrix(kosak_matrix, cols=c("grey80",kcols))
dev.off()
print(imageName)

logName <- paste0(outPath,gsub(".fas","",sname),"_kosak_summary.txt")
write("** 2022-03-14_find_motif.R **", file=logName, append=FALSE)
write(paste("dataPath=",dataPath), file=logName, append=TRUE)
write(paste("outPath=",outPath), file=logName, append=TRUE)
write(paste("sname=",sname), file=logName, append=TRUE)
write(paste("nseqs=",nseqs), file=logName, append=TRUE)
write(paste("npos=",npos), file=logName, append=TRUE)
write("----------------------", file=logName, append=TRUE)
write(paste(1:length(motif_name),"=",motif_name,"=",motif_list,";          offset=",motif_offset),
file=logName, append=TRUE)
write("----------------------", file=logName, append=TRUE)
write("List of sites found in whole file:", file=logName, append=TRUE)
# checking site lists
for (k in 1:4) {
  pos <- which(apply(kosak_matrix==k,2,sum)>0)
  txt <-paste(motif_name[k],":",paste(pos,collapse=" "))
  print(txt)
  write(txt, file=logName, append=TRUE)
}
write("----------------------", file=logName, append=TRUE)


#########################################################################
# Part 4: Read and process data by Host - you DO NOT need to change this
#########################################################################

print("** Part 4 - By Host **")

# get host names from taxa labels, these are in position 5 - using custom function getEl.R (see
at top)
all_host  <- apply(as.matrix(taxa), 1, getEl, ind=5, sep="\\|")

# TO CHECK
# grouping hosts - please check these and alter if nec
umainHost <- c("Homo sapiens","Sus scrofa","Bos taurus","Gallus gallus")
host      <- all_host
jj        <- which(!is.finite(match(host,umainHost)))
```

```
host[jj]  <- "Other"
uhost     <- c(umainHost,"Other")
host      <- factor(host, levels=uhost)
htbl      <- table(host)

write("** By Host **", file=logName, append=TRUE)
write("Host composition:", file=logName, append=TRUE)
write(paste(names(htbl),"=",htbl), file=logName, append=TRUE)

dateTxt   <- apply(as.matrix(taxa), 1, getEl, ind=1, fromEnd=TRUE, sep="\\|")
year      <- apply(as.matrix(dateTxt), 1, getEl, ind=1, sep="/")
year[which(dateTxt=="")]      <- "Unknown"
year[which(dateTxt=="feces")] <- "Unknown"

year_host_tbl <- table(year,host)
write.csv(year_host_host, file=paste0(outPath,gsub(".fas","",sname),"_year_host_tbl.csv"))

## CHECK THIS - set best_pos to be the sites you are particularly interested in
all_pos         <- which(apply(kosak_matrix>0,2,sum)>0)
best_pos            <- which(apply(kosak_matrix>0,2,sum)>(0.3*nseqs)) # at least 30% of the
sequences have this position

#human_pval      <- array(1,length(all_pos))
#for (i in 1:length(all_pos)) {
#      human_pval[i]  <-  fisher.test(  table(host=="Homo  sapiens",kosak_matrix[,all_pos[i]])
)$p.value
#}

write(paste("Plotting at positions:",paste(best_pos)), file=logName, append=TRUE)

kcols <- rainbow(4,s=0.7,v=0.7)[4:1]

for (h in 1:length(uhost)) {
  hinds          <- which(host==uhost[h])
  hk_matrix      <- kosak_matrix[hinds,]
  fract_conserve <- matrix(0, length(motif_name), npos)
  rownames(fract_conserve) <- motif_name
  for (k in 1:length(motif_name)) {
    fract_conserve[k,] <- apply(hk_matrix==k, 2, mean)
  }
  fp <- fract_conserve[,best_pos]
  colnames(fp) <- best_pos
  fp <- fp*100

  imageName                                                                                <-
paste0(outPath,gsub(".fas","",sname),"_",uhost[h],"_selected_kosak_barplot.png")
  png(file=imageName, height=6*300, width=12*300, res=300)
  barplot(fp, col=kcols,border=FALSE,ylim=c(0,110), xlab="Site", ylab="Sequences %")
  legend("top",motif_name,pch=22,pt.bg=kcols,col=kcols,bty="n", horiz=TRUE)
  title(paste0(uhost[h]," n=",length(hinds)))
  dev.off()
  print(imageName)
}
write("----------------------", file=logName, append=TRUE)
```

```
print(paste("Log of info to:",logName))
print("** END **")
```

8.2 Appendix B: AlphaFold parameters (courtesy of Dr Barbara Shih)

```
#!/bin/bash
# Grid Engine options (lines prefixed with #$)
#$ -l h_rt=200:00:00
# Set working directory to the directory where the job is submitted from:
#$ -cwd
#$ -pe sharedmem 2
#$ -l h_vmem=1000G

# Initialise the environment modules and load CUDA
. /etc/profile.d/modules.sh

module add singularity/3.5.3
module add roslin/python/3.6.8
module add igmm/apps/alphafold/2.0.0

FASTA_LIST=${fasta_list} # file with file names of fasta in the IN_FASTA_DIR
IN_FASTA_DIR=${in_fasta_dir} # Please use absolute path. This is the folder with the input fasta files. I use the directory name as sub directory for output
MODEL_PRESENT=${model_present} # Please set this to monomer, monomer_casp14, monomer_ptm, or multimer

IN_DIR_BASENAME=$(basename $IN_FASTA_DIR)
OUT_DIR=${out_dir}/${IN_DIR_BASENAME} # Please use absolute path

mkdir -p $(dirname $OUT_DIR) $OUT_DIR

IN_FASTA=$(awk "NR==$SGE_TASK_ID" $FASTA_LIST)
#IN_FASTA=${IN_FASTA_DIR}/${IN_FASTA}

# please use full path
DATABASE_DIR="/exports/cmvm/eddie/eb/groups/EEID_Mareks_IBV/members/roslin_bioinformatics/2021-07-23-_9707_EEID_AlphaFold_setup/databases"
ALPHAFOLD_SINGULARITY="/exports/cmvm/eddie/eb/groups/EEID_Mareks_IBV/members/roslin_bioinformatics/2021-07-23-_9707_EEID_AlphaFold_setup/alphafold/alphafold2.sif"
```

```
ALPHA_FOLD_GIT="/exports/cmvm/eddie/eb/groups/EEID_Mareks_IBV/members/roslin_bioinformatics/2021-07-23-_9707_EEID_AlphaFold_setup/alphafold"
MAX_TEMPLATE_DATE="2021-08-11" # change this to a consistant date for your own project

ALPHAFOLD_MODELS_DIR="${DATABASE_DIR}/params"
WORKING_DIR=$(pwd)

singularity run \
 -B $DATABASE_DIR:/data \
 -B $IN_FASTA_DIR:/fasta \
 -B $ALPHAFOLD_MODELS_DIR \
 -B .:/etc \
 --pwd $ALPHA_FOLD_GIT $ALPHAFOLD_SINGULARITY \
 --fasta_paths=/fasta/${IN_FASTA} \
 --uniref90_database_path=/data/uniref90/uniref90.fasta \
 --data_dir=/data \
 --mgnify_database_path=/data/mgnify/mgy_clusters_2018_12.fa \
 --
bfd_database_path=/data/bfd/bfd_metaclust_clu_complete_id30_c90_final_seq.sorted_opt \
 --uniclust30_database_path=/data/uniclust30/uniclust30_2018_08/uniclust30_2018_08 \
 --pdb70_database_path=/data/pdb70/pdb70 \
 --template_mmcif_dir=/data/pdb_mmcif/mmcif_files \
 --obsolete_pdbs_path=/data/pdb_mmcif/obsolete.dat \
 --max_template_date=$MAX_TEMPLATE_DATE \
 --model_preset="${MODEL_PRESENT}" \
 --output_dir=$OUT_DIR \
 --db_preset=full_dbs
```

8.3 Appendix C: Identified AUG codons per frame for RV strains RF and SA11.

RF

Segment	Gene	AUG#	Positions (start .. stop)	Length (nt / aa)	Frame	Position -3	Position +4	Kozak
1	**VP1**	**AUG1 (Canonical)**	19..3285	3267 | 1088	1	**A**	**G**	Strong
		AUG2	209..286	78 | 25	2	**G**	T	Moderate
		AUG3	214..3285	3072 | 1023	1	**G**	**G**	Strong
		AUG4	221..286	66 | 21	2	**A**	C	Moderate
		AUG5	406..3285	2880 | 959	1	T	**G**	Moderate
		AUG6	454..3285	2832 | 943	1	**G**	T	Moderate
		AUG7	604..3285	2682 | 893	1	**G**	A	Moderate
		AUG8	655..3285	2631 | 876	1	**G**	C	Moderate
		AUG9	901..3285	2385 | 794	1	**A**	A	Moderate
		AUG10	934..3285	2352 | 783	1	**A**	A	Moderate
		AUG11	982..3285	2304 | 767	1	T	**G**	Moderate
		AUG12	1033..3285	2253 | 750	1	**A**	T	Moderate
		AUG13	1093..3285	2193 | 730	1	**G**	T	Moderate
		AUG14	1111..3285	2175 | 724	1	**A**	C	Moderate
		AUG15	1138..3285	2148 | 715	1	**A**	C	Moderate
		AUG16	1160..1243	84 | 27	2	**A**	A	Moderate
		AUG17	1163..1243	81 | 26	2	**A**	A	Moderate
		AUG18	1282..3285	2004 | 667	1	**A**	C	Moderate
		AUG19	1291..3285	1995 | 664	1	**G**	**G**	Strong
		AUG20	1300..3285	1986 | 661	1	**G**	**G**	Strong
		AUG21	1456..3285	1830 | 609	1	**A**	C	Moderate
		AUG22	1487..1633	147 | 48	2	**A**	C	Moderate
		AUG23	1526..1633	108 | 35	2	C	**G**	Moderate

		AUG24	1532..1633	102 | 33	2	G	T	Moderate
		AUG25	1561..3285	1725 | 574	1	T	G	Moderate
		AUG26	1636..3285	1650 | 549	1	A	G	Strong
		AUG27	1648..3285	1638 | 545	1	G	C	Moderate
		AUG28	1660..3285	1626 | 541	1	A	A	Moderate
		AUG29	1723..3285	1563 | 520	1	C	G	Moderate
		AUG30	1957..3285	1329 | 443	1	C	G	Moderate
		AUG31	2227..3285	1059 | 352	1	A	C	Moderate
		AUG32	2342..2413	72 | 23	2	A	G	Strong
		AUG33	2596..3285	690 | 229	1	A	C	Moderate
		AUG34	3205..3285	81 | 26	1	G	A	Moderate
		AUG35	3216..>3302	87 | 28	3	A	G	Strong
2	**VP2**	**AUG1 (Canonical)**	17..2659	2643 | 880	2	T	G	Moderate
		AUG2	66..161	96 | 31	3	A	A	Moderate
		AUG4	87..161	75 | 24	3	A	A	Moderate
		AUG5	116..2659	2544 | 847	2	A	C	Moderate
		AUG6	545..2659	2115 | 704	2	G	C	Moderate
		AUG7	572..2659	2088 | 695	2	G	G	Strong
		AUG8	600..680	81 | 26	3	G	C	Moderate
		AUG9	698..2659	1962 | 653	2	G	A	Moderate
		AUG10	878..2659	1782 | 593	2	A	G	Strong
		AUG11	1061..2659	1599 | 532	2	A	T	Moderate
		AUG12	1172..2659	1488 | 495	2	G	T	Moderate
		AUG13	1190..2659	1470 | 489	2	A	T	Moderate
		AUG15	1238..2659	1422 | 473	2	G	T	Moderate
		AUG16	1268..2659	1392 | 463	2	G	T	Moderate
		AUG18	1340..2659	1320 | 439	2	G	C	Moderate
		AUG20	1535..2659	1125 | 374	2	T	G	Moderate

		AUG22	1577..2659	1083 | 360	2	**A**	C	Moderate
		AUG23	1688..2659	972 | 323	2	C	**G**	Moderate
		AUG24	1693..1830	138 | 45	1	**G**	T	Moderate
		AUG25	1703..2659	957 | 318	2	**A**	A	Moderate
		AUG26	1709..2659	951 | 316	2	**A**	C	Moderate
		AUG28	1785..1889	105 | 34	3	**G**	C	Moderate
		AUG29	1928..2659	732 | 243	2	**A**	A	Moderate
		AUG32	2081..2659	579 | 192	2	**A**	**G**	Strong
		AUG33	2144..2659	516 | 171	2	**G**	C	Moderate
		AUG35	2264..2659	396 | 131	2	**A**	C	Moderate
		AUG36	2460..2543	84 | 27	3	**A**	A	Moderate
3	**VP3**	**AUG1 (Canonical)**	50..2557	2508 | 835	2	**A**	A	Moderate
		AUG5	444..512	69 | 22	3	**G**	A	Moderate
		AUG6	578..2557	1980 | 559	2	**A**	A	Moderate
		AUG8	833..2557	1725 | 574	2	**A**	T	Moderate
		AUG9	872..2557	1686 | 561	2	**G**	T	Moderate
		AUG10	899..2557	1659 | 552	2	T	**G**	Moderate
		AUG11	1199..2557	1359 | 452	2	**A**	A	Moderate
		AUG12	1208..2557	1350 | 449	2	**G**	**G**	Strong
		AUG13	1286..2557	1272 | 423	2	**A**	**G**	Strong
		AUG14	1413..1505	93 | 30	3	**G**	A	Moderate
		AUG15	1419..1505	87 | 28	3	**A**	T	Moderate
		AUG16	1536..1613	78 | 25	3	**A**	A	Moderate
		AUG17	1640..2557	918 | 305	2	**G**	T	Moderate
		AUG22	2237..2557	321 | 106	2	T	**G**	Moderate
		AUG25	2291..2557	267 | 88	2	**A**	**G**	Strong
		AUG26	2426..2557	132 | 43	2	**A**	T	Moderate

4	**VP4**	**AUG1 (Canonical)**	10..2340	2331 | 776	1	**A**	**G**	Strong
		AUG5	511..2340	1830 | 609	1	**G**	A	Moderate
		AUG6	518..601	84 | 27	2	**A**	**G**	Strong
		AUG11	604..2340	1737 | 578	1	**A**	A	Moderate
		AUG12	802..2340	1539 | 512	1	**G**	C	Moderate
		AUG13	1102..2340	1239 | 412	1	**A**	**G**	Strong
		AUG15	1144..2340	1197 | 398	1	**G**	T	Moderate
		AUG11	1198..2340	1143 | 380	1	**G**	A	Moderate
		AUG18	1549..2340	792 | 263	1	**G**	T	Moderate
		AUG19	1588..2340	753 | 250	1	**G**	T	Moderate
		AUG21	1639..2340	702 | 233	1	T	**G**	Moderate
		AUG22	1654..2340	687 | 228	1	**G**	A	Moderate
		AUG23	1742..1819	78 | 25	2	**G**	C	Moderate
		AUG24	1879..2340	462 | 153	1	**G**	**G**	Strong
		AUG25	1900..2340	441 | 146	1	**G**	A	Moderate
		AUG28	2069..2143	75 | 24	2	C	**G**	Moderate
5	**NSP1**	**AUG1 (Canonical)**	32..1507	1476 | 491	2	**G**	**G**	Strong
		AUG3	151..240	90 | 29	1	**A**	**G**	Strong
		AUG4	278..1507	1230 | 409	2	**A**	C	Moderate
		AUG5	332..1507	1176 | 391	2	**A**	T	Moderate
		AUG7	489..566	78 | 25	3	**A**	T	Moderate
		AUG8	521..1507	987 | 328	2	T	**G**	Moderate
		AUG9	611..1507	897 | 298	2	**A**	T	Moderate
		AUG10	728..1507	780 | 259	2	C	**G**	Moderate
		AUG12	857..1507	651 | 216	2	A	A	Moderate
		AUG13	1076..1507	432 | 143	2	**G**	**G**	Strong
		AUG15	1107..1172	66 | 21	3	C	**G**	Moderate
		AUG16	1208..1507	300 | 99	2	**G**	T	Moderate

		AUG17	1301..1507	207 | 68	2	G	C	Moderate
		AUG18	1310..1507	198 | 65	2	A	G	Strong
		AUG19	1340..1507	168 | 55	2	A	A	Moderate
		AUG20	1370..1507	138 | 45	2	G	A	Moderate
6	**VP6**	**AUG1 (Canonical)**	24..1217	1194 | 397	3	A	G	Strong
		AUG4	144..1217	1074 | 357	3	A	A	Moderate
		AUG5	148..237	90 | 29	1	T	G	Moderate
		AUG6	154..237	84 | 27	1	G	A	Moderate
		AUG7	304..375	72 | 23	1	A	T	Moderate
		AUG8	312..1217	906 | 301	3	T	G	Moderate
		AUG9	321..1217	897 | 298	3	G	G	Strong
		AUG10	552..1217	666 | 221	3	T	G	Moderate
		AUG11	561..1217	657 | 218	3	A	T	Moderate
		AUG13	921..1217	297 | 98	3	A	A	Moderate
		AUG14	1047..1217	171 | 56	3	A	C	Moderate
		AUG15	1116..1217	102 | 33	3	G	A	Moderate
7	**VP7**	**AUG1 (Canonical)**	49..1029	981 | 326	1	T	T	Weak
		AUG3	136..1029	894 | 297	1	A	G	Strong
		AUG4	235..1029	795 | 264	1	T	G	Moderate
		AUG5	339..404	66 | 21	3	T	G	Moderate
		AUG7	502..1029	528 | 175	1	G	T	Moderate
		AUG8	550..1029	480 | 159	1	C	G	Moderate
		AUG9	898..1029	132 | 43	1	A	A	Moderate
		AUG10	901..1029	129 | 42	1	A	C	Moderate
		AUG11	921..992	72 | 23	3	A	G	Strong
8	**NSP2**	**AUG1 (Canonical)**	47..1000	954 | 317	2	G	G	Strong

		AUG3	192..257	66 | 21	3	T	G	Moderate
		AUG4	254..1000	747 | 248	2	G	A	Moderate
		AUG6	542..1000	459 | 152	2	A	T	Moderate
		AUG9	923..1000	78 | 25	2	A	A	Moderate
9	**NSP3**	**AUG1 (canonical)**	26..967	942 | 313	2	T	C	Weak
		AUG2	35..967	933 | 310	2	A	G	Strong
		AUG3	53..967	915 | 304	2	C	G	Moderate
		AUG4	122..967	846 | 281	2	C	G	Moderate
		AUG5	185..967	783 | 260	2	G	G	Strong
		AUG6	249..317	69 | 22	3	T	G	Moderate
		AUG7	287..967	681 | 226	2	T	G	Moderate
		AUG8	341..967	627 | 208	2	A	A	Moderate
		AUG9	344..967	624 | 207	2	A	C	Moderate
		AUG10	374..967	594 | 197	2	A	A	Moderate
		AUG13	512..967	456 | 151	2	A	G	Strong
		AUG15	641..967	327 | 108	2	A	T	Moderate
		AUG16	779..967	189 | 62	2	T	G	Moderate
10	**NSP4**	**AUG1 (Canonical)**	42..569	528 | 175	3	A	G	Strong
		AUG3	129..569	441 | 146	3	G	G	Strong
		AUG4	201..569	369 | 122	3	A	A	Moderate
		AUG5	336..569	234 | 77	3	C	G	Moderate
		AUG6	357..569	213 | 70	3	G	A	Moderate
		AUG7	375..569	195 | 64	3	G	A	Moderate
		AUG9	471..569	99 | 32	3	G	A	Moderate
11	**NSP5**	**AUG1 (Canonical)**	22..618	597 | 198	1	G	T	Moderate
		AUG7	292..618	327 | 108	1	T	G	Moderate

	AUG8	373..618	246 | 81	1	**A**	A	Moderate
NSP6	AUG2	80..376	297 | 98	2	**A**	A	Moderate

SA11

Segment	Gene	AUG#	Positions (start .. stop)	Length (nt / aa)	Frame	Position -3	Position +4	Kozak
1	**VP1**	**AUG1 (Canonical)**	19..3285	3267 | 1088	1	**A**	**G**	Strong
		AUG4	221..289	69 | 22	2	**A**	C	Moderate
		AUG5	406..3285	2880 | 959	1	C	**G**	Moderate
		AUG6	454..3285	2832 | 943	1	**G**	T	Moderate
		AUG7	604..3285	2682 | 893	1	**G**	A	Moderate
		AUG8	655..3285	2631 | 876	1	**G**	T	Moderate
		AUG9	901..3285	2385 | 794	1	**A**	A	Moderate
		AUG10	934..3285	2352 | 783	1	**A**	A	Moderate
		AUG11	982..3285	2304 | 767	1	T	**G**	Moderate
		AUG12	1033..3285	2253 | 750	1	**A**	C	Moderate
		AUG13	1093..3285	2193 | 730	1	**G**	T	Moderate
		AUG14	1111..3285	2175 | 724	1	**A**	T	Moderate
		AUG15	1138..3285	2148 | 715	1	**A**	C	Moderate
		AUG18	1282..3285	2004 | 667	1	**A**	C	Moderate
		AUG19	1291..3285	1995 | 664	1	**G**	**G**	Strong
		AUG20	1300..3285	1986 | 661	1	**G**	**G**	Strong
		AUG21	1456..3285	1830 | 609	1	**A**	T	Moderate
		AUG25	1561..3285	1725 | 574	1	**A**	**G**	Strong
		AUG26	1636..3285	1650 | 549	1	**A**	**G**	Strong
		AUG28	1660..3285	1626 | 541	1	**A**	A	Moderate
		AUG29	1723..3285	1563 | 520	1	C	**G**	Moderate
		AUG30	2111..2176	66 | 21	2	**A**	A	Moderate
		AUG31	2227..3285	1059 | 352	1	**A**	C	Moderate

		AUG33	2596..3285	690 | 229	1	**A**	T	Moderate
		AUG33	2684..2749	66 | 21	2	**G**	C	Moderate
		AUG34	3205..3285	81 | 26	1	**G**	A	Moderate
2	**VP2**	**AUG1 (canonical)**	17..2665	2649 | 882	2	T	**G**	Moderate
		AUG2	66..164	99 | 32	3	**A**	A	Moderate
		AUG4	285..347	63 | 20	3	**A**	A	Moderate
		AUG6	545..2665	2121 | 706	2	T	**G**	Moderate
		AUG7	551..2665	2115 | 704	2	**G**	C	Moderate
		AUG8	578..2665	2088 | 695	2	**G**	**G**	Strong
		AUG9	704..2665	1962 | 653	2	**G**	A	Moderate
		AUG10	884..2665	1782 | 593	2	**A**	**G**	Strong
		AUG11	1067..2665	1599 | 532	2	**A**	T	Moderate
		AUG12	1178..2665	1488 | 495	2	**G**	T	Moderate
		AUG13	1196..2665	1470 |489	2	**A**	T	Moderate
		AUG15	1244..2665	1422 | 473	2	**G**	T	Moderate
		AUG16	1274..2665	1392 | 463	2	**G**	T	Moderate
		AUG17	1300..1365	66 | 21	1	**A**	T	Moderate
		AUG18	1346..2665	1320 | 439	2	**G**	C	Moderate
		AUG19	1355..2665	1311 | 436	2	**A**	C	Moderate
		AUG20	1541..2665	1125 | 374	2	T	**G**	Moderate
		AUG22	1583..2665	1083 | 360	2	**A**	C	Moderate
		AUG23	1694..2665	972 | 323	2	T	**G**	Moderate
		AUG26	1709..2665	957 | 318	2	**A**	A	Moderate
		AUG27	1715..2665	951 | 316	2	**A**	C	Moderate
		AUG28	1774..1836	63 | 20	1	**A**	T	Moderate
		AUG29	1934..2665	732 | 243	2	**A**	A	Moderate
		AUG32	2081..2665	585 | 194	2	**G**	A	Moderate
		AUG33	2087..2665	579 | 192	2	**A**	G	Moderate

		AUG34	2150..2665	516 | 171	2	G	C	Moderate
		AUG35	2171..2665	495 | 164	2	G	T	Moderate
		AUG36	2270..2665	396 | 131	2	A	C	Moderate
		AUG36	2466..2549	84 | 27	3	A	A	Moderate
3	**VP3**	**AUG1 (Canonical)**	50..2557	2508 | 835	2	A	A	Moderate
		AUG4	438..509	72 | 23	3	A	G	Strong
		AUG6	578..2557	1980 | 559	2	A	A	Moderate
		AUG8	833..2557	1725 | 574	2	A	C	Moderate
		AUG11	1199..2557	1359 | 452	2	A	A	Moderate
		AUG12	1208..2557	1350 | 449	2	G	G	Strong
		AUG16	1536..1613	78 | 25	3	A	A	Moderate
		AUG22	2237..2557	321 | 106	2	A	G	Strong
		AUG23	2270..2557	288 | 95	2	A	A	Moderate
		AUG24	2277..2342	66 | 21	3	A	A	Moderate
		AUG25	2291..2557	267 | 88	2	A	G	Strong
		AUG26	2426..2557	132 | 43	2	A	T	Moderate
		AUG27	2462..2557	96 | 31	2	A	A	Moderate
4	**VP4**	**AUG1 (Canonical)**	10..2340	2331 | 776	1	A	G	Strong
		AUG3	104..283	180 | 59	2	A	T	Moderate
		AUG4	143..283	141 | 46	2	G	C	Moderate
		AUG6	206..283	78 | 25	2	T	G	Moderate
		AUG5	511..2340	1830 | 609	1	G	A	Moderate
		AUG9	521..607	87 | 28	2	A	A	Moderate
		AUG11	604..2340	1737 | 578	1	A	A	Moderate
		AUG12	802..2340	1539 | 512	1	G	C	Moderate

		AUG13	971..1042	72 | 23	2	T	G	Moderate
		AUG15	1144..2340	1197 | 398	1	G	T	Moderate
		AUG16	1188..1289	102 | 33	3	A	G	Strong
		AUG17	1352..1429	78 | 25	2	T	G	Moderate
		AUG18	1549..2340	792 | 263	1	G	T	Moderate
		AUG19	1588..2340	753 | 250	1	G	T	Moderate
		AUG21	1639..2340	702 | 233	1	T	G	Moderate
		AUG22	1654..2340	687 | 228	1	G	A	Moderate
		AUG24	1879..2340	462 | 153	1	G	G	Strong
		AUG25	1900..2340	441 | 146	1	G	A	Moderate
		AUG26	2045..2143	99 | 32	2	A	A	Moderate
		AUG27	2048..2143	96 | 31	2	A	T	Moderate
		AUG28	2069..2143	75 | 24	2	T	G	Moderate
5	**NSP1**	**AUG1 (Canonical)**	31..1521	1491 | 496	1	G	G	Strong
		AUG3	194..289	96 | 31	2	G	G	Strong
		AUG5	385..1521	1137 | 378	1	A	A	Moderate
		AUG8	500..568	69 | 22	2	A	T	Moderate
		AUG10	613..1521	909 | 302	1	A	T	Moderate
		AUG11	859..1521	663 | 220	1	G	A	Moderate
		AUG12	862..1521	660 | 219	1	A	A	Moderate
		AUG13	919..1521	603 | 200	1	A	A	Moderate
		AUG14	1060..1521	462 | 153	1	G	T	Moderate
		AUG15	1075..1521	447 | 148	1	T	G	Moderate
		AUG16	1189..1521	333 | 110	1	C	G	Moderate
		AUG18	1258..1521	264 | 87	1	A	T	Moderate
		AUG19	1330..1521	192 | 63	1	C	G	Moderate
		AUG21	1450..1521	72 | 23	1	G	A	Moderate

		AUG22	1508..1609	102 | 33	2	A	A	Moderate
		AUG23	1514..1609	96 | 31	2	A	A	Moderate
		AUG24	1526..1609	84 | 27	2	A	T	Moderate
6	**VP6**	**AUG1 (Canonical)**	24..1217	1194 | 397	3	A	G	Strong
		AUG4	144..1217	1074 | 357	3	A	A	Moderate
		AUG5	148..237	90 | 29	1	T	G	Moderate
		AUG6	154..237	84 | 27	1	G	A	Moderate
		AUG7	304..372	69 | 22	1	A	T	Moderate
		AUG8	312..1217	906 | 301	3	T	G	Moderate
		AUG9	321..1217	897 | 298	3	G	G	Strong
		AUG10	552..1217	666 | 221	3	T	G	Moderate
		AUG11	561..1217	657 | 218	3	A	T	Moderate
		AUG13	921..1217	297 | 98	3	A	A	Moderate
		AUG15	1116..1217	102 | 33	3	G	A	Moderate
7	**VP7**	**AUG1 (Canonical)**	50..1030	981 | 326	2	T	T	Weak
		AUG2	137..1030	894 | 297	2	A	G	Strong
		AUG3	207..329	123 | 40	3	A	G	Strong
		AUG4	236..1030	795 | 264	2	T	G	Moderate
		AUG5	340..405	66 | 21	1	T	G	Moderate
		AUG7	503..1030	528 | 175	2	G	T	Moderate
		AUG8	551..1030	480 | 159	2	C	G	Moderate
		AUG9	608..1030	423 | 140	2	T	G	Moderate
		AUG10	741..824	84 | 27	3	T	G	Moderate
		AUG13	922..1041	120 | 39	1	A	G	Strong

8	**NSP2**	**AUG1 (Canonical)**	47..1000	954 | 317	2	**G**	**G**	Strong
		AUG2	134..1000	867 | 288	2	**G**	C	Moderate
		AUG3	164..1000	837 | 278	2	**G**	**G**	Strong
		AUG4	254..1000	747 | 248	2	**G**	A	Moderate
		AUG5	272..1000	729 | 242	2	**A**	T	Moderate
		AUG6	290..1000	711 | 236	2	**G**	T	Moderate
		AUG6	542..1000	459 | 152	2	**A**	T	Moderate
		AUG10	789..875	87 | 28	3	**G**	T	Moderate
		AUG11	810..875	66 | 21	3	**A**	A	Moderate
		AUG9	923..1000	78 | 25	2	**A**	A	Moderate
9	**NSP3**	**AUG1 (Canonical)**	26..973	948 | 315	2	**T**	**C**	Weak
		AUG2	35..978	939 | 312	2	**A**	**G**	Strong
		AUG3	53..973	921 | 306	2	C	**G**	Moderate
		AUG4	122..973	852 | 283	2	**A**	**G**	Strong
		AUG5	185..973	789 | 262	2	**G**	**G**	Strong
		AUG8	341..973	633 | 210	2	**A**	A	Moderate
		AUG9	344..973	630 | 209	2	**A**	T	Moderate
		AUG10	374..973	600 | 199	2	**A**	A	Moderate
		AUG13	512..973	462 | 153	2	**A**	**G**	Strong
		AUG14	629..973	345 | 114	2	**A**	A	Moderate
		AUG15	641..973	333 | 110	2	**A**	C	Moderate
		AUG16	779..973	195 | 64	2	T	**G**	Moderate
10	**NSP4**	**AUG1 (Canonical)**	42..569	528 | 175	3	**A**	**G**	Strong
		AUG3	129..569	441 | 146	3	**G**	**G**	Strong
		AUG4	201..569	369 | 122	3	**A**	A	Moderate

		AUG5	336..569	234 | 77	3	C	**G**	Moderate
		AUG6	357..569	213 | 70	3	**G**	A	Moderate
		AUG7	375..569	195 | 64	3	**G**	A	Moderate
11	**NSP5**	**AUG1 (Canonical)**	22..618	597 | 198	1	**G**	T	Moderate
		AUG7	292..618	327 | 108	1	T	**G**	Moderate
		AUG8	497..571	75 | 24	2	**A**	T	Moderate
	NSP6	AUG2	80..358	279 | 92	2	**A**	A	Moderate