

Murine Hepatitis Virus Regulation of Nucleotide Selectivity and Fidelity by the RNA-dependent  
RNA Polymerase and the 3'-5' Exoribonuclease

By

Nicole Rose Sexton

Dissertation

Submitted to the Faculty of the  
Graduate School of Vanderbilt University  
in partial fulfillment of the requirements  
for the degree of

DOCTOR OF PHILOSOPHY

in

Microbiology and Immunology

May, 2017

Nashville, Tennessee

Approved:

Mark R. Denison, M.D.

Christopher R. Aiken, Ph.D.

David K. Cortez, Ph.D.

Kristen M. Ogden, Ph.D.

H. Earl Ruley, Ph.D.

To my grandma and grandpa Markey for your genuine and perceptive encouragement of my academic pursuits and to my Aiden, may you achieve all of your aspirations.



## ACKNOWLEDGEMENTS

Who we are and what we accomplish is an accumulation of the people and places that we experience along the way. So many seemingly small moments have profound impacts on trajectory, and I thank all those who have touched my life. I would like to draw attention to some of those who have most notably helped to shape my journey toward a PhD.

### The Denison Lab

Firstly, I must thank my advisor, Mark Denison, for cultivating an environment that allowed me to grow substantially as both a scientist and a person. Mark provided me with the resources and freedom to openly explore my area of research, and for that I am extremely grateful.

Additionally, Mark provided guidance that forced me to clearly articulate and present scientific messages. My presentations have improved immeasurably because of this direction, and I benefit from this development across all aspects of my life.

All of the members of the Denison lab, as well as those who just passed through, have taught me something, but none were as essential to my time at Vanderbilt as Clint Smith. Clint's fearless and meticulous approach to science is truly inspiring. At times Clint served as a valued mentor assisting me in how to frame my work into a cohesive story, as a critical collaborator working through problems which paralleled my own, and as a thoughtful friend listening to, and helping me work through, both personal and professional problems. I cannot adequately state the gratitude I have for all the support and consideration.

Most members of the Denison lab are transient, passing through on their way to something else. However, Xiaotao Lu is the exception. She is the glue that holds together a cohesive lab environment and assures that procedural knowledge is not lost through generations. Assuredly, every project that has been generated in the Denison lab has been touched by Xiaotao and made better for it. My projects have been no different, and I appreciate all she does for the lab. On a more personal level, Xiaotao took on the role of my "lab mom" while I was pregnant with my son. Not a day went by where she did not verify that I was eating well and not lifting anything remotely heavy. This is one of the sweetest things anyone has ever done for me, and I cannot fully express how much I cherished those check-ins.

To the graduate students who started before me: Megan Culler-Freeman, Dia Beachboard and Chris Stobart, thank you for guiding me through the PhD process particularly in regards to the qualifying exam. To Megan in particular, thank you for sharing so much of what Nashville has to offer, for the many social gatherings, great food, rollerblading and gardening. My time in Nashville and in lab would not have been nearly as enjoyable without you. To Brett Case, thank you for the camaraderie, though you joined the year after me, we shared many of the same trials with shifts in the department, our first papers going out nearly simultaneously, BSL-3 trainings/inspections and the split of VU and VUMC. I'm grateful I had someone to share in these experiences. To Michelle Becker, thank you for coordinating the lab resources and for training me in the BSL-3. To Maria Agostini and Kevin Graepel thank you for taking some of my knowledge to build your projects and work through your own PhDs. At the end of my PhD, in addition to a large body of knowledge I will continue to use, I have amassed a great deal of specialized knowledge that I may never use again. I have greatly appreciated every time you have taken my advice or been excited about a paper I was able to share with you. To Erica Andres, first I wish you the best of luck in your career as a medical doctor. You will be amazing. Part of why you will be amazing is that you are one of the most compassionate people I have ever met. I genuinely have more faith in all of humanity having met you. Thank you for sharing a lab space with me. I am so grateful I found such a great group to share my time with.

#### The Simmons Lab

My first real lab experience was in the laboratory of Graham Simmons at the Blood Systems Research Institute (BSRI) in San Francisco. Although I began my internship believing I wanted to do bench science, I left knowing. Again, many people there helped shaped my journey and deserve thanks. To Graham, thank you for taking a chance on me. I had no previous experience, no internal contacts at BSRI, and I lived 40 miles away yet you selected me as your intern and for this I am so grateful. Additionally, you allowed me to work on a diverse set of projects providing me with a solid scientific base to move forward to anything I could image. Although I enjoyed it at the time, I didn't realize until later the enormous value of this experience. Finally, you completely changed my trajectory when you convinced me I would be better off working on your Chikungunya virus project instead of the Chagas Disease project. I may never have known how fascinating viruses are if I had followed my original plan of studying eukaryotic parasites.

To Beatriz (Bea) Salvador-Esteban, thank you for my scientific foundation. I could not have asked for my first science mentor to be more thorough, patient, or kind. You took the time and taught me both the how and why of every step of any experiment I performed. I now know how rare this is; therefore, I have to thank you for saving me from so much wasted time on trial and error. I am absolutely a better scientist because I worked with you. On a practical level, I must also thank you for showing me the joys of Mac over PC, electronic lab notebooks and the great improvement of PRISM over Excel for figures. On the last two, not only I, but my fellow members of the Denison lab thank you. Finally, thank you for teaching me the importance of lunchtime and how to make sure lunch is not forgotten while being maximally productive.

#### The Pediatrics Infectious Disease Group

Although my degree-granting department is Microbiology and Immunology, the Denison lab is physically located in the heart of the PedsID department, and therefore, I have had the great pleasure of being part of this close group. I am so grateful for the administrative support, the many department parties, and the wonderful lunch group I was a part of. Before they moved on to new opportunities the Dermody and Williams labs were functionally available as extended lab families full of disparate knowledge on a variety of topics for which we've all benefited greatly. Janet, Eve, Heather and Adam have all been a wonderful help achieving a variety of goals and I am truly grateful.

#### My Thesis Committee

To Chris Aiken, Dave Cortez, Terry Dermody, Kristen Ogden, and Earl Ruley thank you for all of your guidance, encouragement and support. I have greatly appreciated all of your comments and questions; I feel that I have grown immeasurably through this process and that each of you has assisted in that growth.

#### My Family and Friends

The support of family and friends is essential to accomplishing any strenuous goal and I am very fortunate to have a large group who has supported me. In particular, I must thank my husband, Ryan Sexton. I'm certain that your life-plan didn't involve packing up everything you had and moving more than half way across the country to place where you didn't have a job waiting. You

certainly didn't plan on having to work temporarily delivering pizzas, only to be robbed at gunpoint. Yet, you followed me here, supported my goals, and never complained. I'm so grateful for everything you do for me and for our son, Aiden. Thank you. To my parents, Mary Jo and Bill Markey, thank you for helping to encourage my inquisitive nature. To Lou Ann and Rik Wijssenbeek, Michael Sexton, and Tess Sexton thank you for your support and family visits. Finally, to Katie and Ryan Hebron, thank you for being Ryan's and my family away from home. You made our Nashville holidays ones we will always cherish and provided an extended family for our Aiden.

## TABLE OF CONTENTS

	Page
DEDICATION .....	ii
ACKNOWLEDGEMENTS .....	iii
LIST OF TABLES .....	xi
LIST OF FIGURES .....	xii
Chapter	
I. BACKGROUND AND LITERATURE REVIEW .....	1
Introduction .....	1
Coronavirus genome organization and replication strategy .....	3
Coronavirus diversity and emergence .....	6
Generation of RNA virus diversity through mutation rates and population sizes.....	9
Lethal mutagenesis and error catastrophe .....	12
Structure of, and nucleotide selection by, the RdRps of RNA viruses .....	13
The CoV replication-transcription complex.....	16
Fidelity proteins of the replication-transcription complex .....	17
Additional functions of coronavirus replication-transcription complex proteins .....	19
Co-evolution of interactions within the RTC and potential for additional determinants of fidelity .....	20
Conclusion.....	21
II. HOMOLOGY-BASED IDENTIFICATION OF A MUTATION IN THE CORONAVIRUS RNA-DEPENDENT RNA POLYMERASE THAT CONFERS RESISTANCE TO MULTIPLE MUTAGENS .....	23
Introduction .....	23
Homology modeling of MHV nsp12-RdRp polymerase core domain predicts putative fidelity determinants .....	24
Recovery of mutant viruses in the MHV nsp14-ExoN(+) and nsp14-ExoN(-) isogenic backgrounds.....	26
Resistance of recovered mutant viruses to the base analog 5-FU .....	27
Replication kinetics of nsp12-V553I and nsp12-M611F mutant viruses in the WT and nsp14- ExoN(-) background.....	28
Specific infectivity of nsp12-V553I and nsp12-M611F.....	31
Fitness costs of nsp12-V553I and nsp12-M611F in WT and nsp14-ExoN(-) backgrounds ....	31
Resistance of recovered mutant viruses to the base analog 5-AZC .....	36
The nsp12-V533I mutation results in a decrease in the accumulation of mutations.....	36
Discussion .....	40

Determinants of nucleotide selectivity and fidelity in CoVs may be conserved with other RNA viruses .....	41
Coronavirus nsp12-RdRp and nsp14-ExoN cooperate to optimize both fidelity and replication kinetics.....	42
Conclusion.....	43

### III. IDENTIFICATION OF COMPENSATORY MUTATIONS IN CORONAVIRUS NSP12 AND NSP14 AFTER PASSAGE IN THE PRESENCE OF MUTAGEN .....

45

Introduction .....	45
Identification of minority variants in SARS-CoV populations after treatment with 5-FU .....	46
Engineered viruses result in no change in selectivity for nucleotide analogs 5-fluorouracil or 2'-C-methyladenosine.....	47
Replication kinetics of nsp12-E519N, nsp12-Q520L and nsp12-S900C mutant viruses in the WT and nsp14-ExoN(-) backgrounds .....	50
The effect on fitness of nsp12-E519N and nsp12-S900C in the nsp14-ExoN(-) background. ....	53
Identification of consensus sequence mutations in MHV nsp12 and nsp14 after passage in the presence of the mutagen 5-FU.....	53
Resistance of 5-FU-treated virus populations to the mutagens 5-FU and 5-AZC .....	58
Replication kinetics of nsp14-ExoN(-) viruses treated with 5-FU result in an increased lag phase as well as decreased titers.....	61
Engineering and recovery of a subset of mutant viruses identified by selection in the WT and nsp14-ExoN(-) background.....	63
Discussion .....	64
Some minority variants of mutagen-treated CoVs require the presence of alternative residues for viability .....	64
Minority variants result in subtle to no alterations of replication phenotypes .....	65
Selection for mutations that result in resistance to 5-FU identifies residues involved in nucleotide selectivity, but also increases titers by an alternate mechanism .....	66
Conclusion.....	68

### IV. DEVELOPMENT OF A ROBUST AND QUANTITATIVE MEASURE OF FIDELITY FOR CORONAVIRUSES .....

69

Introduction .....	69
TOPO TA cloning, combined with Sanger sequencing of a 1kb region of the genome results in a small dynamic range.....	72
Selectivity for nucleotide analogs varies between mutant viruses and nucleotide analogs .....	74
Coronaviruses are not sensitive to guanidine hydrochloride at concentrations that inhibit picornaviruses.....	75
Developing an rgRNA/FnCas9 assay for investigating fidelity in CoVs.....	75
Identification of targets in the MHV genome.....	75
HEK293T cells transiently transfected with CEACAM1 are permissive for MHV replication .....	78

Translation of MHV is inhibited in the presence of MHV targeted rgRNAs and FnCas9..	80
MHV viral replication is inhibited in the presence of MHV frame shift targeted rgRNAs and FnCas9 .....	80
FnCas9 transfection increases MHV translation in a concentration dependent manner, and MHV targeted rgRNAs decrease translation without FnCas9.....	83
Changes to WT MHV translation resulting from the presence of rgRNAs and FnCas9 in DBT cells are similar to those in HEK293T cells .....	87
Generation of a dual expression plasmid for stable expression of FnCas9 and rgRNAs in cell lines .....	89
Discussion .....	91
V. MATERIALS AND METHODS .....	94
Virus and cell culture .....	94
Sequence analysis and homology modeling of CoV MHV nsp12-RdRp and nsp14.....	94
Cloning, recovery and verification of mutant viruses .....	94
Compounds and drug sensitivity studies .....	95
Virus replication and RNA synthesis assays .....	95
Determination of Specific Infectivity .....	96
One-step RT-qPCR for determining supernatant genome copies for specific infectivity assay.....	96
Competitive fitness of mutant viruses .....	97
Passage reversion analysis.....	97
Preparation of amplicons for deep sequencing of full viral genomes .....	97
Deep sequencing sample preparation and analysis .....	97
Passaging in the presence and absence of 5-FU.....	98
Mutation frequency by TOPO TA cloning.....	98
Guanidine hydrochloride sensitivity assay.....	99
Transient-transfection followed by infection of HEK293T and DBT cells .....	99
Puromycin treatment of DBT cells.....	100
Plasmid cloning for the generation of stable cell lines.....	100
Statistical analysis .....	100
Primers generated for this dissertation research.....	101
VI. SUMMARY AND FUTURE DIRECTIONS .....	104
Introduction .....	104
CoV's may be capable of selecting for nsp12-RdRps specialized for local nucleotide pools	105
A WT nsp14-ExoN is required for the recovery of many mutant nsp12-RdRp viruses .....	107
Nsp12 and nsp14 have epistatic relationships .....	107
Multiple replicase proteins regulate CoV replication fidelity .....	108
MHV replication is inhibited by targeted rgRNAs and FnCas9 .....	110
Conclusion.....	111

Appendix

A. HOMOLOGY-BASED IDENTIFICATION OF A MUTATION IN THE CORONAVIRUS  
RNA-DEPENDENT RNA POLYMERASE THAT CONFERS RESISTANCE TO MULTIPLE  
MUTAGENS ..... 113

B. THINKING OUTSIDE THE TRIANGLE: REPLICATION FIDELITY OF THE LARGEST  
RNA VIRUSES ..... 128

REFERENCES ..... 151



## LIST OF TABLES

Table	Page
1. Recovery of mutant viruses using site-directed mutagenesis .....	27
2. Recovery of minority-variant mutant viruses using site-directed mutagenesis .....	46
3. Recovery of mutant viruses from 5-FU selection using site-directed mutagenesis .....	63
4. Primers generated for this dissertation research .....	101

## LIST OF FIGURES

Figure	Page
1. Coronavirus virion structure and EM .....	2
2. Coronavirus genome organization, replication and transcription .....	4
3. Timeline for the emergence of human coronaviruses .....	7
4. Comparison of mutation rates between RNA viruses and cellular DNA .....	11
5. Model of the MHV RdRp core domain .....	15
6. Model of the CoV replication-transcription complex (RTC) .....	18
7. Homology modeling of CoV nsp12-RdRp and identification of residues that potentially regulate fidelity based on CVB3 structure .....	25
8. Resistance of MHV nsp12-RdRp mutant viruses to 5-FU in the WT and nsp14-ExoN(-) background .....	29
9. Replication kinetics of MHV nsp12-RdRp mutant viruses .....	30
10. Specific infectivity is increased in both nsp12-V553I and nsp12-M611F mutants .....	32
11. Competitive fitness analysis in the WT and nsp14-ExoN(-) Background .....	34
12. The nsp12-V553I mutation is stable across passage; however, nsp12-M611F is vulnerable to reversion .....	35
13. Resistance of MHV nsp12-RdRp V553I and M611F mutant viruses to 5-AZC in the nsp14-ExoN(-) background .....	37
14. The nsp12-V553I mutation confers decreased accumulation of mutations in the nsp14-ExoN(-) background with no bias toward the exclusion of specific nucleotides .....	38
15. Selectivity for 5-FU or 2'-C-MeA of MHV engineered 5-FU-treated CoV minority variant viruses .....	48
16. Selectivity for 2'-C-MeA of MHV nsp12-RdRp mutant viruses identified by homology modeling .....	49
17. Replication kinetics of 5-FU-treated, minority-variant engineered MHV nsp12-RdRp mutant viruses .....	51
18. RNA accumulation by MHV nsp12-RdRp mutant viruses .....	52
19. Competitive fitness analysis of MHV nsp12-RdRp mutant viruses in the nsp14-ExoN(-) Background .....	54

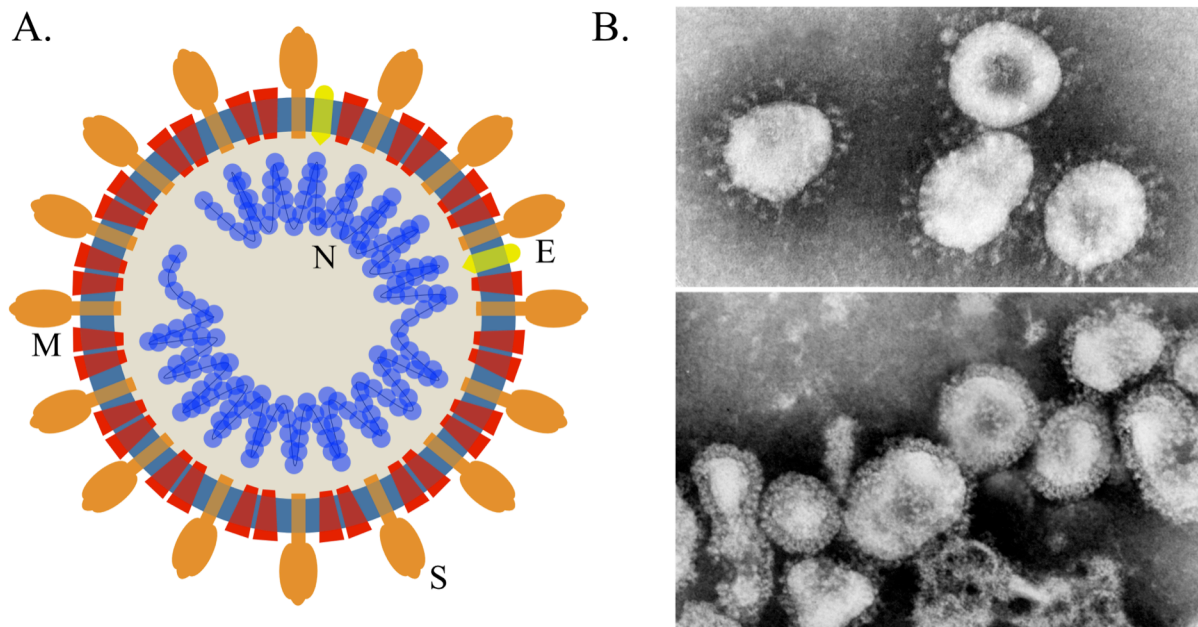
20. Identification of mutations in nsp12 and nsp14 selected for in the presence of 5-FU.....	56
21. Conservation of selection-identified residues across CoVs.....	57
22. Location of identified residues on modeled MHV structures of nsp12-RdRp and nsp14 .....	59
23. Resistance of 5-FU-treated population to 5-FU or 5-AZC .....	60
24. Replication kinetics of WT and nsp14-ExoN(-) viruses in the presence of 5-FU .....	62
25. Mutation frequency determined by TOPO TA cloning .....	73
26. Selectivity of mutant viruses in the nsp14-ExoN(-) background for a panel of nucleotide analogs .....	76
27. Treatment of WT and nsp14-ExoN(-) viruses with GUA-HCl .....	77
28. Structured regions of the MHV genome identified for targeting of FnCas9 by rgRNAs.....	79
29. MHV viruses are able to infect HEK293T cells when CEACAM1 is transiently expressed .	81
30. Translation of WT and nsp14-ExoN(-) MHV-FFL viruses is reduced when rgRNA targets structured regions and FnCas9 is expressed .....	82
31. Viral propagation of WT MHV-FFL viruses is reduced when rgRNA targets to the frame shift and FnCas9 is expressed .....	84
32. WT MHV-FFL translation increases with increasing transfection of FnCas9 plasmid, however this does not result in increased titer .....	85
33. WT MHV-FFL translation is inhibited by rgRNAs targeted to structured regions of the genome even in the absence of FnCas9 .....	86
34. Translation in DBT cells and comparison between WT and an nsp12-RdRp double mutant	88
35. Plasmid for the generation of stable cell lines .....	90

# CHAPTER I

## BACKGROUND AND LITERATURE REVIEW

### Introduction

Positive-sense RNA viruses are the cause of a large percentage of severe human infections. Many positive-sense RNA ((+)RNA) viruses (such as Chikungunya, Dengue, Zika, West Nile, SARS-CoV and MERS-CoV viruses) have recently emerged in human populations (Bolles et al., 2011; Hilgenfeld and Peiris, 2013; Murray et al., 2010; Shuaib et al., 2016; Tsetsarkin et al., 2007). SARS-CoV and MERS-CoV viruses belong to the group of viruses called coronaviruses (CoVs). Beyond these well-known, highly pathogenic, human CoVs, four CoVs are endemic in human populations, often resulting in symptoms of the “common cold”, bronchiolitis, or pneumonia (Graham et al., 2013). CoVs are enveloped, (+)RNA viruses belonging to the order *Nidovirales* in the family *Coronaviridae*. CoVs are named for their crown-like appearance in electron microscopy (EM) images caused by the protrusion of spike fusion proteins across the surface of virions (Masters, 2006) (Figure 1). CoVs encode the largest genomes of any known (+)RNA virus, with genomes ranging from 27-34kb (Gorbalenya *et al.*, 2006; Stenglein *et al.*, 2014). This size is striking since most (+)RNA viruses encode genomes of 12kb and below (Gorbalenya *et al.*, 2006), and the only known (+)RNA viruses with genomes >20kb are in the order *Nidovirales*. Similar to all other (+)RNA viruses, CoVs must encode an RNA-dependent RNA polymerase (RdRp) to facilitate genome replication. Replication fidelity is the relative number of correct vs. incorrect nucleotides incorporated into nascent viral genomes during replication. RdRps – as well as other polymerases without encoded proofreading – incorporate nucleotides with low-fidelity, resulting in high viral mutation rates. However, in contrast to all other (+)RNA viruses, coronaviruses and the other large nidoviruses encode a proofreading 3'-5' exoribonuclease (ExoN) (Eckerle et al., 2010; Minskaia et al., 2006; Smith et al., 2013). In addition to ExoN, non-structural protein 10 (nsp10), a small non-enzymatic protein, is involved in CoV replication fidelity (Bouvet *et al.*, 2014; Smith *et al.*, 2015). All biochemical and genetic data over the past decade support the hypothesis that CoVs encode a multi-component polymerase-fidelity complex and regulate their fidelity by mechanisms in addition to the RdRp.



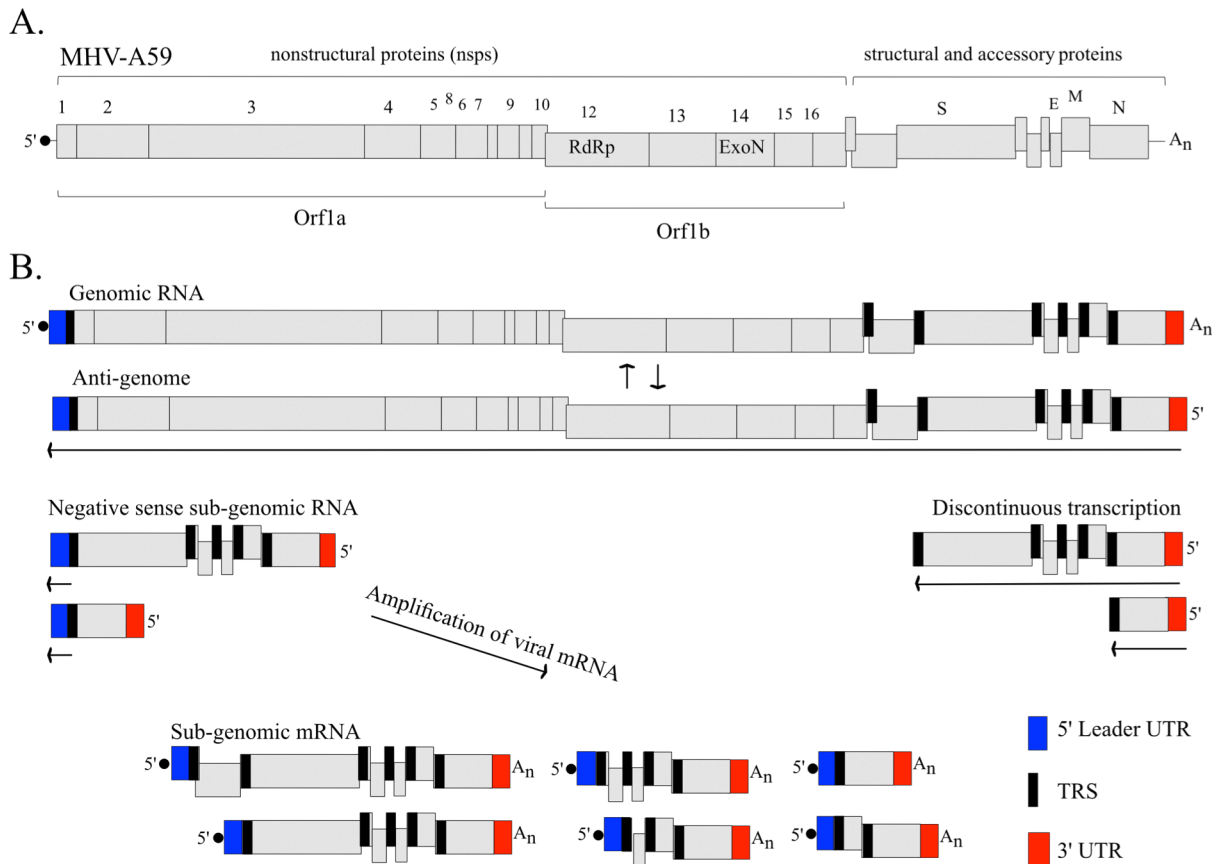
**Figure 1. Coronavirus virion structure and EM**

Coronaviruses are enveloped viruses that encode and package structural proteins in the virions they release. **A.** A representative coronavirus particle where S is the spike fusion protein, M is the membrane protein and E is the envelope protein. Within the particle is the (+)RNA genome coated with nucleocapsid protein, N. Image modified from Clint Smith. **B.** Representative negative-contrast electron microscopic (EM) images of HCoV-229E, top, and Murine Hepatitis virus (MHV), bottom (from (Murphy, 2012) with permission). The particle decoration with spike protein results in the appearance of crown, giving coronaviruses their name.

For most (+)RNA viruses the RdRp has been the only protein implicated in regulation of fidelity. However, in work published since this dissertation research began, a viral helicase has also been shown to be capable of influencing the fidelity of Chikungunya virus (Stapleford *et al.*, 2015). This study demonstrates the potential for other virus-encoded proteins to be involved in fidelity regulation. Similarly, in *Mycobacterium tuberculosis* the  $\beta 2$  clamp, which is a processivity factor, encodes determinants of fidelity (Gu *et al.*, 2016). In CoVs, like other (+)RNA viruses, the RdRp is essential to nucleotide incorporation and therefore necessarily functions in fidelity regulation. However, in the presence of the proofreading ExoN and putative multi-protein replication-transcription complex (RTC) an important question is how the RdRp participates in replication fidelity. In this dissertation research, I investigate the phenotypic relationship between the proofreading ExoN and the RdRp of CoVs. I identify RdRp determinants that regulate nucleotide selectivity and, for one determinant, support increased overall fidelity. I demonstrate that proofreading by the murine hepatitis virus (MHV) ExoN, encoded in nonstructural protein 14 (nsp14-ExoN), is epistatic to nucleotide selectivity by the RdRp, defined here as a situation where the phenotype of one viral protein masks the phenotype of genetic variants of another viral protein. I provide evidence that determinants of fidelity within the CoV RdRp can be predicted based on RdRp structure-function studies from other (+)RNA families, and thus may be common across viral orders. Additionally, I demonstrate that selection for resistance to a specific nucleotide in the CoV RdRp can occur while maintaining baseline susceptibility, or even increased susceptibility, to other nucleotides, demonstrating that nucleotide selectivity by the CoV RdRp is quite adaptable. Finally, I present progress toward the development of a quantitative assay to define the fidelity of CoVs during replication in cells without exogenous nucleoside analogs.

### **Coronavirus genome organization and replication strategy**

Coronaviruses possess the largest known RNA genomes ranging from 27-34kb (Gorbalenya *et al.*, 2006; Stenglein *et al.*, 2014). Upon entering a host cell the genome functions as messenger RNA (mRNA), complete with a 5' cap and a polyA tail (Shi and Lai, 2005). These features protect the genome from degradation by the cell and allow use of host translational machinery. At either end of the genome are untranslated regions (UTRs), which include highly structured regions of RNA (Goebel *et al.*, 2007; Hsue and Masters, 1997; Kang *et al.*, 2006). The remainder of the genome can be divided into nonstructural proteins (nsps) and structural proteins. CoVs



**Figure 2. Coronavirus genome organization, replication and transcription**

**A.** The genomes of coronaviruses can be divided into nonstructural and structural proteins. The nonstructural proteins (nsps) are encoded in the first Orf, which is expressed as either Orf1a including nsps1-11 or Orf1ab including nsps 1-10 and 12-16. Coronavirus genomes are complete mRNAs with 5'-caps and polyA tails. The RdRp of coronaviruses is encoded in nsp12 and is responsible for transcription and replication of genomic and sub-genomic RNAs. **B.** Genomes are replicated by continuous transcription and through negative-sense anti-genome intermediates. Sub-genomic mRNAs are generated by discontinuous transcription linking body TRS's to the 5'-leader TRS. This results in identical 5' and 3' ends for all viral RNAs of the same polarity. Amplification of positive-sense sub-genomic mRNAs occurs from the negative-sense sub-genomic RNAs. This figure is adapted from (Perlman and Netland, 2009; Raaben et al., 2008) and (Smith et al., 2014).

encode 16 nsps all within the first open reading frame (Orf) (Figure 2a). The first Orf (Orf1ab) is further broken into Orf1a and Orf1b and requires use of a programmed -1 ribosomal frameshift guided by a structured RNA region in order to express all 16 nsps (Bredenbeek *et al.*, 1990). This slippery sequence results in a frameshift roughly 40% of the time it is translated (Irigoyen *et al.*, 2016), thereby allowing for around 2.5 times as many protein products of Orf1a as Orf1b. In either instance the full Orf1a or Orf1ab is translated as one large polyprotein. The next step in replication is the proteolytic maturation of intermediate and mature nsps from the replicase polyprotein by CoV encoded proteases within the polyprotein. The specific number of proteases is virus specific. In the case of MHV - the most well studied of the coronaviruses, three separate proteases are required to process all 16 nsps (Shi and Lai, 2005).

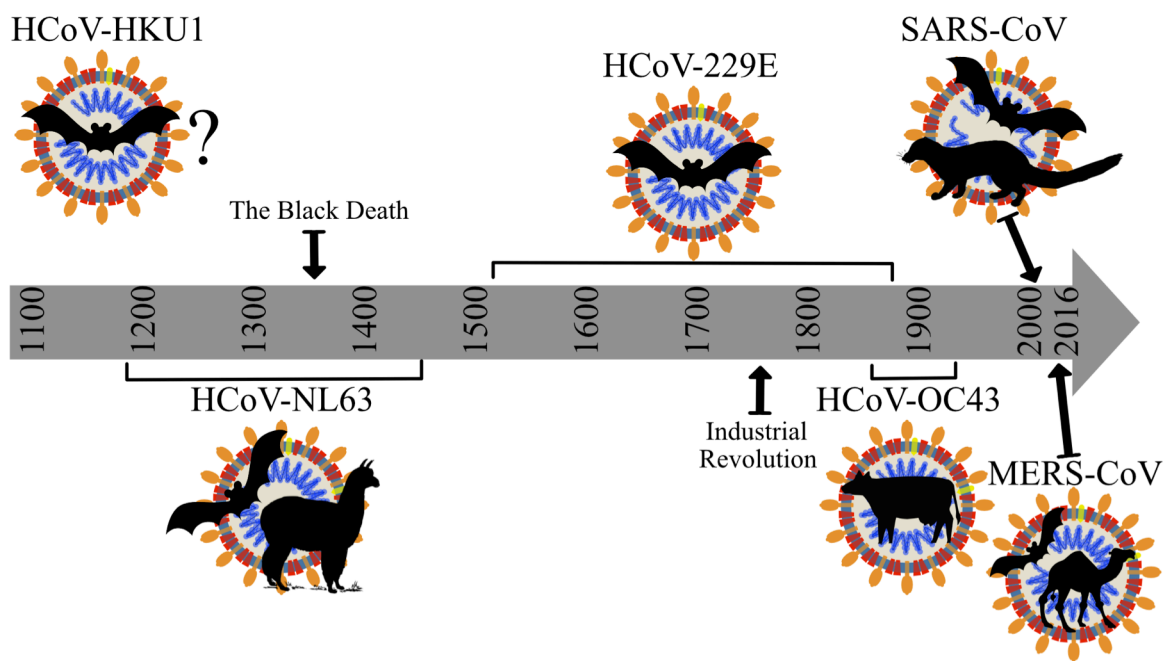
Structural proteins are encoded within the remaining Orfs, the number of which varies by virus. Functionally, this provides both a temporal delay in the expression of these proteins, as the subgenomic mRNAs are not immediately available to the host machinery for translation, as well as differential expression levels (Irigoyen *et al.*, 2016). Once the Orf1ab nonstructural proteins are expressed and processed they form RTCs, which transcribe subgenomic mRNAs and replicate the full-length genome. Transcription occurs in two forms, continuous and discontinuous, the result of which is both full-length negative-sense RNA genomes and a nested set of sub-genomic negative-sense RNAs with lengths corresponding to various portions of sequence downstream of Orf1ab (Perlman and Netland, 2009). Many specific details of the mechanism behind the generation of sub-genomic RNA require further study. What is known is that a transcriptional regulator sequence (TRS) is present in the 5'-UTR and at the beginning of each Orf. During the process of generating subgenomic RNA, discontinuous replication occurs producing negative-sense RNAs which all possess identical 5'-UTRs and one TRS (Figure 2b). Therefore, both genomic and subgenomic RNAs have identical 5' and 3' ends. The CoV RNA-dependent RNA polymerase encoded in nsp12 is responsible for both transcription and replication of viral RNA (Ahn *et al.*, 2012). However, some early evidence suggested that there are two separate polymerase activities, one that mainly resulted in negative RNA synthesis and one in positive (Brayton *et al.*, 1982). Therefore, the RdRp may function within complexes of two different compositions and/or may be present in two different forms during replication (Ahn *et al.*, 2012; Cheng *et al.*, 2005).



The nonstructural proteins of coronaviruses are mainly involved in transcription and replication of the genome and protecting these viral processes from cellular innate sensing of viral components. In contrast, the structural proteins are mainly involved in packaging and dissemination of viral particles. Spike is the receptor-binding and fusion protein of CoVs, which largely defines host species and tropism within a host. CoVs utilize a wide range of cellular receptors. The MHV spike has a receptor-binding domain (RBD) that recognizes the murine CEACAM1 protein (Dveksler *et al.*, 1991), which is required and sufficient for MHV infection of cells in culture. Similarly, the RBD of MERS-CoV spike recognizes human DPP4 and the presence of cell surface DPP4 is necessary and sufficient for infection of various cell types (Raj *et al.*, 2013). However, additional factors can enhance infection, as is the case during MHV or MERS-CoV infections (Chan *et al.*, 2016; Thorp and Gallagher, 2004), and these cofactors may be important for infections *in vivo*.

### **Coronavirus diversity and emergence**

Coronaviruses infect a wide array of species and have emerged as highly virulent human pathogens twice in this century. In humans, symptoms range from the common cold to severe acute respiratory syndromes, resulting in severe morbidity and death. Currently, five CoVs are known to circulate in humans (HCoVs): HCoV-NL63, -229E, -OC43, -HKU1 and MERS-CoV. The sixth HCoV, SARS-CoV, is no longer circulating in human populations. However, a bat SARS-like CoV that utilizes the SARS-CoV receptor, ACE2, was discovered in Chinese horseshoe bats (*Rhinolophus sinicus*) (Ge *et al.*, 2013), showing that an ancestral virus is still circulating in bats. CoVs are adept at transcending species barriers and thriving in new host environments. All HCoVs are proposed to have emerged as zoonotic infections. Molecular clock analysis predicts that HCoV-NL63 emerged in the human population between 1190 and 1449 CE (Huynh *et al.*, 2012), HCoV-229E between 1501-1883 CE (Pfefferle *et al.*, 2009) (or more recently (Crossley *et al.*, 2012)), and OC43 between 1866-1918 CE (Vijgen *et al.*, 2005). The date of emergence for HKU1 is unknown (Woo *et al.*, 2005). SARS-CoV was identified in 2003 (Peiris *et al.*, 2003) and MERS-CoV in 2012 (Zaki *et al.*, 2012), both within a year of entering the human population (Figure 3). This demonstrates that emergence of novel CoVs into the human population is not an uncommon phenomenon; although, it may be a (relatively) recent trend correlating with the beginning of the most recent era of human population expansion. However, the mechanisms behind CoV evolution and adaptation have not been fully elucidated.



**Figure 3. Timeline for the emergence of human coronaviruses**

Molecular clock analysis suggests that all current HCoVs emerged in human populations with the last millennium, roughly corresponding to the most dramatic human population expansion. SARS-CoV and MERS-CoV emerged during the time of surveillance and so have specific years ascribed to their entry into the human population. HKU1 is the only HCoV for which a timeframe of zoonotic transmission has yet to be identified. All HCoVs have been linked to bats except for HCoV-OC43. Several have evidence for transmission to humans through intermediate hosts such as camels for MERS-CoV, civet cats for SARS-CoV and alpacas for HCoV-NL63.

Coronaviruses are found in a plethora of species from birds to bats to humans (Woo *et al.*, 2009). The majority of the diversity of *Alphacoronavirus* and *Betacoronavirus* species, which exclusively infect mammals, exists in bats (Drexler *et al.*, 2013). CoVs are predicted to have co-evolved with bats for millions of years, possibly since their evolutionary split from birds (Wertheim *et al.*, 2013). Bats are optimal species for the emergence of novel viruses into humans and other mammalian species for several reasons including their long life spans, social structure, ability to fly and proximity to the base of the mammalian phylogenetic tree (Calisher *et al.*, 2006). All of the known HCoV have proposed bat origins with the exception of HCoV-OC43, which emerged from a bovine CoV (Huynh *et al.*, 2012). However, emergence into the human population often requires passage through an intermediate host. This was the case for SARS-CoV, whose transmission pattern is well documented (thoroughly reviewed in (Hilgenfeld and Peiris, 2013)), and involved cyclical passage between civet cats and humans. Similarly, the closest known relative of 229E was recently discovered in alpacas, suggesting interspecies transmission to humans might have been through alpacas (Crossley *et al.*, 2012) instead of directly through bats (Pfefferle *et al.*, 2009). Finally, MERS-CoV likely infects humans through dromedary camels as intermediate hosts (Reusken *et al.*, 2016) (Figure 3).

MERS-CoV was originally isolated in June 2012 from a 60-year-old male Saudi Arabian. He presented with acute respiratory distress syndrome (ARDS) and multiple organ dysfunction syndrome (MODS) (van Boheemen *et al.*, 2012), which eventually resulted in death. As of December 5<sup>th</sup>, 2016, there have been 1841 WHO laboratory-confirmed cases with 652 known deaths, resulting in a case fatality rate of 35%. Interestingly, the case fatality rate for secondary cases is closer to 20% (Cauchemez *et al.*, 2014). Therefore, it is possible that there are many symptomatic cases that have occurred, but have not been recorded. The  $R_0$  (the number of individuals likely to be infected by a single infected individual) for uninterrupted MERS-CoV infections is between 0.7-1.5; however, with current intervention protocols in place,  $R_0$  is predicted to remain below 1, indicating human-to-human transmission is not currently sustainable (Brebant *et al.*, 2013; Cauchemez *et al.*, 2014). Supporting this, human cases have been sporadic, originating from diverse foci within the Middle East, with abruptly terminating transmission cycles. Imported cases have been reported throughout Europe and in Tunisia, also with only very limited spread (to close family members and healthcare workers) (Penttinen *et al.*, 2013). An outbreak of MERS-CoV in South Korea was the only exception to this, which resulted

in 186 laboratory-confirmed cases during 2015. However, this outbreak spread almost entirely by nosocomial transmission, with 178 cases contracted in the healthcare setting (Lee et al., 2016). Together, these observations suggest a recurrent introduction to the human population through alternative hosts.

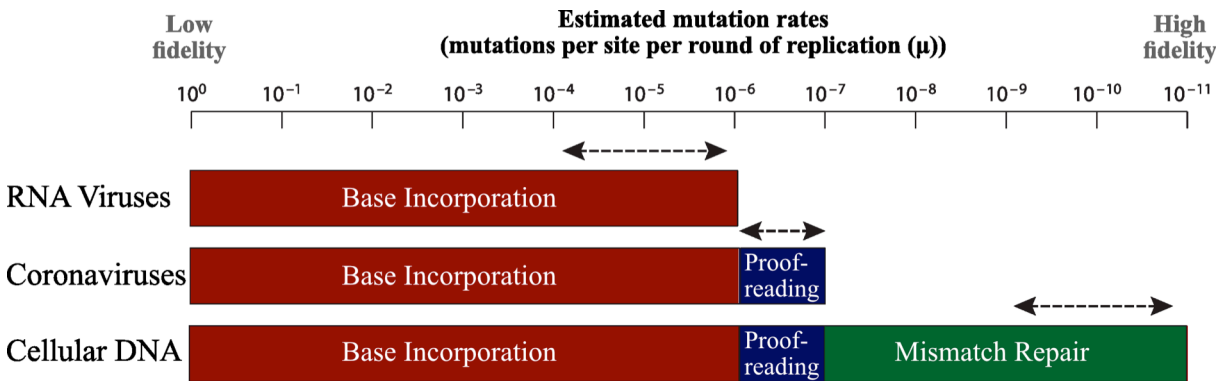
Current evidence suggests that dromedary camels are an intermediate host in the transmission of MERS-CoV to humans (Reusken *et al.*, 2016). It was first noticed that multiple individuals infected with MERS-CoV infections were associated with close contact with dromedary camels. Antibodies to MERS-CoV and MERS-CoV infectious genomes have been found in dromedary camels throughout the Middle East (Briese et al., 2014; Hemida et al., 2013). However, bats may also be involved in some cases where interaction with camels or sick humans is not identified; partial MERS-CoV sequences with 100% sequence identity to human MERS-CoV have been identified in *Taphozous perforatus* bats in Saudi Arabia (Memish *et al.*, 2013). No evidence has been collected that suggests close contact with bats for any human MERS-CoV case to date. However, some evidence has been provided that suggests oral-fecal transmission could be a route of spread (Goh et al., 2013); when combined with demonstrated high environmental stability (van Doremalen et al., 2013), contact with guano becomes a potential source of infection. Further investigation is needed to fully define transmission patterns of MERS-CoV. With infections continuing to be identified in 2016, it is likely that MERS-CoV will be a persistent problem in the Middle East until an intervention can be identified and implemented. The case of MERS-CoV demonstrates that the emergence of a highly pathogenic HCoV was not a one-time event and that understanding the history of CoV evolution and identifying universal CoV targets for prevention and treatment are critically important goals.

### **Generation of RNA virus diversity through mutation rates and population sizes**

RNA viruses are proposed to be among the most diverse replicative units currently in existence (Mokili et al., 2012). They infect an exceptional range of hosts, and regularly cross species barriers to further expand this range. An example of this was described in a previous section with MERS-CoV. RNA virus replication results in the incorporation of a relatively high number of mutations, ranging from  $10^{-4}$  to  $10^{-6}$  mutations per site per round of replication ( $\mu$ ) (Castro et al., 2005; Crotty et al., 2001; Sanjuán et al., 2010; Schaaper, 1993; Smith and Denison, 2012). This is low-fidelity replication when compared with the DNA replication of cellular life forms, which

are observed to be between  $10^{-9}$  to  $10^{-11}$   $\mu$  (Drake et al., 1998; Schaaper, 1993) (Figure 4). It is often stated that the low-fidelity replication of RNA viruses is due to the poor nucleotide selectivity of RNA-dependent RNA polymerases (RdRps) vs. that of DNA polymerases. Although the mutation rate of RNA viruses is driven mainly by the fidelity of nucleotide incorporation by the RdRp, data suggest that this base misincorporation rate is on average equivalent to that of DNA polymerases (Sanjuán et al., 2010; Schaaper, 1993). The difference between the observed mutation rates of RNA and DNA replication is instead accounted for by the availability of proofreading and mismatch repair in DNA based organisms. No known RdRp has a proofreading domain as part of the protein and only coronaviruses, along with the other large nidoviruses, are known or predicted to encode a proofreading exoribonuclease (Gorbalenya *et al.*, 2006). Arenaviruses encode exoribonucleases; however, to date the activity of these 3'-5' exoribonucleases has only been implicated in immune regulation (Hastie *et al.*, 2011; Huang *et al.*, 2015). Studies from our lab show that CoVs display increased fidelity, in the range of  $10^{-6}$  to  $10^{-7}$   $\mu$  comparable to DNA replication with proofreading but lacking mismatch repair (Eckerle et al., 2010; 2007; Smith et al., 2014) (Figure 4).

Mutations provide the raw material for adaptation and emergence, so low-fidelity replication is likely responsible for the capacity of RNA viruses to adapt to new host species and ever changing environmental pressures (Domingo, 2010; Eigen, 1993; Sanjuán, 2012). Considering the mutation rates described along with genome sizes averaging around 12,000 base pairs, RNA viruses generally incorporate between 0.4 and 1.1 mutation per genome (Domingo *et al.*, 1996; Drake *et al.*, 1998). Adding to this source of variation, RNA viruses rarely exist as individuals. Instead, RNA viruses undergo logistic growth, resulting in large genetically diverse populations that could, theoretically, include a substitution at every site in their genome (Lauring and Andino, 2010). Therefore, RNA viruses exist as a population of heterogeneous yet related viruses (mutant swarms) rather than a group of genetically identical clones. Mutant swarms have been shown to display phenotypes that differ from that of their consensus sequence alone (Ojosnegros *et al.*, 2011). In other words, mutant swarms seem to be more than a simple sum of their parts. Increasing fidelity results in decreased pathogenesis and spread, either through a decrease in replication speed, the ability to adapt, or both (Lauring and Andino, 2010; Pfeiffer and Kirkegaard, 2005; Regoes et al., 2012; Vignuzzi et al., 2008). One of the most notable examples of this phenomenon was described in poliovirus. Wild type poliovirus is neurovirulent



**Figure 4. Comparison of mutation rates between RNA viruses and cellular DNA polymerases**

All forms of nucleic acid-based replication result in errors. RNA or DNA polymerases generally manifest incorporation errors of roughly  $10^{-4}$  to  $10^{-6}$  mutations per site per round of replication ( $\mu$ ). Coronaviruses, and the other large nidoviruses, are unique among RNA viruses in encoding a proofreading 3'-5' exoribonuclease, which decreases the error rate to between  $10^{-6}$  to  $10^{-7}$   $\mu$ . Cellular DNA replication results in the lowest number of misincorporations.  $10^{-9}$  to  $10^{-11}$   $\mu$ : mismatch-excision repair pathways in addition to proofreading permit these low mutation rates. Adapted from (Smith et al., 2014).

in mice but high-fidelity poliovirus populations were unable to target the brains of mice. However, the identical high-fidelity poliovirus strain forced toward greater population diversity (through the use of mutagens) was able to infect the brain (Vignuzzi *et al.*, 2005b). No single mutation was identified that conferred this expanded tropism. The ability of populations of viruses to act collectively is likely dependent on the fact that multiple viruses often infect a single cell (Cheng *et al.*, 2005). Multi-particle infections can allow genomic products to be shared during infection (Combe *et al.*, 2015; Lauring *et al.*, 2013; Ojosnegros *et al.*, 2011). In some cases this may lead to novel functions by mutated proteins that are normally intolerant to mutation because of their essential role in viral replication. Overall, mutations and large population sizes combine to enhance the adaptability of viruses.

### **Lethal mutagenesis and error catastrophe**

Although the generation of mutations is essential for adaptation, one of the prevailing misconceptions regarding viral mutation rates is that increasing the number of mutations proportionally enhances the rate at which viral populations adapt (Elena and Sanjuán, 2005; Furio *et al.*, 2005). This belief has contributed to a perceived danger of mutator variants that contradicts the evidence; instead, what is observed is that increasing the mutation rate beyond that of the wild type virus often results in a decrease in fitness (Gnädig *et al.*, 2012; Graham *et al.*, 2012; Lee *et al.*, 1997). There are two main reasons for this decrease in fitness: 1) the majority of viral mutations are deleterious (Drake and Holland, 1999; Lauring and Andino, 2010; Lee *et al.*, 1997), and 2) RNA viruses seem to replicate at or around the error threshold. The error threshold is the maximum mutations a replicating entity can tolerate before no longer being able to replicate resulting in lethal mutagenesis (Lauring and Andino, 2010). Marginal increases in mutation rates could result in too many deleterious mutations within the population (Domingo *et al.*, 2012). The error threshold was first demonstrated by treatment of poliovirus with 5-FU which resulted in a set increase in mutation frequency no matter the amount of mutagen applied but with decreasing surviving virus (Holland *et al.*, 1990). This suggested that the virus could only tolerate increased mutations up to the error threshold before ceasing to replicate. This has also been demonstrated for coronaviruses with an inactivated nsp14-ExoN (nsp14-ExoN(-)) (Smith *et al.*, 2013). Thus, RNA viruses exist within a narrow sequence space where fidelity and infidelity are balanced. The enzymatic functions of RNA viruses must be balanced to achieve optimal replication, spread and survival.

As stated previously, CoVs encode the largest (+)RNA genomes, and increased fidelity in CoVs is likely required for maintenance of these large genomes (Smith *et al.*, 2014). These extensive genomes encode for an array of virus-specific proteins, which direct the mechanisms CoVs use to enter cells and usurp the metabolic machinery. Larger genomes are accompanied by fundamental difficulties not present for organisms of smaller size. First, there is a limit to the number of mutations that are tolerable within a single genome while maintaining function (Chen and Shakhnovich, 2009; Crotty and Andino, 2002; Graci *et al.*, 2012), and given a similar propensity for mutations by the polymerase, errors increase as the length of a genome increases. Next, polymerases have inherent limits to processivity, and replication requires generation of full genomes. The generation of full genomes is impeded when length increases beyond the processivity of a polymerase. Transcription of the genome similarly requires mechanisms for efficiently replicating RNA through all genes. Additionally, products of viral replication are potential targets of the innate immune system, which the virus must regulate in order to optimize replication. So, as the size of the genome increases so do the substrates that must be managed. However, many of these issues are directly addressed through the increased genome size of CoVs. Models have been proposed that envision large CoV replication-transcription complexes (RTCs) including nsp7-14 (Ma *et al.*, 2015; Smith *et al.*, 2014). Many of these proteins may directly address the size of the CoV genome. For instance nsp7 and nsp8 likely function together as a processivity clamp and nsp14-ExoN, with enhancement from the non-enzymatic nsp10 protein, is able to correct errors introduced by the polymerase. Interestingly, these models - based on a combination of available structural, biochemical and biological data - more closely resemble the replication complexes observed in DNA-based organisms rather than those of other RNA-encoded viruses.

### **Structures of, and nucleotide selection by, the RdRps of RNA viruses**

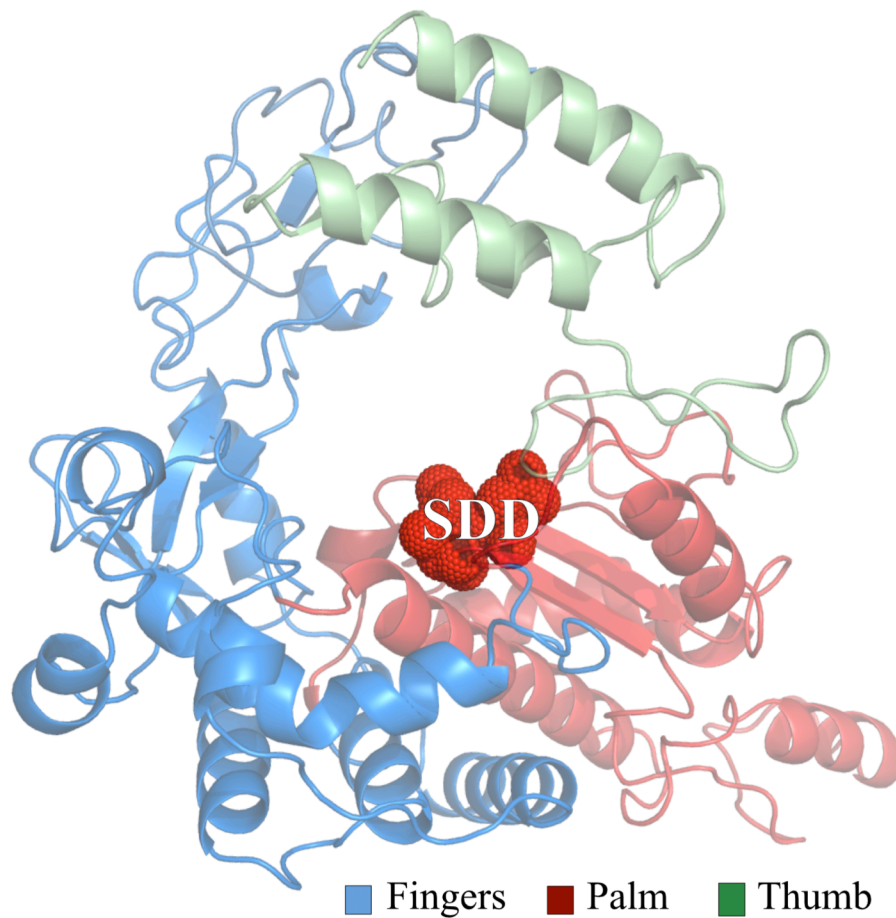
The RdRp is central to the replication of RNA viruses and is generally the key regulator of nucleotide selectivity and fidelity (Arnold *et al.*, 2005; Campagnola *et al.*, 2015). The CoV RdRp must similarly be essential to the regulation of nucleotide selectivity and overall fidelity of CoV replication. However, CoVs, and other large nidoviruses, replicate with higher fidelity than all other known (+)RNA viruses (Eckerle *et al.*, 2010; Smith *et al.*, 2013) and this is attributed to proofreading by nsp14-ExoN. All known RNA viruses encode for RdRps, and most also encode for a helicase. Generally, the replication complexes of RNA viruses are limited in size. However,



CoVs encode larger RTCs with their extensive genomes. Some of the encoded proteins are ubiquitous throughout the RNA virus world such as the RdRp. Others are unique to members of the *Nidovirales* – including CoVs – such as nsp14-ExoN and nsp10, a modulator of nsp14-ExoN activity (Bouvet *et al.*, 2012; Smith and Denison, 2013). Still, others have functions that are not yet fully understood within the context of infection such as the uridylate-specific endonuclease activity of nsp15. With so many CoV-specific proteins involved in replication, it is possible that selective pressures acting on the CoV RdRp differ compared to the RdRps of other RNA viruses.

All described polymerase structures (including RdRps) resemble a “cupped right hand” with fingers, palm and thumb domains (Ng *et al.*, 2008) (Figure 5). The fingers form a channel that allows entry of the template RNA and ribonucleoside triphosphates (rNTPs) and assist in proper positioning of incoming nucleotides in the active site (Ferrer-Orta *et al.*, 2007). The palm contains the active site, and the thumb functions in contacting exiting nascent RNA (Ferrer-Orta *et al.*, 2007; 2015; Gong and Peersen, 2010). The characteristic Gly-Asp-Asp (GDD) active site sequence of RdRps sits within the palm domain; CoVs instead encode SDD. All polymerases have “open” and “closed” conformations that are essential to polymerization and processivity. In most polymerases, transition from one to the other involves separation of the fingers and thumb domain while the palm domain remains in an active conformation. Conversely, all RdRps with solved structures display strong interactions between the fingers and thumb domains that are not released during polymerization (Ferrer-Orta *et al.*, 2007). This is likely the case for CoV as well, although there are no solved structures for any CoV RdRps to date. Supporting this, CoV RdRp core domains can be modeled against solved structures of other viral RdRps (Sexton *et al.*, 2016; Xu *et al.*, 2003) (Figure 5).

Structural studies have begun to elucidate the mechanisms of ribonucleotide addition by RdRps. Residues around the active site are responsible for correct selection of nucleotide bases and sugars. They are also responsible for coordinating the cations required for catalyzing new phosphodiester bonds and catalysis itself. However, ribonucleotide addition is not a static process; as stated above, the active site is found in both open and closed conformations (Gong and Peersen, 2010). Active site closure involves shifting amino acid positions throughout the RdRp structure providing a mechanism by which residues distant from the active site alter



**Figure 5. Model of the MHV RdRp core domain**

All RdRps share a structurally homologous core that can be subdivided into fingers, palm and thumb domains. Unique to RdRps is a strong interaction between the fingers and thumb domains, which prevents the two domains from separating during shifts from open to closed conformations. The active site of all RdRps is located in the palm domain and typically contains a central catalytic GDD motif; however, CoVs instead encode an SDD catalytic triad, highlighted in the above structural model.

elongation rates and nucleotide selection (Arnold *et al.*, 2005). The extensive interactions between the fingers and thumb domains of RdRps prevent the fingers from swinging out when transitioning between open and closed conformations, as is the case for other polymerases (Gong and Peersen, 2010). Therefore, any residue in the fingers or thumb that alters the flexibility required for shifting between conformations can affect fidelity. Additionally, residues of the fingertips make first contact with incoming nucleotides so substitutions in these residues can alter fidelity of nucleotide selection (Curti and Jaeger, 2013).

### **The CoV replication-transcription complex (RTC)**

All RdRps share the above characteristics, however, there is also diversity among RdRps. Additional domains are often present that encode a variety of functions such as methyltransferase, endonuclease, polyribonucleotidyl-transferase, guanylyltransferase, membrane targeting, protein-protein binding or protein-RNA binding activities (Ferrer-Orta *et al.*, 2007; 2015; Gong and Peersen, 2010; Perlman and Netland, 2009). CoVs similarly encode a CoV-specific domain within the same protein as the RdRp core domain. This CoV-specific domain has low predicted structural homology to known proteins, so, cannot be modeled with high confidence (data not shown). However, the CoV-specific domain is highly conserved across all currently sequenced CoVs suggesting an important role in CoV biology. Currently, little is known about the role the CoV-specific domain plays in the virus life cycle. However, a conserved and essential nucleotidylating activity was elucidated within the CoV-specific domain. This sub-domain was coined the nidovirus RdRp-associated nucleotidyltransferase (NiRAN) and has no known viral or cellular homologs (Lehmann *et al.*, 2015). The NiRAN sub-domain could be involved in capping or other RTC functions; however, more data are required to determine its specific role. Additionally, a small sequence within the CoV-specific domain of the RdRp protein is essential for localizing the protein to the replication complexes (Brockway *et al.*, 2003)

In many (+)RNA viruses, the RdRp is the only protein required for replication of a template *in vitro*, and the core of the CoV RdRp can be modeled with strong agreement on solved structures of other viral RdRps (Sexton *et al.*, 2016). However, unlike the RdRps of most other (+)RNA viruses, the CoV RdRp directly interacts with, and requires the presence of certain ancillary factors, nsp7 and nsp8, for efficient binding of RNA (Subissi *et al.*, 2014). Together with nsp7 and nsp8 the CoV RdRp demonstrates high processivity, further supporting the cooperative

function of nsp7 and nsp8 as a processivity clamp reminiscent of DNA-based systems as suggested previously (Smith et al., 2014; Subissi et al., 2014). An additional function of nsp8 may be primase activity, and nsp12-RdRp likely requires the presence of primers for replication and transcription (Imbert et al., 2006; Sevajol et al., 2014; Xiao et al., 2012). Finally, although not essential for replication in a biochemical setting, nsp13 encodes the viral helicase/NTPase, and is likely essential for replication in the context of infection (Ahn et al., 2012; Seybert et al., 2005; Subissi et al., 2014; Velthuis et al., 2009). Interaction studies demonstrate direct interactions between nsp12 and nsp13, and most RNA viruses encode helicases (Adedeji et al., 2012; Imbert et al., 2008). Therefore, the core replication complex for CoVs may legitimately be defined as the combination of nsp7, 8, 12 and 13 (Figure 6).

### **Fidelity proteins of the replication-transcription complex**

Additional proteins are present in the replication complex beyond those directly essential to RNA replication. One of the most intriguing functions conferred by these additional proteins is the ability to increase the fidelity of replication over that of the polymerase alone. This proofreading activity is attributed to the 3'-5' exoribonuclease (ExoN) encoded in the nonstructural protein 14 (nsp14) of CoVs (Figure 2). The role of nsp14-ExoN proofreading in regulating fidelity has been extensively reviewed in (Smith et al., 2014). Many studies to date have demonstrated the ability of CoV nsp14-ExoN to remove mismatched nucleotides and increase the overall fidelity of CoVs up to 20-fold (Bouvet et al., 2012; Eckerle et al., 2010; Graham et al., 2012; Smith et al., 2013). Additionally, the small non-enzymatic protein nsp10 interacts with nsp14-ExoN and increases the activity of nsp14-ExoN by up to 35-fold *in vitro* (Bouvet *et al.*, 2012). Enhancement of nsp14-ExoN proofreading by nsp10 has also been demonstrated *in vitro* and during viral replication (Bouvet et al., 2014; Smith et al., 2015; Subissi et al., 2014). Mutations in nsp10 at the site of interaction between nsp10 and nsp14 increase susceptibility of CoVs to a panel of nucleoside analogs, and this sensitivity is increased with elevated temperatures during viral replication. However, ExoN-inactivated viruses are not subject to fidelity modulation mediated by nsp10 (Smith *et al.*, 2015). Enhancement of ExoN activity is likely a combination effect of direct stimulation of activity by nsp10 as well as the ability of nsp10 to stabilize nsp14, reminiscent of the *E. coli* DNA polymerase III holoenzyme (Maki and Kornberg, 1987; Scheuermann and Echols, 1984; Smith et al., 2014). Regardless, binding of nsp10 to nsp14 seems to directly correlate with increased ExoN activity



(Bouvet *et al.*, 2014). Similar to the case with *E. coli*, CoV nsp14 required the presence of nsp10 for crystallization. Interestingly, an nsp10-nsp14 co-crystal structure demonstrated that the active site residues for nsp14-ExoN are DEEDh instead of the previously presumed DEDDh. However, the overall structural elements of nsp14-ExoN still suggest it should be categorized with the DEDD superfamily of exonucleases (Ma *et al.*, 2015). Previously, nsp10 was co-crystallized with nsp16, the 2'-O-methyltransferase, and it was determined that the sites of interaction between nsp10-nsp16 and nsp10-nsp14 were similar (Chen *et al.*, 2011). More recent data demonstrated that the interaction face between nsp10-nsp14 is more extensive than that of the nsp10-nsp16 interface, and the two interfaces do overlap (Bouvet *et al.*, 2014; Ma *et al.*, 2015).

Nsp14-ExoN and nsp10 are key proteins for CoV replication fidelity; however, in other RNA viruses the RdRp is the main determinant of fidelity. Other viral RdRps can be mutated resulting in up to a 4-fold change in fidelity while still producing viable viruses (Smith *et al.*, 2014). For CoVs, it could easily be imagined that the presence or absence of nsp14-ExoN activity alters the phenotypic outcome of mutations in the RdRp domain, considering the much larger fold change in fidelity ensuing from mutations in nsp14-ExoN (Eckerle *et al.*, 2007; 2010; Smith *et al.*, 2013). Additionally, CoV nsp12 directly interacts with nsp14, bringing the RdRp and ExoN domains into close contact (Subissi *et al.*, 2014). Together, these data suggest that at minimum nsp12-RdRp, nsp14-ExoN and nsp10 are determinants of fidelity for CoVs. Additionally, with known interactions between nsp7/nsp8 with nsp12, nsp13 with nsp12, and nsp10 with nsp16, it is certainly possible that additional replication complex proteins alter fidelity in CoVs either on a host-species basis or potentially even with changes in the cellular environment.

#### **Additional functions of coronavirus replication-transcription complex proteins**

Many CoV proteins are multi-functional. RNA capping is important for efficient translation in the absence of an alternative mechanism such as an internal ribosome entry site (IRES) and for escape from host innate immune responses. Capping in CoVs is well reviewed in (Sevajol *et al.*, 2014). In addition to encoding ExoN proofreading activity at the N-terminal end, an N7-guanine methyltransferase (N7-MTase) is encoded at the C-terminal end of nsp14 (Subissi *et al.*, 2014). The interface between the two domains of nsp14 is highly conserved, hydrophobic and extensive, suggesting the subunits function largely in concert with each other and that manipulation of one domain may affect the other (Ma *et al.*, 2015). *In vitro*, when in complex

with nsp7/8, nsp12 and nsp10, nsp14 maintains N7-MTase activity, demonstrating that N7-MTase activity is potentially occurring within the context of the full complex (Subissi *et al.*, 2014). Interestingly, this context may influence the activity of N7-MTase as both the aspartate and glycine of the DxG motif of the S-adenosyl-L-methionine binding pocket appear to be essential to binding *in vitro* but aspartate appears have no effect on viral replication (Chen *et al.*, 2009). In contrast the glycine residue of the DxG motif was shown to regulate sensitivity to IFN- $\beta$  as well as translation efficiency in the context of viral infection (Case *et al.*, 2016). Altering nsp14-N7-MT activity does not appear to affect nsp14-ExoN activity. Nsp10 additionally plays a dual role in fidelity and capping by interacting with the nsp16 the 2'-O-methyltransferase, and nsp13 contains NTPase activity in addition to helicase activity (Decroly *et al.*, 2011; Sevajol *et al.*, 2014). It is possible that nsp16 itself is also a member of the RTC but experimental evidence is needed to support this conjecture.

Involvement in capping may not be the only mechanism by which nsp14 influences innate immune sensing. Using transmissible gastroenteritis virus (TGEV), an *Alphacoronavirus*, it was determined that a mutation in the N-terminal-most zinc finger of ExoN (ZF1) resulted in decreased plaque size, decreased activation of IFN- $\beta$ , and decreased accumulation of dsRNA intermediates (Becares *et al.*, 2016). With the complex nature of nsp14 it is difficult to determine by what specific mechanism the decreased IFN- $\beta$  response to TGEV is occurring. However, it is conceivable that nsp14-ExoN may serve a function outside of proofreading. Interestingly, so far no *Alphacoronavirus* has been successfully recovered with an inactivated nsp14-ExoN. In fact, only *Betacoronavirus* species have been observed to tolerate this inactivation to date. It will be interesting to see whether inactivation of nsp14-ExoN in a *Betacoronavirus* will similarly result in altered innate immune responses. Future research in this area may elucidate additional functions of nsp14-ExoN.

### **Co-evolution of interactions within the RTC and potential for additional determinants of fidelity**

The RTC of CoVs contains at minimum nsp7-14 and may include nsp15 and nsp16 as well (Brunn *et al.*, 2007; Imbert *et al.*, 2008; Smith *et al.*, 2014). Many of these proteins have multiple domains with disparate functions involved in the production of complete viral RNAs. Due to the extensive interactions between these proteins, the members of the complex must co-evolve in a

cooperative manner. It would be expected that there would be heavy overall negative, or purifying, selection exerted on these proteins in order to maintain essential functions, not only in individual proteins but also in their co-factors and binding partners. This has been observed for lineage C *Betacoronavirus* species with greater than 25 percent of the residues in nsp 7, 8, 12, 13 and 14 undergoing negative selection (Forni *et al.*, 2016). With close interactions among the proteins of the RTC, it is conceivable that mutations outside of nsp12-RdRp, nsp14-ExoN and nsp10 could also influence fidelity and nucleotide selectivity. This has been demonstrated for other organisms. For instance, a mutation in the helicase of Chikungunya virus was shown to increase the overall fidelity of that virus in an RdRp-independent manor (Stapleford *et al.*, 2015). The *Mycobacterium tuberculosis* DNA polymerase III holoenzyme replication complex, a proposed analogue of the CoV RTC, is composed of  $\alpha$ , the polymerase subunit (containing exonuclease activity);  $\epsilon$ , an additional exonuclease; and  $\beta_2$ , a processivity clamp. The  $\beta_2$  clamp of *M. tuberculosis* DNA polymerase III, was demonstrated to increase the exonuclease activity in the  $\alpha$  subunit. Interestingly, in contrast, the  $\beta_2$  clamp reduced the exonuclease activity of the full  $\alpha\epsilon\beta_2$  complex, which includes the external exonuclease (Gu *et al.*, 2016). Similarly, increased DNA binding by the processivity factor of Herpes Simplex virus resulted in an increase in mutation frequencies (Jiang *et al.*, 2009). Together, these data suggest that nsp7/nsp8 in complex could modulate fidelity in CoVs. However, in distinction to CoVs and *E. coli*, the  $\alpha$  and  $\epsilon$  subunits of *M. tuberculosis* likely function together through an intermediate interaction with the  $\beta_2$  clamp instead of directly with one another (Gu *et al.*, 2016; Smith *et al.*, 2014). Finally, one could imagine that the 2'-O-methyltransferase affects fidelity by altering its interaction with nsp10, thus controlling availability of this stimulatory co-factor to nsp14-ExoN. Overall, altering individual functions of proteins in the RTC likely has a reverberating effect, altering interactions and functions of other proteins. These adjustments may serve to balance fidelity under a variety of conditions *in vitro* and *in vivo*.

### Conclusion

Currently, there are no available vaccines or antiviral treatments against any human CoV. Many of the non-structural proteins encoded by coronaviruses are present only among members of the *Nidovirales*. Therefore, understanding these proteins, specifically those of the RTC could lead to identification of novel drug targets. Additionally, altering viral replication fidelity is attenuating



in most circumstances. Understanding the proteins involved in maintaining viral fidelity may foster development of vaccine candidates for CoVs.

At the start of this dissertation research, evidence had been presented implicating only nsp14-ExoN in fidelity regulation in the context of viral replication. In this dissertation, I present data demonstrating for the first time the involvement of nsp12, particularly nsp12-RdRp, in the regulation of fidelity for CoVs. Moreover, I demonstrate that the functional relationship between the activities of nsp14-ExoN and nsp12-RdRp is complex. There are many proteins functioning together within the RTC and these likely also play a role in fidelity regulation. Since this research began, nsp10 has been shown to influence replication fidelity during viral infection, supporting the previous biochemical evidence. Additionally, other proteins and processes have been implicated in the regulation of fidelity of other RNA viruses beyond simply nucleotide selection and proofreading. Therefore, it is possible that other members of the RTC contribute to the overall fidelity of CoVs. CoVs offer a unique platform to interrogate the contributions of different replication components to overall fidelity. However, due to size and complexity of CoVs, fidelity studies have not been amenable to many of the techniques used in other viral systems. Therefore, in this dissertation I also present data toward the development of novel techniques for investigating fidelity in CoVs that also should be applicable to other viral systems. This is important, as another current limitation to viral fidelity research is the inability to accurately compare data across viruses.

In Chapter II, I present work identifying and characterizing residues in the CoV RdRp with predicted involvement in nucleotide selectivity by homology to picornaviruses (Sexton *et al.*, 2016). In Chapter III, I describe the identification of fidelity-regulating mutations in nsp12 and nsp14 of CoVs through adaptive selection methods as a complement to work in Chapter II. In Chapter IV, I present results from different techniques designed to investigate fidelity and demonstrate the potential for a novel approach. Chapter V summarizes the materials and methods used throughout this dissertation. Chapter VI explores the implications and future direction of this work.

## CHAPTER II

# HOMOLOGY-BASED IDENTIFICATION OF A MUTATION IN THE CORONAVIRUS RNA-DEPENDENT RNA POLYMERASE THAT CONFERS RESISTANCE TO MULTIPLE MUTAGENS

### Introduction

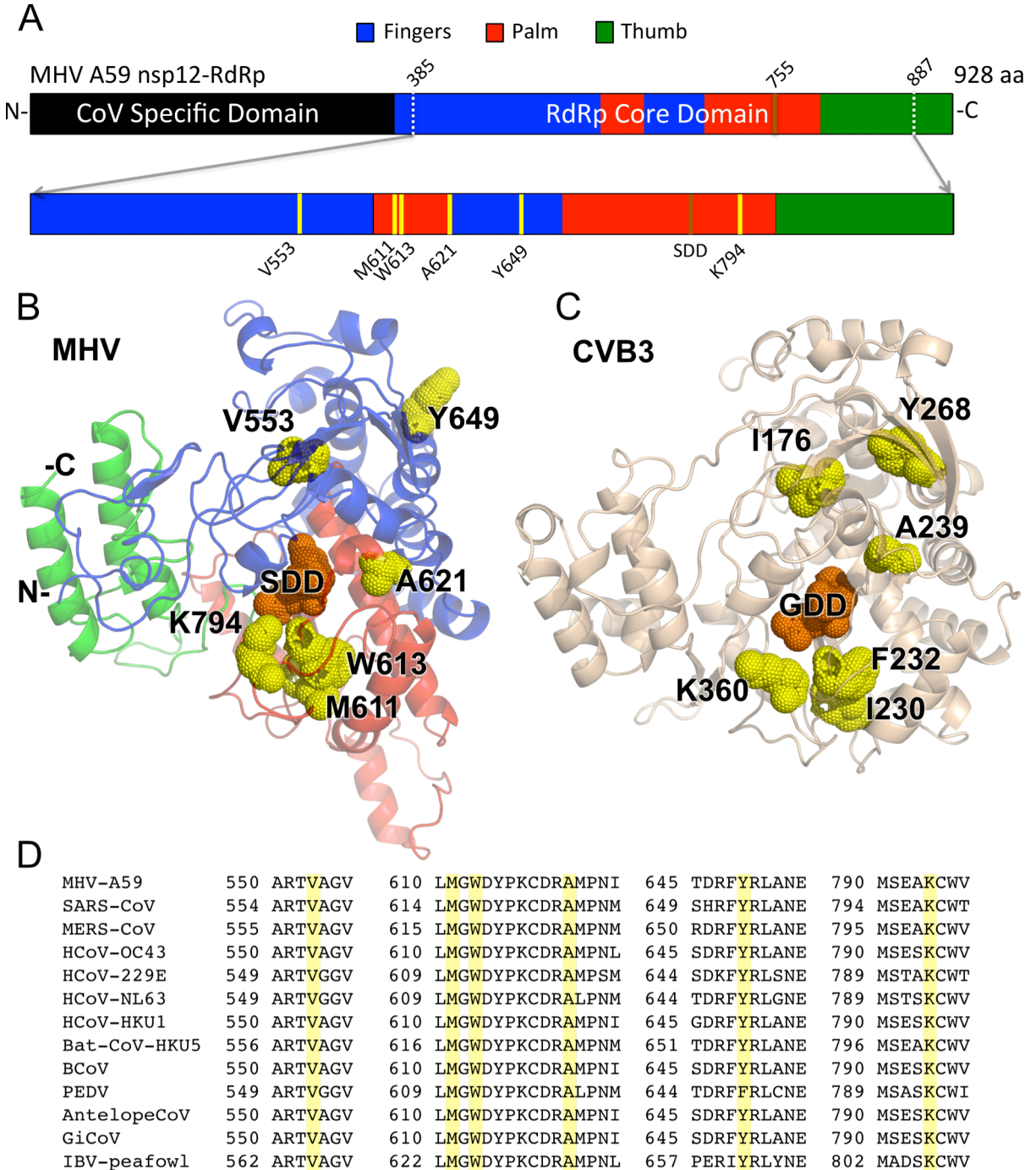
At the start of this dissertation research, coronavirus fidelity had been investigated only in regards to the 3'-5' exoribonuclease, encoded in non-structural protein 14 (nsp14-ExoN). It had been demonstrated that inactivation of nsp14-ExoN activity by mutating the DE of the DEED active site to AA resulted in an increase in mutations and increased susceptibility to multiple mutagens. The data suggested that inactivation of nsp14-ExoN resulted in up to a 20-fold decrease in fidelity (Denison et al., 2011; Eckerle et al., 2007; 2010; Graham et al., 2012). Up to this discovery, all fidelity determinants in positive-sense RNA ((+)RNA) viruses had mapped to the RNA-dependent RNA polymerase. Studies of CVB3, polio, HIV-1 and other viruses demonstrate that viable viruses are only recoverable within a four-fold range of RdRp fidelity (Dapp et al., 2013; Gnädig et al., 2012; Smith et al., 2014; Vignuzzi et al., 2005a). In most cases altered RdRp fidelity decreases fitness relative to WT viruses; this has been demonstrated for changes as small as a 1.2 fold difference in the accumulation of mutations (Pfeiffer and Kirkegaard, 2005; Severson et al., 2003; Smith et al., 2014; Vignuzzi et al., 2005a). There are no solved structures for any CoV nsp12-RdRp, but the presence of conserved RdRp motifs and modeling of the C-terminal half of nsp12 predicts an RdRp domain that is structurally similar to other RNA viruses (Lehmann et al., 2015; Xu et al., 2003).

The demonstrated function of nsp14-ExoN in high fidelity CoV replication raised the question of whether and how nsp12-RdRp participates in fidelity regulation. I sought to determine whether nsp12-RdRp could modulate nucleotide selectivity independently or in association with the proofreading nsp14-ExoN. I modeled the RdRp domain of CoV nsp12 on coxsackievirus virus B3 (CVB3) and poliovirus polymerase structures and predicted residues important for fidelity based on prior results from those virus systems. Substitution mutations at these residues were introduced in the isogenic recombinant genome of the  $\beta$ -CoV, murine hepatitis virus (MHV-A59). In this chapter I demonstrate that two of these mutations, nsp12-V553I and nsp12-M611F,

confer resistance to the mutagen 5-fluorouracil (5-FU) and one, nsp12-V553I, also results in resistance to the mutagen 5-azacytidine (5-AZC) and demonstrates a decreased accumulation of mutations. Increased mutagen resistance and decreased accumulation of mutations was only observed in viruses with an inactivated nsp14-ExoN, demonstrating that nsp14-ExoN proofreading activity is epistatic to the nucleotide selectivity of nsp12-RdRp. For the project I define epistasis as a situation where the phenotype of one gene or viral protein masks the phenotype of genetic variants of another viral protein. This result is consistent with a primary role for nsp14-ExoN in error recognition and removal. However, introduction of RdRp mutations within the WT-MHV background decreased fitness relative to WT. Together the results suggest that nsp12-RdRp shares common determinants of nucleotide selectivity with RdRps from other RNA virus families. Further, the CoV RdRp has likely evolved to function in cooperation with nsp14-ExoN rather than independently. I performed all experiments and final analysis for the data in this chapter. My coauthors provided assistance in the following experiments: Clint Smith developed the qRT-PCR protocol used in the specific infectivity assays, Hervé Blanc in Marco Vignuzzi's lab prepared the samples I sent them for deep sequencing, ran the samples and sent the data through the ViVan bioinformatics pipeline for initial cleanup and analysis, and Olve Peersen provided guidance on the homology modeling.

### **Homology modeling of MHV nsp12-RdRp polymerase core domain predicts putative fidelity determinants**

Mutations that alter nucleotide selectivity have been identified across multiple RNA virus RdRps (Castro et al., 2005; Gnädig et al., 2012; Levi et al., 2010; Pfeiffer and Kirkegaard, 2003; Verdaguer and Ferrer-Orta, 2012); however, whether these residues are conserved across virus families is unknown. I sought to determine whether residues within nsp12-RdRp that are structurally homologous to known RNA virus fidelity determinants would have similar effects on nucleotide selectivity when introduced into the MHV background. To do this, we modeled the structure of MHV nsp12-RdRp using Phyre2 software (Kelley *et al.*, 2015). A series of nsp12-RdRp truncations was assessed, and the highest confidence model was used for further study. This region corresponded with a published model for the SARS-CoV nsp12-RdRp (Xu *et al.*, 2003) and included residues 385-887 of the MHV nsp12 protein, referred to here as the RdRp core domain (Figure 7a, b). Deletion of the CoV-specific domain (residues 1-384), and a small C-terminal portion of the thumb domain (888-928), was required to establish this high-



**Figure 7. Homology modeling of CoV nsp12-RdRp and identification of residues that potentially regulate fidelity based on CVB3 structure**

Phyre2 software was used to model **A**, a subsection of the MHV nsp12-RdRp core domain (expanded from nsp12-RdRp full schematic). This **B**, modeled MHV RdRp structure was aligned with the **C**, solved CVB3 RdRp structure. Residues chosen for site-directed mutagenesis (**A and B**) were selected by comparing previously determined fidelity altering mutations of picornavirus RdRps. Amino acid alignments **D**, across CoVs show residues are almost completely conserved.

confidence model (Figure 7a). The model was resolved by highest probability similarity to human rhinovirus serotype 1b (PDB ID: 1XR5a), rabbit hemorrhagic disease virus (PDB ID: 1KHV) and EV71 (PDB ID: 3N6M). The Phyre2 confidence, i.e. the probability of true homology, for the RdRp core domain to these structures was >99% while the percent identity was only 14-20%. Having generated a structural model for the MHV nsp12-RdRp core domain I next sought to predict residues involved in nucleotide selectivity. The nsp12-RdRp core domain model was aligned with the solved structure of coxsackievirus B3 (CVB3, PDB ID: 3DDK) using PyMol (Figure 7c). A series of CVB3 RdRp mutations have been shown to result in decreased fidelity (Campagnola et al., 2015; Gnädig et al., 2012; Yang et al., 2012). The CVB3 fidelity determinants were compared with the MHV nsp12-RdRp core domain model. Those that aligned well structurally and by amino acid similarity were further investigated: MHV nsp12-V553, M611, W613, A621, Y649 and K794 (Figure 7). Finally, the nsp12 amino acid sequences of 27 different  $\alpha$ -,  $\beta$ - and  $\gamma$ -CoVs were aligned, including SARS-CoV and MERS-CoV. All six identified residues were conserved across these CoVs (Figure 7d). Analysis of similarity and types of residues in the picornaviruses were then used to determine the specific amino acid changes that would be introduced at the identified MHV residues. The resulting substitution mutations were engineered in the isogenic cloned MHV genome: nsp12-V553A/I, M611F, W613Y, A621G, Y694H/W and K794R (Table 1).

### **Recovery of mutant viruses in the MHV nsp14-ExoN(+) and nsp14-ExoN(-) isogenic backgrounds**

I next tested whether viable viruses could be recovered with substitutions at the identified residues. Virus recovery was attempted a maximum of 3 times, resulting in recovery of 6 of the 8 mutant viruses in the WT background: nsp12-V553I, M611F, W613Y, A621G, Y649H, and K794R. The time required for recovery of mutant viruses in the WT background ranged from 24 to 48 hours. No other mutations were identified across nsp12 sequences in recovered viruses. The A621G mutant was not further studied as it demonstrated rapid primary reversion even in the recovery (P0) supernatant. Since our goal was to understand the relationship of nsp12-RdRp and nsp14-ExoN in fidelity regulation, I additionally attempted recovery of the WT background viable mutants in the setting of inactivated nsp14-ExoN (nsp14-ExoN(-)). As nsp14-ExoN(-) results in delays in replication and decreased titers we did not attempt recovery of mutant viruses with mutation non-recoverable in the WT background. In contrast to mutant viruses recovered in

the WT background, I only recovered 2 of the 5 mutants in the nsp14-ExoN(-) background: nsp12-V553I/nsp14-ExoN(-) and nsp12-M611F/nsp14-ExoN(-) (Table 1). The time to recovery for nsp12-V553I/nsp14-ExoN(-) was 84 hours and for nsp12-M611F/nsp14-ExoN(-) was 96 hours. Working stocks of all viruses were made by infecting D9s with an MOI of 0.01 and recovering stocks at around 24 h.p.i for WT viruses, or between 32 and 48 h.p.i for mutant viruses in the nsp14-ExoN(-) background. Therefore 4-10 replication cycles were required to generate stocks. Working stocks were sequenced to verify that the introduced mutations were still present.

**Table 1. Recovery of mutant viruses using site-directed mutagenesis**

nsp12-RdRp Region	CVB3 engineered substitutions (nsp12) (8)	MHV engineered substitutions (nsp12)	Recovery: WT (nsp14-ExoN(+))	Recovery: nsp14-ExoN(-)
Fingers	I176V	V553A	No	Not attempted
		V553I	Yes	Yes
	Y268H	Y649H	Yes	No
	Y268W	Y649W	No	Not attempted
Palm	I230F	M611F	Yes	Yes
	F232Y	W613Y	Yes	No
	A239G	A621G	Revertant	Not attempted
	K360R	K794R	Yes	No

### **Resistance of recovered mutant viruses to the base analog 5-fluorouracil**

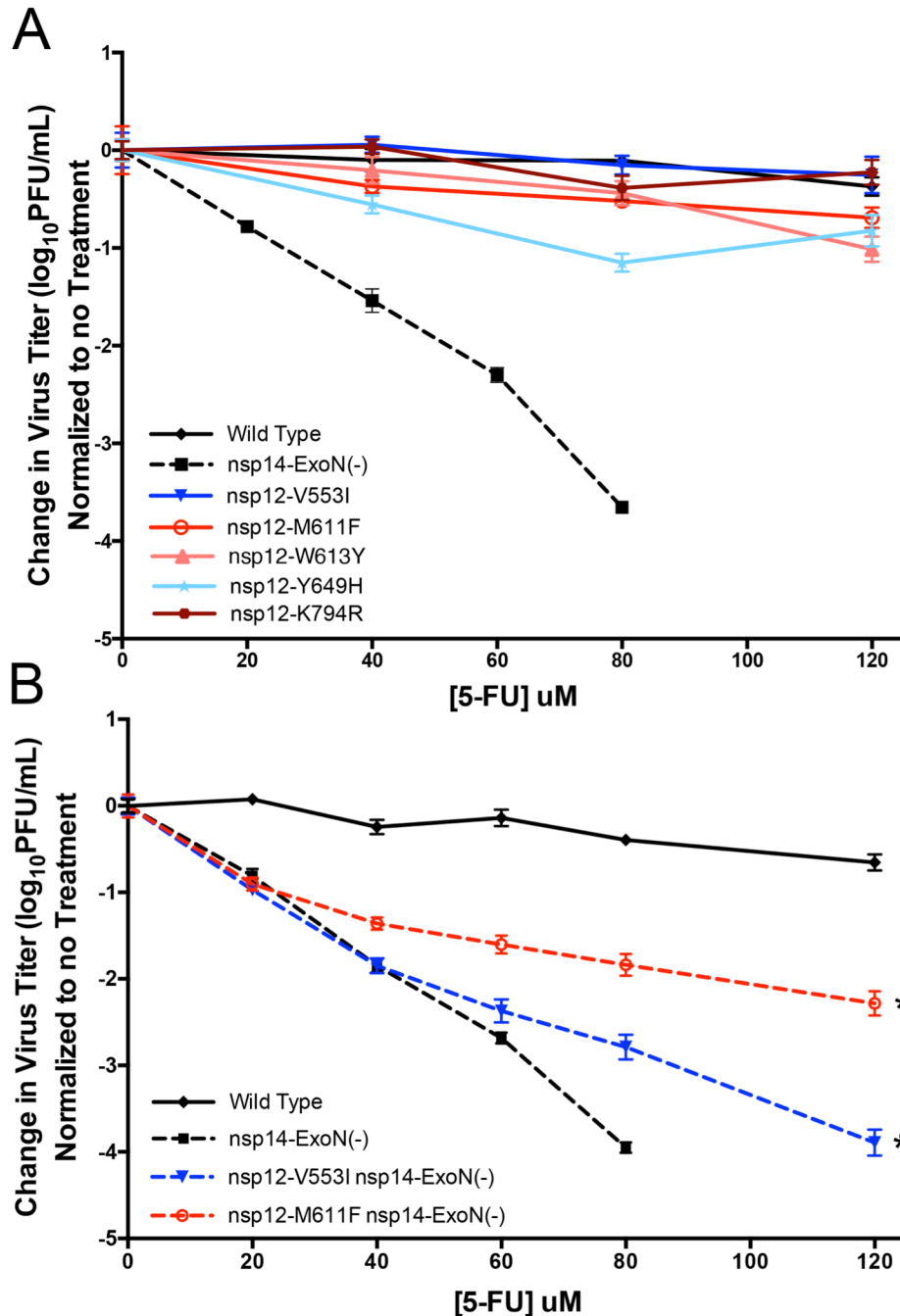
I next tested our panel of recovered mutant viruses for resistance to the RNA mutagen 5-fluorouracil (5-FU). 5-FU selectivity has been used to reveal changes in fidelity based on increased or decreased sensitivity to incorporation and virus inhibition for picornaviruses, influenza viruses, vesicular stomatitis viruses and others (Gnädig et al., 2012; Holland et al., 1990; Pauly and Luring, 2015; Smith et al., 2013). WT CoVs (MHV, SARS-CoV) are resistant to 5-FU, while nsp14-ExoN(-) mutants are profoundly sensitive to 5-FU inhibition, consistent

with nsp14-ExoN-mediated removal of misincorporated 5-FU. The effect of 5-FU on DBT cell viability was previously tested with no effect observed up to 400uM (Smith *et al.*, 2013). I compared WT-MHV with the nsp12 mutants in both the WT and nsp14-ExoN(-) backgrounds. There was no significant change in sensitivity to 5-FU compared to WT for any of the nsp12-RdRp mutant viruses in the WT background at up to 120uM 5-FU although the mutation Y649H does appear to decrease resistance slightly (Figure 8a). In contrast, the nsp12-M611F/nsp14-ExoN(-) and nsp12-V553I/nsp14-ExoN(-) mutant viruses both were significantly less sensitive to 5-FU than nsp14-ExoN(-) alone, with both populations persisting when treated with 120uM 5-FU where nsp14-ExoN(-) was not detectable beyond 80uM 5-FU (Figure 8b). These data demonstrate that both nsp12-RdRp mutations V553I and M611F confer resistance to 5-FU. This suggests a couple possibilities: that it is not possible to increase the exclusion of 5-FU beyond the high level dictated by nsp14-ExoN or that selectivity for native nucleotides over 5-FU is in fact increased by nsp12 mutations, but at a low level that is not detectable by changes in virus titer.

Three of the mutations (nsp12-Y649H, W613Y, and K794R) were viable in the nsp14-ExoN(+) background, but failed to grow in the absence of proofreading. The observation that they only grew in an nsp14-ExoN(+) background indicates that the mutations retained sufficient polymerase function to support virus replication, but it was critically dependent on having proofreading functionality, suggesting these mutations may have given rise to low fidelity variants. Unfortunately, the low titer from the nsp14-ExoN(-) background precluded direct sequencing analysis and I therefore cannot definitively show this is the case.

### **Replication kinetics of nsp12-V553I and nsp12-M611F mutant viruses in the WT and nsp14-ExoN(-) backgrounds**

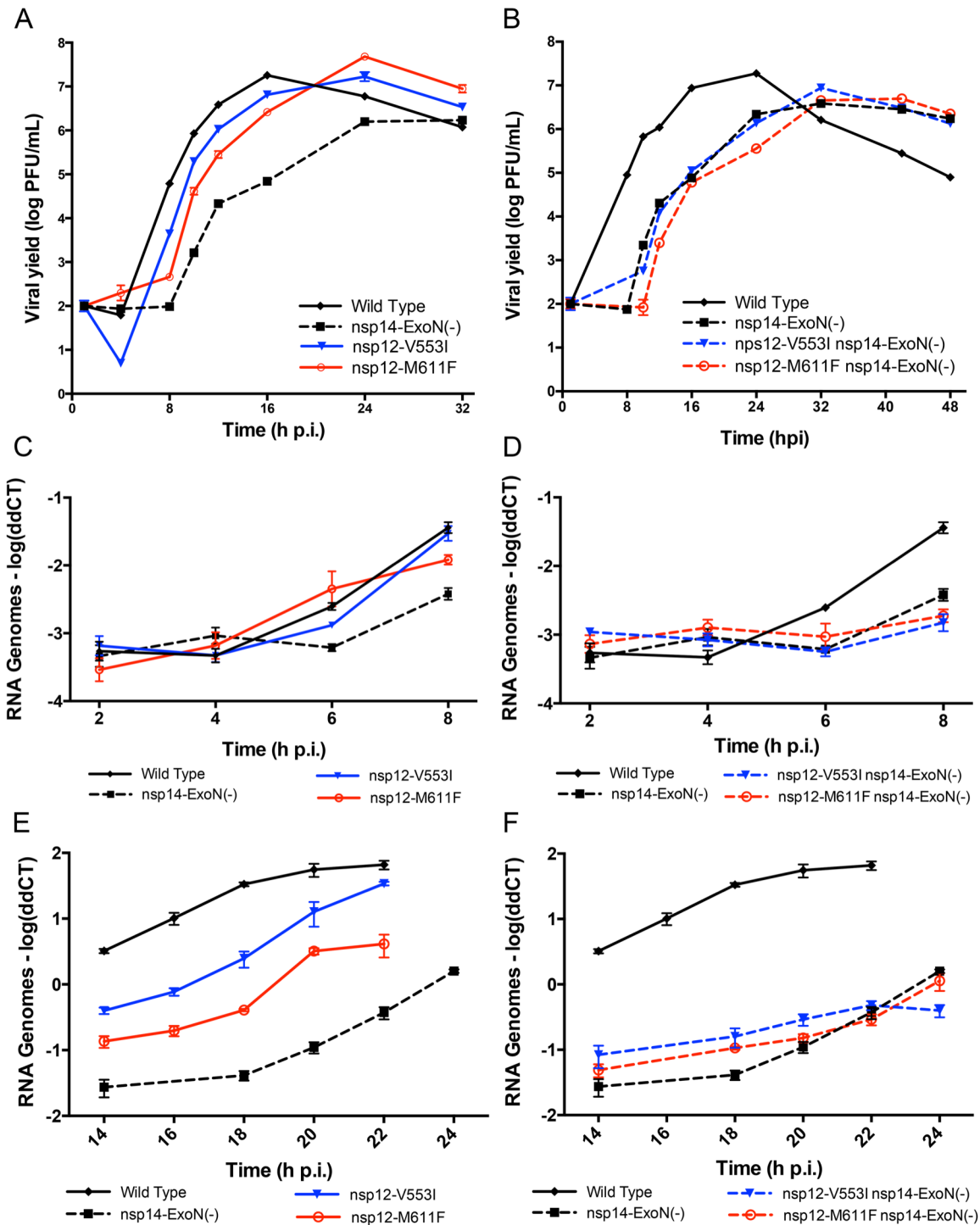
Since I was interested in mutations that potentially confer altered fidelity, I prioritized the nsp12-V553I and nsp12-M611F mutant viruses for further analysis. I next sought to determine how the nsp12-V553I and nsp12-M611F viruses replicated in comparison to their isogenic background (Figure 9). In the wild-type background both mutant viruses had slightly delayed exponential replication, but eventually reached similar peak titers comparable to WT. In contrast, in the nsp14-ExoN(-) background both nsp12-V553I and nsp12-M611F mutant viruses displayed similar replication kinetics to the isogenic nsp14-ExoN(-) background. I also assessed RNA



**Figure 8. Resistance of MHV nsp12-RdRp mutant viruses to 5-fluorouracil in the WT and nsp14-ExoN(-) background**

Domain location of mutations is indicated by: fingers (blues) and palm (reds). DBT cells were pre-treated with different concentrations of 5-FU for 30 min. Treatment was removed and cells were infected with indicated viruses in the **A**. WT background or **B**. nsp14-ExoN(-) background at an MOI of 0.01 PFU/cell. Media containing 5-FU was replaced 30 min p.i. Virus samples were taken at 24 (WT) or 32 (nsp14-ExoN(-)) h.p.i and titer was determined by plaque assay. Data represents 3 independent experiments, each with 2 replicates. Error bars represent SEM (\* $P < 0.05$  by Wilcoxon test and are relative to nsp14-ExoN(-)).





**Figure 9. Replication kinetics of MHV nsp12-RdRp mutant viruses**

Mutation location is indicated by: fingers (blue) and palm (red). DBT cells were infected with the viruses indicated in the **A, C and E**. WT background or **B, D and F**. nsp14-ExoN(-) background at an MOI of 0.01 PFU/cell. Supernatant aliquots were taken at indicated times p.i. and titer determined by plaque assay. Total RNA was taken at indicated times p.i. and RT-qPCR was performed. Data represents 3 independent experiments. Error bars represent SEM.

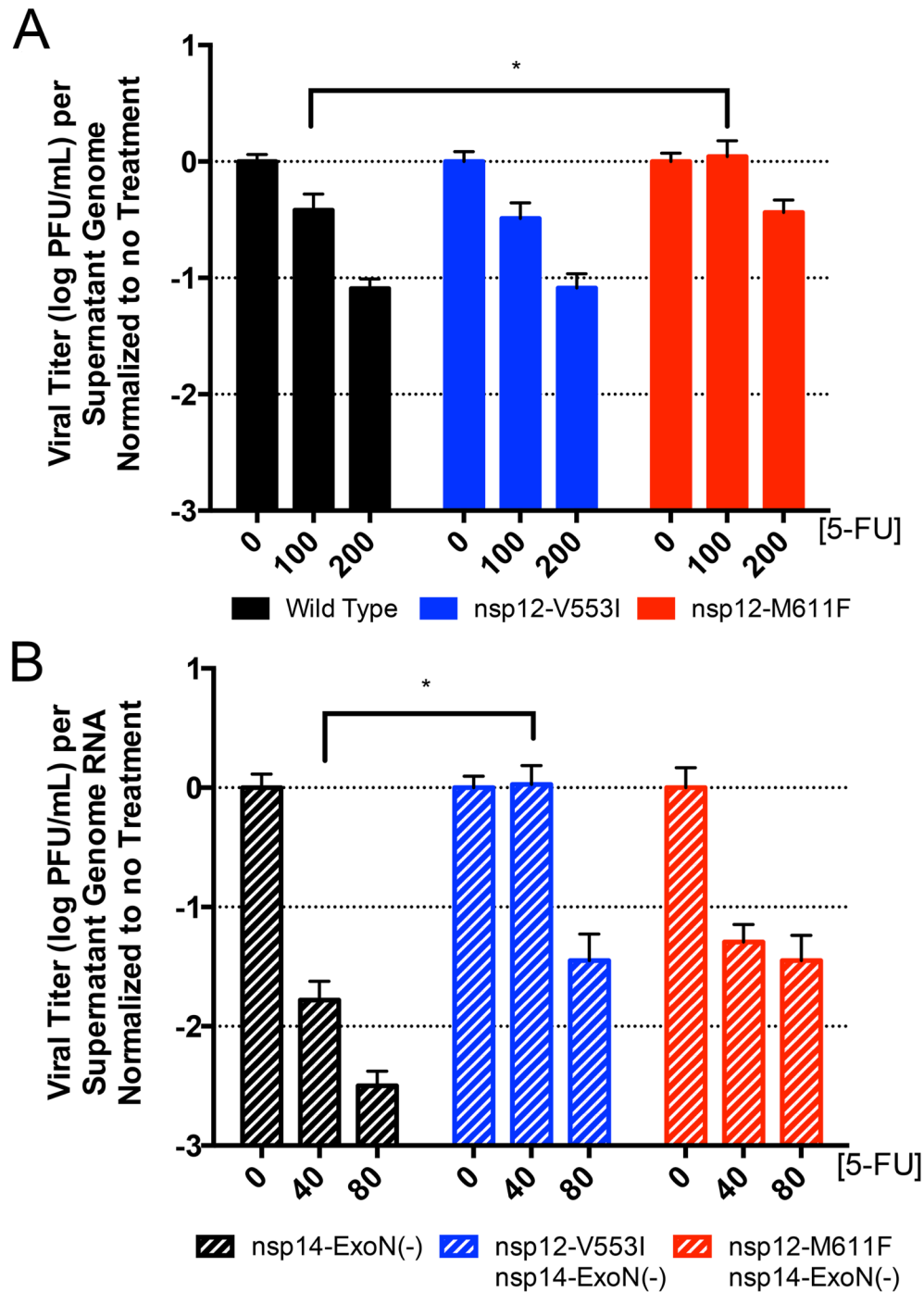
synthesis for nsp12-V553I and nsp12-M611F by RT-qPCR. Measured genomic RNA levels were consistent with the virus replication kinetics data, and I observed delayed and decreased genome RNA levels in the WT background after multiple rounds of replication (Figure 9e). However, at early time points there was no difference in RNA accumulation suggesting that decreased RNA is a result of steps post RNA synthesis (Figure 9c). In the nsp14-ExoN(-) background, RNA synthesis levels were indistinguishable from the nsp14-ExoN(-) background for both nsp12-V553I and nsp12-M611F, with no additional RNA synthesis defects detectable (Figure 9d and f). The results, along with my recovery of several of the nsp12-RdRp mutants only in the nsp14-ExoN(+) background, support the hypothesis that nsp14-ExoN and nsp12-RdRp may have an epistatic relationship. However, for replication kinetics the effects of nsp12-V553I and nsp12-M611F are not observable in the presence of an inactive nsp14-ExoN, demonstrating that the replication phenotype of an inactive nsp14-ExoN is epistatic to replication variants encoded in nsp12-RdRp.

#### **Specific infectivity of nsp12-V553I and nsp12-M611F**

Having identified two mutant viruses with resistance to 5-FU I wanted to further test whether resistance was due to decreased incorporation of the mutagen. Measurement of specific infectivity has been useful for determining lethal mutagenesis for MHV and other RNA viruses (Rozen-Gagnon et al., 2014; Smith et al., 2013). I tested the nsp12-M611F and nsp12-V553I mutants, in both the WT and nsp14-ExoN(-) backgrounds, for changes in specific infectivity when infected at an MOI of 0.01 and treated with 5-FU (Figure 10). Both nsp12-M611F and nsp12-V553I resulted in an increased ratio of infectious particles (PFU/ml) to total particles (RNA genomes) in the nsp14-ExoN(-) background (Figure 10b). Similarly, in the WT background nsp12-M611F demonstrated an increase in the ratio of infectious particles to RNA genomes (Figure 10a). Thus specific infectivity may be a more sensitive measure of lower level changes in nucleotide selectivity in the setting of nsp14-ExoN(-).

#### **Fitness cost of nsp12-V553I and nsp12-M611F in WT and nsp14-ExoN(-) backgrounds**

In multiple RNA viruses, including CoVs, both increased and decreased fidelity have been reported to have a fitness cost (Smith et al., 2014). I therefore sought to determine whether nsp12-V553I or nsp12-M611F conferred any cost in fitness, defined as the ability to directly compete during co-infection. In the WT background, both nsp12-V553I and nsp12-M611F

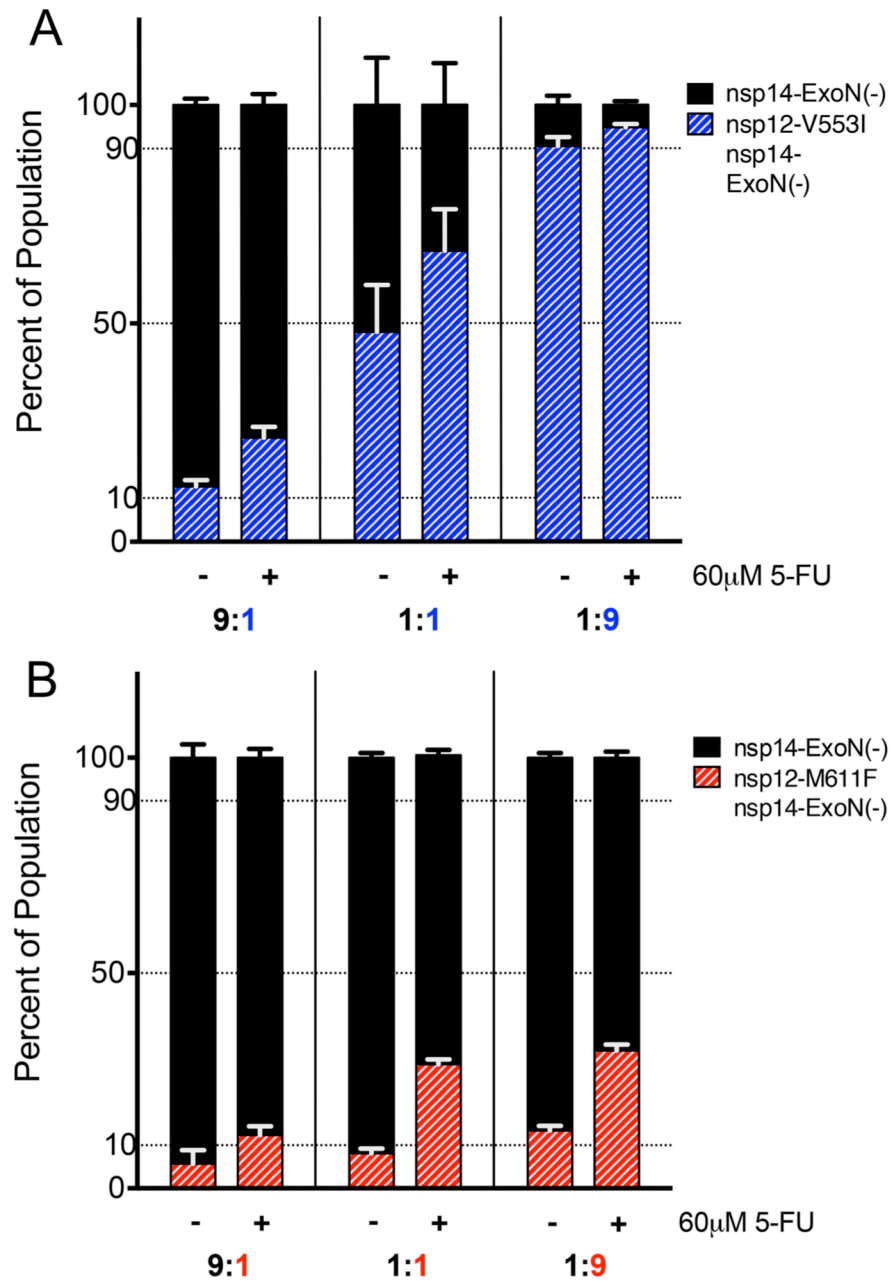


**Figure 10. Specific infectivity is increased in both nsp12-V553I and nsp12-M611F mutants**

DBT cells were pre-treated with increasing concentrations of 5-FU for 30 min. Treatment was removed and cells were infected with indicated viruses in the **A**. WT background or **B**. nsp14-ExoN(-) background at an MOI of 0.01. Media containing 5-FU was replaced 60 min p.i. Virus samples were taken at 20 and 24 h.p.i respectively. Titer was determined by plaque assay and number of supernatant genomes was determined using one-step RT-qPCR. Data represents 2 independent experiments, each with 3 replicates. Error bars represent SEM (\* $P < 0.05$  by 2way ANOVA using the Bonferroni correction for multiple comparisons).

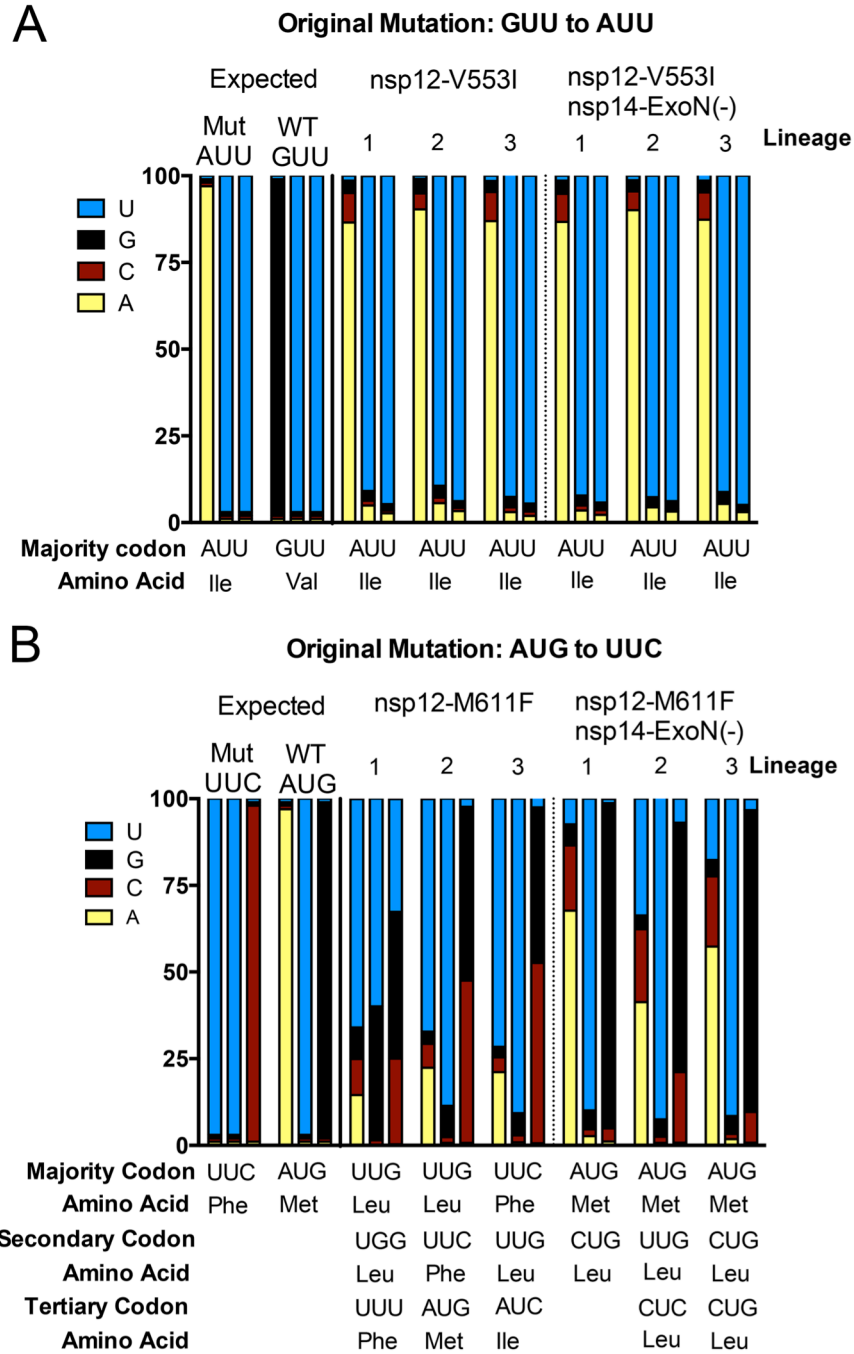
demonstrated delays in replication and impaired RNA accumulation. However, since there were not any observed additional defects in replication or RNA synthesis for nsp12-V553I and nsp12-M611F when introduced in the nsp14-ExoN(-) background (Figure 9b, d), I tested for any additional fitness cost of nsp12-V553I/nsp14-ExoN(-) or nsp12-M611F/nsp14-ExoN(-) viruses compared with nsp14-ExoN(-) alone. When co-infected with nsp14-ExoN(-) at ratios from 1:9 to 9:1, nsp12-V553I/nsp14-ExoN(-) maintained the input ratio compared with nsp14-ExoN(-) (Figure 11a). A small advantage to nsp12-V553I/nsp14-ExoN(-) was observed when the co-infected cultures were treated with 60uM 5-FU, consistent with a conferred advantage for 5-FU resistance (Figure 11a). Thus there appeared to be no additional fitness cost of nsp12-V553I/nsp14-ExoN(-) compared to nsp14-ExoN(-) alone. In contrast, nsp12-M611F/nsp14-ExoN(-) was not able to compete with nsp14-ExoN(-) at any ratio (Figure 11b). Treatment with 60uM 5-FU again favored nsp12-M611F/nsp14-ExoN(-); however, even then the percent of the population made up of nsp12-M611F/nsp14-ExoN(-) only remained at around 30% of the population when initially given a 9-fold advantage (Figure 11b).

I next tested whether the relative differences in fitness cost resulted in selective pressure for reversion of nsp12-V553I and nsp12-M611F (Figure 12). DBT cells were infected with nsp12-V553I, nsp12-V553I/nsp14-ExoN(-), nsp12-M611F, or nsp12-M611F/nsp14-ExoN(-) at an initial MOI of 0.01. After 5 passages the viruses were analyzed for retention of original mutations using di-deoxy (Sanger) sequencing. The nsp12-V553I mutation was stable after passage in both WT and nsp14-ExoN(-) backgrounds, maintaining the mutated AUU codon (Figure 12a), suggesting minimal selective pressure on that nucleotide, codon, or amino acid. In contrast, the nsp12-M611F mutation demonstrated significant change over passage in both the WT and the nsp14-ExoN(-) backgrounds. The original mutation, UUC, was no longer the majority codon in the nsp14-ExoN(-) background and was less than 52% of the population in all WT background lineages (Figure 12b). The nsp12-M611F/nsp14-ExoN(-) population resulted in a mix of two nucleotide changes resulting in reversion to AUG (methionine) ( $\leq 68\%$  of the population), single nucleotide changes that result in mutation it to Leu ( $\leq 55\%$  of the population), a somewhat smaller but still hydrophobic residue, or retaining a Phe substitution ( $\leq 27\%$  of the population). Thus the fitness cost of nsp12-M611F results in significant selective pressure for changes at that residue during passage in absence of 5-FU.



**Figure 11. Competitive fitness analysis in the nsp14-ExoN(-) background**

DBT cells were pre-treated with media alone or containing 60µM 5-FU for 30 min. Treatment was removed and cells were co-infected, at total MOI of 0.01 PFU/cell, with nsp14-ExoN(-) and **A.** nsp12-V553I/nsp14-ExoN(-) or **B.** nsp12-M611F/nsp14-ExoN(-) at a ratio of 9:1, 1:1 or 1:9. Media alone or containing 60µM or 300µM 5-FU was replaced 30 min p.i. Total RNA was taken at 24 h.p.i. Sequencing was performed across a 1.7kb region of nsp12-RdRp that included both mutations. Data represents 3 independent experiments, each with 2 replicates. Error bars represent SEM.



**Figure 12. The nsp12-V553I mutation is stable across passage; however, nsp12-M611F is vulnerable to reversion**

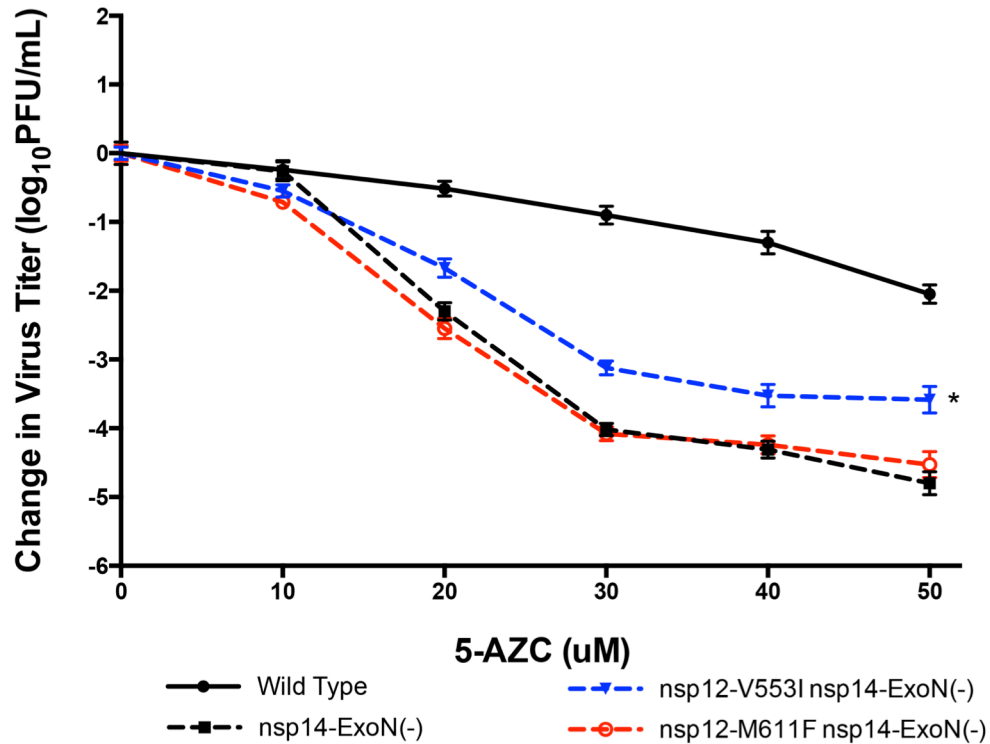
DBT cells were infected with an initial MOI of 0.01 then blind passaged in triplicate for 5 passages. Total RNA was taken and sequencing was performed across a 1.7kb region of nsp12-RdRp that included both mutations. Percentage of each nucleotide present in each of the triplicate lineages after 5 passages is shown. Mutant viruses in the WT (solid bars) and nsp14-ExoN(-) (slashed bars) backgrounds are both shown. The original mutation for each of the viruses is shown above the graph and the likely majority, secondary and tertiary codons present in the population are shown below the graph.

### **Resistance of recovered mutant viruses to the base analog 5-azacytidine**

Having shown that both nsp12-M611F and nsp12-V553I mutations conferred resistance to 5-FU in the nsp14-ExoN(-) background but that nsp12-M611F reverted quickly and was at a fitness disadvantage in both the WT and nsp14-ExoN(-) background, I next wanted to test whether the mutations conferred resistance specifically to 5-FU or broadly to base analogs and therefore were likely determinants of fidelity. A study by Arias et al demonstrated that for foot-and-mouth disease virus (FMDV) resistance to a specific mutagen can result from a point mutation while conferring the opposite overall fidelity (Arias *et al.*, 2008). I therefore tested for resistance to the additional mutagen 5-azacytidine (5-AZC). Similar to the 5-FU results, the nsp12-V553I/nsp14-ExoN(-) mutant virus was more resistant to 5-AZC than nsp14-ExoN(-) alone, maintaining approximately one log higher titers from 20-50uM 5-AZC (Figure 13). However, the nsp12-M611F/nsp14-ExoN(-) mutant virus showed no difference in resistance to 5-AZC compared with nsp14-ExoN(-) alone (Figure 13). These data suggest that nsp12-V553I is likely a fidelity determinant where as nsp12-M611F confers specific resistance to 5-FU and has an unknown overall fidelity. This result is also consistent with the rapid reversion observed for the nsp12-M611F mutation.

### **The nsp12-V533I mutation results in a decrease in the accumulation of mutations**

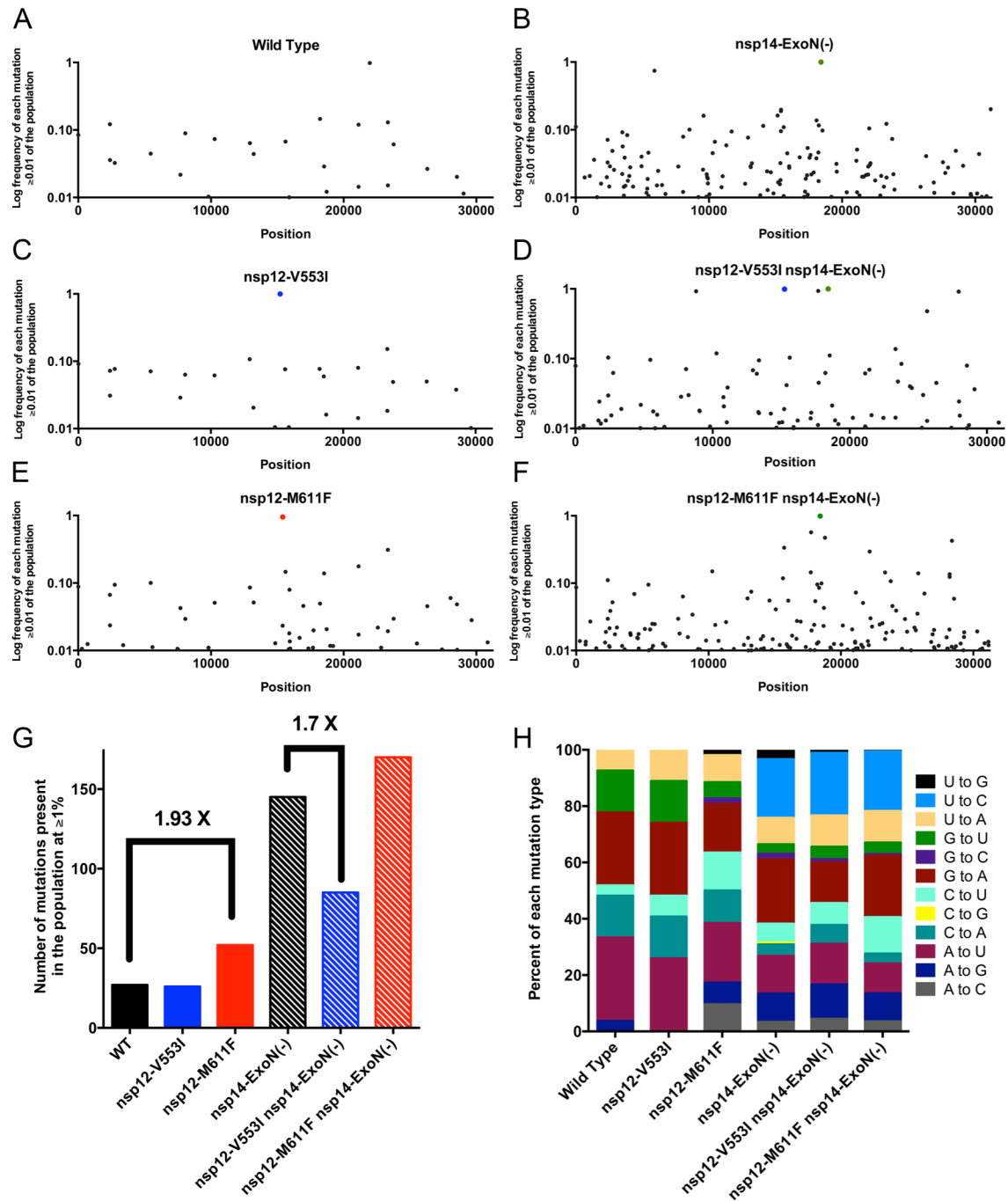
Having shown that the nsp12-V553I and nsp12-M611F mutations resulted in resistance to 5-FU, and that nsp12-V553I additionally conferred resistance to 5-AZC, I sought to directly determine whether either of these mutations resulted in a change in the number of mutations accumulated in viral RNA. DBT cells were infected with wild type, nsp12-V553I, nsp12-M611F, nsp14-ExoN(-), nsp12-V553I/nsp14-ExoN(-) or nsp12-M611F/nsp14-ExoN(-) at an MOI of 0.01 and RNA was collected at 20 h. pi. My collaborators in Marco Vignuzzi's laboratory then prepared these samples for Illumina next generation sequencing (NGS) across the full genome and analyzed them using the ViVan analysis pipeline (Isakov *et al.*, 2015). I graphed mutations present as 1% or more of the population by frequency and position in the genome. Engineered mutations are depicted with colored dots. Most non-engineered new mutations were present at 10% or less of the population and were distributed across the genome with no detectable "hot-spots" (Figure 14a-f). For nsp12-V553I in the WT background, no difference was observed in the number of mutations accumulated to 1% or greater of the population compared to WT alone. In contrast, in the nsp14-ExoN(-) background, nsp12-V553I was associated with a 1.7-fold decrease in the



**Figure 13. Resistance of MHV nsp12-RdRp V553I and M611F mutant viruses to 5-azacytidine in the nsp14-ExoN(-) background**

Domain location of mutations is indicated by: fingers (blues) and palm (reds). DBT cells were pre-treated with different concentrations of 5-AZC for 30 min. Treatment was removed and cells were infected with indicated viruses at an MOI of 0.01. Media containing 5-AZC was replaced 30 min p.i. Virus samples were taken at 32 h.p.i and titer was determined by plaque assay. Data represents 5 independent experiments, each with 2 replicates. Error bars represent SEM (\*P < 0.05 by ratio paired t test relative to nsp14-ExoN(-) virus).





**Figure 14. The nsp12-V553I mutation confers decreased accumulation of mutations in the nsp14-ExoN(-) background with no bias toward the exclusion of specific nucleotides**

DBTs were infected with an MOI of 0.01 and total RNA collected. Deep sequencing was performed on these samples. The statistically significant mutations present as greater than or equal to 1% of the total population are shown for wild type, nsp14-ExoN and nsp12-V553I or nsp12-M611F in both backgrounds. These were graphed (A-F) according to their distribution across the genome with intentionally introduced mutations (B-F) shown with circles colored blue (nsp12-V553I), red (nsp12-M611F) or green (nsp14-ExoN(-)), as (G) the total number of mutations present in the population and as (H) the percent of specific mutations present.

frequency of mutations compared to nsp14-ExoN(-) alone (Figure 14g), again with no change in the distribution of mutations across the genome (Figure 14c, d). These results are consistent with both the 5-FU and 5-AZC data in suggesting increased fidelity. Results from the nsp12-M611F mutant viruses were more complicated. To our surprise, the nsp12-M611F/nsp14-ExoN(-) virus fully reverted at both engineered nsp12-M611F nucleotides during the low MOI infection, resulting in a viral population that was nsp14-ExoN(-) alone (Figure 14f). This reversion made the results for the nsp12-M611F/nsp14-ExoN(-) virus interpretable. However, since accumulation of mutations to over 1% of the population is a combination of all replication cycles from initial recovery to the final sample (roughly 11 total for nsp12-M611F/nsp14-ExoN(-)) I included this data in Figure 14g and 14h. In the WT background the nsp12-M611F mutations were still present as 100% of the population and resulted in a 1.93 fold increase in the total number of accumulated mutations. I observed only a slight increase in the number of mutations accumulated in the nsp12-M611F/nsp14-ExoN(-) sample over those of nsp14-ExoN(-), which was not surprising seeing as the M611F mutation was no longer present. Neither sample appeared to have mutations concentrated in specific locations across the genome suggesting the accumulation of mutations was due to random generation of mutations rather than strong selection in particular locations or proteins (Figure 14e and f). Of note, one mutation in nsp3 of the nsp14-ExoN(-) sample and a mutation in nsp3, nsp13 and the E proteins of nsp12-V553I/nsp14-ExoN(-) reached nearly 100% of the sample population. None of these mutations were present in the fragments used for recovery and the nsp3 mutation, which arose in the nsp14-ExoN(-) population, was not present prior to the final low MOI infection. It is possible that these mutations provide some benefit to these viruses since they were fixed so rapidly in the population.

I next determined whether either of the mutations resulted in a change in the types of mutations occurring during replication (Figure 14h). Consistent with our previous studies, there were differences in the types of mutations incorporated when comparing WT and nsp14-ExoN(-) backgrounds (Smith *et al.*, 2013). However, the addition of nsp12-V553I or nsp12-M611F did not alter these patterns in either WT or nsp14-ExoN(-) backgrounds. Thus, nsp12-V553I results in an overall decrease in the accumulation of mutations over passage while nsp12-M611F seems to increase the number of mutations accumulated. These results confirm by sequence analysis the results from the 5-FU and 5-AZC resistance experiments for nsp12-V553I, specifically that the

effects of nsp12-V553I are dependent on inactivation of nsp14-ExoN for their detection and that nsp12-V553I likely confers broad resistance to the incorporation of incorrect nucleotides. In contrast, these results further increase the complexity of the nsp12-M611F mutation in relation to incorporation of nucleotides and their analogs. I conclude that nsp12-M611F confers resistance to 5-FU but that this resistance is not likely to be due to a broad resistance to the incorporation of alternate nucleotides.

### Discussion

RdRp structures of divergent RNA viruses are structurally conserved and likely have common determinants of activity in the finger, palm and thumb domains. Positive strand RNA virus polymerases appear to utilize a common palm domain-based mechanism for active site closure (Gong and Peersen, 2010), and associated molecular determinants of fidelity in different RdRp domains have been proposed based on biochemical and mutagenesis studies (Campagnola *et al.*, 2015). To date there are no solved crystal structures of any CoV RdRp and thus direct comparison with other virus RdRp structures has not been possible. Further, regulation of CoV fidelity is likely dependent on multiple proteins, including the RdRp and proofreading nsp14-ExoN. Thus it was not clear that a CoV would phenotypically exhibit effects from mutating fidelity-determining residues located in the RdRp itself. In this chapter I sought to determine whether I could use structure and mutagenesis data from distantly related RNA viruses to identify determinants of CoV nsp12-RdRp fidelity. Our results suggest that CoV RdRps do in fact participate in fidelity regulation, at residues orthologous to those in the picornaviruses, and likely in other RNA virus RdRps. The results also define, for the first time, a CoV RdRp determinant that increases resistance to multiple mutagens, decreases the accumulation of mutations over time and so, likely increases overall fidelity. Additionally, the data in this paper suggest that CoV RdRp-mediated increased fidelity is only detectable when nsp14-ExoN is inactive, and is only partially compensating for the loss of nsp14-ExoN high fidelity. Finally, both the nsp12-V553I and M611F mutations confer a replication cost in the WT background but only nsp12-M611F confers a fitness disadvantage in the nsp14-ExoN(-) background. Together the results suggest that nsp14-ExoN proofreading activity is epistatic to nsp12-RdRp fidelity but that in contrast the replication defects of an inactive nsp14-ExoN are epistatic to replication defects in nsp12-RdRp.

## **Determinants of nucleotide selectivity and fidelity in CoVs may be conserved with other RNA viruses**

Picornavirus functional RdRps contain only RdRp domains (Campagnola et al., 2008; Hansen et al., 1997). This is not the case for many viral RdRps, CoVs included (Lehmann et al., 2015; Velthuis, 2014). In addition to the predicted RdRp core domain, all CoV nsp12 proteins contain a “CoV-specific” domain of over 350 amino acids at the N terminus of the protein. A nucleotidyltransferase activity was identified in this CoV-specific domain and has been shown to be important in SARS-CoV replication, but the specific function in replication remains to be determined (Lehmann *et al.*, 2015). The N-terminal CoV-specific domain could not be modeled due to lack of evolutionary homologues of known structure but, the predicted RdRp core domain could be modeled with high confidence using bioinformatics approaches. The structures that provided the best models for the CoV RdRp were from picornaviruses, including enterovirus 71, foot and mouth disease virus and coxsackievirus B3. This allowed for direct alignment and comparison of the known fidelity determinants in CVB3 with the MHV nsp12 core domain, across both fingers and palm domains. The recovery of mutations in these structurally conserved residues and their participation in CoV nucleotide selectivity supports the hypothesis that there are determinants of base specificity conserved between CoVs and distantly related RNA viruses, specifically the picornaviruses. This supports the idea that all RdRps function similarly due to their structural conservation and despite the low level of sequence similarity and attached domains (Yang *et al.*, 2012). It also suggests that the CoV RdRp domain folds in a manner similar to other RdRps, likely separate from the CoV-specific domain.

RdRp fidelity and nucleotide selectivity has been investigated extensively in picornaviruses, especially poliovirus (Hobdey et al., 2010; Korneeva and Cameron, 2007; Pfeiffer and Kirkegaard, 2003; Verdaguer and Ferrer-Orta, 2012; Vignuzzi et al., 2005a). However, even between picornaviruses, the impact on fidelity of changes at identical or similar residues can vary dramatically (Campagnola *et al.*, 2015). Mutations at the same residues in the RdRp of poliovirus and CVB3 affect fidelity differently; poliovirus mutations generally result in increased fidelity and CVB3 mutations in decreased fidelity (11). Specifically, the structurally orthologous residue to nsp12-M611 in poliovirus is F230 and in CVB3 it is I230; when F230 is mutated to an Ile in poliovirus it results in an increase in fidelity, but when I230 in CVB3 is mutated to a Phe, Trp or Val it results in a decrease in fidelity (Campagnola *et al.*, 2015). Similarly, the

orthologous residue to nsp12-V553 in CVB3 is I176; when I176 is mutated to a Val in CVB3 it results in a decrease in fidelity (Campagnola et al., 2015; Gnädig et al., 2012). Although the specific substitutions differ, when nsp12-V553I and nsp12-M611F are introduced into the MHV genome the resulting 5-FU resistance, and for nsp12-V553I a likely increase in fidelity, mimics results seen in poliovirus rather than CVB3. However, nsp12-M611F seems to accumulate more mutations over time than controls, and many of the predicted mutations were only recoverable in the WT background. Therefore, changes in nucleotide selectivity and fidelity were not tested, so it remains possible that the mutations that were non-viable in the nsp14-ExoN(-) background had decreased fidelity and that nsp14-ExoN(-) treated with 5-FU defines the error threshold for CoVs.

### **Coronavirus nsp12-RdRp and nsp14-ExoN cooperate to optimize both fidelity and replication kinetics**

In addition to nsp12-RdRp, CoVs encode nsp14-ExoN, which functions as a proofreading enzyme (Bouvet et al., 2012; Eckerle et al., 2007; 2010; Smith et al., 2013). Beyond nsp14-ExoN, there are additional CoV-encoded nsps that potentially contribute to overall fidelity, such as: the small molecule modulator of nsp14-ExoN encoded in nsp10 (Bouvet et al., 2012; Chen et al., 2011), nsp7 and nsp8 that together function as an elongation factor (Velthuis et al., 2012) and a primase (Imbert et al., 2006; Xiao et al., 2012), and nsp13 that functions as a helicase (Adedeji *et al.*, 2012). These proteins interact with each other and likely function as a multi-protein replication/fidelity complex (Smith and Denison, 2013; Subissi *et al.*, 2014). The results of this study show that while determinants in nsp12-RdRp are likely capable of increasing fidelity, those changes are only detectable in the setting of loss of nsp14-ExoN proofreading, and do not completely compensate for the impaired fidelity associated with nsp14-ExoN(-). If altering nucleotide selectivity determinants within nsp12-RdRp no longer significantly affect the overall nucleotide selectivity of the WT virus, then the evolutionary pressure on nsp12-RdRp may be more heavily weighted toward other aspects of replication, such as speed. Recent evidence suggests that RdRp fidelity and speed have an inverse relationship (Campagnola et al., 2015; Regoes et al., 2012). So, if fidelity regulation by nsp14-ExoN occurs more rapidly than correct nucleotide selectivity by the RdRp then CoV nsp12-RdRp may have been selected specifically for replication speed. Our data may support this hypothesis, as in the WT background both nsp12-V553I and nsp12-M611F resulted in increased replication lag phases (Figure 9a and e),

though early RNA synthesis was not observably different in our system (Figure 9c). However, this was not seen in the nsp14-ExoN(-) background, which could be explained by our data indicating that replication is already slowed when nsp14-ExoN is inactivated. One possible mechanism for this could be that the nsp14-ExoN(-) protein is trying to remove incorrect nucleotides and stalling replication due to its inability to do so. In this case nsp14-ExoN(-) would be epistatic to nsp12-RdRp in relation to speed, thereby obscuring decreases in replication speed caused by mutations in nsp12-RdRp itself.

## Conclusion

My results support the hypothesis that determinants of nucleotide selectivity are conserved across viral orders, identify the first likely increased fidelity determinant for CoV nsp12-RdRp, and demonstrate that nsp14-ExoN proofreading activity is epistatic to nsp12-RdRp nucleotide selectivity. Knowing that some fidelity determinants may be conserved across viral orders is an exciting discovery, as many fidelity determinants identified so far have resulted in attenuation (Coffey *et al.*, 2011; Gnädig *et al.*, 2012; Lauring *et al.*, 2010; Pfeiffer and Kirkegaard, 2005; Vignuzzi *et al.*, 2008; Yang *et al.*, 2012). It would be interesting to determine whether nsp12-V553I is also attenuated. I would predict that it would be, based on the fitness cost *in vitro*. However, there is also no clear increase in nucleotide selectivity or fidelity in the WT background. This may be due to a very minimal change that is just not measurable even with the deep sequencing technology used in this study. Alternatively, it could be that any change is simply overwhelmed by the high fidelity of intact nsp14-ExoN proofreading. In any case, these mutations or other changes at these residues may allow for selection of viruses that replicate with normal kinetics *in vitro* and *in vivo*, yet confer attenuation in an animal setting. We know that the nsp14-ExoN(-) mutations confer genotypically and phenotypically stable attenuation *in vivo* (Graham *et al.*, 2012). However, the concept of a high level mutator as a mechanism for attenuation in live viruses may be problematic. The identification of increased fidelity mutations in the RdRp that can partially, or potentially completely, compensate for the fidelity impairment of nsp14-ExoN(-) viruses, may allow for development of approaches that can benefit from the stability of the nsp14-ExoN(-) mutator phenotype while allowing more stability to the input genomes. Finally, these results combined with those from previous work (Bouvet *et al.*, 2012; Smith *et al.*, 2015) suggest that CoVs encode at least three proteins involved in fidelity (nsp12-RdRp, nsp14-ExoN and nsp10) supporting the assembly of a multi-protein replicase / fidelity

complex as described previously (Subissi *et al.*, 2014). This increases the importance of establishing a biochemical model of the multi-protein complex to directly test the interactions of fidelity determinants as well as potential inhibitors of each or all of these functions.

## CHAPTER III

### IDENTIFICATION OF COMPENSATORY MUTATIONS IN CORONAVIRUS NSP12 AND NSP14 AFTER PASSAGE IN THE PRESENCE OF MUTAGEN

#### Introduction

In the previous chapter, I identified residues involved in nucleotide selectivity by structural analysis and homology modeling of the murine hepatitis virus (MHV) RNA-dependent RNA polymerase (RdRp) encoded in nonstructural protein 12 (nsp12). However, the classically used method for identifying fidelity determinants is selection for resistance to mutagens. Additionally, at the start of this dissertation research, there were solved structures for neither any CoV nsp12 nor nsp14, which encodes for the proofreading 3'-5' exoribonuclease (ExoN), thus limiting the domains I could investigate by homology to the highly conserved nsp12-RdRp domain. In contrast, allowing the virus to select for mutations in the presence of mutagen provides an opportunity to investigate all domains including those with no known structure or function. Many mutagens are available to select for resistance mutations. Ribavirin is one of the most commonly used mutagens; however, it is known to have additional antiviral effects (Graci and Cameron, 2006; Parker, 2005). Another common mutagen is 5-FU, used in RNA viral studies and currently not known to have additional roles affecting RNA virus replication (Longley et al., 2003). Additional nucleoside analogs are available that provide strong selection against CoVs, such as the non-obligate chain terminator 2'-C-methyladenosine (2'-C-MeA). However, 2'-C-MeA is a sugar analog instead of a base analog and so is less likely to identify mutations that would confer altered base-pairing fidelity in a natural system (Carroll *et al.*, 2003).

In this chapter, I investigate the result of mutations in MHV-A59 nsp12 and nsp14 on nucleotide selectivity and replication. I demonstrate that minority variants present in 5-FU-treated virus populations are not likely candidates for increased fidelity variants, but variants that fix in 5-FU-treated virus populations typically result in resistance to the nucleotide analog. For wild type (WT) viruses, resistance may be specific to 5-FU, whereas in nsp14-ExoN(-) viruses selection may be skewed towards a preference for increased overall fidelity and/or replication. Xiaotao Lu assisted in the preparation and sending of sequencing samples across nsp12 and nsp14 and the final recovery steps of some of the viruses used. Additionally, Dr. Everett Clinton Smith



performed the 5-FU replication kinetics experiments. I performed all other experiments reported in this chapter.

### **Identification of minority variants in SARS-CoV populations after treatment with 5-FU**

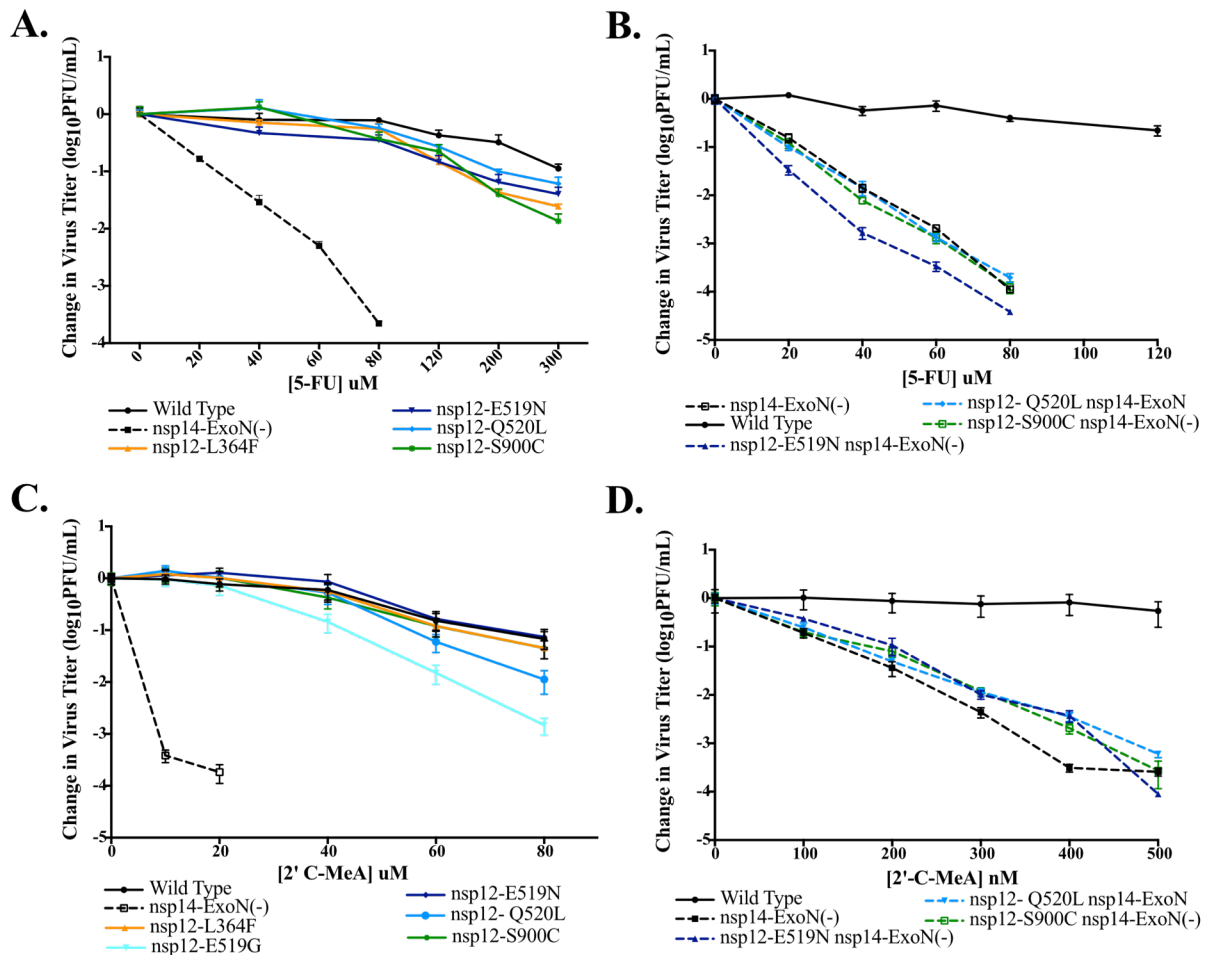
In previously published work from our lab, SARS-CoV was treated with 400uM 5-FU over one 0.01 MOI infection and total RNA was sent for deep sequencing (Smith *et al.*, 2013). Analysis of the minority variants present in these populations identified a panel of mutations in SARS-CoV nsp12. From this data, I further identified mutations that were non-synonymous, present at 1.5% or more of the population, and where the original amino acid was conserved between SARS-CoV and MHV-A59. This analysis resulted in a final panel of six nsp12 mutations: nsp12-L364F, P374S, E519G, E519N, Q520L and S900C. These mutations made up 1.5, 4.8, 5.7, 25, 6.8 and 2 percent of the population, respectively. The identified residues span all domains of nsp12 except the palm domain (Table 2). I next attempted to recover viruses with the identified mutations in isogenic backgrounds. Virus recovery was attempted a maximum of three times, resulting in recovery of five out of the six attempted mutant viruses in the WT background; only P374S was not recovered. All five engineered viruses had no additional mutations present across nsp12. Since results from the previous chapter demonstrated that proofreading by nsp14-ExoN is epistatic to nucleotide selectivity by the RdRp, I attempted recovery of the WT-background viable mutants in the setting of and inactivated nsp14-ExoN (nsp14-ExoN(-)). Only three of the five mutants recoverable in the WT background were also recoverable in the nsp14-ExoN(-) background: nsp12-E519N/nsp14-ExoN(-), nsp12-Q520L/nsp14-ExoN(-), and nsp12-S900C/nsp14-ExoN(-) (Table 2).

**Table 2. Recovery of minority-variant mutant viruses using site-directed mutagenesis.**

nsp12-RdRp Region	MHV engineered substitutions (nsp12)	Percent of treated population	Recovery: WT (nsp14-ExoN(+))	Recovery: nsp14-ExoN(-)
CoV Specific Domain	L364F	1.5	Yes	No
	P374S	4.8	No	Not attempted
Fingers	E519G	5.7	Yes	No
	E519N	25	Yes	Yes
	Q520L	6.8	Yes	Yes
Thumb	S900C	2	Yes	Yes

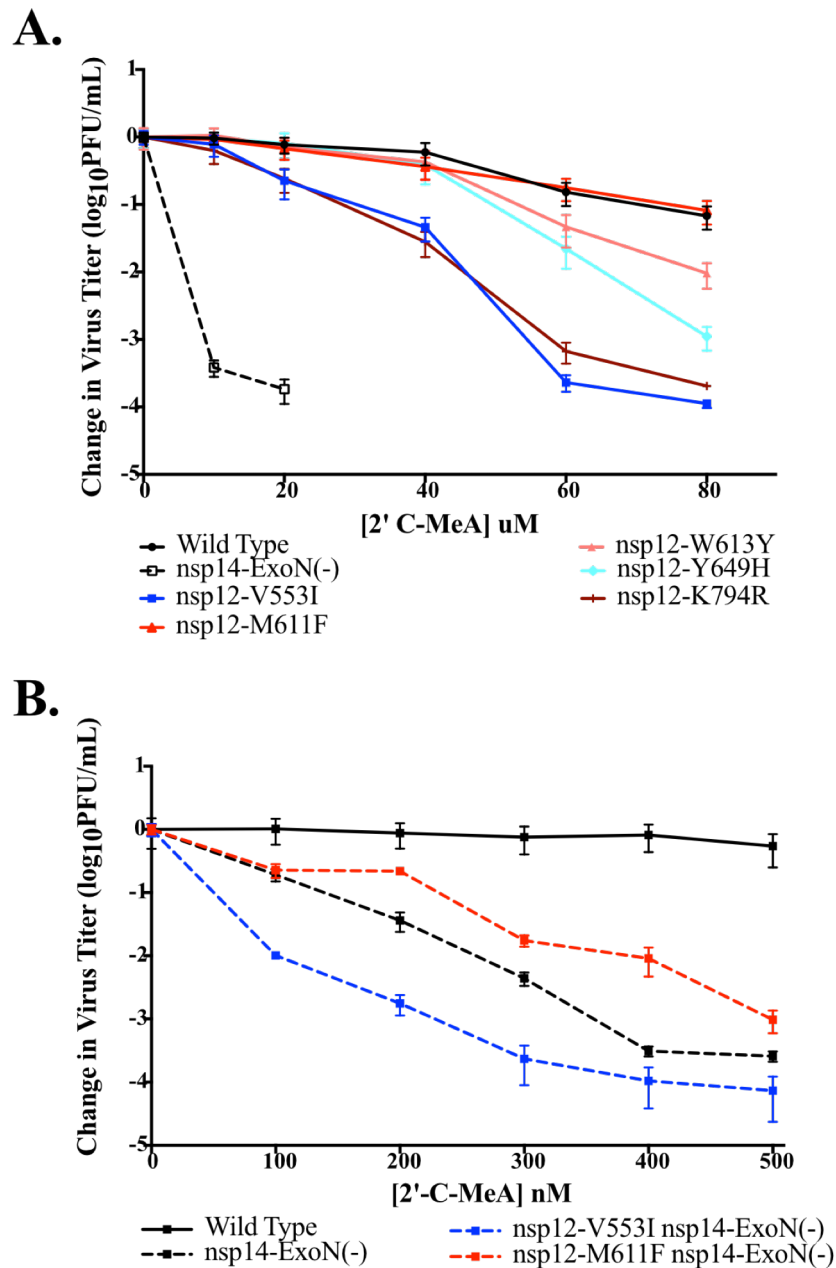
## **Engineered viruses result in no change in selectivity for the nucleotide analogs 5-fluorouracil or 2'-C-methyladenosine**

I next wanted to test whether engineered viruses resulted in altered nucleotide selectivity for the RNA mutagen 5-FU. 5-FU is routinely utilized as a screen to identify altered fidelity. WT CoVs (MHV, SARS-CoV) are resistant to 5-FU, while nsp14-ExoN(-) mutants are profoundly sensitive to 5-FU inhibition, consistent with nsp14-ExoN-mediated removal of misincorporated 5-FU (Sexton et al., 2016; Smith et al., 2013). Likely owing to efficient removal of 5-FU by nsp14-ExoN, nsp12-RdRp mutations often have no effect on overall viral sensitivity or resistance to 5-FU when present in the WT background (Figure 8a). Therefore, I compared WT viruses with nsp12 mutant viruses in both the WT-nsp14 and nsp14-ExoN(-) backgrounds. Similar to what was observed for mutations identified by homology modeling, there was no significant change in selectivity for 5-FU for any of the nsp12-RdRp mutant viruses in the WT background at up to 300uM 5-FU when compared to WT viruses (Figure 15a). However, the same was observed for nsp12 variant viruses in the nsp14-ExoN(-) background, for which all were non-recoverable at 120uM 5-FU (Figure 15b). These data demonstrate that 5-FU-treated minority variants do not confer resistance across CoVs to the mutagen 5-FU and suggest that they are not fidelity determinants. However, it is possible that these variants confer resistance to 5-FU in SARS-CoV. As mutations in nsp12-RdRp can result in differential selectivity between nucleotide analogs (Sexton *et al.*, 2016), I next tested engineered viruses for altered selectivity for the non-obligate chain terminator 2'-C-methyladenosine (2'-C-MeA). Again, no change in selectivity for 2'-C-MeA was observed for any of the viruses tested (Figure 15c, d). One mutation, nsp12-E519G, resulted in a virus that was more sensitive to 2'-C-MeA in the WT-nsp14 background, however this mutation was not recoverable in the nsp14-ExoN(-) background and therefore was not pursued further (Figure 15c). Of note, almost all viruses engineered after identification by homology modeling displayed altered selectivity for 2'-C-MeA, even in the WT-nsp14 background (Figure 16). Together, these data suggest that the minority variants present in mutagen-treated populations are not good candidates for residues involved in the regulation of fidelity but are instead beneficial in another way or are an accident of a bottlenecking event.



**Figure 15. Selectivity for 5-FU or 2'-C-MeA of MHV engineered 5-fluorouracil-treated CoV minority variant viruses**

Domain location of mutations is indicated by: CoV-specific domain (yellow), fingers (blues) and thumb (green). DBT cells were pre-treated with different concentrations of 5-FU (**A and B**) or 2'-C-MeA (**C and D**) for 30 min. Treatment was removed and cells were infected with indicated viruses in the WT background (**A and C**) or nsp14-ExoN(-) background (**B and D**) at an MOI of 0.01. Media containing 5-FU or 2'-C-MeA was replaced 30 min p.i. Virus samples were taken at 24 (WT) or 32 (nsp14-ExoN(-)) h.p.i and titer was determined by plaque assay. Data represents 2 independent experiments, each with 2 replicates, except panel A, which represents 1 independent experiment with 2 replicates. Error bars represent SEM.



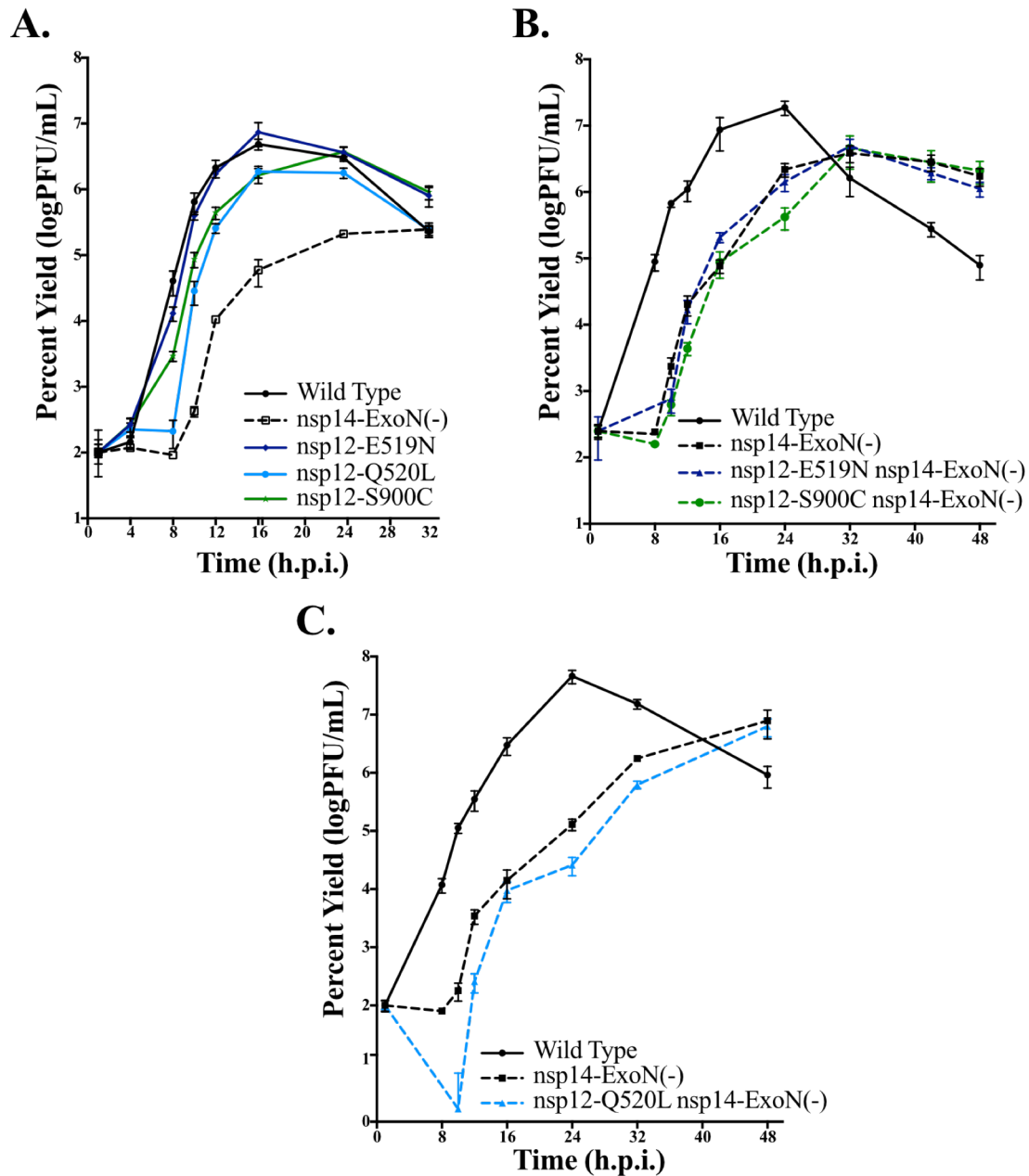
**Figure 16. Selectivity for 2'-C-MeA of MHV nsp12-RdRp mutant viruses identified by homology modeling**

Domain location of mutations is indicated by: fingers (blues) and palm (reds). DBT cells were pre-treated with indicated concentrations of 2'-C-MeA for 30 min. Treatment was removed and cells were infected with indicated viruses in **A.** the WT background or **B.** nsp14-ExoN(-) background at an MOI of 0.01. Media containing 2'-C-MeA was replaced 30 min p.i. Virus samples were taken at 24 (WT) or 32 (nsp14-ExoN(-)) h.p.i and titer was determined by plaque assay. Data represents 2 independent experiments, each with 2 replicates. Error bars represent SEM.

## **Replication kinetics of nsp12-E519N, nsp12-Q520L and nsp12-S900C mutant viruses in the WT and nsp14-ExoN(-) backgrounds**

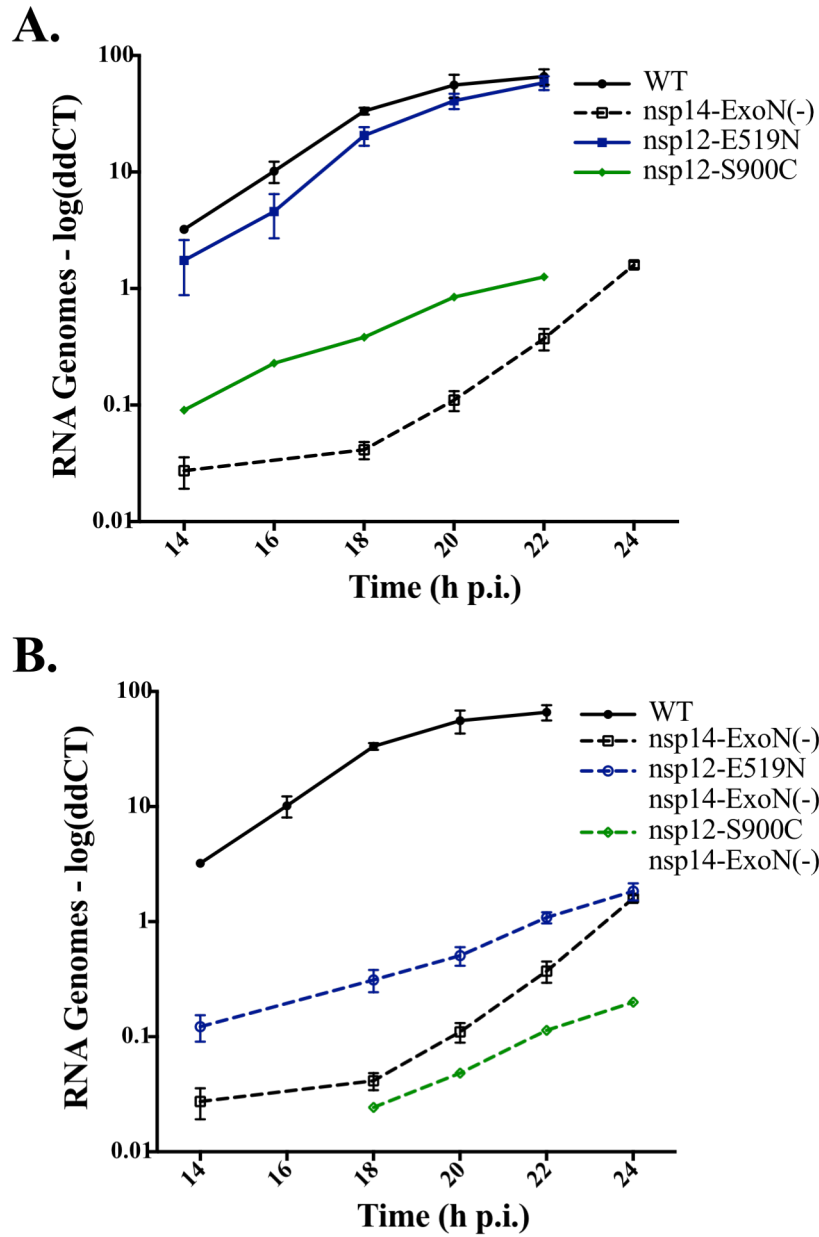
Although mutations identified in 5-FU-treated nsp14-ExoN(-) CoV populations didn't confer observable changes in nucleotide selectivity, alternative replicative advantages could have allowed for selection of these minority variants. To address this, I next sought to determine if nsp12-E519N, nsp12-Q520L or nsp12-S900C viruses resulted in altered replication kinetics. In the WT-nsp14 background, nsp12-E519N replicated with kinetics mimicking that of the WT virus alone, whereas both nsp12-Q520L and nsp12-S900C resulted in a delay in replication and never reached peak titers. Strikingly, nsp12-Q520L resulted in a lag phase of at least eight hours, similar to that seen for the nsp14-ExoN(-) virus. However, at eight hours, titers increased to 2-logs more than nsp14-ExoN(-) viruses (Figure 17a). Similar to what was observed with mutant viruses identified by homology modeling (Chapter II), replication of both nsp12-E519N and nsp12-Q520L viruses in the nsp14-ExoN(-) background displayed similar replication kinetics to nsp14-ExoN(-) (Figure 17b). However, nsp12-Q520L/nsp14-ExoN(-) resulted in a longer lag phase from which it did not recover, but instead maintained the delay throughout the exponential replication phase (Figure 17c). Overall, these data do not support the hypothesis that these mutations are present in the mutagenized population to help increase replication kinetics of the virus. In the case of nsp12-Q520L, these data also demonstrate that this variant delays replication kinetics beyond that of nsp14-ExoN(-).

Next, I assessed RNA synthesis for nsp12-E519N and nsp12-S900C by RT-qPCR. Similar to previous data, virus replication kinetics and genomic RNA levels of nsp12 mutant viruses were consistent within the WT-nsp14 background. I observed no difference in genomic RNA accumulation for nsp12-E519N but delayed and decreased genome RNA accumulation for nsp12-S900C in the WT-nsp14 background after multiple rounds of replication (Figure 18a). In contrast, in the nsp14-ExoN(-) background, RNA synthesis levels for nsp12-E519N/nsp14-ExoN(-) surpassed those of nsp14-ExoN(-) and the opposite was true for nsp12-S900C/nsp14-ExoN(-) (Figure 18b). These data suggest that the increase in the percentage of nsp12-E519N in the 5-FU-treated population may be partially attributable to an increase in the accumulation of RNA by this virus.



**Figure 17. Replication kinetics of 5-fluorouracil-treated, minority-variant engineered MHV nsp12-RdRp mutant viruses**

Mutation location is indicated by: fingers (blues) and thumb (green). DBT cells were infected with the viruses indicated in the **A**. WT background at an MOI of 0.1 PFU/cell or **B and C**. Nsp14-ExoN(-) background at an MOI of 0.01 PFU/cell. Supernatant aliquots were taken at indicated times p.i. and titer determined by plaque assay. Data represents 3 independent experiments. Error bars represent SEM.



**Figure 18. RNA accumulation by MHV nsp12-RdRp mutant viruses**

Mutation location is indicated by: fingers (blues) and thumb (green). DBT cells were infected with the viruses indicated in the **A.** WT background or **B.** nsp14-ExoN(-) background at an MOI of 0.01 PFU/cell. Total RNA was taken at indicated times p.i. and RT-qPCR was performed. Data represents 2 independent experiments. Error bars represent SEM.

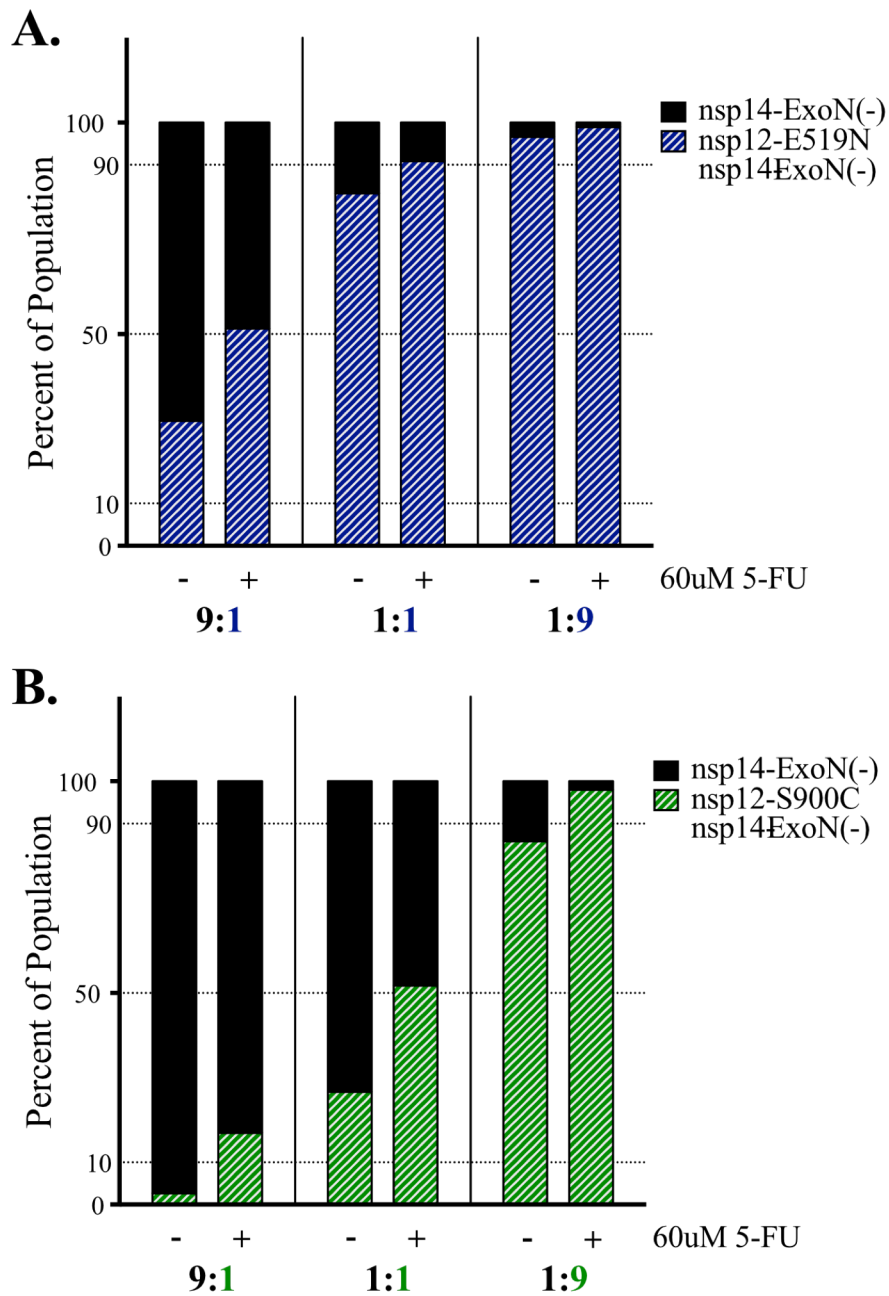
### **The effect on fitness of nsp12-E519N and nsp12-S900C in the nsp14-ExoN(-) background**

Although there was no change in replication kinetics for nsp12-E519N/nsp14-ExoN(-) or nsp12-S900C/nsp14-ExoN(-), my previous results demonstrated that subtle changes in fitness were not detected in the replication data. Therefore, I sought to determine if nsp12-E519N/nsp14-ExoN(-) or nsp12-S900C/nsp14-ExoN(-) would result in any change in fitness compared to nsp14-ExoN(-) alone. Here, fitness is defined as the ability of one virus to outgrow another in co-infections. Supporting the finding that the presence of nsp12-E519N/nsp14-ExoN(-) is 25% of the originally sequenced population, when co-infected with nsp14-ExoN(-) at ratios from 1:9 to 9:1, nsp12-E519N/nsp14-ExoN(-) was present at a larger than input percentage of the population at every ratio. Additionally, when infected in the presence of 60uM 5-FU, the ratio further increased in the favor of nsp12-E519N/nsp14-ExoN(-). Specifically, when infected at identical initial MOIs nsp12-E519N/nsp14-ExoN(-) was present at approximately 80% of the resulting population untreated and 90% treated (Figure 19a). A similar result was observed for nsp12-S900C/nsp14-ExoN(-) in the presence of 60uM 5-FU, trending toward an increase in the ratio of nsp12-S900C/nsp14-ExoN(-) rather than nsp14-ExoN(-) alone. However, in the case of nsp12-S900C/nsp14-ExoN(-), the virus was less fit than nsp14-ExoN(-) when 5-FU was not present, and fitness only matched that of nsp14-ExoN(-) when 5-FU was present (Figure 19b). These data suggest that both nsp12-E519N and nsp12-S900C result in a minor fitness advantage for nsp14-ExoN(-) viruses when replicating in the presence of the mutagen 5-FU and that their greater abundance as minority variants of the population may be a result of this selective advantage. Additionally, the fitness advantage observed for each of the variants correlates with the percent of the deep sequenced population they comprised (Figure 19 and table 2).

### **Identification of consensus sequence mutations in MHV nsp12 and nsp14 after passage in the presence of the mutagen 5-FU**

Although co-infection results suggested that in the nsp14-ExoN(-) background two of the variants identified may confer subtle effects on viral fitness in the presence of 5-FU, investigating these variants further will require a robust biochemical system. In the absence of a robust biochemical system, I next sought to identify residues resulting from selection for resistance to 5-FU by passaging. WT and nsp14-ExoN(-) viruses were blind passaged in media alone or in sub-lethal doses of 5-FU for seven total passages (Figure 20a). Nsp14-ExoN(-) viruses were not able to replicate beyond the seventh passage even when cells were infected in





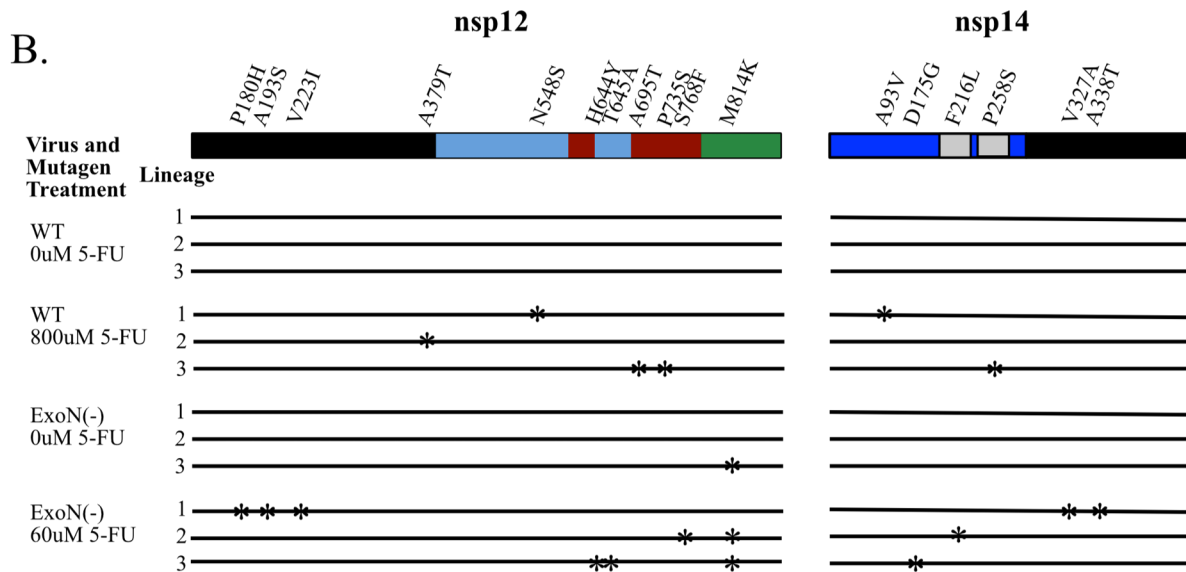
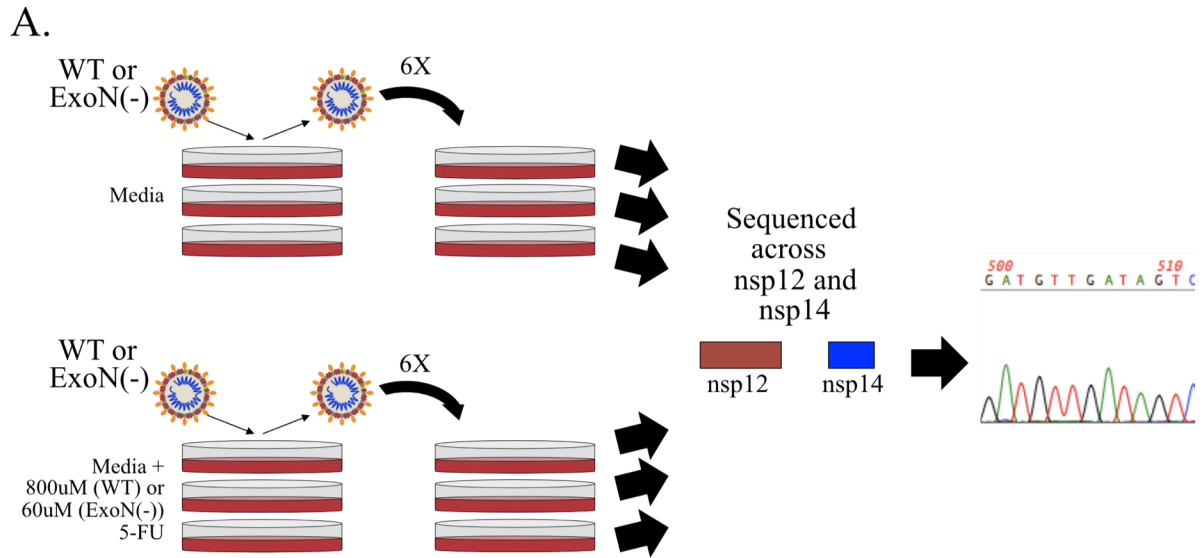
**Figure 19. Competitive fitness analysis of MHV nsp12-RdRp mutant viruses in the nsp14-ExoN(-) background**

DBT cells were pre-treated with media alone or containing 60uM 5-FU for 30 min. Treatment was removed and cells were co-infected, at total MOI of 0.01, with nsp14-ExoN(-) and **A.** nsp12-V553I/nsp14-ExoN(-) or **B.** nsp12-M611F/nsp14-ExoN(-) at a ratio of 9:1, 1:1 or 1:9. Media alone or containing 60uM or 300uM 5-FU was replaced 30 min p.i. Total RNA was taken at 24 h.p.i. Sequencing was performed across a 1.7kb region of nsp12-RdRp that included both mutations. Data represents 2 independent experiments.

the absence of 5-FU, suggestive of lethal mutagenesis. The inability to continue passage suggests that any mutations that arose were not enough to compensate for the population damage, which occurred during extended treatment with 5-FU. In contrast, cells infected with WT virus in the presence of 5-FU became fully involved in syncytia progressively more quickly with less transferred volume as passaging continued. This suggested that mutations arose within the mutagenized WT population that confer resistance to 5-FU. Both WT and nsp14-ExoN(-) viruses passaged in the absence of 5-FU maintained relatively consistent passaging conditions, and nsp14-ExoN(-) remained viable at the end of the passage series. Total RNA was taken from the final passage of all lineages and subjected to Sanger sequencing across all of nsp12 and nsp14.

Sequencing results from passaged populations across nsp12 and nsp14 were analyzed for non-synonymous mutations present at 50% or greater of the total population. No mutations were observed in nsp12 or nsp14 for any of the three lineages of WT virus passaged in the absence of 5-FU. Similarly, nsp14-ExoN(-) populations resulted in no observed non-synonymous mutations in nsp14 and only one in nsp12 (M814K). Nsp12-M814K was found in one of the three untreated populations and is a residue of the nsp12-RdRp thumb domain. At least one mutation was observed in nsp12 for every lineage of WT virus passaged in the presence of 5-FU, and lineages one and three also selected for individual mutations within nsp14-ExoN. Finally, every lineage of nsp14-ExoN(-) virus passaged in the presence of 5-FU selected for non-synonymous mutations in both nsp12 and nsp14. All mutations identified were unique to their lineage, except nsp12-M814K, which was observed in one untreated nsp14-ExoN lineage and two treated nsp14-ExoN(-) lineages. In total, 11 unique mutations were identified across nsp12. Of these, four mutations were in the nsp12-CoV-specific domain: nsp12-P180H, A193S, V223I and A379T; and seven were in the nsp12-RdRp domain: nsp12-N548S, H644Y, T645A, A695T, P735S, S768F and M814K. There were six unique mutations identified across nsp14. The residues were also divided between the nsp14-ExoN(-) domain: nsp14-A93V, D175G, F216L and P258S; and the nsp14-N7-methyltransferase (N7-MT) domain: nsp14-V327A and A338T (Figure 20b). The residues identified were variably conserved across CoVs with 100% nsp12-A695 conservation across 13 CoVs and no nsp14-P258 conservation to the proline of MHV-A59.

I next sought to compare the location of mutations identified by selection with the structural locations of known catalytic residues, as well as the residues previously predicted by homology



**Figure 20. Identification of mutations in nsp12 and nsp14 selected for in the presence of 5-FU**

**A.** DBT cells were infected with at an initial MOI of 0.01 PFU/cell with either WT or nsp14-ExoN(-) viruses in triplicate lineages. Infections were performed in media alone or containing 5-FU (60uM for nsp14-ExoN(-) and 800uM for WT). Populations were then blind passaged for 6 additional passages. Total RNA was taken and prepared for Sanger sequencing. Sequencing was performed across all of nsp12 and nsp14. **B.** Domains of nsp12 are indicated by: CoV specific domain (black), fingers (blues), palm (red) and thumb (green). Domains of nsp14 are indicated by: ExoN (blue), N7-MT (black) and zinc fingers within ExoN (gray). Non-synonymous majority mutations were identified across nsp12 and nsp14.

	nsp12					
	P180H	A193S	V223I	A379T	N548S	
MHV-A59	KKLGP <sup>High</sup> IFN	AKFADALVE	DFGDFV <sup>High</sup> KTVPG	ASASAL	ISAKNR <sup>High</sup> RART	
OC43	KKLGP <sup>High</sup> IFN	TEFADKLVE	DFGDYV <sup>High</sup> IAAPG	ASASAL	ISAKNR <sup>High</sup> RART	
HCoV-229E	AALGK <sup>High</sup> VVA	VAFCD <sup>High</sup> EMVL	DFGDFV <sup>High</sup> LCPPG	ASSPAL	ISGKER <sup>High</sup> RART	
HCoV-NL63	ASLGK <sup>High</sup> IVA	VALCD <sup>High</sup> AMVA	DFGDFV <sup>High</sup> VSLPN	ASSPAL	ISGKER <sup>High</sup> RART	
CoV-HKU1	KKLGP <sup>High</sup> IFN	VIFADTLVE	DFGDFI <sup>High</sup> QTAPG	ASASAL	ISAKNR <sup>High</sup> RART	
SARS	ANLGER <sup>High</sup> VR	VQFCD <sup>High</sup> AMRD	DFGDFV <sup>High</sup> QVAPG	ASGNLL	ISAKNR <sup>High</sup> RART	
MERS	HKLGER <sup>High</sup> VR	VKFC <sup>High</sup> DMVK	DFGDFV <sup>High</sup> ITQPG	ASSNAF	ISAKNR <sup>High</sup> RART	
BAT-CoV-HKUS	HKLGER <sup>High</sup> IR	VKFC <sup>High</sup> DQMK	DFGDFV <sup>High</sup> ITQPG	ASASAL	ISAKNR <sup>High</sup> RART	
BCoV	KKLGP <sup>High</sup> IFN	TEFADKLVE	DFGDYV <sup>High</sup> IAAPG	ASASAL	ISAKNR <sup>High</sup> RART	
PEDV	ALLGT <sup>High</sup> IVA	VKFC <sup>High</sup> DAMVE	DFGDFT <sup>High</sup> C <sup>High</sup> SIKG	ASSPAL	ISGKER <sup>High</sup> RART	
AntelopeCoV	KKLGP <sup>High</sup> IFN	TEFADKLVE	DFGDYV <sup>High</sup> IAAPG	ASASAL	ISAKNR <sup>High</sup> RART	
GiCoV	KKLGP <sup>High</sup> IFN	TEFADKLVE	DFGDYV <sup>High</sup> IAAPG	ASASAL	ISAKNR <sup>High</sup> RART	
IBV-peafowl	AKMGP <sup>High</sup> IVR	IEFG <sup>High</sup> NLMVE	DFGDFQ <sup>High</sup> KTAPG	GTSNKL	ISAKNR <sup>High</sup> RART	

	nsp12					
	H644Y T645A	A695T	P735S	S768F	M814K	
MHV-A59	CSH <sup>High</sup> TDRFYR	NICQAV <sup>High</sup> SANV	HVDPA <sup>High</sup> FVSE	CYNSE <sup>High</sup> FASK	CSQHT <sup>High</sup> MLV <sup>High</sup> KM	
OC43	CSQ <sup>High</sup> SDRFYR	NICQAV <sup>High</sup> SANV	KVDST <sup>High</sup> FVTE	CYN <sup>High</sup> SDYASK	CSQHT <sup>High</sup> MLV <sup>High</sup> KM	
HCoV-229E	CTA <sup>High</sup> SDKFYR	NIFQAV <sup>High</sup> SSNI	NVDES <sup>High</sup> SVDD	CYNK <sup>High</sup> TYAGL	CSQHT <sup>High</sup> MQI <sup>High</sup> VD	
HCoV-NL63	CTA <sup>High</sup> TDRFYR	NIFQAV <sup>High</sup> SSNI	SVEES <sup>High</sup> FI <sup>High</sup> DD	CYNK <sup>High</sup> DY <sup>High</sup> AEL	CSQHT <sup>High</sup> MQI <sup>High</sup> VD	
CoV-HKU1	CSH <sup>High</sup> GDRFYR	NICQAV <sup>High</sup> TANV	YVDY <sup>High</sup> TFVNE	CYN <sup>High</sup> SDYASK	CSQHT <sup>High</sup> MLV <sup>High</sup> KI	
SARS	CNLS <sup>High</sup> HRFYR	NICQAV <sup>High</sup> TANV	DVDHE <sup>High</sup> FVDE	CYN <sup>High</sup> SNYAAQ	CSQHT <sup>High</sup> MLV <sup>High</sup> KQ	
MERS	CTT <sup>High</sup> RDRFYR	NILQAT <sup>High</sup> TANV	SPDPK <sup>High</sup> FV <sup>High</sup> DK	CYN <sup>High</sup> SDYAAK	CSQHT <sup>High</sup> LT <sup>High</sup> IKD	
BAT-CoV-HKUS	CTN <sup>High</sup> TDRFYR	NILQAT <sup>High</sup> TANV	QDPK <sup>High</sup> FV <sup>High</sup> DR	CYN <sup>High</sup> SDYATK	CSQHT <sup>High</sup> LF <sup>High</sup> IKD	
BCoV	CSQ <sup>High</sup> SDRFYR	NICQAV <sup>High</sup> SANV	MVDST <sup>High</sup> FVTE	CYN <sup>High</sup> SDYASK	CSQHT <sup>High</sup> MLV <sup>High</sup> KM	
PEDV	CSST <sup>High</sup> DRFFR	NIFQAV <sup>High</sup> SANV	IVDD <sup>High</sup> QFVVE	CYN <sup>High</sup> NDYASL	CSQHT <sup>High</sup> MQI <sup>High</sup> VD	
AntelopeCoV	CSQ <sup>High</sup> SDRFYR	NICQAV <sup>High</sup> SANV	MVDTT <sup>High</sup> FVTE	CYN <sup>High</sup> SDYASK	CSQHT <sup>High</sup> MLV <sup>High</sup> KM	
GiCoV	CSQ <sup>High</sup> SDRFYR	NICQAV <sup>High</sup> SANV	MVDTT <sup>High</sup> FVTE	CYN <sup>High</sup> SDYASK	CSQHT <sup>High</sup> MLV <sup>High</sup> KM	
IBV-peafowl	CTW <sup>High</sup> PERIYR	NIQAT <sup>High</sup> SANV	NFDS <sup>High</sup> AFVEK	CYN <sup>High</sup> NTLAKQ	CSQHT <sup>High</sup> MLV <sup>High</sup> EVE	

	nsp14					
	A93V	D175G	F216L	P258S	V327A	A338T
MHV-A59	DAEGA <sup>High</sup> H	SDHLA <sup>High</sup> DL	RATCF <sup>High</sup> NSR	HDPI <sup>High</sup> CSV	RYDV <sup>High</sup> CY	KGLA <sup>High</sup> CV
HCoV-OC43	DAEGA <sup>High</sup> H	ADHLI <sup>High</sup> DL	RATVY <sup>High</sup> NSR	HDLY <sup>High</sup> CSV	RYTLCY	KAIACV
HCoV-229E	DVEGA <sup>High</sup> H	ADFLA <sup>High</sup> GS	VATCY <sup>High</sup> NSV	HHAIC <sup>High</sup> NV	NPKAI <sup>High</sup> H	KGIRCA
HCoV-NL63	DVES <sup>High</sup> AH	SDYLS <sup>High</sup> NL	SATCY <sup>High</sup> NSV	HHTFC <sup>High</sup> NI	KPSVI <sup>High</sup> H	KGVRCA
HCoV-HKU1	DVEGA <sup>High</sup> H	SDYLL <sup>High</sup> DL	RATCY <sup>High</sup> NSR	HDII <sup>High</sup> CNV	RYNLCY	KGLACV
SARS-CoV	DVEG <sup>High</sup> CH	SDTLK <sup>High</sup> GL	RATCF <sup>High</sup> S <sup>High</sup> TS	HDQHC <sup>High</sup> QV	KFPV <sup>High</sup> LH	KAIKCV
Bat-CoV-HKU5	DVEGA <sup>High</sup> H	SDTLK <sup>High</sup> DL	RASTY <sup>High</sup> SSP	HDRY <sup>High</sup> CSV	RFERYV	KGIFIV
BCoV	DAEGA <sup>High</sup> H	ADHLI <sup>High</sup> DL	RATAY <sup>High</sup> NSR	HDLY <sup>High</sup> CSV	RYTLCY	KAIACV
PEDV	DVEGA <sup>High</sup> H	SDYLA <sup>High</sup> NL	VATCY <sup>High</sup> NSA	HHEHC <sup>High</sup> NV	NPKAI <sup>High</sup> Y	KGIRCA
AntelopeCoV	DAEGA <sup>High</sup> H	ADHLI <sup>High</sup> DL	RATAY <sup>High</sup> NSR	HDLY <sup>High</sup> CSV	RYTLCY	KAIACV
GiCoV	DAEGA <sup>High</sup> H	ADHLI <sup>High</sup> DL	RATAY <sup>High</sup> NSR	HDLY <sup>High</sup> CSV	RYTLCY	KAIACV
IBV-peafowl	DVEA <sup>High</sup> TH	ADNLC <sup>High</sup> NV	RATTF <sup>High</sup> NSY	HDLHC <sup>High</sup> NV	KVNV <sup>High</sup> VY	KGIKCV

Conservation across CoVs  
■ High ■ Medium ■ Low

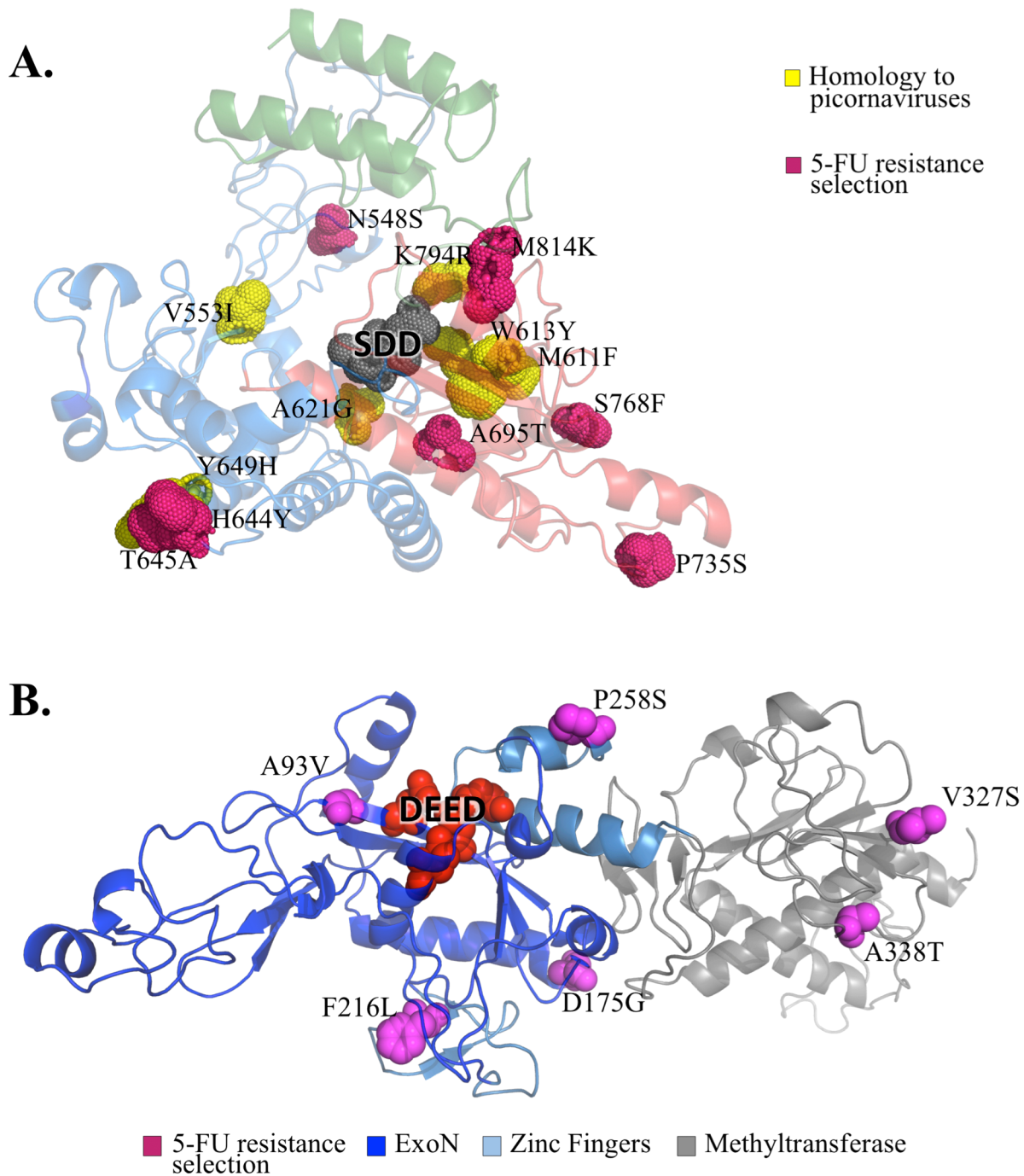
Figure 21. Conservation of selection-identified residues across CoVs

Amino acid alignments of nsp12 and nsp14 for a representative set of CoVs. Residues where mutations occurred are highlighted with colors indicating the degree of conservation across CoVs as high (green), medium (yellow), and low (red). Conservation ranged from completely conserved (nsp12-A695) to no residues matching that of MHV-A59 (nsp14-P258).

modeling. Mutations that arose by passage in nsp12-RdRp were mapped onto the modeled structure from Chapter II (Sexton *et al.*, 2016). All of the mutations identified by selection are distinct from those previously identified by homology, although nsp12-T645A and nsp12-H644Y (in the fingers domain) are adjacent to nsp12-Y649H. One additional mutation, nsp12-N548S, was identified in the fingers domain. Three mutations: nsp12-A695T, S768F and P735S were found in the palm domain but all were distant from the SDD active site. Finally, the only mutation present in multiple lineages, nsp12-M814K, is within the thumb domain. However, modeling suggests that nsp12-M814K folds into a pocket behind the active site where many of the residues identified by homology modeling also reside (Figure 22a). In order to investigate the structural layout of mutations that arose in nsp14, a model of nsp14 needed to be generated. The SARS-CoV nsp14 structure was solved and published during this dissertation research, providing the opportunity to model MHV nsp14 using Phyre2 software as described for nsp12-RdRp core domain in Chapter II (Kelley *et al.*, 2015). Residues identified by selection were mapped onto the modeled structure. One of the mutations, nsp14-A93V, was located just behind the DEED catalytic residues. Two of the mutations were in zinc fingers of nsp14-ExoN: nsp14-F216L and P258S. Another, nsp14-D175G, resides at the interface between nsp14-ExoN and nsp14-N7-MT. Finally, nsp14-V327S and nsp14-A338T were identified in the nsp14-N7-MT domain (Figure 22b). Overall, most mutations were predictably identified in the RdRp domain of nsp12 and in the ExoN domain of nsp14. However, a few mutations were identified in the nsp12 CoV-specific domain and in the nsp14-N7-MT domain. Furthermore, of the mutations identified in the nsp12-RdRp domain, most were in either the fingers or the palm, and the only mutation identified in the thumb domain appears to be structurally within the predicted fidelity-modulating pocket behind the SDD motif.

### **Resistance of 5-FU-treated virus populations to the mutagens 5-FU and 5-azacitidine**

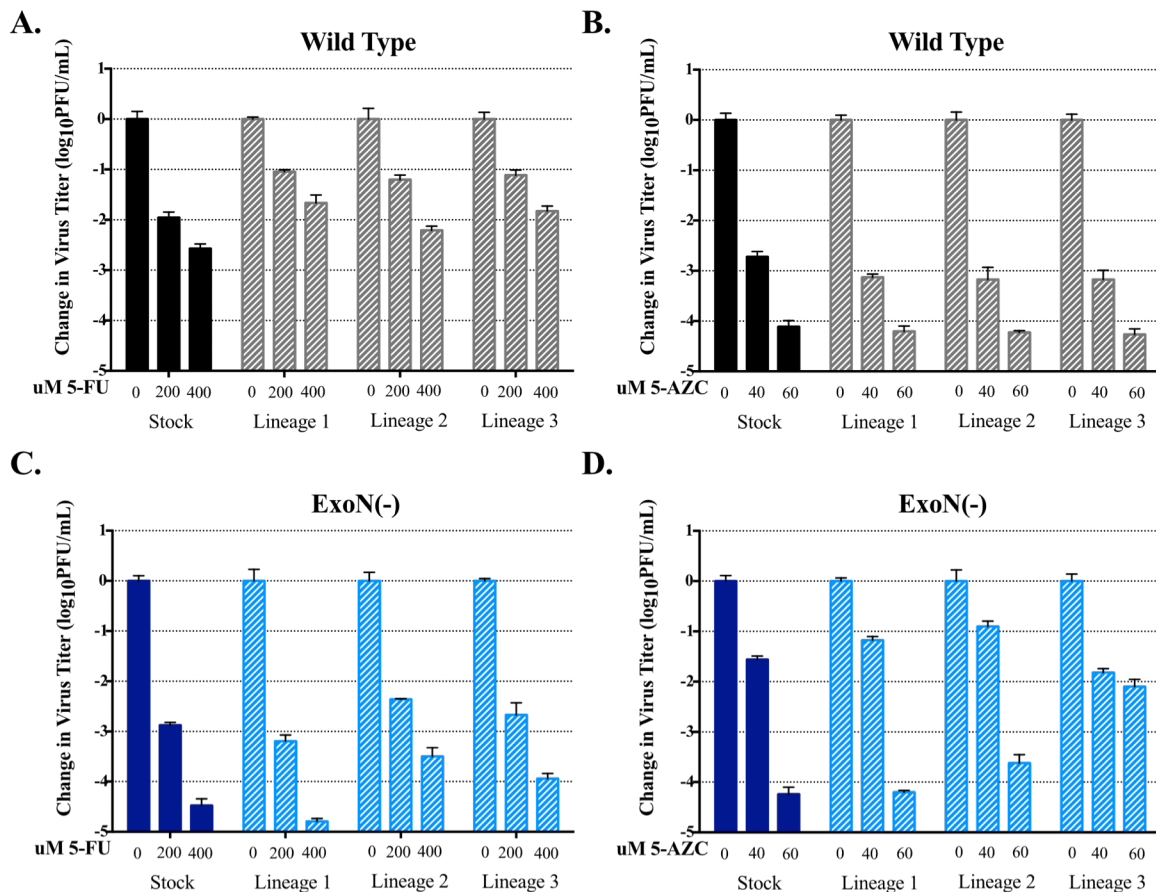
The goal of passaging WT and nsp14-ExoN(-) viruses in the presence of 5-FU was to identify residues that confer increased fidelity; therefore, I next sought to determine if the treated and passaged virus populations displayed increased resistance to mutagens. I compared resistance to either 5-FU or the mutagen 5-azacitidine (5-AZC) between WT and nsp14-ExoN(-) stock viruses and 5-FU-passaged virus populations. When treated with 5-FU, all WT lineages result in increased resistance to 5-FU at 200 and 400uM 5-FU when compared to the original virus stock (Figure 23a). In contrast, no difference in resistance to 5-AZC was observed for any WT lineage



**Figure 22. Location of identified residues on modeled MHV-A59 structures on nsp12-RdRp and nsp14**

Phyre2 software was used to model **A.** the core RdRp domain of nsp12 and **B.** nsp14. Mutations that arose within these structures were mapped and compared with previously identified mutations, known domains and catalytic residues. The RdRp domain contains fingers (blue), palm (red) and thumb (green) domains.





**Figure 23. Resistance of 5-fluorouracil (5-FU) treated population to 5-FU or 5-azacytidine (5-AZC)**

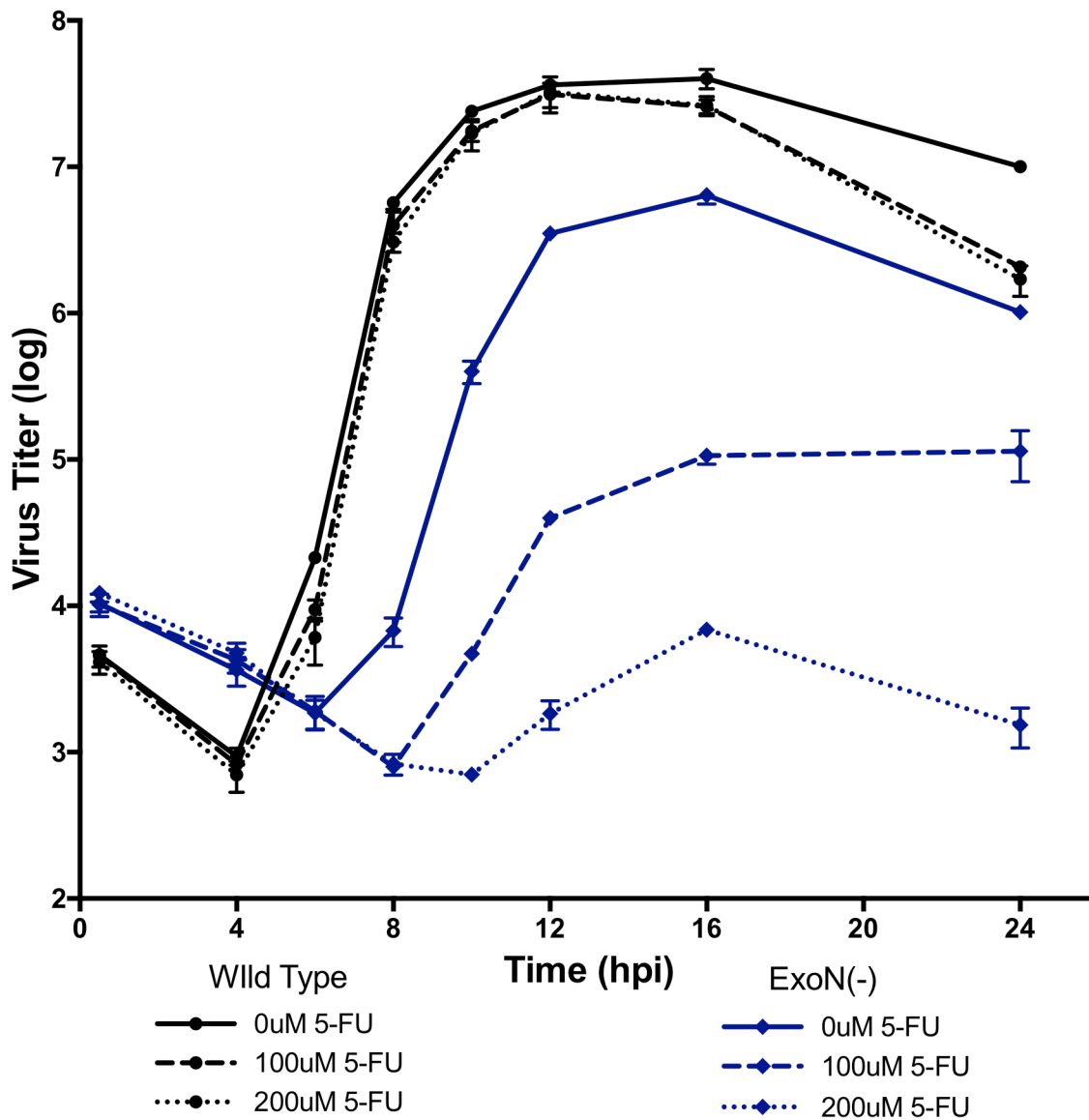
DBT cells were pre-treated with indicated concentrations of 5-FU (**A and C**) or 5-AZC (**C and D**) for 30 min. Treatment was removed and cells were infected with stock viruses used for initiating passages or passage 7 5-FU-treated populations in the WT background (**A and B**) or passage 6 5-FU-treated populations nsp14-ExoN(-) background (**C and D**) at an MOI of 0.01. Passaged populations were from the final passage 7 supernatants except for Media containing 5-FU or 5-AZC was replaced 30 min p.i. Virus samples were taken at 24 (WT) or 32 (nsp14-ExoN(-)) h.p.i and titer was determined by plaque assay. Data represents 2 independent experiments, each with 2 replicates. Error bars represent SEM.

(Figure 23b). I also tested nsp14-ExoN(-) populations passaged in the presence of 5-FU for resistance to 5-FU and 5-AZC. However, as the final passage seven populations were no longer capable of replicating they could not be used for these studies. Therefore, I tested the passage six populations instead. Nsp14-ExoN(-) lineages two and three resulted in increased resistance to 5-FU or 5-AZC (Figure 23c and d). However, the magnitude of increased resistance to 5-FU observed was smaller than for WT lineages (Figure 23a and c). Interestingly, nsp14-ExoN(-) lineage one did not demonstrate any increase in resistance to 5-FU or 5-AZC (Figure 23c and d). This suggests that the mutations selected for in this lineage benefited the virus by a mechanism unrelated to nucleotide selectivity. This idea is further supported by the fact that all of the mutations selected for in lineage one are found outside of the nsp12-RdRp and nsp14-ExoN domains (Figure 20b). Replication assays demonstrated that nsp14-ExoN(-) viruses are delayed in replication (Figure 17). Therefore, passaging nsp14-ExoN(-) viruses may result in compensatory mutations not directly related to the proofreading function of nsp14-ExoN.

#### **Replication kinetics of nsp14-ExoN(-) viruses treated with 5-FU result in an increased lag phase as well as decreased titers**

Since MHV 5-FU passaged populations resulted in differential resistance to mutagens between WT and nsp14-ExoN(-) lineages as well as within nsp14-ExoN(-) lineages, I wanted to determine the effect of 5-FU treatment on replication kinetics. Clint Smith infected DBT cells with WT or nsp14-ExoN(-) viruses (at an MOI of 1) in the presence of 0, 100 or 200uM 5-FU. Supernatant samples were collected at 4, 6, 8, 10, 12, 16 and 24 hours post infection and titered by plaque assay. Replication kinetics for WT viruses were similar for treated and untreated samples with a similar lag phase, exponential phase, and peak titers. The only difference observed between treated and untreated replication curves were half-log decreases in titers observed at 24 hours post infection. These decreased titers were identical for WT viruses treated with 100uM 5-FU or 200uM 5-FU (Figure 24). Nsp14-ExoN(-) viruses result in an increased lag phase over WT-nsp14 viruses and never reached peak titers. This phenotype was exaggerated in a dose dependent manner when nsp14-ExoN(-) viruses were treated with 5-FU. When treated with 100uM 5-FU, the lag phase increased from six hours to eight hours and peak titers were 2-logs lower than replication without 5-FU. The lag phase extended to 10 hours and titers dropped an additional 1.5-logs when nsp14-ExoN(-) viruses were treated with 200uM 5-FU (Figure 24). These data support the hypothesis that WT viruses are highly resistant to 5-FU due to the





**Figure 24. Replication kinetics of WT and ExoN(-) viruses in the presence of 5-FU**

DBT cells were infected with the viruses indicated at an MOI of 1 PFU/cell. Supernatant aliquots were taken at indicated times p.i. and titer determined by plaque assay. Data represents 3 independent experiments. Error bars represent SEM.

presence of nsp14-ExoN proofreading activity. In addition they demonstrate how mutations that alter phenotypes other than fidelity could compensate for the effects of 5-FU treatment.

**Engineering and recovery of a subset of mutant viruses identified by selection in the WT and nsp14-ExoN(-) background**

Having investigated the involvement of residues in the nsp12-RdRp fingers and palm domains on fidelity (Chapter II), I next wanted to determine if residues outside of these regions regulate fidelity. Therefore, I engineered and attempted recovery of mutant viruses for a sub-set of the mutations identified by selection. Mutant viruses with the following mutations were engineered: nsp12-P180H/A193S/V223I (nsp12-PH-AS-VI) of the CoV-specific domain, nsp12-N548S of the fingers domain, nsp12-M814K of the thumb domain, nsp14-A93V located just outside of the DEED catalytic site, nsp14-D175G found at the interface between nsp14-ExoN and the nsp14-N7-MT, and nsp14-P258S in the second zinc finger of nsp14-ExoN. Recovery of mutant viruses was attempted in the WT and nsp14-ExoN(-) backgrounds a maximum of three times. In contrast to residue identification by homology modeling or 5-FU population minority variants, all engineered viruses were recoverable in both the WT and nsp14-ExoN(-) background, with only one exception, nsp14-D175G, which was only recovered in the WT background (Table 3).

**Table 3. Recovery of mutant viruses from 5-FU selection using site-directed mutagenesis.**

nsp12 Region	MHV engineered substitutions (nsp12)	Recovery: WT (nsp14-ExoN(+))	Recovery: nsp14-ExoN(-)
CoV Specific Domain	P180H/A193S/V223I (PHASVI)	Yes	Yes
	A379T	Yes	Yes
Fingers	N548S	Yes	Yes
Thumb	M814K	Yes	Yes
nsp14 Region	MHV engineered substitutions (nsp14)	Recovery: WT (nsp14-ExoN(+))	Recovery: nsp14-ExoN(-)
ExoN	A93V	Yes	Yes
	D175G	Yes	No
	P258S	Yes	Yes

This is consistent with identification by selection, which should result in highly fit viruses. In the case of nsp14-D175G it is likely that this mutation is dependent on one of the other mutations selected for in this passage series. Further supporting this hypothesis, nsp14-D175G was present as only 60% of the population, while the other mutation in nsp12 and nsp14 for lineage 3 were already fixed in the population. Nsp14-D175G may have arisen later during passaging once a complementary mutation was fixed. No additional mutations were identified in nsp14 but three were identified in nsp12-RdRp (Figure 20).

## **Discussion**

At the start of this dissertation research, no mutations in CoV nsp12 had been identified that altered nucleotide selectivity and only the DEED active site of nsp14 had been altered resulting in decreased fidelity (Eckerle *et al.*, 2007; 2010; Graham *et al.*, 2012; Smith *et al.*, 2013). Additionally, interactions between nsp12 and nsp14 activities had not been investigated. Chapter II of this dissertation identified residues involved in nucleotide selectivity within the nsp12-RdRp by homology (Sexton *et al.*, 2016). However, nsp12 also encodes a CoV-specific domain and nsp14 influences fidelity. Further, as regulation of CoV fidelity is dependent on multiple proteins, disrupting or strengthening interactions between these proteins could alter CoV fidelity (Eckerle *et al.*, 2007; Sexton *et al.*, 2016; Smith *et al.*, 2015). Chapter III of this dissertation built upon the data presented in Chapter II, with the goal of identifying mutations across nsp12 or nsp14, which alter nucleotide selectivity and fidelity. In this work, I sought to determine whether the residues identified by selection would be unique or would mimic those identified by homology modeling. Additionally, I sought to determine if minority variants (defined as present at greater than 1% of the population) of nsp14-ExoN(-) populations treated with 5-FU confer increased resistance to 5-FU or demonstrate other measureable benefits to the virus. Finally, I sought to investigate the interaction between the phenotypic effects of mutations in nsp12 and nsp14.

### **Some minority variants of mutagen-treated CoVs require the presence of alternative residues for viability**

RNA viruses are known to replicate with high mutation rates resulting in mutant swarms: populations with a common consensus sequence where most individuals vary from the consensus (Domingo *et al.*, 1996; Lauring and Andino, 2010; Vignuzzi *et al.*, 2005b). Mutant swarms

influence the overall phenotype of viral populations (Vignuzzi et al., 2005b). However, individual minority-variant mutations from CoV populations had not been previously investigated. In this study, I identified minority variants present in nsp12 and recovered viruses containing identified mutations in an isogenic MHV background. Considering that these mutations were identified from nsp14-ExoN(-) populations that arose in the presence of 5-FU, it was possible that these variants would result in increased resistance to 5-FU. Additionally, since these mutations were present at a frequency of 1% of the population or more, it seemed likely that they would be recoverable as variant viruses in an isogenic background. As expected, the majority of mutations identified were recoverable in the WT background. Only nsp12-P374S was non-recoverable. However, in the nsp14-ExoN(-) background two additional mutations were non-recoverable (Table 2). Since the mutations arose originally in the nsp14-ExoN(-) background, this was particularly surprising. There are many possible explanations for the non-recovery of these variant viruses. In the case of nsp12-P374S, it may be that this virus would be viable in the nsp14-ExoN(-) background. This could be the case if the mutation is beneficial in the context of an inactivated nsp14-ExoN but detrimental in the WT-nsp14 background this could be the case. One way this could occur is if the mutation results in an increase in the fidelity of nsp12-RdRp. For variants non-recoverable in the nsp14-ExoN(-) background (nsp12-L364F, nsp12-E519G and potentially nsp12-P374S), several possibilities exist to explain these data. The simplest explanation is that MHV and SARS-CoV are different viruses and these variants may affect replication differently between the two. However, in most cases attempted, conserved mutations between MHV and SARS-CoV have resulted in similar phenotypes (Sexton et al., 2016; Smith et al., 2013). It is also possible that the observed minority variant viruses require co-infection with WT viruses or the presence of additional mutations within the viral clone in order to persist in a population.

### **Minority variants result in subtle to no alternations of replication phenotypes**

The minority variants present in a population could represent mutations that are likely to be fixed in subsequent rounds of replication under similar conditions or could be present only when complemented by the mutant swarm. None of the mutations investigated in this study resulted in a significant change in resistance to the nucleotide analogs 5-FU or 2'-C-MeA (Figure 15). Similarly, none resulted in any increase in replication kinetics or peak titers, and some even resulted in decreases in both (Figure 17). Therefore, it is unlikely that the majority of 5-FU-

treated population minority variants are present in the mutant swarms resist incorporation of 5-FU into viral RNA or increase replication kinetics. However, in co-infection experiments, nsp12-S900C/nsp14-ExoN(-) resulted in equal fitness to nsp14-ExoN(-) in the presence of 5-FU, despite replication and RNA accumulation defects in the absence of 5-FU (Figures 17a, 18 and 19b). Therefore, some of the minority variants present in mutant swarms may represent viruses that are neutral compared with WT viruses in the specific selective environment. These viruses are likely ones which benefit from minor to major bottlenecking events and so are able to persist in the population under specific conditions. Additionally, a threshold may exist, dependent on the specific conditions, which defines what fraction of the population is likely to consist of these types of minority variants. Above the bottlenecking threshold may be where mutations with measurable outcomes for virus replication can be identified. In this set of variants, nsp12-E519N was present as 25% of the original population (Table 2). This is also the only variant that resulted in a measurable beneficial phenotype, specifically in the nsp14-ExoN(-) background, RNA accumulation was increased over np14-ExoN(-) alone and in co-infection experiments nsp12-E519N/nsp14-ExoN(-) increased in the population over input (Figures 18b and 19a). However, even in the case of nsp12-E519N the phenotypes were relatively subtle. Together, these data suggest that minority variants arising in the presence of a selective pressure are not strong targets for overcoming the specific applied selective pressure.

### **Selection for mutations that result in resistance to 5-FU identifies residues involved in nucleotide selectivity, but also increases titers by an alternate mechanism**

In contrast to identifying minority variants of 5-FU treated populations, selection for resistance to a mutagen followed by the identification of majority variants by Sanger sequencing is a reliable source of residues involved in fidelity regulation (Arias et al., 2008; Beaucourt et al., 2011; Coffey et al., 2011; Levi et al., 2010; Zeng et al., 2014). In 5-FU treated populations, mutations arose within the expected domains of nsp12 and nsp14, as expected, the nsp12-RdRp finger and palm domains and the nsp14-ExoN domain (Campagnola et al., 2015; Korneeva and Cameron, 2007). Yet, mutations were also identified in all other domains (nsp12-RdRp thumb, nsp12-CoV-specific and nsp14-N7-MT). None of the mutations identified were the same as those identified by homology modeling and only nsp12-M814K was identified in multiple lineages (Figures 20, 22). In fact, nsp12-M814K was also the first mutation to arise in a separate nsp14-ExoN(-) passaging experiment performed by Xiaotao Lu in our lab (unpublished). This

separate passaging experiment was performed in the absence of any intentional, external selective pressures. The repeated selection for the variant nsp12-M814K in the nsp14-ExoN(-) background (whether treated or untreated) suggests that this variant provides significant compensation for the replication defects conferred by the inactivation of nsp14-ExoN. Yet the facts that none of the mutations flagged by selection were identified by homology modeling, and only nsp12-M814K was identified in multiple lineages and in alternative passaging experiments, suggests that CoVs exist in a robust mutational landscape, where many viable mutations are generated and available for selection and can compensate for treatment with 5-FU and likely other mutagens. To determine the breadth of mutations available to overcome challenge with 5-FU, it would be interesting to see how many lineages it would take before repetitive mutations were identified.

The 5-FU-treated WT populations resulted in increased resistance specifically to 5-FU rather than broadly to mutagens; whereas, the 5-FU-treated nsp14-ExoN(-) populations resulted in increased resistance to multiple mutagens or had no change in resistance (Figure 23). These data suggest that there are different pressures being applied to WT vs. nsp14-ExoN(-) viruses when treated with 5-FU. Replication kinetics experiments for WT and nsp14-ExoN(-) viruses further support this observation, WT viruses exhibited no change in replication kinetics when treated with 5-FU, but nsp14-ExoN(-) viruses had longer lag phases and reached decreased peak titers in the presence of increasing concentrations of 5-FU (Figure 24). It is important to remember that nsp14-ExoN(-) viruses encode a full nsp14 protein, which is inactivated by substituting two of the active site residues. Therefore, it is possible that 5-FU effects on nsp14-ExoN(-) replication kinetics can be explained if the mutant nsp14-ExoN still recognizes and attempts to remove 5-FU nucleotides incorporated by nsp12-RdRp. If this is the case, nsp14-ExoN may stall the replication-transcription complex (RTC) each time 5-FU is incorporated and ultimately result in an increased lag phase. If this hypothesis is correct, then in the context of an inactivated nsp14-ExoN, reducing or eliminating the interaction between nsp12-RdRp and nsp14-ExoN may be beneficial and thus selected for. Generally, any mutations that increase the speed of replication would be selected for, unless it is otherwise deleterious to fitness. It will be interesting to determine if nsp12-PH-AS-VI, nsp12-M814K, nsp14-D175G or nsp14-F216L result in increased replication speed in the context of an nsp14-ExoN(-) virus.

## Conclusion

The results described in this chapter demonstrate that (1) minority variants found in 5-FU treated populations provide subtle to no ability to resist 5-FU; (2) that selection identifies a diverse panel of mutations most of which are likely involved in nucleotide selectivity; and (3) that nsp14-ExoN(-) viruses may be compensated for by mutations involved in increasing replication kinetics. The increased lag phase for MHV nsp14-ExoN(-) viruses is a consistently observed phenotype (Eckerle et al., 2007; Sexton et al., 2016; Smith et al., 2013). Defects in viral replication are generally attributed exclusively, and specifically, to the loss of proofreading. However, a recent paper demonstrated that mutating one of the nsp14-ExoN zinc fingers results in an altered innate immune response (Becares *et al.*, 2016). The results of that study suggest that nsp14-ExoN may have additional functions. They also suggest it may be possible to unlink the proofreading defect from that of the larger replication lag phase. This could greatly simplify fidelity studies in the future and provide a mechanism for further understanding the interactions between nsp14 and other CoV proteins involved in fidelity, particularly nsp12. Finally, the results of this chapter again highlight the need for a robust and reproducible biochemical assay for the CoV RTC.

## CHAPTER IV

### DEVELOPMENT OF A ROBUST AND QUANTITATIVE MEASURE OF FIDELITY FOR CORONAVIRUSES

#### Introduction

Fidelity of viral replication has been studied most extensively for the picornaviruses (Campagnola *et al.*, 2015; Gnädig *et al.*, 2012; Gohara *et al.*, 2004; Korneeva and Cameron, 2007; Moustafa *et al.*, 2011; Pfeiffer and Kirkegaard, 2003; 2005; Vignuzzi *et al.*, 2005a). Many tools have been developed for the picornavirus field to investigate the incorporation of incorrect nucleotides and the rate of misincorporations. The RNA-dependent RNA polymerase (RdRp) of coronaviruses (CoVs) modeled closely to the RdRps of picornaviruses (Sexton *et al.*, 2016). However, CoV genomes are roughly three times the length of picornavirus genomes and encode many proteins, which interact to form the CoV replication-transcription complex (RTC). Therefore, many of the methods used to investigate fidelity in picornaviruses may not be effective for CoVs. For instance, there are no robust biochemical systems available for CoV replication. The RdRp of CoVs, located in non-structural protein 12 (nsp12), is often unstable when expressed alone (Velthuis *et al.*, 2009). In fact, nsp12-RdRp requires the presence of (at minimum) nsp7 and nsp8 for stable and processive polymerization (Subissi *et al.*, 2014).

Testing for changes in selectivity to one or more mutagens is a regularly used screening strategy for identifying RdRp variant viruses that alter replication fidelity. To determine if RdRp variant viruses that exhibit altered selectivity for mutagens have altered fidelity, a fraction of RdRp variant virus genomes are cloned and Sanger sequenced to determine any change in the accumulation of mutations (Gnädig *et al.*, 2012; Rozen-Gagnon *et al.*, 2014; Zeng *et al.*, 2014). In the case of poliovirus a specific fidelity assay was developed. The poliovirus 2C protein (an ATPase) is sensitive to guanidine hydrochloride (GUA-HCl), and resistance is conferred by specific mutations. Therefore, plaque reduction in the presence of GUA-HCl was assayed to quantify escape mutants in a population (Pfeiffer and Kirkegaard, 2003). *Orthoreoviruses* are also sensitive to GUA-HCl, demonstrating that sensitivity is not picornavirus specific (Murray and Nibert, 2007). After identifying RdRp fidelity determinants, deep sequencing of RdRp-variant virus populations was performed to further describe the types and number of mutations



arising compared to WT populations. However, current deep sequencing platforms result in a relatively high number of errors. These errors arise during each step of sample preparation and can confound data interpretation. New techniques have recently been developed to limit the influence of such processing errors such as circle resequencing and duplex sequencing (Acevedo and Andino, 2014; Kennedy et al., 2014; Lou et al., 2013; Schmitt et al., 2012). Circle resequencing and duplex sequencing require a large amount of RNA, which can be a limiting factor for some virus variants. In particular, CoV nsp12-RdRp variants in the nsp14-ExoN(-) background result in greatly reduced total RNA and RNA genomes (Figure 9d, f and 18b). Additionally, determining changes in fidelity by deep sequencing is difficult because of selection. Beneficial viral mutations are selected for and those that are detrimental are selected against, providing a bias in mutation frequencies. A bioinformatics approach was developed to determine changes in fidelity by deep sequencing. This method only counts stop codons and known lethal mutations introduced in viral genomes (Xiao et al., 2016); therefore, it requires an extensive knowledge of the genome, which is not always available.

A novel method for investigating RNA virus fidelity was introduced by Marco Vignuzzi et al. Briefly, a microRNA target was cloned into poliovirus genomes encoding WT or fidelity-altered RdRps. Viruses were grown in cell lines known to produce complementary microRNA thus, virus replication was restricted except when mutations were present in the microRNA target that could evade detection (Vignuzzi et al., 2008). MicroRNA targeting as a fidelity assay could be applied to many viruses and isn't affected by genome size or RNA quantity. However, it requires a reverse genetics system and knowledge of microRNAs produced by cell lines permissive for infection. In the case of CoVs, a reverse genetics system is available for MHV, SARS-CoV and MERS-CoV but viral recovery is cumbersome and time intensive. Additionally, the cell line typically used for MHV studies, murine delayed brain tumor (DBT), is not well characterized (Chen and Baric, 1996). Finally, CoVs are known for high rates of recombination and rapid deletion of nonfunctional genes (Makino *et al.*, 1986), which may prevent observation of fidelity data when applying this system to CoVs. In this chapter, I present a strategy for the development of a broadly applicable assay for investigating fidelity in CoVs. The CRISPR/Cas based assay I propose utilizes similar principles to that of the above microRNA method but doesn't require engineering microRNA targets into viruses.

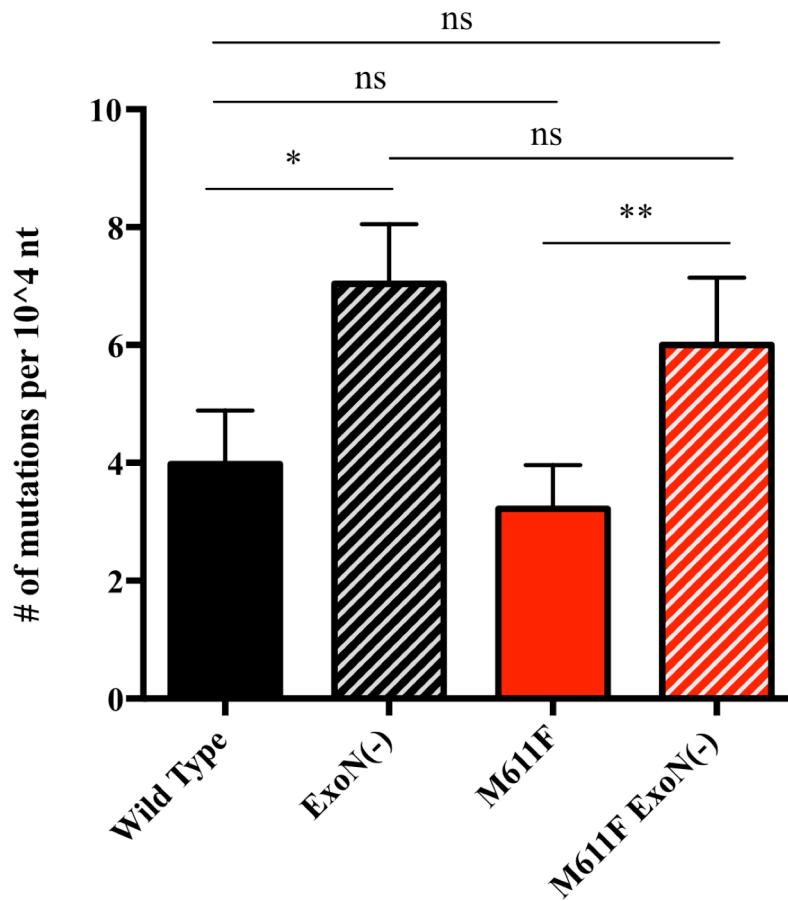
CRISPR/CAS is a prokaryotic system for acquiring immunity to foreign nucleic acids. There are many CRISPR (clustered regularly interspaced short palindromic repeats) associated system (CAS) proteins encoded for in prokaryotic systems, discussed here is the Cas9 endonuclease. The targeting of foreign DNA using CRISPR/Cas9 has been well elucidated. Briefly, fragments of invading DNA are inserted between CRISPR elements and are expressed as small non-coding RNAs, called CRISPR RNAs (crRNAs). When bound to adapter RNAs called transactivating CRISPR RNAs (tracrRNA), crRNAs function to target Cas9 endonucleases to cleave reintroduced foreign DNA (Mohanraju *et al.*, 2016). The CRISPR/Cas9 system was developed as a general DNA editing tool. In these editing systems, crRNAs and tracrRNAs are expressed in a single RNA construct and called sgRNAs (single guide RNAs) (Le Cong *et al.*, 2013; Mali *et al.*, 2013). Recently, it was discovered that CRISPR/CAS type systems could target mRNAs. One mechanism for targeting mRNA is through expression of an alternate Cas9-like protein, such as C2c2, which cleaves RNA instead of DNA. Once activated C2c2 cleaves RNAs nonspecifically and is toxic to the bacterial cell in which it is expressed (Abudayyeh *et al.*, 2016; East-Seletsky *et al.*, 2016). Alternatively, Cas9 itself can be targeted to RNA instead of DNA. RNA targeting occurs through an alternative RNA, the small CRISPR/CAS-associate RNA (scaRNA). Interestingly, when Cas9 interacts with mRNA it likely doesn't cleave the RNA, as all of the known catalytic domains of Cas9 can be mutated without consequence to mRNA down regulation. However, mutating the RNA interacting arginine-rich motif (ARM) of Cas9 resulted in the restoration of WT levels of targeted mRNAs (Sampson *et al.*, 2013). RNA-targeted Cas9 was adapted to inhibit the RNA virus HCV. The tracrRNA and scaRNA were combined into a single construct, coined RNA-targeting guide RNA (rgRNA), with a 20bp region complementary to HCV (Price *et al.*, 2015). In both the endogenous bacterial system and the bioengineered system, RNA with secondary structure was targeted. Unexpectedly, only the 2 nucleotides at the 3' end of the HCV genome target needed to be a perfect match in the rgRNA for effective inhibition of viral translation (Price *et al.*, 2015). Considering the specificity of targeting by this system, RNA targeted Cas9 may be adaptable as a quantitative fidelity assay for CoVs.

In this chapter, I describe a strategy to develop an rgRNA/Cas9-based system for investigating fidelity in CoVs and report preliminary studies toward this goal. I first demonstrate that methods used in picornaviruses aren't effective screens or measures for CoVs. I then provide evidence that MHV translation and replication can be targeted by rgRNA/Cas9. I show that transfection of

the Cas9 vector into HEK293T cells unexpectedly increases translation of MHV but doesn't increase viral titer. Further, I present preliminary data demonstrating that rgRNAs targeted to MHV alone may inhibit translation of MHV. Finally, I present additional preliminary data suggesting that small differences in escape may be observable between WT and nsp12-RdRp mutant viruses. I performed all of the work described in this chapter.

### **TOPO TA cloning, combined with Sanger sequencing of a 1kb region of the genome results in a small dynamic range**

Cloning subsets of viral genomes and counting mutations present in individual bacterial clones has been routinely used to determine changes in fidelity between viruses (Chung *et al.*, 2007; Gnädig *et al.*, 2012; Rozen-Gagnon *et al.*, 2014). Moreover, in a paper from the Denison lab, full genomes of individual clones were sequenced for WT and nsp14-ExoN(-) MHV viruses and mutations present were identified by Sanger sequencing. This resulted in a 15-fold change in mutations present between WT and nsp14-ExoN(-) viruses (Eckerle *et al.*, 2007). Therefore, I first attempted to assess fidelity for CoVs by TOPO TA cloning a 1kb region across nsp2. The nsp2 protein was chosen because nsp2 is dispensable for replication in cell culture and has no known role in the regulation of fidelity (Freeman *et al.*, 2014). Since nsp2 is dispensable for replication, it is one of the proteins most likely to experience neutral selection and thus would be capable of carrying mutations but would not artificially enhance the number of mutations observed due to positive selection. Additionally, nsp2 is likely to maintain neutral selection in the context of an altered fidelity virus or the introduction of a nucleoside analog. WT, nsp12-M611F (from homology modeling, Chapter II), nsp14-ExoN(-), and nsp12-M611F/nsp14-ExoN(-) viruses were cultured in triplicate for 5 passages. The nsp2 region was cloned into the TOPO TA vector and 75 to 100 transformed bacterial clones were Sanger sequenced. Four mutations per  $10^4$  nucleotides were recorded for WT MHV (Figure 25). This is much higher than the number of mutations recorded by full genome sequencing, 0.29 (Eckerle *et al.*, 2007). However, the full genome sequencing only measured fixed mutations where as TOPO cloning allowed for observation of minority variants. Additionally, initial infections were performed with MOIs of 0.1 and 0.01, respectively. Therefore, differential results were not unexpected. The nsp14-ExoN(-) virus resulted in just over 7 mutations per  $10^4$  nucleotides (Figure 25), again more than the 4.4 mutations per observed for full genome sequencing (Eckerle *et al.*, 2007). The mutation frequency for nsp12-M611F viruses was similar to those of WT and nsp14-ExoN(-)



**Figure 25. Mutation frequency determined by TOPO TA cloning**

DBTs were infected with an initial MOI of 0.01 then blind passaged in triplicate for 5 passages. Total RNA was collected and cloned into TOPO TA vector. A 1kb region in nsp2 was sequenced for between 75 and 100 clones for each virus. Average mutation accumulation per 10<sup>4</sup> nucleotides is shown. (Error bars represent SEM; ns, not significant; \*P < 0.05; \*\*P < 0.01 by Mann Whitney test).

alone, with 3.2 and 6 mutations observed, respectively (Figure 25). The fold change in mutations observed, when nsp14-ExoN was inactivated, was 1.8 for the TOPO TA experiment and 15.2 for the full genome sequencing. A two-fold change in mutations observed between WT and nsp14-ExoN(-) mutations is uncharacteristic for CoVs, which have an average reported fold change of 14 (Eckerle et al., 2007; 2010; Graham et al., 2012; Sexton et al., 2016; Smith et al., 2013). No statistically significant difference between WT and nsp12-M611F viruses was observed, which was not surprising considering the relatively minor difference observed between WT and nsp14-ExoN(-) viruses. Additionally, reversion data demonstrated that nsp12-M611F/nsp14-ExoN(-) had nearly completely reverted to WT-nsp12 by passage 5, so little to no difference between these samples would be expected (Figure 12b). Together these data suggest that TOPO TA cloning is not an effective way to measure changes in mutation frequency conferred by mutations in either the nsp12-RdRp or nsp14-ExoN of CoVs. However, utilizing a different region of the genome could result in frequencies with enough resolution to observe fidelity changes for nsp12-RdRp mutant viruses and that are consistent with previously observed differences between WT and nsp14-ExoN(-) viruses.

### **Selectivity for nucleotide analogs varies between mutant viruses and nucleoside analogs**

When my dissertation research began, all published reports of fidelity-altered RNA viruses (except one) demonstrated that resistance to nucleoside analogs, specifically mutagens, correlated with changes in fidelity. An exception was a foot and mouth disease RdRp variant virus, selected in the presence of ribavirin, which exhibited in decreased fidelity for natural nucleotides while specifically excluding ribavirin incorporation (Arias *et al.*, 2008). However, since the majority of reported variants demonstrated consistency between fidelity and nucleoside analog selectivity, I tested nsp12-RdRp mutant viruses for altered sensitivity to a panel of nucleoside analogs. I included in this panel the mutagens 5-FU, a uracil analog; 5-AZC, a cytidine analog; and ribavirin, a guanosine analog. I additionally tested for altered sensitivity to the non-obligate chain terminator: 2'-C-MeA, a cytidine analog with an altered sugar. WT, nsp14-ExoN(-), nsp12-M611F/nsp14-ExoN(-) or nsp12-V553I/nsp14-ExoN(-) viruses were assessed for altered selectivity for any of the nucleoside analogs described. As presented in Chapter II, nsp12-V553I/nsp14-ExoN(-) viruses were resistant to both 5-FU and 5-AZC (Figure 26a, b). However, nsp12-V553I/nsp14-ExoN(-) was more sensitive to 2'-C-MeA and didn't demonstrate any change in sensitivity to ribavirin (Figure 26c, d). The nsp12-M611F/nsp14-

ExoN(-) mutant virus demonstrated similarly inconsistent results, demonstrating resistance to 5-FU and 2'-C-MeA but no change in selectivity for either 5-AZC or ribavirin (Figure 26). The only virus that resulted in predictable sensitivity to nucleoside analogs was nsp14-ExoN(-) when compared to WT viruses; in all cases observed nsp14-ExoN(-) viruses are more sensitive to nucleoside analogs – consistent with their known proofreading defect (Figure 26a-c). This result was observed for ribavirin as well (data not shown). Together this suggests that mutating nsp12-RdRp may result in complex changes to nucleotide selectivity, where each specific nucleotide is differentially incorporated. Therefore, mutagen treatment assays do not define the overall fidelity of a virus.

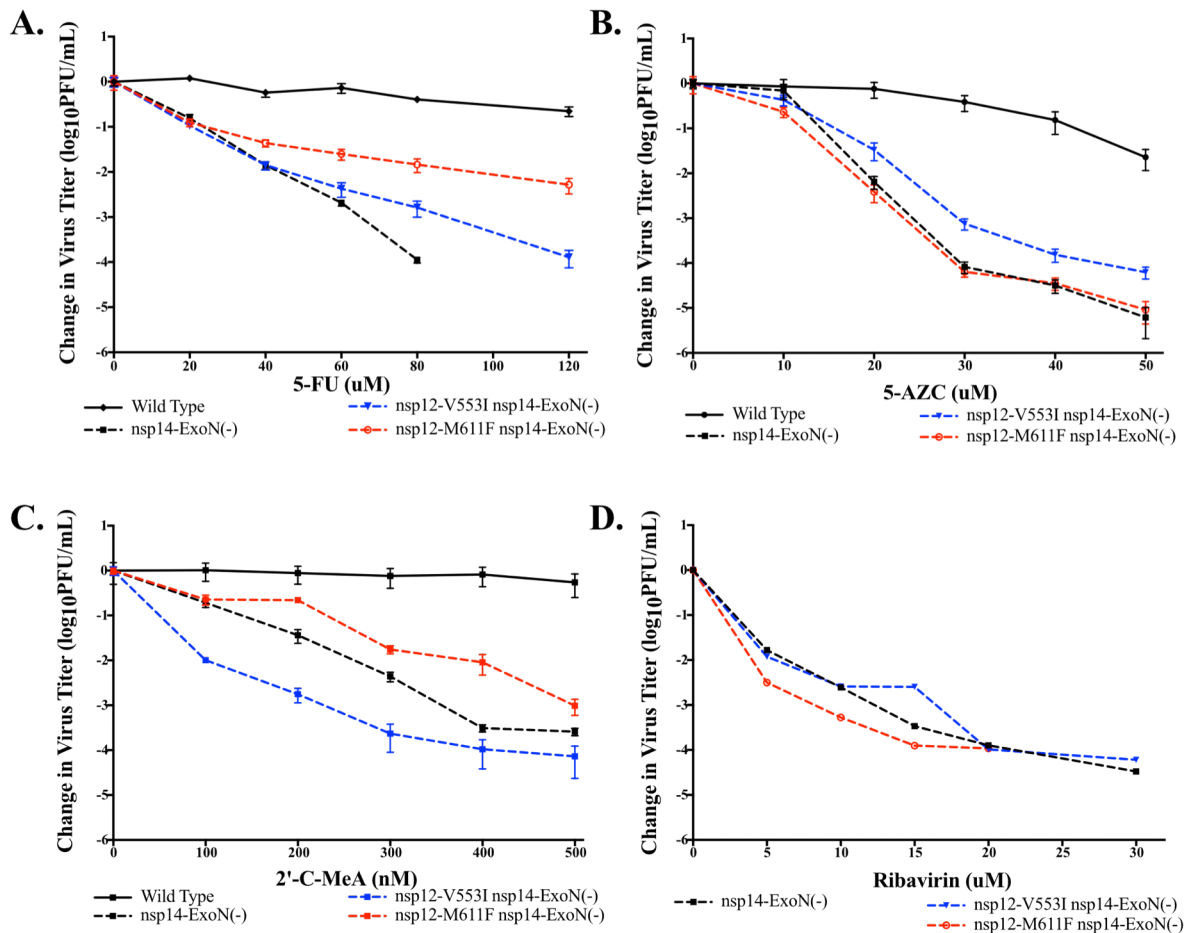
### **Coronaviruses are not sensitive to guanidine hydrochloride at concentrations that inhibit picornaviruses**

Since CoV proteins show homology with picornavirus proteins such as the RdRp, and 3C protease, I next sought to determine if CoVs were sensitive to guanidine hydrochloride (GUA HCl) similar to picornaviruses, and if I could utilize this sensitivity to develop a fidelity assay similar to that used in picornaviral systems. In picornaviruses, GUA HCl specifically inhibits the 2C protein and mutations in 2C that result in escape from GUA HCl inhibition have been characterized. To test for inhibition in CoVs, cultures were inoculated with 200 plaque forming units (PFU) of either WT or nsp14-ExoN(-) viruses. Next, cells were overlaid with agar containing up to 2mM GUA HCl. For all conditions tested approximately 200 PFU were observed, and GUA HCl-treated WT virus plaques appeared to be homogeneous and of a similar size to the untreated ones. Similarly, plaques formed by nsp14-ExoN(-) viruses treated with GUA HCl were similar to plaques formed in its absence (Figure 27). Together this shows that MHV is not sensitive to GUA HCl at the concentrations tested, and treatment with GUA HCl will not serve as a CoV fidelity assay.

### **Developing an rgRNA/FnCas9 assay for investigating fidelity in CoVs**

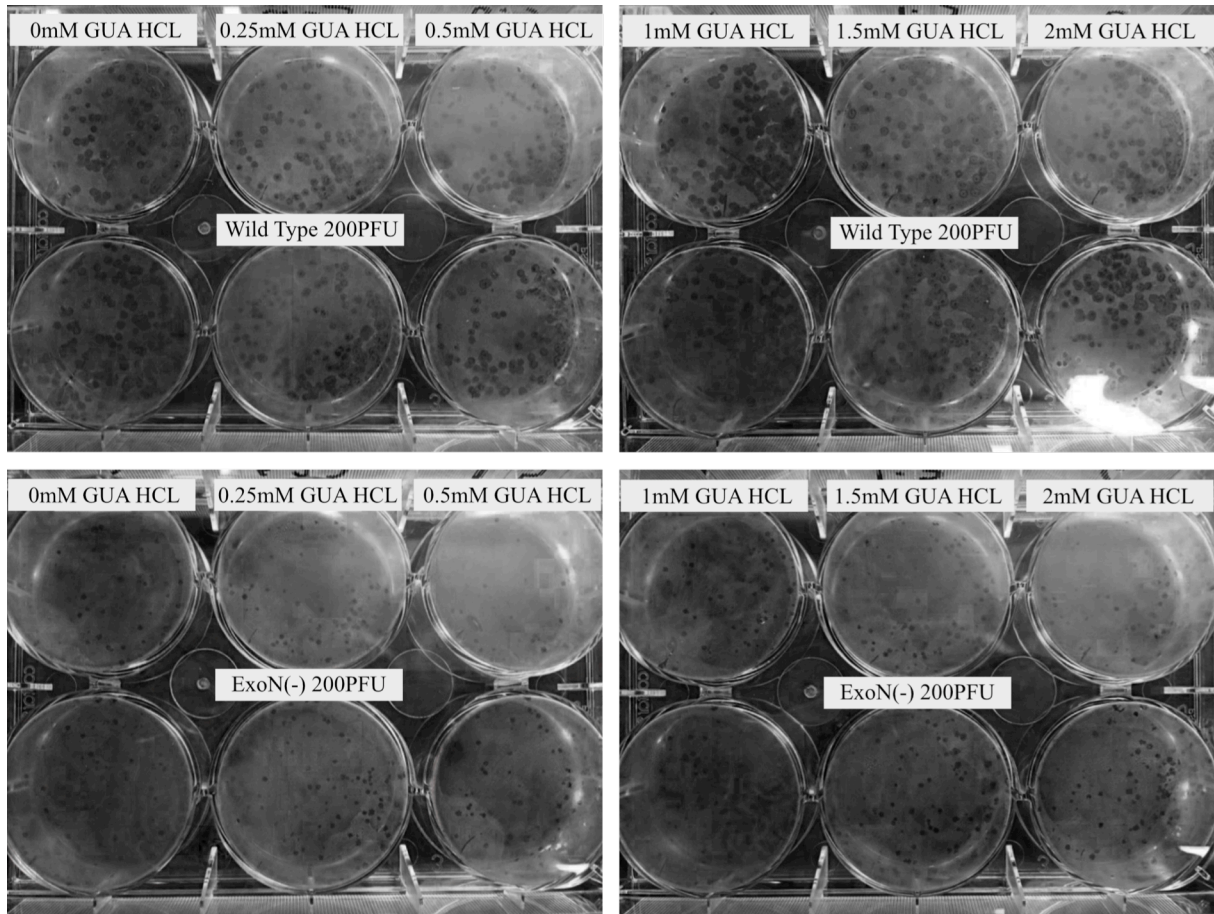
#### **Identification of targets in the MHV genome**

In order to investigate the contributions of CoV fidelity regulating proteins beyond ExoN, and particularly nsp12, a sensitive and rapid assay is required. Therefore, I attempted to assess the ability of RNA-targeted Cas9 endonuclease to inhibit MHV. Cas9 from *Francisella novicida* (FnCas9) was used for this study as it is currently the best characterized RNA-targeted



**Figure 26. Selectivity of mutant viruses in the nsp14-ExoN(-) background for a panel of nucleotide analogs**

Domain location of mutations is indicated by: fingers (blues) and palm (reds). DBT cells were pre-treated with indicated concentrations of 5-FU, 5-AZC, 2'-C-MeA or Ribavirin for 30 min. Treatment was removed and cells were infected with indicated viruses in **A.** the WT background or **B.** nsp14-ExoN(-) background at an MOI of 0.01. Media containing 5-FU, 5-AZC, 2'-C-MeA or Ribavirin was replaced 30 min p.i. Virus samples were taken at 32 h.p.i and titer was determined by plaque assay. Data represents 3 independent experiments, each with 2 replicates (except ribavirin, which is only one experiment with two replicates). Error bars represent SEM.



**Figure 27. Treatment of WT and nsp14-ExoN(-) viruses with GUA-HCl**

DBT-9 cells were infected with 200 PFU per well of WT or nsp14-ExoN(-) viruses in 6-well plates. Overlays were placed with indicated concentrations of GUA-HCl (0mM – 2mM left to right) and allowed to incubate overnight. Plates were fixed and imaged.

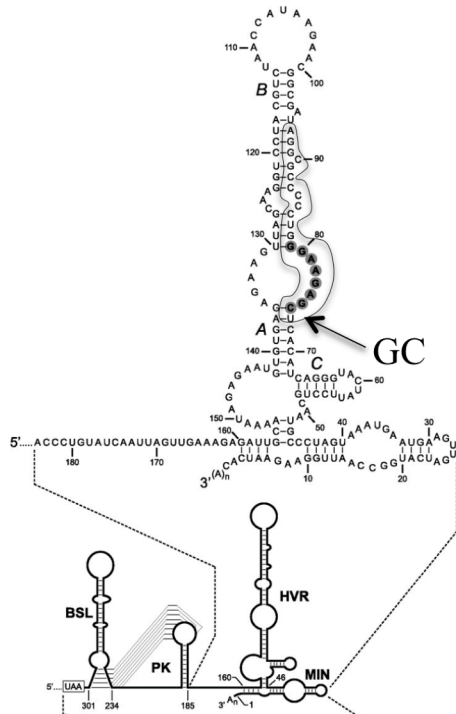


Cas9, and the only Cas9 with a demonstrated ability to specifically require the two 3' most nucleotides for inhibition (Price *et al.*, 2015). The pFnCas9 and pU6rgRNA plasmids used in (Price *et al.*, 2015) were acquired from David Weiss's lab. The acquired pU6rgRNA plasmid targets the E2 protein of murine herpes virus 68 (MHV68) – a DNA virus – and is used in this study as a control rgRNA, which doesn't target the CoV MHV genome. I next identified three structured RNA regions of the MHV genome to target rgRNA/FnCas9 to (Figure 28). I selected 20 base pair regions of the 5'-UTR, the Orf1a/Orf1b -1 ribosomal frame shift and the 3'-UTR. The 5'-UTR site, 5'-AGUUUAUAAACGGCACUCC-3' (TRS-SL4), starts with the last four nucleotides of the transcription-regulating sequence (TRS) and ends in an open loop of stem-loop 4 (SL4) (Kang *et al.*, 2006). The Orf1a/Orf1b -1 ribosomal frame shift site, 5'-CCCGUCUUGUACCCUGUGCC-3' (Frame Shift) begins four nucleotides after the stop codon of Orf1a and continues through the stem-loop of the pseudoknot (Bredenbeek *et al.*, 1990). The final 3'-CC nucleotides are not involved in nucleotide binding to form the pseudoknot but are the final cytosines of a GCC codon, which encodes for an alanine in the CoV-specific domain of nsp12. All changes to the final C retain an alanine, whereas substitutions at the second positions result in a valine, aspartic acid or glycine. This suggests that the virus should tolerate mutations in at least the final cytosine. Finally, the 3'-UTR site, 5'-AGGCGCCCCUGGGAAGAGC-3' (HVR-Oct), exists entirely within the hypervariable region (HVR) of the 3'-UTR and the final GC are located in the conserved octanucleotide (Oct), which makes up an open loop. All mutations are tolerated at all positions of the Oct (Goebel *et al.*, 2007). pU6rgRNA plasmids were engineered to target each of the described regions of the MHV genome.

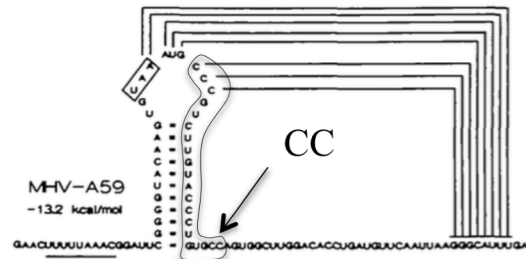
### **HEK293T cells transiently transfected with CEACAM1 are permissive for MHV replication**

Having identified targets, I next needed to develop a system for transient transfection of cells followed by MHV infection. The commonly used cell line for MHV studies is the murine delayed brain tumor (DBT) cell line; however, data in our lab suggest that DBT cells are rarely permissive to infection after transfection (unpublished). Therefore, I chose to use HEK293T cells, well described for efficient transfection. However, HEK293T cells do not express the MHV receptor, murine CEACAM1. I first determined if HEK293T cells transiently transfected with CEACAM1 were permissive for MHV replication. WT or nsp14-ExoN(-) viruses with the firefly luciferase (FFL) gene inserted at the 5' end of nsp2 were used to infect HEK293T cells

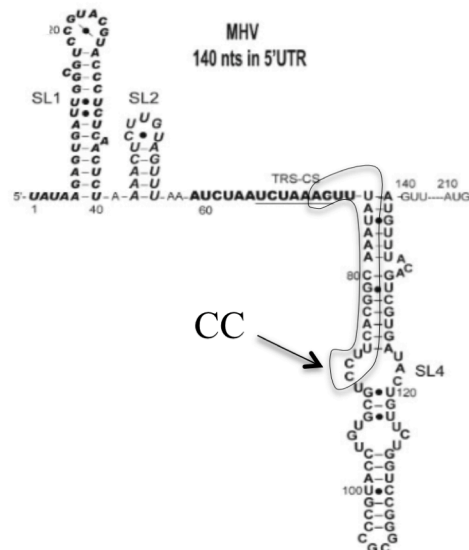
**A** The hypervariable region (HVR) of the 3'-UTR



**B** The Orf1a/Orf1b -1 ribosomal frame shift



**C** The transcriptional regulatory sequence (TRS) of the 5'-UTR



**Figure 28. Structured regions of the MHV genome identified for targeting of FnCas9 by rgRNAs**

Three structured RNA regions of the MHV genome A) the 5'-UTR, B) the Orf1a/Orf1b -1 ribosomal frame shift and C) the 3'-UTR. The 3'-UTR site (HVR-Oct) exists entirely within the hypervariable region (HVR) of the 3'-UTR and the final GC are located in the conserved octanucleotide (Oct) (Goebel *et al.*, 2007). The Orf1a/Orf1b -1 ribosomal frame shift site (Frame Shift) begins four nucleotides after the stop codon of Orf1a and continues through the stem-loop of the pseudoknot (Bredenbeek *et al.*, 1990). The 5'-UTR site (TRS-SL4) starts with the last four nucleotides of the transcription-regulating sequence (TRS) and ends in an open loop of stem-loop 4 (SL4) (Kang *et al.*, 2006).

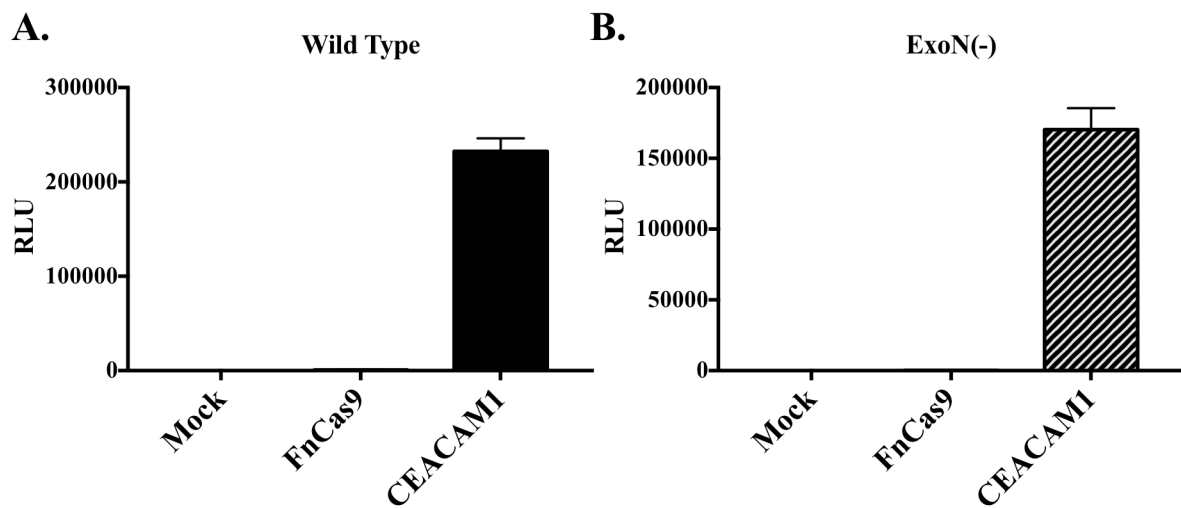
(Freeman *et al.*, 2014), which were mock transfected or transfected with CEACAM1 or FnCas9. No FFL expression was measured for mock transfected or FnCas9 transfected cells, however, cells transfected with CEACAM1 resulted in high FFL expression for both WT and nsp14-ExoN(-) viruses (Figure 29).

### **Translation of MHV is inhibited in the presence of MHV targeted rgRNAs and FnCas9**

Having demonstrated that HEK293T cells support MHV translation, I next wanted to determine if rgRNA/FnCas9 targeted to the MHV genome could inhibit translation of MHV. To test this, I transiently transfected cells with CEACAM1 and either rgRNAs alone or with FnCas9. 24 hours after transfection, I infected cells with WT-FFL or nsp14-ExoN(-)-FFL viruses at an MOI of 0.1 PFU/cell. Cell lysates were measured for the expression of FFL. WT or nsp14-ExoN(-) viral translation was inhibited when FnCas9 was expressed in combination with any of the three MHV targeted rgRNAs (rgRNA<sub>HVR-Oct</sub>, rgRNA<sub>TRS-SL4</sub> and rgRNA<sub>Frame Shift</sub>) (Figure 30). The efficiency of translation of WT viruses, as measured by relative FFL expression, for rgRNA<sub>HVR-Oct</sub> and rgRNA<sub>TRS-SL4</sub> co-transfected with FnCas9 (rgRNA<sub>HVR-Oct</sub>/FnCas9 and rgRNA<sub>TRS-SL4</sub>/FnCas9) was 54 and 49% of no-rgRNA/FnCas9, respectively. Whereas, rgRNA<sub>Frame Shift</sub>/FnCas9 resulted in almost complete abolishment of FFL expression from WT viruses (Figure 29b). Nsp14-ExoN(-) viruses were similarly inhibited (Figure 30d). Surprisingly, rgRNA<sub>MHV68 E2</sub> increased translation compared with both the no-rgRNA and rgRNA<sub>HVR-Oct</sub> controls, in the absence of FnCas9. However, it didn't increase translation over no-rgRNA/FnCas9. Similarly, cells transfected with FnCas9 resulted in increased translation compared to CEACAM1 alone (Figure 30a, c). Together, these data suggest that rgRNA targeting of FnCas9 to the MHV genome can specifically inhibit translation. However, it also suggests that FnCas9 alone and rgRNA<sub>MHV68 E2</sub> enhance translation of MHV proteins by an unknown mechanism.

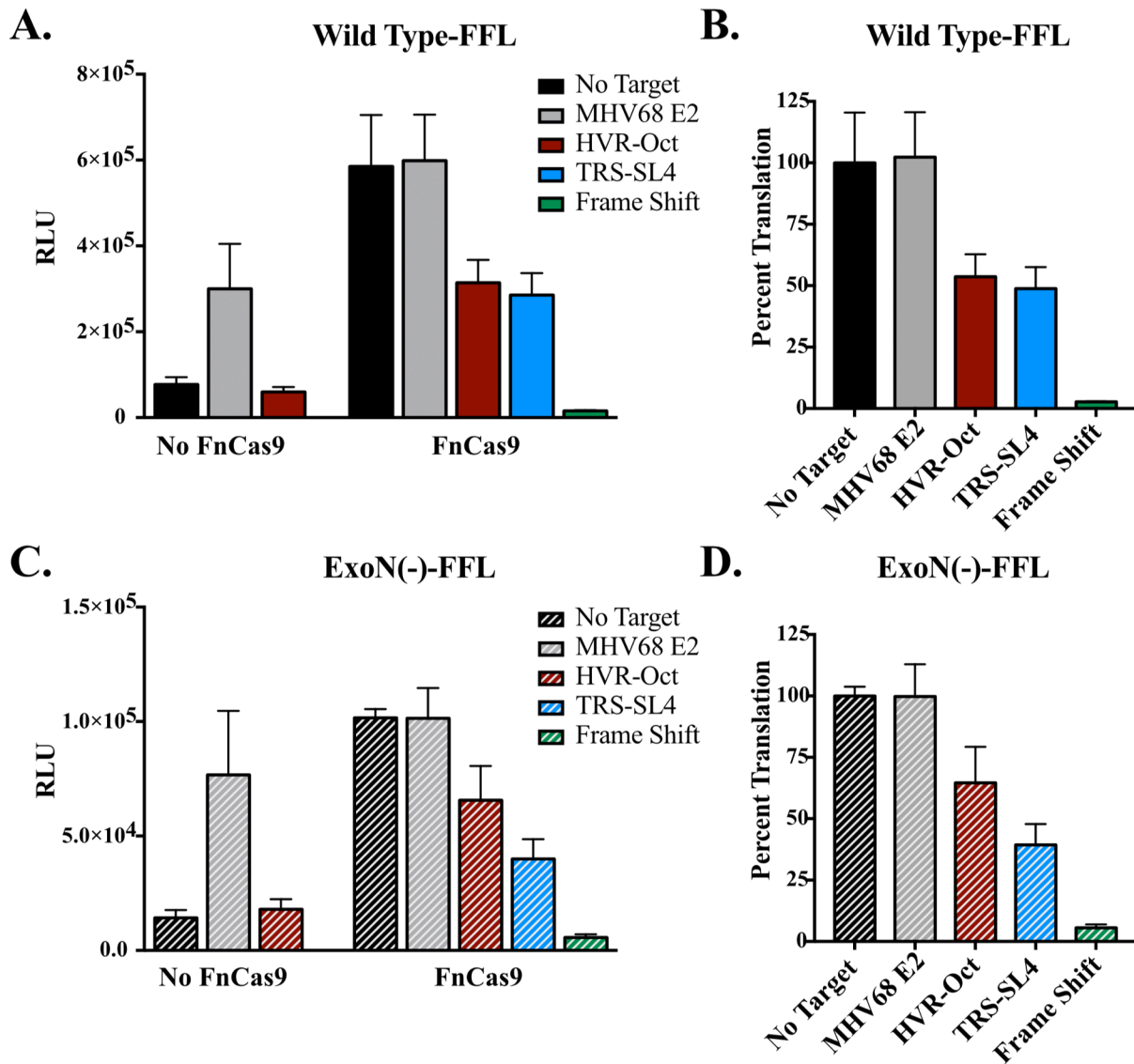
### **MHV viral replication is inhibited in the presence of MHV frame shift targeted rgRNAs and FnCas9**

Having shown that translation can be inhibited by rgRNA/FnCas9 targeted to the MHV genome, I next wanted to investigate whether viral replication was also inhibited. Similar to the previously described experiments, cells were transiently transfected with CEACAM1 and rgRNAs with or without FnCas9. 24-hours later cells were infected with WT-FFL or nsp14-ExoN(-)-FFL viruses and supernatants were titered. HEK293Ts form weak attachment to plates



**Figure 29. MHV viruses are able to infect HEK293T cells when CEACAM1 is transiently expressed**

HEK293T cells were transfected with the MHV receptor: CEACAM1, FnCas9, or mock transfected. 24 hours after transfection cells were infected with WT or ExoN(-) MHV viruses encoding firefly luciferase (FFL) at the end of nsp2 at an MOI of 1. Infection was allowed to proceed for 16 hours at which time monolayers were lysed and assessed for expression of luciferase. Data represents 3 replicates. Error bars represent SEM.



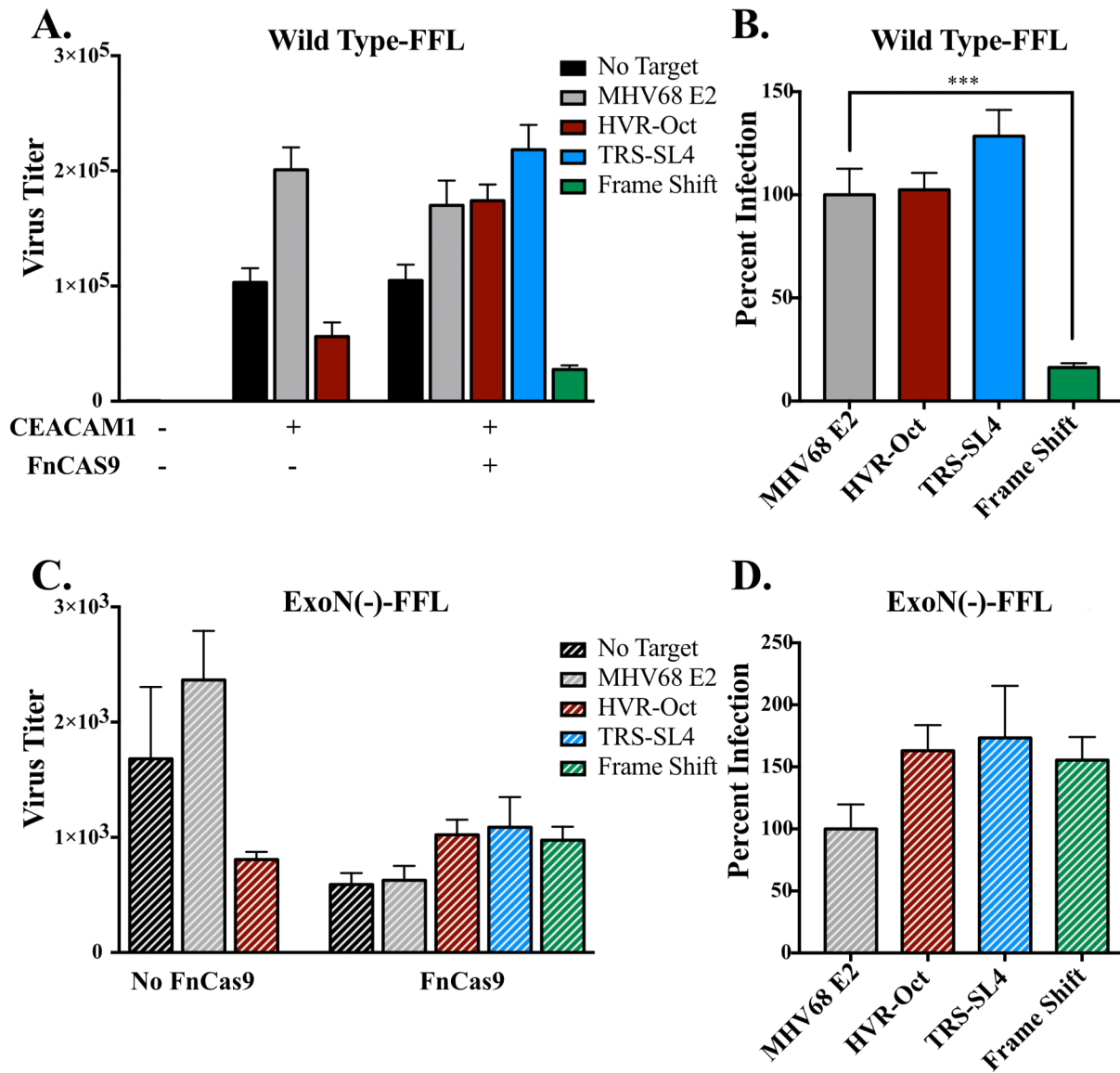
**Figure 30. Translation of WT and ExoN(-) MHV-FFL viruses is reduced when rgRNA targets structured regions and FnCas9 is expressed**

HEK293T cells were mock transfected or transiently transfected with the MHV receptor: CEACAM1, FnCas9, and/or rgRNAs targeted to either the hyper-variable region (HVR) of the 3'-UTR, the translational regulatory sequence (TRS) of the 5'-UTR, or the -1 ribosomal frame shift at the Orf1a/Orf1b junction in the nsp12-RdRp coding region at equal molar concentrations of each plasmid. 24 hours after transfection cells were infected with WT or ExoN(-) MHV viruses encoding firefly luciferase (FFL) at the end of nsp2 at an MOI of 0.1 PFU/cell. Infection was allowed to proceed for 24 hours at which time monolayers were lysed and assessed for expression of luciferase. Data represents three independent experiments, with three replicates. Error bars represent SEM.

and so were not washed after virus infections but WT viruses resulted in no detectable virus when CEACAM1 was not present in HEK293Ts cells demonstrating that infectious virus was not present in the supernatant (Figure 31a). In contrast to the translation data, titers were not lower in rgRNA<sub>HVR-Oct</sub>/FnCas9 or rgRNA<sub>TRS-SL4</sub>/FnCas9 samples (Figure 31a and b). However, rgRNA<sub>Frame Shift</sub>/FnCas9 resulted in decreased titers similar to the decreased in translation (Figure 31b). Nsp14-ExoN(-) virus titers did not decrease in cells expressing rgRNA<sub>HVR-Oct</sub>/FnCas9, rgRNA<sub>TRS-SL4</sub>/FnCas9, or rgRNA<sub>Frame Shift</sub>/FnCas9 (Figure 31c and d). However, nsp14-ExoN(-) viruses were not assessed for viruses persisting in the supernatant; so, it is possible that the measured viruses were those remaining from initial inoculation. Interestingly, FnCas9 did not increase WT titers; yet, rgRNA<sub>MHV68 E2</sub> did (Figure 31a). Finally, although introduction of rgRNA<sub>HVR-Oct</sub>/FnCas9 didn't result in decreased WT titers, rgRNA<sub>HVR-Oct</sub> alone did reduce titers. RgRNA<sub>HVR-Oct</sub> reduced WT titers by 50% compared to rgRNA and close to 75% compared to rgRNA<sub>MHV68 E2</sub> (Figure 31a and c). Importantly, rgRNA<sub>Frame Shift</sub>/FnCas9 remains a promising target to continue studies.

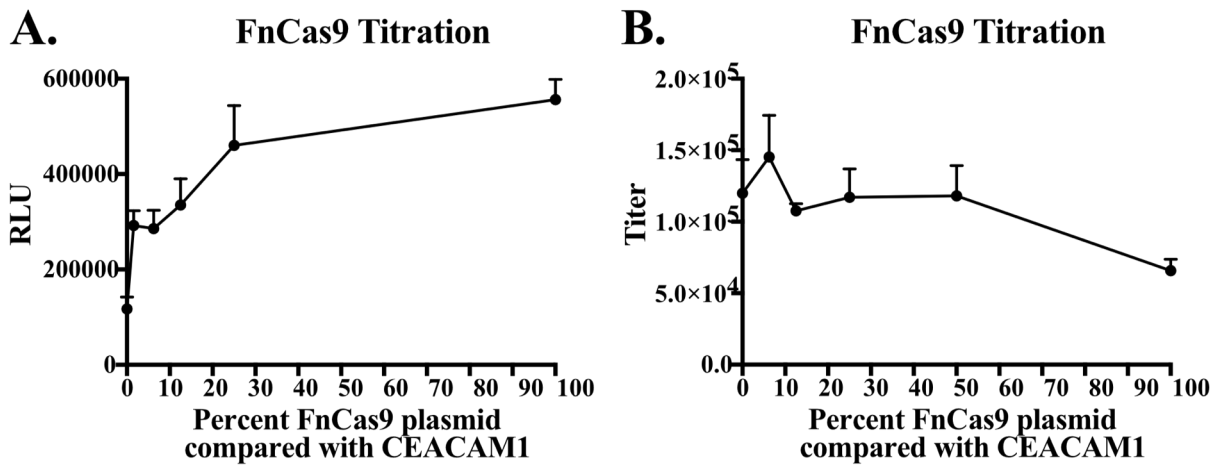
### **FnCas9 transfection increases MHV translation in a concentration dependent manner, and MHV targeted rgRNAs decrease translation without FnCas9**

The discovery that FnCas9 increases translation was unexpected; therefore, I next sought to determine if this was a concentration dependent increase in translation. Serial dilutions of the FnCas9 plasmid were performed starting with equimolar concentrations to that of the CEACAM1 plasmid. Transfected HEK293T cells were infected with WT-FFL viruses and samples were then assessed for expression of FFL in cell lysates and viral titer in supernatants. Confirming the data for the single concentration experiment, expression of FFL increased with increasing concentration of plasmid, yet titer doesn't significantly change with increasing concentration of FnCas9 plasmid (Figure 32). Next, I sought to determine if rgRNA<sub>HVR-Oct</sub> alone truly resulted in decreased FFL expression and viral titers compared with rgRNA controls (Figure 30a and 31a). Therefore, I infected HEK293T cells with WT-FFL after transfected with either rgRNA<sub>MHV68 E2</sub>, rgRNA<sub>HVR-Oct</sub>, rgRNA<sub>TRS-SL4</sub> or rgRNA<sub>Frame Shift</sub> plasmids. Cell lysates were collected and FFL expression measured. Expression of FFL decreased by a similar proportion to that observed when FnCas9 was present (Figure 30a and 33). This suggests that restriction of MHV translation may be dependent solely on the introduction of MHV targeted rgRNA constructs. However, in contrast to previous data the no target control resulted in higher



**Figure 31. Viral propagation of WT MHV-FFL viruses is reduced when rgRNA targets to the frame shift and FnCas9 is expressed**

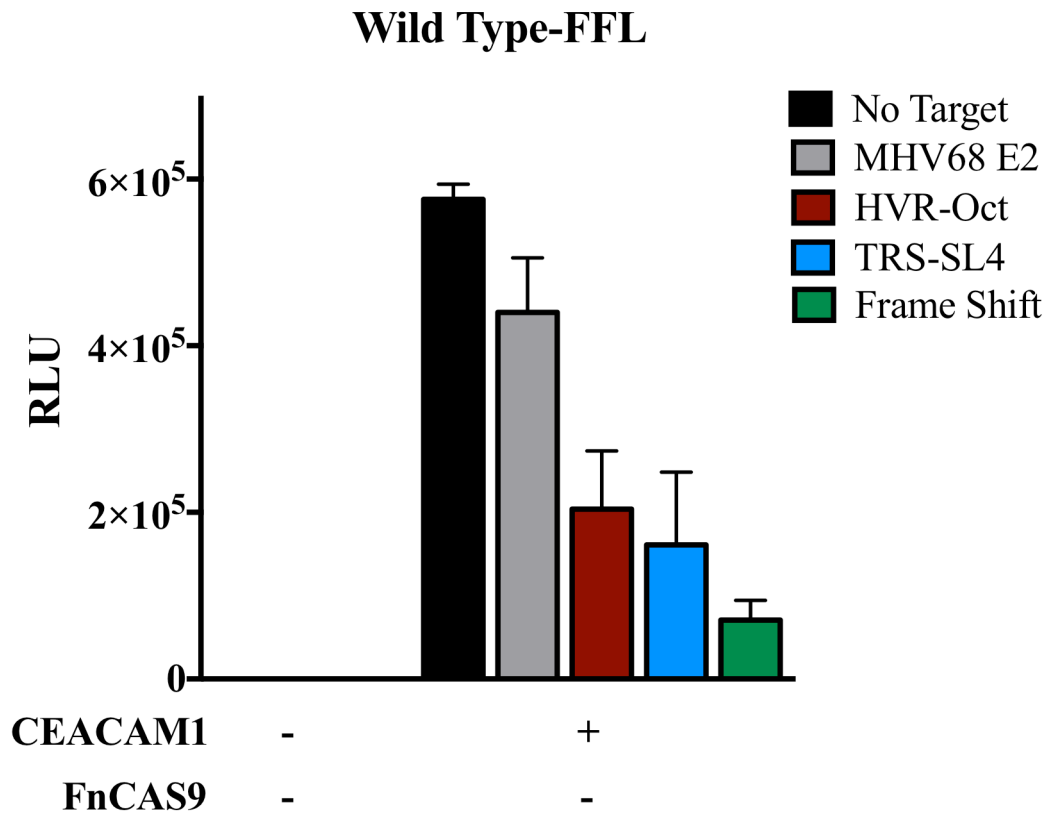
HEK293T cells were mock transfected or transiently transfected with the MHV receptor: CEACAM1, FnCas9, and/or rgRNAs targeted to either the hyper-variable region (HVR) of the 3'-UTR, the translational regulatory sequence (TRS) of the 5'-UTR, or the -1 ribosomal frame shift (Frame Shift) at the Orf1a/Orf1b junction in the nsp12-RdRp coding region at equal molar concentrations of each plasmid. 24 hours after transfection cells were infected with WT or ExoN(-) MHV viruses encoding firefly luciferase (FFL) at the end of nsp2 at an MOI of 0.1 PFU/cell. Infection was allowed to proceed for 24 hours at which time supernatants were collected and tittered by plaque assay. Data represents three independent experiments, with three replicates. Error bars represent SEM.



**Figure 32. WT MHV-FFL translation increases with increasing transfection of FnCas9 plasmid, however this does not result in increased titer**

HEK293T cells were transiently transfected with the MHV receptor: CEACAM1 and FnCas9. The FnCas9 plasmid was serially diluted 2-fold from equimolar with CEACAM. 24 hours after transfection cells were infected with WT MHV viruses encoding firefly luciferase (FFL) at the end of nsp2 at an MOI of 0.1. Infection was allowed to proceed for 24 hours at which time supernatants were collected and tittered by plaque assay and monolayers were lysed and assessed for expression of luciferase. Data represents 3 independent experiments, with three replicates. Error bars represent SEM.





**Figure 33. WT MHV-FFL translation is inhibited by rgRNAs targeted to structured regions of the genome even in the absence of FnCas9**

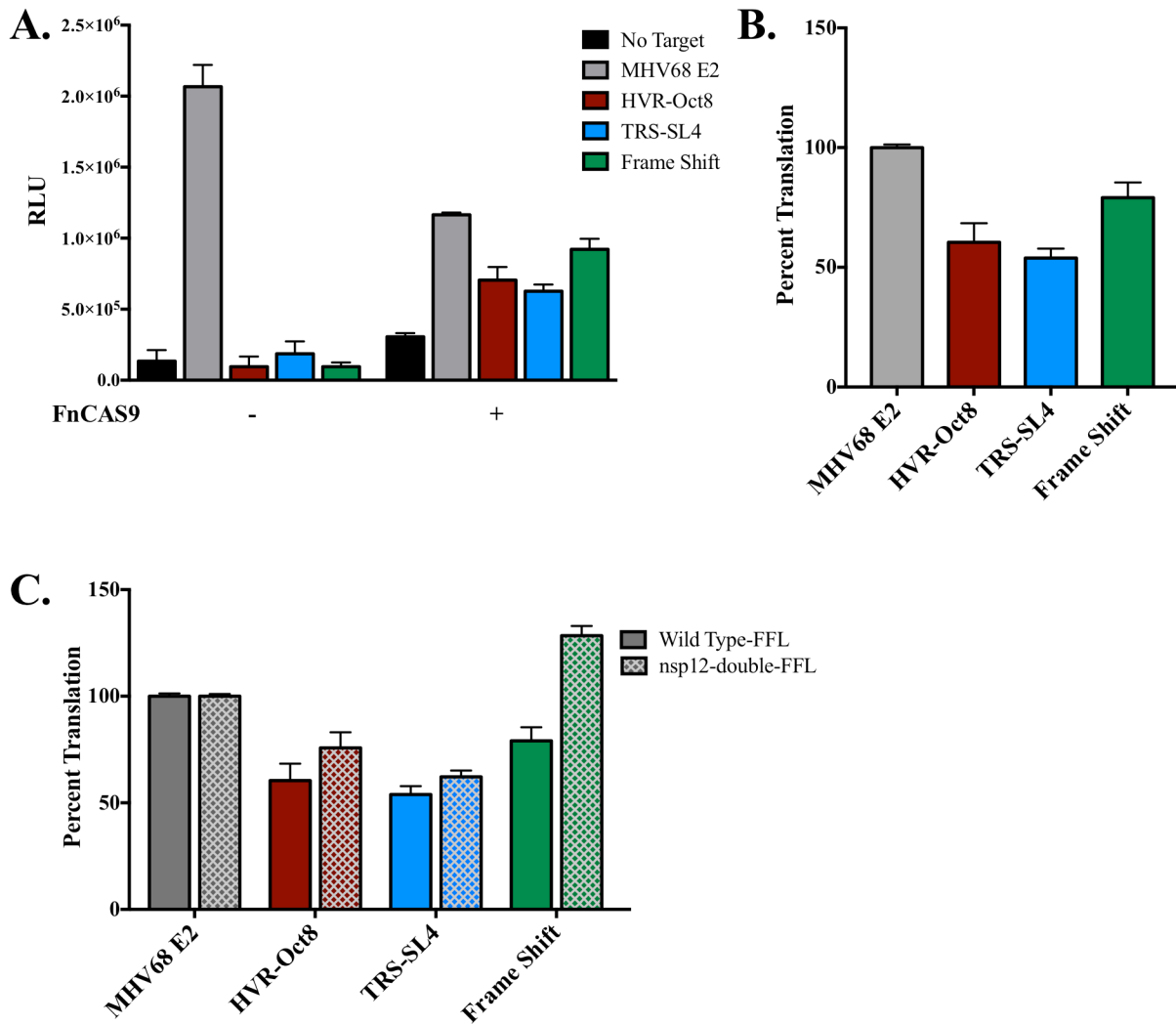
HEK293T cells were mock transfected or transiently transfected with the MHV receptor: CEACAM1 and rgRNAs targeted to the HVR, TRS or the frame shift at equal molar concentrations of each plasmid. 24 hours after transfection cells were infected with WT MHV viruses encoding firefly luciferase (FFL) at the end of nsp2 at an MOI of 0.1 PFU/cell. Infection was allowed to proceed for 24 hours at which time monolayers were lysed and assessed for expression of luciferase. Data represents three replicates. Error bars represent SEM.

expression of FFL than  $\text{rgRNA}_{\text{MHV68 E2}}$ . Additionally, this experiment is only preliminary, and it is therefore uncertain whether inhibition of translation results in inhibition of viral replication.

### **Changes to WT MHV translation resulting from the presence of rgRNAs and FnCas9 in DBT cells are similar to those in HEK293T cells**

Having identified an unexpected increase in WT MHV translation when HEK293T cells transfected with FnCas9 or the control rgRNA,  $\text{rgRNA}_{\text{MHV68 E2}}$ , were infected with MHV, I next wanted to determine if this outcome was cell type dependent. I attempted the same translation experiment described previously for HEK293T cells in DBT cells. In DBT cells transfected with  $\text{rgRNA}_{\text{HVR-Oct}}/\text{FnCas9}$ ,  $\text{rgRNA}_{\text{TRS-SL4}}/\text{FnCas9}$ , and  $\text{rgRNA}_{\text{Frame Shift}}/\text{FnCas9}$  WT translation decreased, resulting in decreased FFL expression (Figure 34). Similar to data in HEK293Ts,  $\text{rgRNA}_{\text{HVR-Oct}}/\text{FnCas9}$  or  $\text{rgRNA}_{\text{TRS-SL4}}/\text{FnCas9}$  resulted in a 60 and 54% decrease in FFL expression, respectively, when compared to  $\text{rgRNA}_{\text{MHV68 E2}}/\text{FnCas9}$ . However, in contrast to the 97% decrease in translation observed in infections performed in HEK293T cells transfected with  $\text{rgRNA}_{\text{Frame Shift}}/\text{FnCas9}$ , in DBT cells  $\text{rgRNA}_{\text{Frame Shift}}/\text{FnCas9}$  resulted in a 21% decrease in WT MHV translation (Figure 34b). Additionally, in DBT cells, there was no apparent effect on translation conferred by  $\text{rgRNA}_{\text{HVR-Oct}}$ ,  $\text{rgRNA}_{\text{TRS-SL4}}$  or  $\text{rgRNA}_{\text{Frame Shift}}$  when compared to no rgRNA. Finally, an increase in FFL expression was observed in DBT cells transfected with  $\text{rgRNA}_{\text{MHV68 E2}}$ , with or without FnCas9, similar to what was observed in HEK293T cells (Figure 34a). Together, these data demonstrate that  $\text{rgRNA}_{\text{MHV68 E2}}$  enhances CoV translation and suggest that rgRNAs targeted to CoV genomes result in restriction of translation when in the presence of FnCas9. However, translation was increased in all cases where an rgRNA and FnCas9 were transfected in cells together over that of no rgRNA. These results are preliminary and require repeating.

I next sought to determine if differences between nsp12-RdRp variant viruses could be measured using transiently transfected FnCas9 and rgRNAs. Therefore, I next tested for changes in FFL expression between WT MHV and an MHV virus with two mutations in the nsp12-RdRp. Erica Andres, a research assistant in the Denison lab, identified two mutations during passage in the presence of GS5734, an uncharacterized nucleoside analog from Gilead Sciences, which resulted in a dramatically less sensitive virus: nsp12-V553L and nsp12-F476L. A variant virus containing these two mutations and the FFL protein at the beginning of nsp2 was engineered and



**Figure 34. Translation in DBT cells and comparison between WT and an nsp12-RdRp-double mutant**

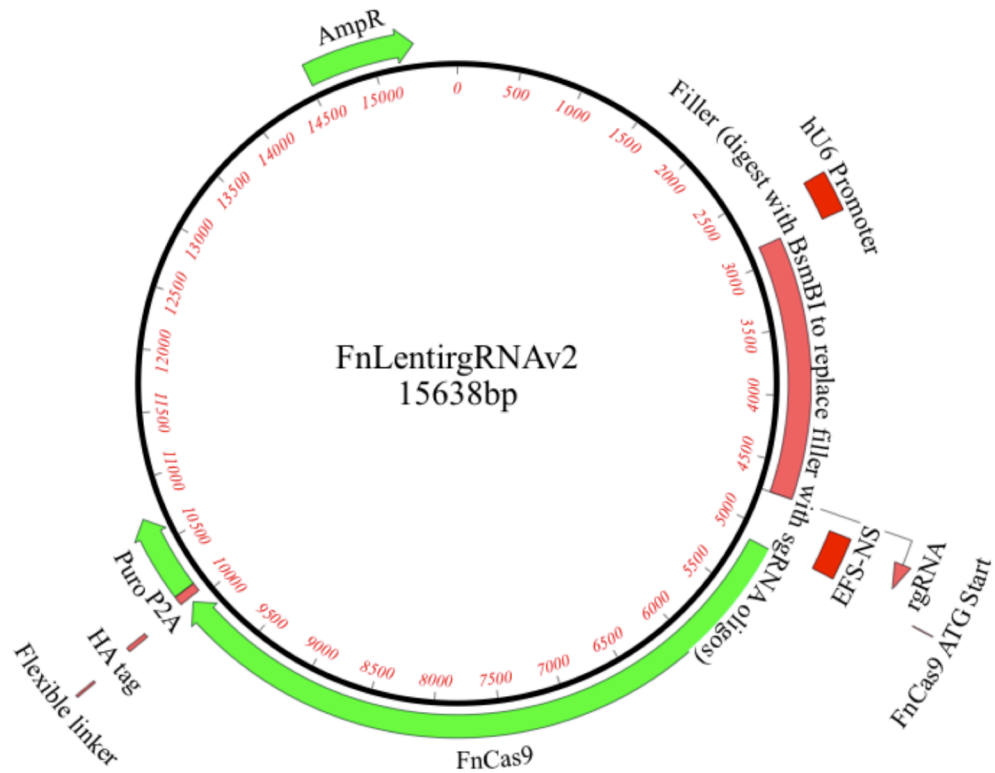
DBT-9 cells were mock transfected or transiently transfected with FnCas9, and/or rgRNAs targeted to either a control rgRNA, the hypervariable region (HVR) of the 3'-UTR, the translational regulatory sequence (TRS) of the 5'-UTR, or the -1 ribosomal frame shift at the Orf1a/Orf1b junction in the nsp12-RdRp coding region at equal molar concentrations of each plasmid. 24 hours after transfection cells were infected with WT or nsp12-RdRp-double mutant MHV viruses encoding firefly luciferase (FFL) at the end of nsp2 at an MOI of 0.1. Infection was allowed to proceed for 24 hours at which time monolayers were lysed and assessed for expression of luciferase. Data represent three replicates. Error bars represent SEM.

recovered (nsp12-double-FFL). It is likely that this virus is altered in fidelity, although this has not yet been determined. DBT cells transfected with rgRNA<sub>MHV68 E2</sub>/FnCas9, rgRNA<sub>HVR-Oct</sub>/FnCas9, rgRNA<sub>TRS-SL4</sub>/FnCas9, or rgRNA<sub>Frame Shift</sub>/FnCas9 were infected with WT-FFL or nsp12-double-FFL viruses. When compared to WT, the nsp12-double-FFL virus resulted in less inhibition of translation in cells expressing rgRNA<sub>HVR-Oct</sub>/FnCas9, rgRNA<sub>TRS-SL4</sub>/FnCas9, or rgRNA<sub>Frame Shift</sub>/FnCas9 (Figure 34c). There was no difference in FFL expression between WT and nsp12-double-FFL viruses in cells transfected with rgRNA<sub>MHV68 E2</sub>/FnCas9 (data not shown). In total these data support the feasibility of an rgRNA/FnCas9-based system to investigate changes in fidelity.

### **Generation of a dual expression plasmid for stable expression of FnCas9 and rgRNAs in cell lines**

Having demonstrated that FnCas9 targeted to the MHV genome by rgRNAs results in decreased translation and that rgRNA<sub>Frame Shift</sub>/FnCas9 results in decreased viral titers, I moved forward with development of a fidelity assay using the rgRNA/FnCas9 system. In order to measure small changes in fidelity FnCas9 and rgRNAs will need to be present in every cell available for infection. Therefore, I next sought to generate a stable cell line expressing both rgRNA and FnCas9 constructs. As screening with CRISPR/Cas9 is common, plasmids are commercially available for dual expression. However, targeting of RNA with Cas9 was only recently discovered, thus reagents are limited. In order to generate stable cell lines with rgRNA/FnCas9, I started with a plasmid for lentiviral transduction of CRISPR/Cas9 (Sanjana et al., 2014). I replaced the Cas9 present in the original plasmid (and removed the nuclear localization signal) with the FnCas9 from the transient transfection plasmid described, including the flexible linker and HA tag (Price *et al.*, 2015). Similarly, I need to replace the DNA-targeting chimeric RNA tail present in the commercially available dual expression plasmid with the rgRNA tail from the pU6rgRNA plasmid used for transient transfection (Figure 35). CoV genome targets will be rapidly introduced using annealed primers with over hanging ends that match those of the vector after digestion with BsmBI. The plasmid includes a puromycin resistance gene that will be expressed as one protein with FnCas9 separated by P2A, a self-cleaving peptide (Sanjana et al., 2014). Stable cell lines will be selected for resistance to puromycin.

DBT cells will be used for the generation of cell lines expression rgRNA/FnCas9 as they are naturally permissive for MHV. Therefore, I next sought to determine the lethal dose of



**Figure 35. Plasmid for the generation of stable cell lines**

Plasmid to be used for the generation of stable cell lines containing Fncas9 with an HA tag adjacent to the self-cleaving peptide P2A, adjacent to the puromycin resistance gene. Additionally, the rgRNA tail is adjacent to a BsmBI cut site where targeting primers with matching overhangs can be inserted after digestion to generate full-length rgRNAs.

puromycin for DBT cells. In duplicate, DBT cells were grown in 0, 0.25, 0.5, 1, 2, 3, 4, 5, 6, 7, 8, or 9ug/mL puromycin for 7 days. Media containing puromycin was replaced every second day. On the 7<sup>th</sup> day media without puromycin was provided and cells were allowed to replicate for two more days. Toxicity resulting from the presence of puromycin was observed on day one post treatment at concentrations as low as 1ug/mL. On day seven, cell numbers were reduced at all concentrations, and 1ug/mL was the lowest concentration that appeared to have no viable cells present. To confirm that all cells were non-viable at 1ug/mL cells were observed at day 9. At 1ug/mL no viable cells were present, whereas in the wells previously treated with 0.25 or 0.5ug/mL puromycin cells had begun to re-populate. These data suggest that 1ug/mL puromycin is an appropriate concentration for selection of cell transduced with the FnLentirgRNAv2 plasmid once generated.

### **Discussion**

CoVs provide a unique RNA virus system to investigate fidelity because they encode a multi-protein replication-transcription complex (RTC) where many of the involved proteins may influence fidelity. Among RNA viruses, only the large Nidoviruses are known to encode a proofreading enzyme. Understanding the relationship between the proteins of the RTC, and identifying fidelity determinants throughout it, may provide future targets for attenuation of current and future HCoV. To date, rapid and reliable measures of fidelity are not available for CoV systems. Some assays used for other viruses may not work because of the size of CoV genomes and potential differential evolution across the large genome (this may be the case for TOPO cloning). Others may not work well for any virus system, such as sensitivity to nucleoside analogs for RdRp variant viruses. Finally, in the case of the GUA-HCl sensitivity assay used for picornaviruses, CoVs may simply not share any proteins similar enough for inhibition to occur. In this chapter, I present work towards the development of a potential assay for the assessment of CoV fidelity. I demonstrate that CoV replication can be restricted by rgRNA/FnCas9 targeted to structured regions of the genome and that small difference in the magnitude of restriction may be observable between nsp12 variant viruses. However, much of this data is preliminary and will need to be further investigated.

The results of the combined CEACAM1/rgRNA<sub>Frameshift</sub>/FnCas9 transient transfection experiments in HEK293T cells result in reduced translation and titers, suggesting feasibility for a

plaque reduction assays in stably expression cells (Figures 30 and 31). However, a few unexpected results occurred with experimental controls. One such result was the observation that both translation and viral propagation increased when cells were transfected with rgRNA<sub>MHV68 E2</sub>. One way this result might occur is if the viral target matched that of RNA native to the cell. Increased translation was observed in both HEK293T and DBT cells. It would be interesting to determine if either or both cell types contain an RNA sequence with high similarity to the viral target DNA. Another way increased CoV translation could occur is if nonspecific rgRNAs somehow divert the cellular immune response toward the rgRNAs and away from CoV replication. More experimentation is needed to address this finding. Similarly, transfection of FnCas9 resulted in increased translation in HEK293T cells in a plasmid concentration-dependent fashion. However, in this case increased translation didn't correlate with increased titers. The introduced FnCas9 does not include a nuclear localization signal, so it is unlikely that it is targeting cellular DNA. It is possible that FnCas9 is able to bind to cellular RNA or proteins (Walsh et al., 2013). Transcription of MHV viral genomes and translation are mutually exclusive processes (Andino *et al.*, 1999). So, one could imagine that FnCas9 may stabilize translational complexes, thus preventing transcription and decreasing titers. However, how binding a cellular RNA or protein would result in increased translation of viral genomes without altering titers is more difficult to explain. It would be interesting to pursue what effects FnCas9 has on cells when introduced in the absence of rgRNAs with or without virus and if any cellular protein or RNA could be pulled down.

The surprising results of transfection controls observed should be considered when moving forward with the development of an FnCas9/rgRNA based assay for the assessment of fidelity in CoV systems. However, the data supporting CoV-specific inhibition is promising. Further study is warranted to determine if escape occurs similarly to HCV, by changing individual nucleotides at the 3' end of the viral target. Additionally, stable cells lines need to be developed and assessed for stable expression of constructs and quantitative plaque reduction between cells containing FnCas9/rgRNA targeted to CoVs or a non-CoV target. Even in the event that restriction is not complete with this method, if viruses with intact targets result in small plaques, it will still be possible to quantify escape mutations. If successful, a fidelity assay based on FnCas9/rgRNA targeting of CoV genomes will be particularly beneficial as it could be edited to target any RNA virus, thus potentially providing a mechanism to compare changes in fidelity across viruses.

Additionally, if the mechanism of escape is consistent across viruses (specifically between HCV and CoVs) then it will provide a mechanism to quantitatively determine fidelity for specific natural nucleotides during viral replication (Price *et al.*, 2015). CoVs provide a unique system to study the influence of altered fidelity on viral biology, yet established methods for investigating viral fidelity have not been effective. This FnCas9/rgRNA assay may provide the opportunity to effectively probe into CoV determinants of fidelity across the proteins of the RTC.



## CHAPTER V

### MATERIALS AND METHODS

#### Virus and cell culture

Murine delayed brain tumor (DBT) cells (Chen and Baric, 1996), baby hamster kidney 21 cells expressing the murine hepatitis virus (MHV) receptor (BHK-R) (Yount *et al.*, 2002) and HEK-293T cells were maintained at 37°C in Dulbecco's modified Eagle medium (DMEM; Gibco) supplemented with 10% FBS (Invitrogen), penicillin, streptomycin (Gibco) and amphotericin B (Corning). BHK-R cells were further supplemented with 0.8mg/mL of G418 (Mediatech). All virus work was performed using recombinant WT-MHV strain MHV-A59 (GenBank accession number AY910861 (Yount *et al.*, 2002)).

#### Sequence analysis and homology modeling of CoV MHV nsp12-RdRp and nsp14

The MHV RdRp domain structure was generated using Phyre2 online program (Kelley *et al.*, 2015) using nsp12 residues 385-887 that correspond to the reported SARS-CoV nsp12-RdRp model (Xu *et al.*, 2003). Briefly, the Phyre2 online available bioinformatics software uses multiple alignments to identify publicly available solved structural domains similar to . First secondary structures are identified (alpha helices and beta sheets). These are then used to create a likely structural backbone, next loops are added based off loop structures available and finally, side chains are incorporated into the modeled structure. The structural model was compared to the X-ray crystal structures of coxsackievirus B3 (CVB3) and poliovirus (PDB accession numbers 3DDK and 1RA7 respectively) using the Pymol Molecular Graphics System (Schrödinger, LLC). The MHV nsp14 structure was also generated using Phyre2; the whole sequence was used. ClustalX multiple-sequence alignments were generated using the program MacVector.

#### Cloning, recovery and verification of mutant viruses

Quick-change mutagenesis was used to generate point mutations in individual MHV genome cDNA fragment plasmids using the previously described MHV infectious clone reverse genetics system (Yount *et al.*, 2002). Mutant viruses were recovered in co-cultured BHK-R and DBT cells (or in BHK-R cells alone) following electroporation of *in vitro* transcribed genome RNA in

BHK-R cells. All viruses that included nsp14-ExoN(-) mutations were generated using the F fragment previously described (Eckerle *et al.*, 2007). Before use in viral recovery all mutagenized plasmids were fully sequenced (GenHunter Corporation, Nashville, TN) to assure no additional mutations were introduced. We also sequence verified engineered mutations in recovered viruses. Viruses in the nsp14-ExoN(-) background took between 84 and 96 hours to reach around 80 percent involvement in syncytia for a P0 stock in contrast to viruses in the WT background, which were frozen at 24-48 h.p.i P1 working stocks were made by infecting D9s at an MOI 0.01 and freezing when 80 percent involved in syncytia, approximately 24 h p.i for WT viruses and 36 h p.i for nsp14-ExoN(-) viruses (2-3 rounds of replication).

### **Compounds and drug sensitivity studies**

5-fluorouracil (5-FU), 5-azacytidine (5-AZC) and Ribavirin were obtained from Sigma. 5-FU and ribavirin were prepared as 200mM stock solution in DMSO. 5-azacytidine (5-AZC) was prepared as a 50mM stocks solution in water. 2'-C-MeA was a kind gift from Gilead Sciences and was prepared as 20mM stocks in DMSO. Sub-confluent DBT cells were pre-treated for 30min with DMEM with the indicated concentrations of 5-FU, ribavirin, 2'-C-MeA or DMSO, 5-AZC or media alone. Treatment was removed, and inocula added and allowed to adsorb for 1hr at 37°C. Inocula was then removed and media with or without drug or DMSO was returned. Infection proceeded for 24hrs for WT or 32hrs for nsp14-ExoN(-) viruses, when supernatants were acquired, frozen and titered by plaque assay as previously described (Eckerle *et al.*, 2007).

### **Virus replication and RNA synthesis assays**

Sub-confluent DBT cell monolayers in triplicate were infected at an MOI of 0.01 PFU/cell. Virus was allowed to adsorb for 30min when inocula were removed and the cells were washed 2X with PBS, followed by addition of pre-warmed media. For replication kinetics assays samples were taken at various time points post infection and an equal volume of media replaced. For replication kinetics in the presence of 5-FU, 5-FU was added 30min prior to time zero and media containing 100 or 200uM 5-FU was restored after PBS washes. Titering was performed by plaque assay as previously described (Eckerle *et al.*, 2007). For analysis of RNA synthesis, total infected-cell RNA was obtained by extraction using TRIzol reagent (Invitrogen) at various times post infection and two-step RT-qPCR was performed as previously described (Smith *et al.*, 2015).

### **Determination of specific infectivity**

Sub-confluent DBT cells were pre-treated for 30min with DMEM with indicated concentrations of 5-FU or DMSO alone. Treatment was removed, and inocula added and allowed to adsorb for 1hr at 37°C. Inocula was then removed and media with drug or DMSO was returned. Infection proceeded for 20hrs for WT or 24hrs for nsp14-ExoN(-) viruses then supernatants were acquired, frozen and titered by plaque assay as previously described (Eckerle *et al.*, 2007). Supernatants were also used for RNA genome isolation by adding 100uL supernatant to 900uL TRIzol reagent (Invitrogen), chloroform extracting by phase separation and using the aqueous layer in the PureLink Mini RNA kit (Ambion) as per the manufacturers protocol. One-step RT-qPCR was performed as below and the ratio of PFU to genomes of the supernatant was determined.

#### **One-step RT-qPCR for determining supernatant genome copies for specific infectivity assay**

An RNA standard was generated using MHV A fragment (Yount *et al.*, 2002) to generate a 931 nucleotide RNA. First cDNA was generated by PCR using the primers: forward 5'-TAATACGACTCACTATAGGGGGCTATGTGGATTGTTGTGG-3' which begins with a T7 promoter and reverse 5'-AATTCTTGACAAGCTCAGGC-3'. RNA for the standard curve was then generated using mMessage mMachine T7 kit (Ambion). An agarose gel with 1% bleach was run and a ~900nt band was observed. RNA was purified using RNeasy Mini kit (Qiagen). Dilutions of the standard curve were made from  $10^3$  to  $10^8$  genome equivalents for use in assay as needed. Primers and probes for one-step RT-qPCR were purchased from BioSearch Tech. Probe is 5'-FAM labeled and 3' BHQ-1 labeled with the sequence 5'-TTCTGACAACGGCTACACCCAACG-3' and made up to 5uM in nuclease free water. The primers used were forward 5'-AGAAGGTTACTGGCAACTG-3' and reverse 5'-TGTCACGGCTAAATCAAAC-3'. Reactions were set-up on ice with enzyme added last. Final volume for reactions was 20uL with 150nM probe, 900nM each primer, 2uL sample RNA and 10uL 2X ToughMix, one-step, low ROX enzyme mix (Quantas) used per reaction. Samples were plated in duplicate and run on the Applied Biosciences 7500 Real-Time PCR System with the conditions 55C for 10min, 95C for 5min, 95C for 30sec and 60C for 1min with the last two steps repeated 40X. The standard curve was graphed and genomes per mL determined.

### **Competitive fitness of mutant viruses**

Sub-confluent DBT monolayers were co-infected at a total MOI of 0.01 PFU/mL with RdRp mutant viruses in the nsp14-ExoN(-) background and nsp14-ExoN(-) at either a 1:1, 1:9 or 9:1 ratio. When 50-70% of the monolayer was involved in syncytia, total RNA was harvested. RNA was then reverse transcribed using SuperScriptIII (Invitrogen) per the manufactures protocol and amplicons were generated using Comp12\_F and Comp12\_R primers covering the region of nsp12 including the codons for both the V553 and M611 residues. Amplicons were sent for sequencing with Comp12\_F and electropherograms were analyzed using MacVector.

### **Passage reversion analysis**

Triplicate monolayers of sub-confluent DBT cells were infected with an initial MOI of 0.01 PFU/mL of nsp12-V553I and nsp12-M611F viruses in both the WT and nsp14-ExoN(-) background. Viruses were then blind passaged in triplicate for 5 passages. Total RNA was sequenced across a 1.7kb region of nsp12-RdRp, using Comp12\_F and Comp12\_R primers, which included both nsp12-RdRp mutations. Electropherograms were analyzed using MacVector.

### **Preparation of amplicons for deep sequencing of full viral genomes**

Sub-confluent DBT cells were infected at an MOI of 0.01 PFU/mL with nsp12-V553I or nsp12-M611F in either the WT or nsp14-ExoN(-) background, nsp14-ExoN(-) alone or WT alone. Infections were allowed to progress for 20 hours then RNA was isolated. RNA was reverse transcribed using SuperScriptIII (Invitrogen) per the manufactures protocol, and 12 amplicons were generated to cover the whole genome and processed as described previously (Smith *et al.*, 2015).

### **Deep sequencing sample preparation and analysis**

Amplicons were subsequently purified via a nucleospin PCR purification kit (Macherey-Nagel), quantified by picogreen, fragmented (Fragmentase) and prepared using the Illumina NextSeq500 Mid Output 150 cycle kits following the standard protocols. Sequences were obtained with an Illumina NextSeq500 machine. Sequencing runs were analyzed using the previously published ViVan bioinformatics pipeline (Isakov *et al.*, 2015). Briefly, the pipeline performs quality filtering and adaptor cleaning was done using fastq-clipper ([http://hannonlab.cshl.edu/fastx\\_toolkit/index.html](http://hannonlab.cshl.edu/fastx_toolkit/index.html)). The 150-nt reads were aligned to the

reference sequence with a maximum 2 mismatches per read, using BWA (Li and Durbin, 2010) and processed using SAMTOOLS (Li *et al.*, 2009) to obtain the nucleotide/base calling at each position. The ViVan pipeline then identifies statistically significant variants above the background noise due to sequencing error, calculated for each nucleotide site: for each position throughout the viral genome, base identity and their quality scores are gathered. Each variant allele's rate is initially modified according to its covering read qualities based on maximum likelihood estimation, and tested for significance using a generalized likelihood-ratio test. Additionally, an allele confidence interval is calculated for each allele. In order to correct for multiple testing, Benjamini-Hochberg false-discovery rate of 5% was set. In all experiments, a minimum coverage of 3000X reads was obtained and the background error frequency at every nucleotide site was always below 0.0001. For analysis we use a conservative frequency cut off of 0.01 consistent with previous studies (Bordería *et al.*, 2015; Pauly and Lauring, 2015; Van Slyke *et al.*, 2015).

#### **Passaging in the presence and absence of 5-FU**

Triplicate monolayers of sub-confluent DBT cells were infected with an initial MOI of 0.01 PFU/cell with either WT or nsp14-ExoN(-) viruses. Infections were performed in media alone or containing sub-lethal concentrations of 5-FU (60uM for nsp14-ExoN(-) and 800uM for WT). Populations were then blind passaged in triplicate for 6 additional passages. Total RNA was taken from cell monolayers and prepared for Sanger sequencing. Sequencing was performed across all of nsp12 and nsp14 as described previously (Smith *et al.*, 2015).

#### **Mutation frequency by TOPO TA cloning**

Triplicate monolayers of sub-confluent DBT cells were infected with an initial MOI of 0.01 PFU/mL of nsp12-M611F viruses in both the WT and nsp14-ExoN(-) background and WT and nps14-ExoN(-) viruses. Viruses were then blind passaged in triplicate for 5 passages. Total RNA was extracted and RT-PCR amplified using the primers sets Freq2\_F and Freq2\_R covering a ~1kb part of nsp2. PCR products were TOPO TA-cloned (Invitrogen), sequenced, and analyzed using MacVector software. Mutation frequency was calculated as the number of mutations identified per virus over the total number of nucleotides sequenced multiplied by  $10^4$ , as described previously (Gnädig *et al.*, 2012).

### **Guanidine hydrochloride sensitivity assay**

Guanidine hydrochloride, GUA-HCl (Sigma) was prepared as a 1M stock solution in water. Monolayers of sub-confluent DBT cells were infected with 200PFU of WT or nsp14-ExoN(-) viruses. Inocula were allowed to incubate for 45min then plaque assay overlays (prepared as previously described (Eckerle *et al.*, 2007)) containing 0, 0.25, 0.5, 1, 1.5 or 2mM GUA HCl were placed. Infection was allowed to continue for 24 hours then plates were fixed as previously described (Eckerle *et al.*, 2007). Images were captured using a 12-megapixel camera (Apple).

### **Transient-transfection followed by infection of HEK293T and DBT cells**

The pcDNA3.3 FnCas9-hu (amp<sup>r</sup>) and control FnrgRNA (kan<sup>r</sup>) plasmids used in this dissertation research for transfection experiments were acquired as kind gifts from David Weiss's lab and were previously described in (Price *et al.*, 2015). The pcineo-mCEACAM1 plasmid was acquired from Brett Case. Finally, the LentiCRISPRv2 plasmid was acquired as a gift from Chris Aiken's lab (Addgene) and is described in (Sanjana *et al.*, 2014). The 20bp control target or the FnrgRNA plasmid was replaced using around-the-horn or exponential-megapriming PCR (Ulrich *et al.*, 2012). MHVrgRNA\_HVROCT8\_F, MHVrgRNA\_TRS\_F, or MHVrgRNA\_Frameshift\_F and MHVrgRNA\_R primers were treated with T4 PNK (NEB), using T4 DNA ligase buffer. PCR was set up with Q5 Hot Start High-Fidelity 2X Master Mix (NEB) was run with the conditions 98°C for 30sec, 98°C for 10sec, 57°C for 30sec and 72°C for 2min with the last three steps repeated 30X then 72°C for 2min and hold at 4°C. PCR product was treated with 1uL DpnI (NEB) to digest template, followed by ligation with T4 ligase (NEB) as per manufacturers protocol. Plasmids were sequenced using U6\_F primer.

HEK293T or DBT cells were plated at  $1 \times 10^4$  cells per well of 96-well plates. 24hrs after plating transfected cells using FuGene 6 Transfection Reagent (Promega). A 4:1 FuGene to DNA ratio was used and an equimolar ratio of plasmids was transfected resulting in a total of 150ng total DNA/well in samples with all three plasmids (mCEACAM1, FnCas9 and an rgRNA). Plasmids were prepared in OptiMEM (Gibco) and allowed to sit for 5min after reagents were mixed before adding 5uL/well. Transfected cells were incubated for an additional 24 hours. Half the media was removed (50uL) so as not to disrupt the cell monolayer. Cells were then infected with 50uL WT, nsp14-ExoN(-) or nsp12-double viruses at an MOI of 0.1. After a final 24-hours supernatants were removed and titered by plaque assay as described previously (Eckerle *et al.*,

2007), and cells were lysed with 80uL/well of PBS+/+ (Gibco) containing 0.5% Triton-X100 (INC Biomedicals Inc.) while rocking for 30 min. 40uL lysate was transferred to opaque plates and assayed using 40uL luciferase assay reagent (Promega) on a Veritas microplate luminometer as described previously (Freeman *et al.*, 2014).

### **Puromycin treatment of DBT cells**

Puromycin (Sigma) was prepared as a 20ug/uL stock solution in water. Duplicate monolayers of sub-confluent DBT cells were grown in DMEM (Gibco) alone or containing 50-1000ug/mL puromycin (changed every other day) for seven days. After seven days cells were bathed in media alone for two days.

### **Plasmid cloning for the generation of stable cell lines**

LentiCRISPRv2 plasmid was digested with XbaI and BamHI to remove SpCas9. The FnCas9, flexible linker and HA tag from the pcDNA3.3 FnCas9-hu (amp<sup>r</sup>) plasmid were amplified using the primers: FnCas9InsXbaI\_F and FnCas9InsBamHI\_R. 7.5nmol cut vector and 20nmol insert were incubated with Gibson Assembly Master Mix (Gibson) at 50C for 1 hour. Resulting plasmid was sequenced using primers LntCRSPRCasSeq\_F and LntCRSPRCasSeq\_R. Plasmid was amplified in chemically competent Stbl3 *E. Coli* (Invitrogen).

### **Statistical analysis**

Statistics were applied as described in figure legends using GraphPad Prism 6 software (La Jolla, CA). The number of replicates performed for each experiment is similarly listed in each respective figure legend. Finally, some of the data was normalized to controls; GraphPad Prism 6 software also performed this analysis. Ratio paired t-tests were used to determine if mutant viruses treated with paired groups of increasing concentrations of mutagens differed from each other, unless values were included in the data set where no virus was observed at which point Wilcoxon tests were applied. Two-way ANOVA was used to determine if 5-FU concentration or mutant virus used in specific infectivity assays resulted in differences between the condition applied and WT viruses without 5-FU. Bonferroni method of correcting for multiple comparisons was used to compare the set of 5-FU-treated WT viruses with the sets of mutant 5-FU-treated viruses. The Mann-Whitney test was applied to TOPO cloned amplicons to determine if differences existed between the mean mutations accumulated for the unpaired viruses tested,

which were larger than expected by random sampling. SEM was used to indicated the probability that the mean of the data presented is representative of the true mean for all data with representative error bars.

### Primers generated for this dissertation research

**Table 4: Primers generated for this dissertation research (IDT)**

Primer Name	Sequence 5'-3'
<b>Quick-change primers for generation of homology mutant viruses</b>	
V553Afor	GGGCCCCGCACCGCTGCTGGTGTCTC
V553Arev	GAGACACCAGCAGCGGTGCGGGCCC
V553Ifor	TAGGGCCCCGCACCATTGCTGGTGTCTC
V553Irev	GAGACACCAGCAATGGTGC GG G C C C T A
M611Ffor	ATGTTGATAGTCCTGTACTCTTCGGTTGGGACTATCCTAAATG
M611Frev	CATTTAGGATAGTCCCAACCGAAGAGTACAGGACTATCAACAT
W613Yfor	GATAGTCCTGTACTCATGGGTTATGACTATCCTAAATGTGATCGT
W613Yrev	ACGATCACATTTAGGATAGTCATAACCCATGAGTACAGGACTATC
A621Gfor	ATCCTAAATGTGATCGTGGTATGCCAAACATACTGCG
A621Grev	CGCAGTATGTTTGGCATAACCACGATCACATTTAGGAT
K794Rfor	GTGTTTATGTCTGAGGCCAGATGTTGGGTAGAAACAGAC
Y649Hfor	CGCATAACGGATAGATTCCATCGTCTTGCGAACGAG
Y649Wfor	TGCTGTTTCGCATACGGATAGATTCTGGCGTCTTGCGAACG
Y649Wrev	CGTTCGCAAGACGCCAGAATCTATCCGTATGCGAACAGCA
Y649Hrev	CTCGTTCGCAAGACGATGGAATCTATCCGTATGCG
K794Rrev	GTCTGTTTCTACCCAACATCTGGCCTCAGACATAAACAC
<b>Quick-change primers for generation of minority variant mutant viruses</b>	
L364Ffor	ACACATCGTTATCGCTTGTCTTTTAAGGACTTGCTTTTGTATG
L364Frev	CATACAAAAGCAAGTCCTTAAAAGACAAGCGATAACGATGTGT
P374Sfor	CTTTTGTATGCTGCAGACAGTGCCTTCATGTGGCGTC
P374Srev	GACGCCACATGAAGGGCACTGTCTGCAGCATAACAAAAG
E519Gfor	GAGGCATTATCATTTGAGGGGCAGGATGAAATTTATGCG
E519Grev	CGCATAAATTTTCATCCTGCCCTCAAATGATAATGCCTC
E519Nfor	CTATTATGAGGCATTATCATTTGAGAATCAGGATGAAATTTATGCGTATACCA
E519Nrev	TGGTATACGCATAAATTTTCATCCTGATTCTCAAATGATAATGCCTCATAATAG
Q520Lfor	AGGCATTATCATTTGAGGAGCTGGATGAAATTTATGCGTATAC
Q520Lrev	GTATACGCATAAATTTTCATCCAGCTCCTCAAATGATAATGCCT
S900Cfor	ATCAGATCCTGGACAGCTACTGTGTTATTTTAAGTACTTGTG
S900Crev	CACAAGTACTTAAAATAACACAGTAGCTGTCCAGGATCTGAT



<b>Quick-change primers for generation of drug passage mutant viruses</b>	
XSP180HF	ATTAATGTGTATAAAAAGCTTGGTCACATATTTAATAGAGCCCTGCTTAAC
XSP180HR	GTTAAGCAGGGCTCTATTAATATGTGACCAAGCTTTTTATACACATTAAT
XSA193SF	CCTGCTTAACACTGCCAAGTTTTGTGACGCATTAGT
XSA193SR	ACTAATGCGTCACAAAACCTGGCAGTGTTAAGCAGG
XSV223IF	GGTCAATGGTATGACTTTGGAGATTTTATTAAGACAGTACCT
XSV223IR	AGGTACTGTCTTTATAAAATCTCCAAAGTCATACCATTGACC
A379TF	CCTGCCCTTCATGTGACCTCTGCTAGTGCAC
A379TR	GTGCACTAGCAGAGGTCACATGAAGGGCAGG
N548SF	ATATGCTATTAGTGCTAAGAGCAGGGCCCGCACC
N548SR	GGTGCGGGCCCTGCTCTTAGCATAATAGCATAT
HT644YAF	AAACATGATTCGTGCTGTTTCGTACGCCGATAGATTCTATCGTCTTG
HT644YAF	CAAGACGATAGAATCTATCGGCGTACGAACAGCACGAATCATGTTT
A695TF	TGCTAATTCTGTGTTTTAACATTTGTCAAACAGTTTCCGCCAATG
A695TR	CATTGGCGGAAACTGTTTGACAAATGTTAAACACAGAATTAGCA
P735SF	TACTCACTAACAAATGCGCTGTCAACATGGTCCGCACG
P735SR	CGTGCGGACCATGTTGACAGCGCATTGTTAGTGAGTA
S768FF	TGTGTTATAATTCAGAGTTTGCCTTTAAGGGTTATATTGCTAATATAAG
S768FR	CTTATATTAGCAATATAACCCTTAAACGCAAACCTCTGAATTATAACACA
XM814KF	TGAATTTTGTCTCAACATACAAAACCTAGTCAAGATGGATGGTG
XM814KR	CACCATCCATCTTGACTAGTTTTGTATGTTGAGAACAAAATTCA
14A93VF	CTTCGATGCAGAAGGTGTGCATGCGATACGTGATA
14A93VR	TATCACGTATCGCATGCACACCTTCTGCATCGAAG
14A93VXF	CTGGGTGGCTTCGCAGCAGCAGGTGTGCATGCGATACGTG
14A93VXR	CACGTATCGCATGCACACCAGCTGCTGCGAAGCCAACCCAG
14F216LF	CCAAGCGTGCACATGTCTAAATTTCTAGAACTGGATACT
14F216LR	AGTATCCAGTTCTAGAATTTAGACATGTGCGACGCTTGG
14D175GF	CAGACCACCTAGCGGGATTGGCAGACAGTGT
14D175GR	ACACTGTCTGCCAATCCCCTAGGTGGTCTG
14P258SF	GATCTTTAACTAGCAATCATGATAGTATTTGCAGCGTGCATAAGGGTG
14P258SR	CACCCTTATGCACGCTGCAAATACTATCATGATTGCTAGTTAAAGATC
<b>Primers for competition assay and reversion analysis</b>	
Comp12_F	AGGGGAGCTCCGTTGATTTG
Comp12_R	AACTACAGACAACGCAGGCA
<b>TOPO-TA cloning primers</b>	
FREQ2_F	ACGCAGTTTGGCTATGTGGA
FREQ2_R	GCCAGCAAGACACACCCTAT

<b>FnCas9/rgRNA cloning and sequencing primers</b>	
MHVrgRNA_HVROCT8_F	GCTCTTCCCAGGGGGCGCCTCTCGTAATTAATAAACCATGAAAG
MHVrgRNA_TRS_F	GGAAGTGCCGTTTATAAAGTCTCGTAATTAATAAACCATGAAAG
MHVrgRNA_Frameshift_F	GGCACAGGGTACAAGACGGGCTCGTAATTAATAAACCATGAAAG
MHVrgRNA_R	CGGTGTTTTCGTCCTTTC
U6_F	CGATACAAGGCTGTTAGAGAG
FnCas9InsXbaI_F	GCCAGAACACAGGACCGGTTCTAGAGCGCTGCCACCATGAACTTTAAGATCCT CCCTATTG
FnCas9InsBamHI_R	GCAGAGAGAAGTTTGTGCGCCGGATCCCGGCGTAGTCAGGCACATC
LntCRSPRCasSeq_F	GTTCTTTTTTCGCAACGGG
LntCRSPRCasSeq_R	CAGAGAGAAGTTTGTTC

## CHAPTER VI

### SUMMARY AND FUTURE DIRECTIONS

#### Introduction

Coronaviruses (CoVs) are the cause of mild to severe disease in humans, as well as economically important infections in livestock and pets (Weiss and Navas-Martin, 2005). CoVs have been emerging as novel infections in human populations for centuries and the two most recent human CoVs (SARS-CoV and MERS-CoV) result in severe disease (Graham et al., 2013). Currently, there are no approved treatments or vaccines for any human CoV. Effective treatments are available for many other RNA viruses utilizing nucleoside analogs (Crotty et al., 2002; Jacobson et al., 2013); however, ribavirin treatment of patients infected with SARS-CoV suggested that the virus was at best resistant, and at worst resulted in enhanced pathogenesis in the presence of the mutagen (Stockman et al., 2006). CoV resistance to ribavirin may be due to increased fidelity as a result of the proofreading activity encoded in the 3'-5' exoribonuclease found in non-structural protein 14 (nsp14-ExoN) (Smith *et al.*, 2013). Fidelity of viral replication has evolved to assure sufficient mutational diversity is available for positive selection during shifts in selective pressures, while concurrently preserving the core sequence essential to virus viability and fitness. This is a critically fine balance, since for RNA viruses studied to date, most mutations resulting in changes to fidelity are attenuating *in vivo* (Smith et al., 2014). Finally, CoVs encode for a large replication-transcription complex (RTC) (Sevajol et al., 2014; Smith and Denison, 2013; Snijder et al., 2016). Thus, CoVs provide a compelling system to investigate determinants of replication fidelity and the interactions between them.

At the start of this dissertation research, the fidelity of CoVs had only been investigated in relation to the inactivation of nsp14-ExoN, by substituting residues of the active site motifs (nsp14-ExoN(-)) (Bouvet et al., 2012; Chen et al., 2007; Eckerle et al., 2007; 2010). The nsp14-ExoN(-) mutant viruses demonstrate an average 14-fold decrease in replication fidelity, impaired replication, increased sensitivity to RNA mutagens, and attenuation *in vivo*. In contrast, for all other RNA viruses outside of the order *Nidovirales*, the RNA-dependent RNA polymerase (RdRp) is the key enzyme regulating replication fidelity. The overall goal of this dissertation research was to test whether determinants in the CoV nonstructural protein 12 (nsp12), including

those in the RdRp (nsp12-RdRp), or in nsp14-ExoN outside of the active site regulate overall CoV replication fidelity. The main objectives of the research design were to test whether nsp12 encodes specific determinants of fidelity, whether nsp12 or nsp14-ExoN are capable of adaptation for increased fidelity, and how nsp12 interacts with nsp14-ExoN to regulate fidelity. In this dissertation research, I investigated three separate methods (homology modeling, analysis of mutagen-treated minority variants and selection for consensus mutations after treatment with mutagen) for identifying determinants of nucleotide selectivity and fidelity in nsp12, emphasizing nsp12-RdRp, and nsp14-ExoN.

The results from this dissertation research demonstrate complex phenotypic interactions between nsp12 and nsp14-ExoN during virus replication, and they support the hypothesis that nsp14-ExoN is epistatic (defined as one protein masking the phenotypic effect of another) to nsp12-RdRp with respect to fidelity. Contrarily, any change to nsp12-RdRp resulted in a decrease in replication capacity in the WT background, but had a variety of effects in the nsp14-ExoN(-) background. This dissertation research also suggests that nsp12-RdRp can be readily selected for alterations in nucleotide pools by natural or experimental approaches. This research also demonstrates that, both dependent and independent of fidelity changes, fitness defects conferred by inactivation of nsp14-ExoN can be partially compensated for by mutations in nsp12-RdRp. Finally, I introduce a potential approach for rapidly assessing the fidelity of CoVs.

**CoVs may be capable of selecting for nsp12-RdRps specialized for local nucleotide pools**  
Viruses often have specific nucleotide biases (van Hemert et al., 2016). For CoVs there is a preference for uracil and against cytosine (Berkhout and van Hemert, 2015). In natural systems it is possible that cells respond to viral infections by altering nucleotide pools (Löffler et al., 2016; Papadopoulou et al., 2015; Wang et al., 2016). In response, viruses may select for RdRps that optimize replication in the new environment. Ultimately, this may lead to altered nucleotide compositions and different codon usage by viral genomes (Berkhout and van Hemert, 2015; Lauring et al., 2012; van Hemert and Berkhout, 2016; van Hemert et al., 2016). One theme that emerged during my dissertation work was that mutations in the CoV nsp12-RdRp result in variable selectivity for nucleotide analogs. These data support the hypothesis that viruses are capable of selection for RdRps that exclude specific nucleotides while preserving the inclusion of others. Interestingly, ribavirin alters nucleotide pools in addition to acting as a mutagen, which

could explain why mutations identified by passaging in ribavirin have typically been reported to result in increased fidelity, rather than only excluding ribavirin specifically (Coffey et al., 2011; Crotty et al., 2000; Graci and Cameron, 2006; Pfeiffer and Kirkegaard, 2003; Zeng et al., 2014). However, passaging in the presence of other nucleoside analogs (generally mutagens) may not result in increased fidelity for natural nucleotides. This is an important consideration since passaging in the presence of mutagens is regularly used when attempting to identify fidelity determinants.

The large Nidoviruses, including CoVs, are unique among (+)RNA viruses in that they encode a proofreading enzyme. So far nsp14-ExoN and nsp10 fidelity-altered viruses have demonstrated consistent changes in sensitivity to nucleoside analogs (Sexton et al., 2016; Smith et al., 2013; 2015). This is likely representative of the different mechanisms used by these proteins to regulate fidelity. Currently, the structural and biochemical mechanisms of nsp14-ExoN proofreading are poorly understood; whereas for nsp12-RdRp, nucleotides must fit properly into the nucleotide binding pocket for efficient catalysis (Gong et al., 2009; Kortus et al., 2012; Moustafa et al., 2011). It is possible that mismatched nucleotides are identified by nsp14-ExoN as larger-scale structural aberrations resulting in a less selective nsp14-ExoN compared with nsp12-RdRp. However, the deep sequencing results presented in Chapter II and (Smith *et al.*, 2013) suggest that nsp14-ExoN may preferentially remove some specific misincorporations over others. In the case of nsp10, it functions in fidelity by enhancing the activity of nsp14-ExoN, therefore, it would likely alter fidelity in the same manner as nsp14-ExoN (Bouvet *et al.*, 2014; Smith *et al.*, 2015).

If CoV nsp12-RdRp, or nsp14-ExoN, can be selected to inhibit incorporation of specific nucleotides while retaining baseline fidelity for others, then the identification of fidelity determinants requires alternative methods to those typical in the RNA virus replication fidelity field. Specifically, passaging in the presence of altered nucleotide pools or multiple mutagens may be a more reliable method for the identification of broad fidelity determinants in the future. Similarly, testing mutant viruses for sensitivity or resistance to altered nucleotide pools, rather than nucleoside analogs, may provide a better screen for fidelity determinants. However, it is possible that selectivity for natural nucleotides is also specific for individual nucleotides. It would be interesting to determine if residues in nsp12-RdRp or nsp14-ExoN can specifically

alter the incorporation of individual natural purine or pyrimidine nucleotides. Questions of specific nucleotide preferences could be addressed with new reduced error deep sequencing technologies or with the development of a reproducible *in vitro* system.

### **A WT nsp14-ExoN is required for the recovery of many mutant nsp12-RdRp viruses**

For the engineered mutations in nsp12-RdRp, many were recoverable only in the WT-MHV background, not the nsp14-ExoN(-) background (Tables 1-3). Additionally, none of the recovered viruses with mutations in nsp12-RdRp, in the nsp14-ExoN(-) background, resulted in increased sensitivity to any of the mutagens tested: 5-FU, 5-AZC, or ribavirin. It is possible that no mutations were identified that result in a decrease in fidelity; however, it is also possible that the nsp14-ExoN(-) virus results in decreased fidelity that is at or near the lowest level tolerated by the virus. Another result of the inactivation of nsp14-ExoN, beyond a change in replication fidelity, is an increase in the lag phase of replication. This lag phase was exacerbated by one of the mutations in nsp12, nsp12-Q520L, and also by growth in the presence of 5-FU. Similarly, only one mutation in nsp12, nsp12-S900C, resulted in an observed decrease in the accumulation of RNA. These data show that viruses are viable when replication is delayed or decreased beyond what results from the inactivation of nsp14-ExoN alone. However, the inability to recover viruses with mutation in nsp12, in the nsp14-ExoN(-) background, may have resulted from larger delays or decreases in RNA synthesis. Once an assay is developed to rapidly and quantitatively investigate changes in fidelity in the WT background it will be interesting to go back to the mutant viruses recovered in this dissertation research and determine whether they result in any changes to overall fidelity.

### **Nsp12 and nsp14 have epistatic relationships**

Comparing the results of nsp12 mutations in the nsp14-WT and nsp14-ExoN(-) background revealed that specific phenotypes are masked depending on the presence or absence of nsp14-ExoN activity. Proofreading by nsp14-ExoN seems to be epistatic to nucleotide selectivity by nsp12-RdRp (Sexton *et al.*, 2016). When one considers that the fold-change in fidelity resulting from the inactivation of nsp14-ExoN is 14, yet RdRp mutations are rarely recoverable outside of a 5-fold change in fidelity, it is not entirely surprising that nsp14-ExoN proofreading might be epistatic to nsp12-RdRp fidelity (Eckerle *et al.*, 2007; 2010; Graham *et al.*, 2012; Sexton *et al.*, 2016; Smith *et al.*, 2013; 2014). What was more surprising to observe was that in many cases

increased lag phases and decreased titers resulting from mutations introduced in nsp12 were only observable in the WT nsp14 background. Viruses with variant nsp12-RdRps in the nsp14-ExoN(-) background generally replicated with identical kinetics to that of nsp14-ExoN(-) alone. While there are many possible explanations these results, I favor the hypothesis, that because nsp14-ExoN(-) viruses still encode the complete nsp14, albeit with a double alanine substitution in the active site, nsp14-ExoN is still able to recognize and attempt removal of incorrect nucleotides. If this is the case, then the RTC may be stalling every time an incorrect nucleotide is inserted and this stall may be more pronounced than any delay resulting from changes in the nsp12-RdRp. It would be interesting to determine if the fidelity and replication defects resulting from the inactivation of nsp14-ExoN can be separated. A knockout of nsp14 has been previously attempted but the virus was non-recoverable (unpublished). Nsp14 also encodes for an N7-methyltransferase domain (N7-MT) (Chen *et al.*, 2009). Additionally, the crystal structure of SARS-CoV nsp14 was solved and demonstrated that there is extensive interaction between nsp14-ExoN and nsp14-N7-MT (Ma *et al.*, 2015). Thus, it would be difficult to remove nsp14-ExoN without disrupting nsp14-N7-MT. In all passaging experiments to date nsp14-ExoN(-) viruses have not reverted (Graham *et al.*, 2012; Sexton *et al.*, 2016). Therefore, if nsp14-ExoN(-) is delaying replication by attempting to remove misincorporated nucleotides, then passaging nsp14-ExoN(-) viruses could result in the identification of mutations that alter or eliminate interactions between nsp12 and nsp14. During 5-FU virus passage, one set of mutations arose in the nsp12-CoV-specific domain that didn't result in an increase in resistance to 5-FU but did increase virus titers. It is possible that this mutant virus, nsp12-PH-AS-VI, reduces the interaction between nsp12 and nsp14. Whether or not this is the case, identifying and weakening the site of interaction between nsp12 and nsp14 should be further pursued in order to fully understand the contributions and mechanisms of action of nsp14-ExoN proofreading in the context of viral infection. Finally, separating out the replication defect and the fidelity defect caused by nsp14-ExoN(-) could facilitate the investigation of other CoV proteins involved in fidelity.

### **Multiple replicase proteins regulate CoV replication fidelity**

CoVs encode multiple proteins that together assemble in the replication-transcription complex (RTC). (Smith *et al.*, 2014). At the start of this dissertation research, only nsp14-ExoN had been shown to regulate CoV fidelity during replication (Eckerle *et al.*, 2007; 2010). In addition, nsp10

is now known to regulate fidelity through enhancement of nsp14-ExoN (Bouvet *et al.*, 2012; Smith *et al.*, 2015). In this dissertation research, I focused on the contributions of nsp12 and nsp14, emphasizing the nsp12-RdRp and nsp14-ExoN domains. The results presented suggest that mutations distant from the nsp12-RdRp active site and nucleotide-binding pocket may alter CoV replication fidelity. There are three main mechanisms with which mutations in the nsp12-RdRp domain are likely to alter fidelity. The first is by altering the nucleotide-binding pocket. This could occur through mutations found directly in the pocket or by altering the overall structure on nsp12 in a way that also altered the pocket. A second way that mutations in nsp12-RdRp could alter the mutation rate of the polymerase would be by altering the speed of catalysis. Evidence in picornaviruses suggest that when catalysis occurs more rapidly incorrect nucleotides are incorporated at a higher rate (Campagnola *et al.*, 2015). The final mechanism for altering fidelity via mutations in nsp12-RdRp would be through interactions with other proteins of the RTC. For instance, biochemical data suggests that nsp7 and nsp8 act together as an essential processivity factor, which directly interacts with nsp12. It is conceivable that changes to processivity could alter viral fidelity and that altering the direct interaction between nsp8 and nsp12 could result in increased or decreased viral fidelity. Additionally, mutations that enhance or relax interactions with nsp14 could result in changes in fidelity. Modulation of fidelity could also occur through mutation in nsp12 that prevent nsp14-ExoN from accessing mispaired nucleotides. One way in which this could occur would be if nucleotide excision by the nsp14-ExoN were occurring upon nsp12-RdRp backtracking after incorrect nucleotide addition similar to DNA replication (Dulin *et al.*, 2015; Ren, 2016). If this is the case then mutations in nsp12-RdRp that result in an inability to backtrack would result in decrease fidelity. Similarly, mutations in proteins that contact nsp14 may influence the overall fidelity of CoVs. Future research into which proteins are involved in regulating CoV fidelity is warranted. In most cases any change to fidelity for RNA viruses has resulted in attenuation where tested (Smith *et al.*, 2014). Similarly, when mice are infected with mouse-adapted SARS-CoV nsp14-ExoN(-) viruses, they survive and are protected from further challenge with WT SARS-CoV (Graham *et al.*, 2012). Future research should determine if any of the nsp12 or nsp14 mutations identified in these studies result in attenuation. Similarly, it would be interesting to determine if fidelity changes in other proteins result in attenuation. Finally, it would be interesting to determine if CoV fidelity is set or if specific cellular conditions result in CoV driven increases or decreases in



fidelity. If instead only nsp12 and nsp14 are involved in fidelity, it will still be beneficial to understand the interface between proteins of the RTC. This knowledge could inform the development of small molecules that interfere with the assembly of CoV RTC, thus leading to novel drugs against CoVs.

### **MHV replication is inhibited by targeted rgRNAs and FnCas9**

An assay to assess changes in CoV fidelity by absolute selection is needed. In this dissertation research, several methods were attempted to assess changes in mutation frequencies for CoVs. Overall, these data illustrate that replication by CoVs differs from that of other RNA viruses. Additionally, these data emphasize the need for a broadly applicable assay that functions to assess changes in fidelity in a wide range of RNA viruses. A broadly applicable assay will improve the overall understanding of the role of fidelity in RNA virus replication and evolution by facilitating comparisons between viruses; current methods are difficult to compare (Sanjuán *et al.*, 2010). A novel approach was introduced in Chapter IV utilizing RNA targeted rgRNAs and FnCas9. There is much work left to determine if this approach will result in an effective, quantitative fidelity assay. However, so far, I was able to show that CoVs are inhibited through the introduction of an rgRNA targeted to the frame shift between Orf1a and Orf1b in the presence of FnCas9. Beyond a fidelity assay it is possible that these data could lead to a novel therapeutic for CoV infections. Since, RNA targeted FnCas9 is likely acting in a manor similar to siRNA (Sampson *et al.*, 2013), which were recently shown to be effective treatments against Ebola when delivered to rhesus monkeys (Thi *et al.*, 2015). Data presented in this dissertation suggest CoV targeted rgRNAs alone inhibit CoV replication. If these data are reproducible then development of a treatment using CoV targeted rgRNA may be more readily achievable. There is still concern that the introduction of Cas9 protein into cells *in vivo* could result in non-specific cleavage of DNA, and the same could be true for targeting non-specific RNAs in the system under development in this work (Abudayyeh *et al.*, 2016). Clinical trials were recently approved for DNA editing in humans and will help to determine if *in vivo* therapeutics utilizing Cas9 proteins will be possible in the future ('First CRISPR clinical trial gets green light from US panel', 2016). While developing the above rgRNA/FnCas9 fidelity assay some unexpected results arose in experimental controls, which may allow new areas of investigation in coronavirus biology. Specifically, an rgRNA targeting murine herpes virus 68 (MHV68) enhanced translation and replication of MHV-A59. It may be informative to determine by what

mechanism targeting of a non-CoV (and not present) DNA virus can enhance CoV replication. A new non-CoV control rgRNA should be developed, which hopefully results in no change in translation or replication in the presence or absence of FnCas9.

One of the most important next steps in the development of an FnCas9 based fidelity assay is the generation of stable cell lines expressing FnCas9 and individual CoV targeted rgRNAs. Stable cell lines will be infected with predicted fidelity-altered viruses and agar overlays will be placed to allow counting on individual escape mutant viruses. WT MHV replication results in  $5 \times 10^{-5}$  mutations per nucleotide per round of replication, and two different sites should result in escape from CoV targeted rgRNA restriction, I estimate that approximately  $4.0 \times 10^4$  PFU will be required to identify an escape mutation. A typical WT MHV virus titer is at least  $10^7$  PFU. So, titers should be high enough to result in a quantifiable assay. The results from the experimental controls for the transient transfection experiments highlight the importance of also generating a stable cell line expressing FnCas9 and a control rgRNA. Additionally, results from nucleotide selectivity experiments throughout this dissertation demonstrate that the key nucleotides at the 3' end of the CoV genome target should be diverse to distinguish between selection for specific nucleotides. Finally, to verify that single mutations at the 3' end of the genome target will allow for viral escape, mutant viruses should be developed containing these changes. Frame shift escape mutant viruses should be utilized to determine if they can restore translation and replication to the levels observed for untargeted FnCas9. Escape is essential to the success of an rgRNA/FnCas9 based fidelity assay. If this approach is successful, fidelity-altered mutant CoVs will be compared by identifying the number of escape mutants arising in a population after a controlled number of replication cycles.

### **Conclusion**

Coronaviruses infect an impressive breath of species and continue to emerge in human populations (Woo *et al.*, 2009). Even during the time this dissertation research, a new human CoV was identified, as were multiple non-human but potentially zoonotic CoVs (Crossley *et al.*, 2012; Shirato *et al.*, 2012; Zaki *et al.*, 2012). Our current understanding of what primes a virus to jump from one species to another and to be (or not to be) successful in that new species is lacking. Additionally, countermeasures against human, zoonotic, agricultural or domestic CoVs (antiviral – vaccines) are minimal, with none currently available against human CoVs. The

research in this dissertation is my contribution to the overall understanding of these much broader principles. I hope that the new knowledge provided by this research can be built upon in the future to provide an intricate picture of how CoVs and other pathogens interact with host species and emerge in new species. Finally, I hope that, as more knowledge is acquired regarding the detailed biology of CoVs, that new targetable mechanisms are illuminated, leading to treatments and preventions against current and future infections.

**APPENDIX A:**

**HOMOLOGY-BASED IDENTIFICATION OF A MUTATION IN THE CORONAVIRUS  
RNA-DEPENDENT RNA POLYMERASE THAT CONFERS RESISTANCE TO  
MULTIPLE MUTAGENS**

# Homology-Based Identification of a Mutation in the Coronavirus RNA-Dependent RNA Polymerase That Confers Resistance to Multiple Mutagens

Nicole R. Sexton,<sup>a,b</sup> Everett Clinton Smith,<sup>b,c</sup> Hervé Blanc,<sup>d</sup> Marco Vignuzzi,<sup>d</sup> Olve B. Peersen,<sup>e</sup> Mark R. Denison<sup>a,b,c</sup>

Department of Pathology, Microbiology, and Immunology,<sup>a</sup> Department of Pediatrics,<sup>b</sup> and Elizabeth B. Lamb Center for Pediatric Research,<sup>c</sup> Vanderbilt University Medical Center, Nashville, Tennessee, USA; Institut Pasteur, Centre National de la Recherche Scientifique UMR 3569, Paris, France<sup>d</sup>; Department of Biochemistry and Molecular Biology, Colorado State University, Fort Collins, Colorado, USA<sup>e</sup>

## ABSTRACT

Positive-sense RNA viruses encode RNA-dependent RNA polymerases (RdRps) essential for genomic replication. With the exception of the large nidoviruses, such as coronaviruses (CoVs), RNA viruses lack proofreading and thus are dependent on RdRps to control nucleotide selectivity and fidelity. CoVs encode a proofreading exonuclease in nonstructural protein 14 (nsp14-ExoN), which confers a greater-than-10-fold increase in fidelity compared to other RNA viruses. It is unknown to what extent the CoV polymerase (nsp12-RdRp) participates in replication fidelity. We sought to determine whether homology modeling could identify putative determinants of nucleotide selectivity and fidelity in CoV RdRps. We modeled the CoV murine hepatitis virus (MHV) nsp12-RdRp structure and superimposed it on solved picornaviral RdRp structures. Fidelity-altering mutations previously identified in coxsackie virus B3 (CVB3) were mapped onto the nsp12-RdRp model structure and then engineered into the MHV genome with [nsp14-ExoN(+)] or without [nsp14-ExoN(-)] ExoN activity. Using this method, we identified two mutations conferring resistance to the mutagen 5-fluorouracil (5-FU): nsp12-M611F and nsp12-V553I. For nsp12-V553I, we also demonstrate resistance to the mutagen 5-azacytidine (5-AZC) and decreased accumulation of mutations. Resistance to 5-FU, and a decreased number of genomic mutations, was effectively masked by nsp14-ExoN proofreading activity. These results indicate that nsp12-RdRp likely functions in fidelity regulation and that, despite low sequence conservation, some determinants of RdRp nucleotide selectivity are conserved across RNA viruses. The results also indicate that, with regard to nucleotide selectivity, nsp14-ExoN is epistatic to nsp12-RdRp, consistent with its proposed role in a multiprotein replicase-proofreading complex.

## IMPORTANCE

RNA viruses have evolutionarily fine-tuned replication fidelity to balance requirements for genetic stability and diversity. Responsibility for replication fidelity in RNA viruses has been attributed to the RNA-dependent RNA polymerases, with mutations in RdRps for multiple RNA viruses shown to alter fidelity and attenuate virus replication and virulence. Coronaviruses (CoVs) are the only known RNA viruses to encode a proofreading exonuclease (nsp14-ExoN), as well as other replicase proteins involved in regulation of fidelity. This report shows that the CoV RdRp (nsp12) likely functions in replication fidelity; that residue determinants of CoV RdRp nucleotide selectivity map to similar structural regions of other, unrelated RNA viral polymerases; and that for CoVs, the proofreading activity of the nsp14-ExoN is epistatic to the function of the RdRp in fidelity.

RNA virus replication results in the incorporation of a relatively high number of mutations, ranging from  $10^{-4}$  to  $10^{-6}$  mutations per site per round of replication (1–5). It is thought that low-fidelity replication is largely responsible for the capacity of RNA viruses to evolve rapidly and adapt to new host species and ever-changing environmental pressures (6–8). RNA-dependent RNA polymerase (RdRp) is central to the replication of RNA viruses and is a key regulator of nucleotide selectivity and fidelity (9, 10). Recent studies of coxsackievirus virus B3 (CVB3), poliovirus, HIV-1, and other viruses have demonstrated that viable viruses are recoverable only within a 4-fold range of RdRp fidelity (11–14). In most cases, altered RdRp fidelity decreases fitness relative to wild-type (WT) viruses; this has been demonstrated for changes as small as a 1.2-fold difference in the accumulation of mutations (12, 14–16). Despite having as little as no amino acid identity outside conserved motifs (11–14, 17–19), all described polymerase structures (including RdRps) resemble a “cupped right hand,” with finger, palm, and thumb domains (20). The fingers form a channel that allows entry of the template RNA and ribonucleotide

triphosphates (rNTPs) and assist in proper positioning of incoming nucleotides in the active site (21). The palm contains the active site, and the thumb functions in contacting exiting nascent RNA (21–23). However, there is diversity in the viral genes that encode RdRps; additional domains that perform a variety of functions, such as methyltransferase, endonuclease, polyribonucleotidyl transferase, guanylyltransferase, membrane targeting, protein-

Received 11 January 2016 Accepted 27 May 2016

Accepted manuscript posted online 8 June 2016

Citation Sexton NR, Smith EC, Blanc H, Vignuzzi M, Peersen OB, Denison MR. 2016. Homology-based identification of a mutation in the coronavirus RNA-dependent RNA polymerase that confers resistance to multiple mutagens. *J Virol* 90:7415–7428. doi:10.1128/JVI.00080-16.

Editor: S. Perlman, University of Iowa

Address correspondence to Mark R. Denison, mark.denison@vanderbilt.edu.

Copyright © 2016, American Society for Microbiology. All Rights Reserved.

protein binding, or protein-RNA binding activities, are often present (24–26).

Coronaviruses (CoVs) infect a wide array of species and have emerged as highly pathogenic human pathogens twice in this century, first with severe acute respiratory syndrome coronavirus (SARS-CoV) in 2003 (27) and then with Middle East respiratory syndrome coronavirus (MERS-CoV) in 2012 (28). CoVs, and other large nidoviruses, replicate with higher fidelity than all other known positive-sense RNA viruses (29, 30). CoVs also have the largest known RNA virus genomes, ranging from 27 to 34 kb (31, 32), and increased fidelity in CoVs is likely required for the maintenance of these large genomes (14). CoV genomes encode 16 nonstructural proteins (nsp1 to nsp16), several of which are known or predicted to function in fidelity regulation, including nsp14-ExoN, a 3′-5′ exonuclease, and nsp10, a modulator of nsp14-ExoN activity (33, 34). Mutating the DE of the nsp14-ExoN active site to AA inactivates the exonuclease, yielding nsp14-ExoN(–) viruses, and nsp14-ExoN(–) viruses exhibit a greater-than-10-fold increase in mutation frequency (29, 35–37). Recent evidence has demonstrated that nsp14 directly interacts with the CoV RdRp encoded in nsp12 (nsp12-RdRp) (38), but the effect of this interaction on nucleotide selectivity and overall fidelity regulation is not known. There are no solved structures for any CoV nsp12-RdRp, but the presence of conserved RdRp motifs and modeling of the C-terminal half of nsp12 predict an RdRp domain that is structurally similar to those of other RNA viruses (39, 40).

The demonstrated function of nsp14-ExoN in high-fidelity CoV replication raises the questions of whether and how nsp12-RdRp participates in fidelity regulation. We sought to determine whether nsp12-RdRp can modulate nucleotide selectivity independently or in association with the proofreading nsp14 exonuclease. We modeled the RdRp domain of CoV nsp12 on CVB3 and poliovirus polymerase structures and predicted the residues important for fidelity based on prior results from those virus systems. Substitution mutations at these residues were introduced in the isogenic recombinant genome of the  $\beta$ -CoV murine hepatitis virus (MHV-A59). We demonstrate that two of these mutations, nsp12-V553I and nsp12-M611F, confer resistance to the mutagen 5-fluorouracil (5-FU) and that one, nsp12-V553I, also results in resistance to the mutagen 5-azacytidine (5-AZC) and demonstrates decreased accumulation of mutations. Increased mutagen resistance and decreased accumulation of mutations were observed only in viruses with an inactivated ExoN, demonstrating that nsp14-ExoN proofreading activity is epistatic to the nucleotide selectivity of nsp12-RdRp. In this paper, we define epistasis as a situation where the phenotype of one gene or viral protein masks the phenotype of genetic variants of another viral protein. This result is consistent with a primary role for nsp14-ExoN in error recognition and removal. However, introduction of RdRp mutations within the WT MHV background decreased fitness relative to WT. Together, the results suggest that nsp12-RdRp shares common determinants of nucleotide selectivity with RdRps from other RNA virus families. Further, the CoV RdRp has likely evolved to function in cooperation with nsp14-ExoN rather than independently.

## MATERIALS AND METHODS

**Virus and cell culture.** Murine delayed brain tumor (DBT) cells (41) and baby hamster kidney 21 cells expressing the MHV receptor (BHK-R) (42) were maintained at 37°C in Dulbecco's modified Eagle medium (DMEM)

(Gibco) supplemented with 10% fetal bovine serum (FBS) (Invitrogen), penicillin, streptomycin (Gibco), and amphotericin B (Corning). The BHK-R cells were further supplemented with 0.8 mg/ml of G418 (Mediatech). All virus work was performed using the recombinant WT MHV strain MHV-A59 (GenBank accession number AY910861 [42]).

**Sequence analysis and homology modeling of CoV MHV nsp12-RdRp.** The MHV RdRp domain structure was generated with the Phyre2 online program (43) using nsp12 residues 385 to 887, which correspond to the reported SARS-CoV nsp12-RdRp model (40). The structural model was compared to the X-ray crystal structures of CVB3 and poliovirus (Protein Data Bank [PDB] accession numbers 3DDK and 1RA7, respectively) using the Pymol Molecular Graphics System (Schrödinger, LLC). ClustalX multiple-sequence alignments were generated using the program MacVector.

**Cloning, recovery, and verification of mutant viruses.** Quick-change mutagenesis was used to generate point mutations in individual MHV genome cDNA fragment plasmids using the previously described MHV infectious clone reverse genetics system (42). Mutant viruses were recovered in cocultured BHK-R and DBT cells following electroporation of *in vitro*-transcribed genome RNA in BHK-R cells. All viruses that included nsp14-ExoN(–) mutations were generated using the F fragment previously described (35). Before use in virus recovery, all mutagenized plasmids were fully sequenced (GenHunter Corporation, Nashville, TN) to ensure no additional mutations were introduced. We also sequence verified engineered mutations in recovered viruses. Viruses in the nsp14-ExoN(–) background took between 84 and 96 h to reach around 80% involvement in syncytia for a passage zero (P0) stock, in contrast to viruses in the WT background, which were frozen at 24 to 48 h postinfection (p.i.). P1 working stocks were made by infecting DBT cells at a multiplicity of infection (MOI) of 0.01 and freezing them when 80% were involved in syncytia, approximately 24 h p.i. for WT viruses and 36 h p.i. for nsp14-ExoN(–) viruses (2 or 3 rounds of replication).

**Compounds and drug sensitivity studies.** 5-FU was obtained from Sigma, and prepared as 200 mM stock solutions in dimethyl sulfoxide (DMSO). 5-AZC was also obtained from Sigma and prepared as 50 mM stock solutions in water. Subconfluent DBT cells were pretreated for 30 min with DMEM with the indicated concentrations of 5-FU or with DMSO, 5-AZC, or medium alone. The treatment was removed, and the inoculum was added and allowed to adsorb for 1 h at 37°C. The inoculum was then removed, and medium with drug or DMSO was returned. Infection proceeded for 24 h for WT or 32 h for nsp14-ExoN(–) viruses, after which the supernatants were acquired and frozen and the titer was determined by plaque assay as previously described (35).

**Virus replication and RNA synthesis assays.** Subconfluent DBT cell monolayers in triplicate were infected at an MOI of 0.01 PFU/cell. The virus was allowed to adsorb for 30 min, after which the inocula were removed and the cells were washed 2 times with phosphate-buffered saline (PBS), followed by addition of prewarmed medium. For replication kinetics assays, samples were taken at various time points postinfection. Titering was performed by plaque assay, as previously described (35). For analysis of RNA synthesis, total infected-cell RNA was obtained with TRIzol reagent (Invitrogen) at various times postinfection, and two-step reverse transcription-quantitative PCR (RT-qPCR) was performed as previously described (44).

**Determination of specific infectivity.** Subconfluent DBT cells were pretreated for 30 min with DMEM with the indicated concentration of 5-FU or with DMSO alone. The treatment was removed, and the inoculum was added and allowed to adsorb for 1 h at 37°C. The inoculum was then removed, and medium with drug or DMSO was returned. Infection proceeded for 20 h for WT or 24 h for nsp14-ExoN(–) virus, and then the supernatants were acquired and frozen, and the titer was determined by plaque assay as previously described (35). The supernatants were also used for RNA genome isolation by adding 100  $\mu$ l supernatant to 900  $\mu$ l TRIzol reagent (Invitrogen), chloroform extraction by phase separation, and using the aqueous layer in the PureLink Mini RNA kit (Ambion) according

to the manufacturer's protocol. One-step RT-qPCR was performed as described below, and the ratio of PFU to genomes of the supernatant was determined.

**One-step RT-qPCR for determining supernatant genome copies for specific infectivity assay.** An RNA standard was generated using the MHV A fragment (42) to generate a 931-nucleotide (nt) RNA. First, cDNA was generated by PCR using the following primers: forward, 5'-TAA TACGACTACTATAGGGGCTATGTGGATTGTGTGG-3', which begins with a T7 promoter, and reverse, 5'-AATTCTTGACAAGCTC AGGC-3'. RNA for the standard curve was then generated using an mMessage mMachine T7 kit (Ambion). An agarose gel with 1% bleach was run, and an ~900-nt band was observed. RNA was purified using an RNeasy minikit (Qiagen). Dilutions of the standard curve were made from  $10^3$  to  $10^8$  genome equivalents for use in assays as needed. Primers and probes for one-step RT-qPCR were purchased from BioSearch Tech. The probe was 5'-6-carboxyfluorescein (FAM) labeled and 3' black hole quencher 1 (BHQ-1) labeled with the sequence 5'-TTCTGACAACGGC TACACCAACG-3' and made up to 5  $\mu$ M in nuclease-free water. The primers used were forward, 5'-AGAAGTTACTGGCAACTG-3', and reverse, 5'-TGTCCACGGCTAAATCAAAC-3'. Reaction mixtures were set up on ice, with enzyme added last. The final volume for reaction mixtures was 20  $\mu$ l, with 150 nM probe, 900 nM each primer, 2  $\mu$ l sample RNA, and 10  $\mu$ l 2 $\times$  ToughMix one-step low ROX enzyme mix (Quantas) used per reaction. Samples were plated in duplicate and run on the Applied Biosciences 7500 real-time PCR system with the following conditions: 55°C for 10 min, 95°C for 5 min, 95°C for 30 s, and 60°C for 1 min, with the last two steps repeated 40 times. The standard curve was graphed, and the number of genomes per milliliter was determined.

**Competitive fitness of mutant viruses.** Subconfluent DBT monolayers were coinoculated at a total MOI of 0.01 PFU/ml with RdRp mutant viruses in the nsp14-ExoN(-) background and nsp14-ExoN(-) at either a 1:1, 1:9, or 9:1 ratio. When 50 to 70% of the monolayer was involved in syncytia, total RNA was harvested. The RNA was then reverse transcribed using SuperScriptIII (Invitrogen) according to the manufacturer's protocol, and amplicons were generated using primers designed to cover the region including the codons for both the V553 and M611 residues. Amplicons were sent for sequencing, and electropherograms were analyzed using MacVector.

**Passage reversion analysis.** Triplicate monolayers of subconfluent DBT cells were infected at an initial MOI of 0.01 PFU/ml of nsp12-V553I and nsp12-M611F viruses in both the WT and nsp14-ExoN(-) backgrounds. The viruses were then blind passaged in triplicate for 5 passages. Total RNA was sequenced across a 1.7-kb region of nsp12-RdRp that included both nsp12-RdRp mutations. Electropherograms were analyzed using MacVector.

**Preparation of amplicons for deep sequencing of full viral genomes.** Subconfluent DBT cells were infected at an MOI of 0.01 PFU/ml with nsp12-V553I or nsp12-M611F in either the WT or nsp14-ExoN(-) background, with nsp14-ExoN(-) alone, or with the WT alone. The infections were allowed to progress for 20 h, and then RNA was isolated. The RNA was reverse transcribed using SuperScriptIII (Invitrogen) according to the manufacturer's protocol, and 12 amplicons were generated to cover the whole genome and processed as described previously (44).

**Deep-sequencing sample preparation and analysis.** Amplicons were subsequently purified with a nucleospin PCR purification kit (Macherey-Nagel), quantified with PicoGreen, fragmented (Fragmentase), and prepared using the Illumina NextSeq500 Mid Output 150-cycle kit following the standard protocols. Sequences were obtained with an Illumina NextSeq500 machine. Sequencing runs were analyzed using the previously published ViVan bioinformatics pipeline (45). Briefly, the pipeline performed quality filtering, and adaptor cleaning was done using fastq-clipper ([http://hannonlab.cshl.edu/fastq\\_toolkit/index.html](http://hannonlab.cshl.edu/fastq_toolkit/index.html)). The 150-nt reads were aligned with the reference sequence with a maximum of 2 mismatches per read using BWA (46) and processed using SAMtools (47) to obtain the nucleotide/base calling at each position. The ViVan pipeline

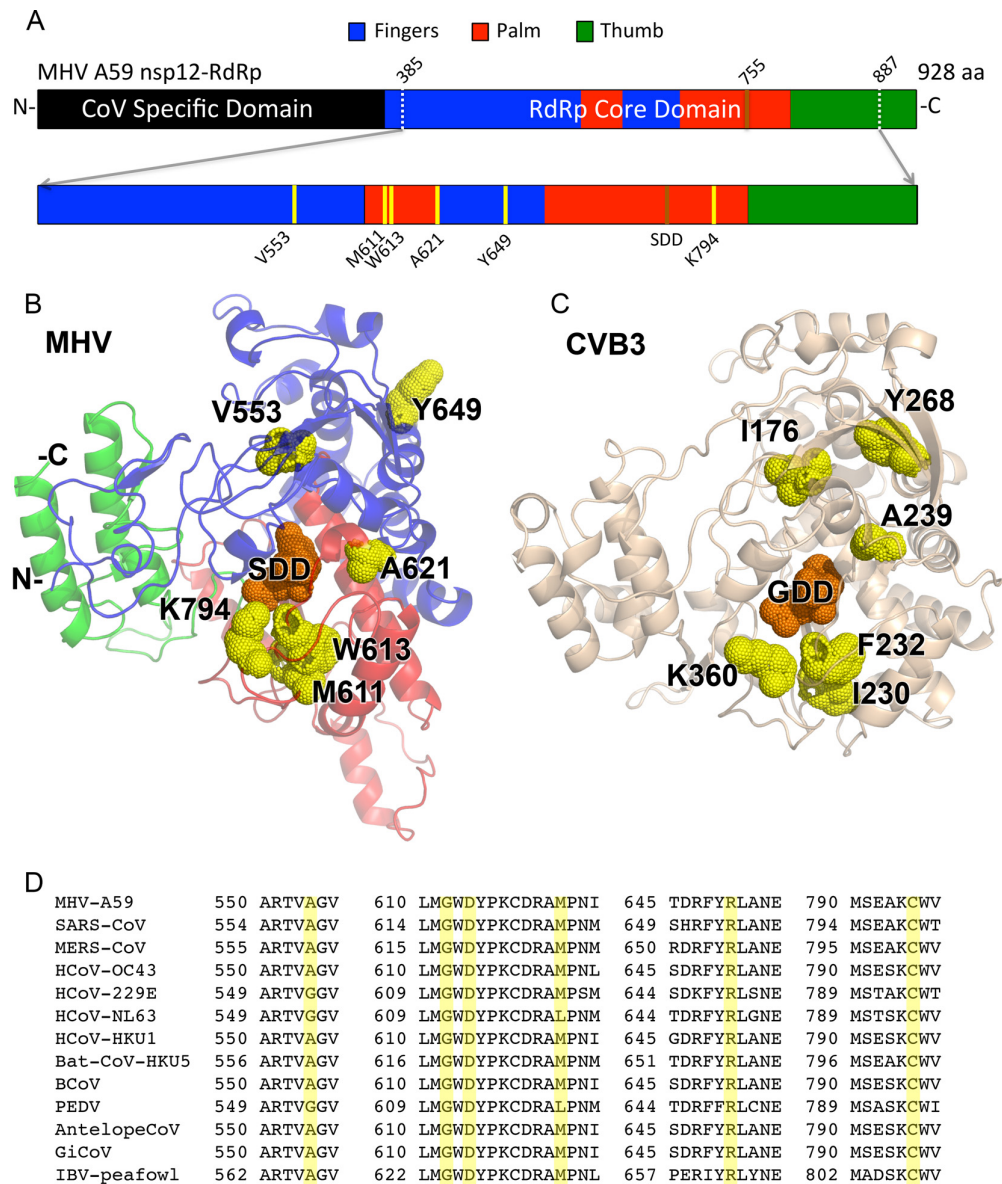
then identified statistically significant variants above the background noise due to sequencing error, calculated for each nucleotide site as follows. For each position throughout the viral genome, base identities and their quality scores were gathered. Each variant allele's rate was initially modified according to its covering read qualities based on a maximum-likelihood estimation and tested for significance using a generalized likelihood ratio test. Additionally, an allele confidence interval was calculated for each allele. In order to correct for multiple testing, a Benjamini-Hochberg false-discovery rate of 5% was set. In all experiments, a minimum coverage of 3,000 reads was obtained, and the background error frequency at every nucleotide site was always below 0.0001. For analysis, we use a conservative frequency cutoff of 0.01, consistent with previous studies (48–50).

**Statistical analysis.** Statistics were applied as described in the figure legends using GraphPad (La Jolla, CA) Prism 6 software. The number of replicates performed for each experiment is similarly provided in each figure legend. Finally, some of the data were normalized to controls; GraphPad Prism 6 software was also used to perform this analysis.

## RESULTS

**Homology modeling of MHV nsp12-RdRp polymerase core domain predicts putative fidelity determinants.** Mutations that alter nucleotide selectivity have been identified across multiple RNA virus RdRps (3, 11, 51–53); however, whether these residues are conserved across virus families is unknown. We sought to determine whether residues within nsp12-RdRp that are structurally homologous to known RNA virus fidelity determinants would have similar effects on nucleotide selectivity when introduced into the MHV background. To do this, we modeled the structure of MHV nsp12-RdRp using Phyre2 software (43). A series of nsp12-RdRp truncations was assessed, and the highest-confidence model was used for further study. This region corresponded to a published model for the SARS nsp12-RdRp (40) and included residues 385 to 887 of the MHV nsp12 protein, referred to here as the RdRp core domain (Fig. 1A and B). Deletion of the CoV-specific domain (residues 1 to 384) and a small C-terminal portion of the thumb domain (residues 888 to 928) was required to establish this high-confidence model (Fig. 1A). The model was resolved by highest-probability similarity to human rhinovirus serotype 14 (PDB ID 1XR5), rabbit hemorrhagic disease virus (PDB ID 1KHV), and enterovirus 71 (EV71) (PDB ID 3N6M). The Phyre2 confidence, i.e., the probability of true homology of the RdRp core domain with these structures, was >99%, while the identity was only 14 to 20%. Having generated a structural model for the MHV nsp12-RdRp core domain, we next sought to predict the residues involved in nucleotide selectivity. The nsp12-RdRp core domain model was aligned with the solved structure of CVB3 (PDB ID 3DDK) using PyMol (Fig. 1C). A series of CVB3 RdRp mutations have been shown to result in decreased fidelity (10, 11, 54). The CVB3 fidelity determinants were compared with the MHV nsp12-RdRp core domain model. Those that aligned well structurally and by amino acid similarity—MHV nsp12-V553I, -M611I, -W613I, -A621I, -Y649I, and -K794I (Fig. 1)—were further investigated. Finally, the nsp12 amino acid sequences of 27 different  $\alpha$ -,  $\beta$ -, and  $\gamma$ -CoVs were aligned, including SARS-CoV and MERS-CoV. All six identified residues were conserved across these CoVs (Fig. 1D). Analyses of similarity and the types of residues in the picornaviruses were then used to determine the specific amino acid changes that would be introduced at the identified MHV residues. The resulting substitution mutations were engineered in





**FIG 1** Homology modeling of CoV nsp12-RdRp and identification of residues that potentially regulate fidelity based on CVB3 structure. (A) Phyre2 software was used to model a subsection of the MHV nsp12-RdRp structure (expanded from the nsp12-RdRp full schematic). (B and C) The modeled MHV RdRp structure (B) was aligned with the solved CVB3 RdRp structure (C). (A and B) The residues chosen for site-directed mutagenesis were selected by comparing previously determined fidelity-altering mutations of picornavirus RdRps. (D) Amino acid alignments across CoVs showing that all residues are almost completely conserved.

the isogenic cloned MHV genome: nsp12-V553A/I, -M611F, -W613Y, -A621G, -Y649H/W, and -K794R (Table 1).

**Recovery of mutant viruses in the MHV nsp14-ExoN(+) (with ExoN activity) and nsp14-ExoN(-) isogenic backgrounds.** We next tested whether viable viruses could be recovered with substitutions at the identified residues. Virus recovery was attempted a maximum of 3 times, resulting in recovery of 6 of the 8 mutant viruses in the WT background: nsp12-V553I, -M611F, -W613Y, -A621G, -Y649H, and -K794R. The time required for recovery of mutant viruses in the WT background ranged from 24

to 48 h. No other mutations were identified across nsp12 sequences in recovered viruses. The A621G mutant was not further studied, as it demonstrated rapid primary reversion even in the recovery (P0) supernatant. Since our goal was to understand the relationship of nsp12-RdRp and nsp14-ExoN in fidelity regulation, we additionally attempted recovery of the WT background viable mutants in the setting of inactivated nsp14-ExoN [nsp14-ExoN(-)]. In contrast to mutant viruses recovered in the WT background, we recovered only 2 of the 5 mutants in the nsp14-ExoN(-) background: nsp12-V553I/nsp14-ExoN(-) and



TABLE 1 Recovery of mutant viruses using site-directed mutagenesis

nsp12-RdRp region	Engineered substitution (nsp12)		Recovery	
	CVB3 <sup>a</sup>	MHV	nsp14-ExoN(+) (WT)	nsp14-ExoN(-)
Fingers	I176V	V553A	No	Not attempted
		V553I	Yes	Yes
	Y268H	Y649H	Yes	No
		Y649W	No	Not attempted
Palm	I230F	M611F	Yes	Yes
	F232Y	W613Y	Yes	No
	A239G	A621G	Revertant	Not attempted
	K360R	K794R	Yes	No

<sup>a</sup> Reference 8.

nsp12-M611F/nsp14-ExoN(-) (Table 1). The time to recovery for nsp12-V553I/nsp14-ExoN(-) was 84 h, and that for nsp12-M611F/nsp14-ExoN(-) was 96 h. Working stocks of all the viruses were made by infecting DBT cells at an MOI of 0.01 and recovering stocks at around 24 h p.i. for WT viruses or between 32 and 48 h p.i. for mutant viruses in the nsp14-ExoN(-) background. Therefore, 4 to 10 replication cycles were required to generate stocks. Working stocks were sequenced to verify that the introduced mutations were still present.

**Resistance of recovered mutant viruses to the base analog 5-fluorouracil.** We next tested our panel of recovered mutant viruses for resistance to the RNA mutagen 5-FU. 5-FU has been used with picornaviruses, influenza viruses, vesicular stomatitis viruses, and others to reflect changes in fidelity based on increased or decreased sensitivity to incorporation and virus inhibition (11, 30, 49, 55). WT CoVs (MHV and SARS-CoV) are resistant to 5-FU, while nsp14-ExoN(-) mutants are profoundly sensitive to 5-FU inhibition, consistent with nsp14-ExoN-mediated removal of misincorporated 5-FU. The effect of 5-FU on DBT cell viability was previously tested, with no effect observed up to 400  $\mu$ M (30). We compared WT-MHV with the nsp12 mutants in both the WT and nsp14-ExoN(-) backgrounds. There was no significant change in sensitivity to 5-FU compared to WT for any of the nsp12-RdRp mutant viruses in the WT background at up to 120  $\mu$ M 5-FU, although the mutation Y649H did appear to decrease resistance slightly (Fig. 2A). In contrast, the nsp12-M611F/nsp14-ExoN(-) and nsp12-V553I/nsp14-ExoN(-) mutant viruses were both significantly less sensitive to 5-FU than nsp14-ExoN(-) alone, with both populations persisting when treated with 120  $\mu$ M 5-FU, where nsp14-ExoN(-) was not detectable beyond 80  $\mu$ M 5-FU (Fig. 2B). These data demonstrate that both nsp12-RdRp mutations, V553I and M611F, confer resistance to 5-FU. This suggests two possibilities: that it is not possible to increase the exclusion of 5-FU beyond the high level dictated by nsp14-ExoN or that selectivity for native nucleotides over 5-FU is in fact increased by nsp12 mutations, but at a low level that is not detectable as changes in the virus titer.

Three of the mutations (nsp12-Y649H, -W613Y, and -K794R) were viable in the nsp14-ExoN(+) background but failed to grow in the absence of proofreading. The observation that they grew only in an nsp14-ExoN(+) background indicated that the mutations retained sufficient polymerase function to support virus rep-

lication but that it was critically dependent on having proofreading functionality, suggesting these mutations may have given rise to low-fidelity variants. Unfortunately, the low titer from the nsp14-ExoN(-) background precluded direct-sequencing analysis, and we therefore cannot definitively show this is the case.

**Replication kinetics of nsp12-V553I and nsp12-M611F mutant viruses in the WT and nsp14-ExoN(-) backgrounds.** Since we were interested in mutations that potentially confer altered fidelity, we prioritized the nsp12-V553I and nsp12-M611F mutant viruses for further analysis. We next sought to determine how the nsp12-V553I and nsp12-M611F viruses replicated in comparison to their isogenic backgrounds (Fig. 3). In the wild-type background, both mutant viruses had slightly delayed exponential replication but eventually reached similar peak titers comparable to

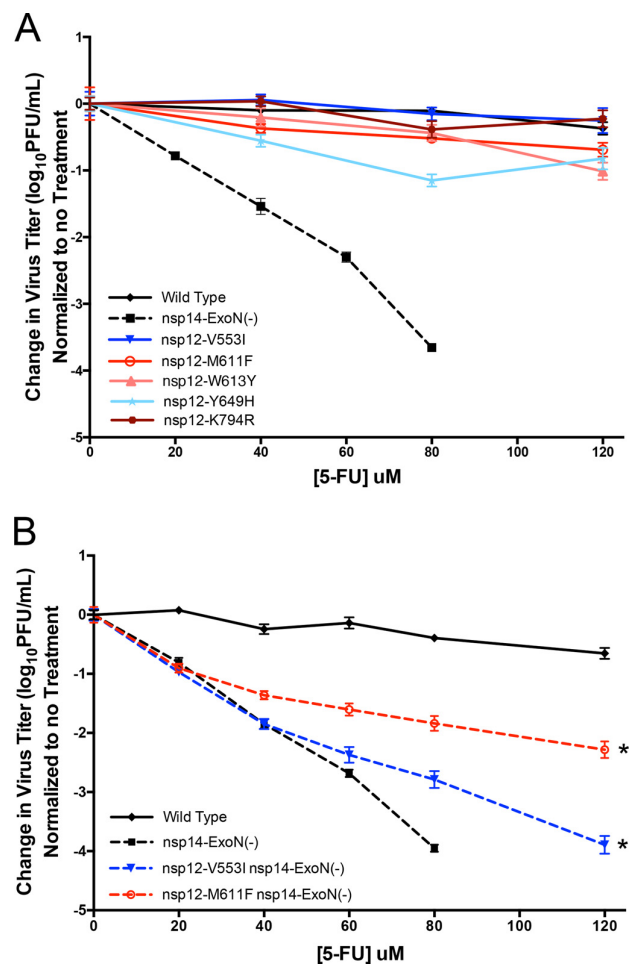


FIG 2 Resistance of MHV nsp12-RdRp mutant viruses to 5-fluorouracil in the WT and nsp14-ExoN(-) backgrounds. The domain locations of mutations are indicated as follows: fingers, blues; palm, reds. DBT cells were pretreated with different concentrations of 5-FU for 30 min. The treatment was removed, and the cells were infected with the indicated viruses in the WT background (A) or the nsp14-ExoN(-) background (B) at an MOI of 0.01. The medium containing 5-FU was replaced 30 min p.i. Virus samples were taken at 24 (WT) or 32 [nsp14-ExoN(-)] h p.i., and the titer was determined by plaque assay. The data represent the results of 3 independent experiments, each with 2 replicates. The error bars represent standard errors of the mean (SEM) \*,  $P < 0.05$  by the Wilcoxon test.

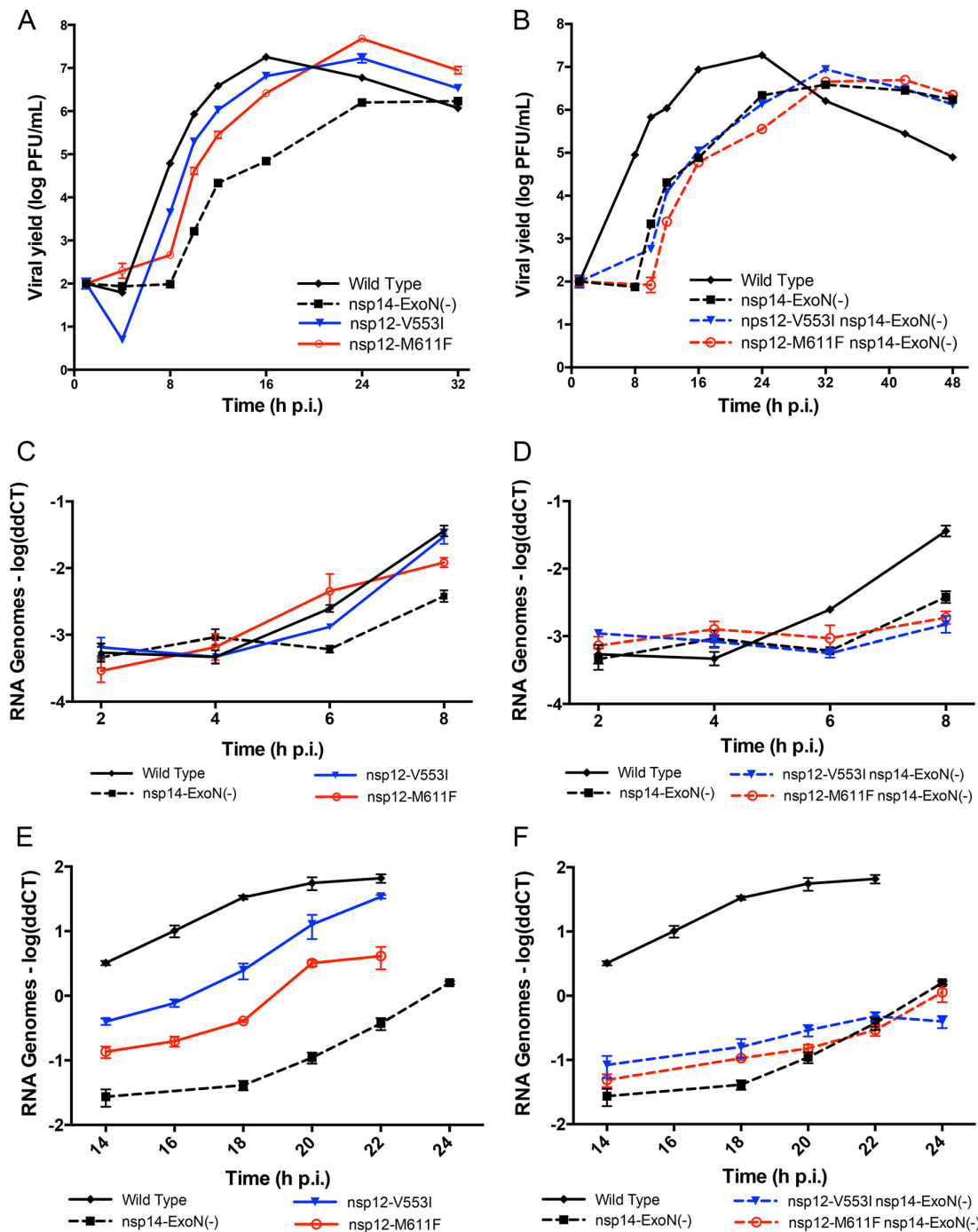


FIG 3 Replication kinetics of MHV nsp12-RdRp mutant viruses. The mutation location is indicated as follows: fingers, blues; palm, red. DBT cells were infected with the viruses indicated in the WT background (A, C, and E) or the nsp14-ExoN(-) background (B, D, and F) at an MOI of 0.01 PFU/cell. Supernatant aliquots were taken at the indicated times p.i., and titers were determined by plaque assay. Total RNA was taken at the indicated times p.i., and RT-qPCR was performed. The data represent the results of 3 independent experiments. The error bars represent SEM.

that of the WT. In contrast, in the nsp14-ExoN(-) background, both nsp12-V553I and nsp12-M611F mutant viruses displayed replication kinetics similar to those of the isogenic nsp14-ExoN(-) background. We also assessed RNA synthesis for nsp12-

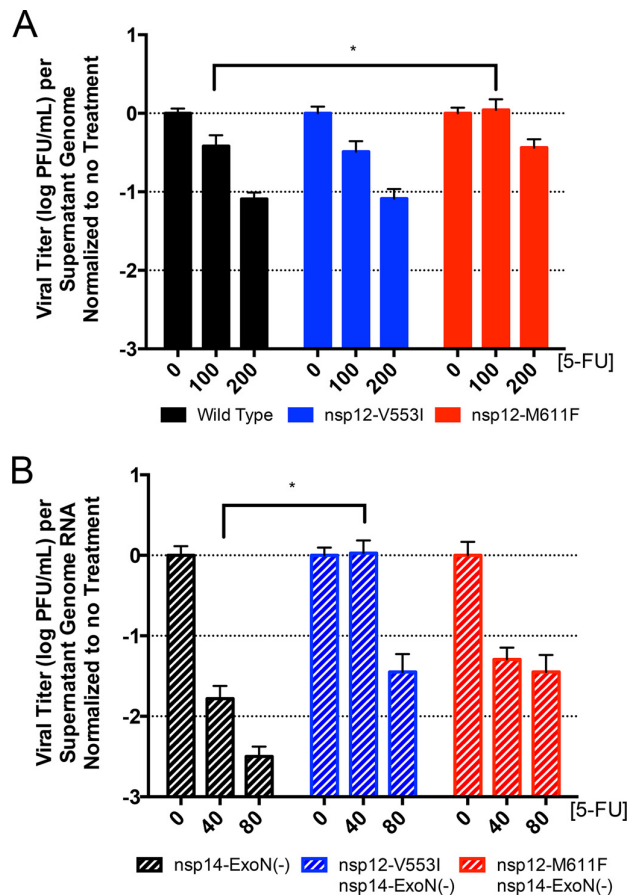
V553I and nsp12-M611F by RT-qPCR. The measured genomic-RNA levels were consistent with the virus replication kinetics data, and we observed delayed and decreased genome RNA synthesis in the WT background after multiple rounds of replication (Fig. 3E).

However, at early time points, there was no difference in RNA accumulation, suggesting that decreased RNA is a result of post-RNA synthesis steps (Fig. 3C). In the nsp14-ExoN(-) background, RNA synthesis levels were indistinguishable from the nsp14-ExoN(-) background for both nsp12-V553I and nsp12-M611F, with no additional RNA synthesis defects detectable (Fig. 3D and F). The results, along with the ability to recover several of the nsp12-RdRp mutants only in the nsp14-ExoN(+) background, support the hypothesis that nsp14-ExoN and nsp12-RdRp may have an epistatic relationship. However, for replication kinetics, the effects of nsp12-V553I and nsp12-M611F were not observable in the presence of inactive nsp14-ExoN, demonstrating that the replication phenotype of inactive nsp14-ExoN is epistatic to replication variants encoded in nsp12-RdRp.

**Specific infectivity of nsp12-V553I and nsp12-M611F.** Having identified two mutant viruses with resistance to 5-FU, we wanted to further test whether the resistance was due to decreased incorporation of the mutagen. Measurement of specific infectivity has been useful for determining lethal mutagenesis for MHV and other RNA viruses (30, 56). We tested the nsp12-M611F and nsp12-V553I mutants in both the WT and nsp14-ExoN(-) backgrounds for changes in specific infectivity when infected at an MOI of 0.01 and treated with 5-FU (Fig. 4). Both nsp12-M611F and nsp12-V553I resulted in an increased ratio of infectious particles (PFU per milliliter) to total particles (RNA genomes) in the nsp14-ExoN(-) background (Fig. 4B). Similarly, in the WT background, nsp12-M611F demonstrated an increase in the ratio of infectious particles to RNA genomes (Fig. 4A). Thus, specific infectivity may be a more sensitive measure of lower-level changes in nucleotide selectivity in the setting of nsp14-ExoN(-).

**Fitness costs of nsp12-V553I and nsp12-M611F in WT and nsp14-ExoN(-) backgrounds.** In multiple RNA viruses, including CoVs, both increased and decreased fidelity have been reported to have a fitness cost (14). We therefore sought to determine whether nsp12-V553I or nsp12-M611F conferred any cost in fitness, defined as the ability to directly compete during coinfection. In the WT background, both nsp12-V553I and nsp12-M611F demonstrated delays in replication and impaired RNA accumulation. However, since there were no observed additional defects in replication or RNA synthesis for nsp12-V553I and nsp12-M611F when introduced into the nsp14-ExoN(-) background (Fig. 3B, D, and F), we tested for any additional fitness cost of nsp12-V553I/nsp14-ExoN(-) or nsp12-M611F/nsp14-ExoN(-) virus compared with nsp14-ExoN(-) alone. When coinfecting with nsp14-ExoN(-) virus at ratios from 1:9 to 9:1, nsp12-V553I/nsp14-ExoN(-) maintained the input ratio compared with nsp14-ExoN(-) (Fig. 5A). A small advantage for nsp12-V553I/nsp14-ExoN(-) was observed when the coinfecting cultures were treated with 60  $\mu$ M 5-FU, consistent with a conferred advantage for 5-FU resistance (Fig. 5A). Thus, there appeared to be no additional fitness cost of nsp12-V553I/nsp14-ExoN(-) compared to nsp14-ExoN(-) alone. In contrast, nsp12-M611F/nsp14-ExoN(-) was not able to compete with nsp14-ExoN(-) at any ratio (Fig. 5B). Treatment with 60  $\mu$ M 5-FU again favored nsp12-M611F/nsp14-ExoN(-); however, even then, the percentage of the population made up of nsp12-M611F/nsp14-ExoN(-) virus remained at only around 30% when initially given a 9-fold advantage (Fig. 5B).

We next tested whether the relative differences in fitness cost resulted in selective pressure for reversion of nsp12-V553I



**FIG 4** Specific infectivity is increased in both nsp12-V553I and nsp12-M611F mutants. DBT cells were pretreated with increasing concentrations of 5-FU for 30 min. The treatment was removed, and the cells were infected with the indicated viruses in the WT background (A) or the nsp14-ExoN(-) background (B) at an MOI of 0.01. The medium containing 5-FU was replaced 60 min p.i. Virus samples were taken at 20 and 24 h p.i. Titers were determined by plaque assay, and the numbers of supernatant genomes were determined using one-step RT-qPCR. The data represent the results of 2 independent experiments, each with 3 replicates. The error bars represent SEM (\*,  $P < 0.05$  by 2-way analysis of variance [ANOVA] using the Bonferroni correction for multiple comparisons).

and nsp12-M611F (Fig. 6). DBT cells were infected with nsp12-V553I, nsp12-V553I/nsp14-ExoN(-), nsp12-M611F, or nsp12-M611F/nsp14-ExoN(-) virus at an initial MOI of 0.01. After 5 passages, the viruses were analyzed for retention of original mutations using dideoxy (Sanger) sequencing. The nsp12-V553I mutation was stable after passage in both WT and nsp14-ExoN(-) backgrounds, maintaining the mutated AUU codon (Fig. 6A), suggesting minimal selective pressure on that nucleotide, codon, or amino acid. In contrast, the nsp12-M611F mutation demonstrated significant change over passage in both the WT and the nsp14-ExoN(-) backgrounds. The original mutation, UUC, was no longer the majority codon in the nsp14-ExoN(-) background and was less than 52% of the population in all WT background lineages (Fig. 6B). The nsp12-M611F/nsp14-ExoN(-) population resulted in a mixture of 2-nucleotide changes, resulting in reversion to AUG (methionine) ( $\leq 68\%$  of the population); single

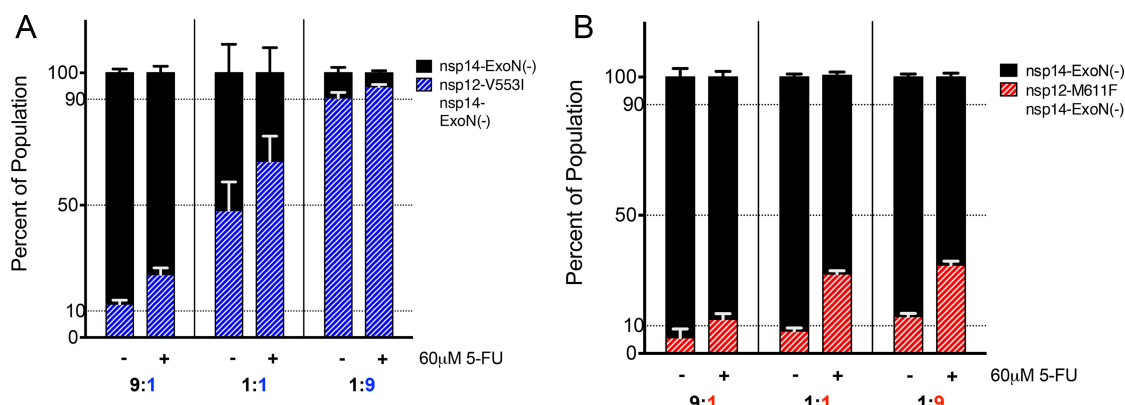


FIG 5 Competitive fitness analysis in the nsp14-ExoN(-) background. DBT cells were pretreated with medium alone or medium containing 60  $\mu$ M 5-FU for 30 min. The treatment was removed, and the cells were coinfectd at a total MOI of 0.01 with nsp14-ExoN(-) and nsp12-V553I/nsp14-ExoN(-) (A) or nsp12-M611F/nsp14-ExoN(-) (B) viruses at a ratio of 9:1, 1:1, or 1:9. Medium alone or containing 60  $\mu$ M 5-FU was replaced 30 min p.i. Total RNA was taken at 24 h p.i. Sequencing was performed across a 1.7-kb region of nsp12-RdRp that included both mutations. The data represent the results of 3 independent experiments, each with 2 replicates. The error bars represent SEM.

nucleotide changes that resulted in mutation to Leu ( $\leq 55\%$  of the population), a somewhat smaller but still hydrophobic residue; or retaining a Phe substitution ( $\leq 27\%$  of the population). Thus, the fitness cost of nsp12-M611F results in significant selective pressure for changes at that residue during passage in the absence of 5-FU.

**Resistance of recovered mutant viruses to the base analog 5-azacytidine.** Having shown that both nsp12-M611F and nsp12-V553I mutations conferred resistance to 5-FU in the nsp14-ExoN(-) background but that nsp12-M611F reverted quickly and was at a fitness disadvantage in both the WT and nsp14-ExoN(-) backgrounds, we next wanted to test whether the mutations conferred resistance specifically to 5-FU or broadly to base analogs and therefore were likely determinants of fidelity. A study by Arias et al. demonstrated that for foot-and-mouth disease virus (FMDV), resistance to a specific mutagen can result from a point mutation while conferring the opposite overall fidelity (57). We therefore tested for resistance to the additional mutagen 5-AZC. Similar to the 5-FU results, the nsp12-V553I/nsp14-ExoN(-) mutant virus was more resistant to 5-AZC than nsp14-ExoN(-) alone, maintaining approximately 1-log-unit-higher titers from 20 to 50  $\mu$ M 5-AZC (Fig. 7). However, the nsp12-M611F/nsp14-ExoN(-) mutant virus showed no difference in resistance to 5-AZC compared with nsp14-ExoN(-) alone (Fig. 7). These data suggest that nsp12-V553I is likely a fidelity determinant, whereas nsp12-M611F confers specific resistance to 5-FU and has unknown overall fidelity. This result is also consistent with the rapid reversion observed for the nsp12-M611F mutation.

**The nsp12-V553I mutation results in a decrease in the accumulation of mutations.** Having shown that the nsp12-V553I and nsp12-M611F mutations resulted in resistance to 5-FU and that nsp12-V553I additionally conferred resistance to 5-AZC, we sought to directly determine whether either of the mutations resulted in a change in the number of mutations accumulated in viral RNA. DBT cells were infected with wild-type, nsp12-V553I, nsp12-M611F, nsp14-ExoN(-), nsp12-V553I/nsp14-ExoN(-), or nsp12-M611F/nsp14-ExoN(-) virus at an MOI of 0.01, and RNA was collected at 20 h p.i. The samples were then prepared for Illumina next-generation sequencing (NGS) across the full ge-

nome and analyzed using the ViVan analysis pipeline (45). Mutations present at 1% or more of the population were graphed by frequency and position in the genome, with engineered mutations depicted with colored dots. Most nonengineered new mutations were present at 10% or less of the population and were distributed across the genome with no detectable hot spots (Fig. 8A to F). For nsp12-V553I in the WT background, no difference was observed in the number of mutations accumulated to 1% or greater of the population compared to the WT alone. In contrast, in the nsp14-ExoN(-) background, nsp12-V553I was associated with a 1.7-fold decrease in the frequency of mutations compared to nsp14-ExoN(-) alone (Fig. 8G), again with no change in the distribution of mutations across the genome (Fig. 8C and D). These results are consistent with both the 5-FU and 5-AZC data in suggesting increased fidelity. The results from the nsp12-M611F mutant viruses were more complicated. To our surprise, the nsp12-M611F/nsp14-ExoN(-) virus fully reverted at both engineered nsp12-M611F nucleotides during the low-MOI infection, resulting in a virus population that was nsp14-ExoN(-) alone (Fig. 8F). This reversion made the results for the nsp12-M611F/nsp14-ExoN(-) virus uninterpretable. However, since accumulation of mutations to over 1% of the population is a combination of all replication cycles from initial recovery to the final sample [roughly 11 in total for nsp12-M611F/nsp14-ExoN(-)], we included these data in Fig. 8G and H. In the WT background, the nsp12-M611F mutations were still present at 100% of the population and resulted in a 1.93-fold increase in the total number of accumulated mutations. We observed only a slight increase in the number of mutations accumulated in the nsp12-M611F/nsp14-ExoN(-) sample over those of nsp14-ExoN(-), which was not surprising, as the M611F mutation was no longer present. Neither sample appeared to have mutations concentrated in specific locations across the genome, suggesting the accumulation of mutations was due to random generation of mutations rather than strong selection in particular locations or proteins (Fig. 8E and F). Of note, one mutation in nsp3 of the nsp14-ExoN(-) sample and a mutation in the nsp3, nsp13, and E proteins of nsp12-V553I/nsp14-ExoN(-) reached nearly 100% of the sample population. None of these mutations were present in the fragments used for recovery, and the nsp3



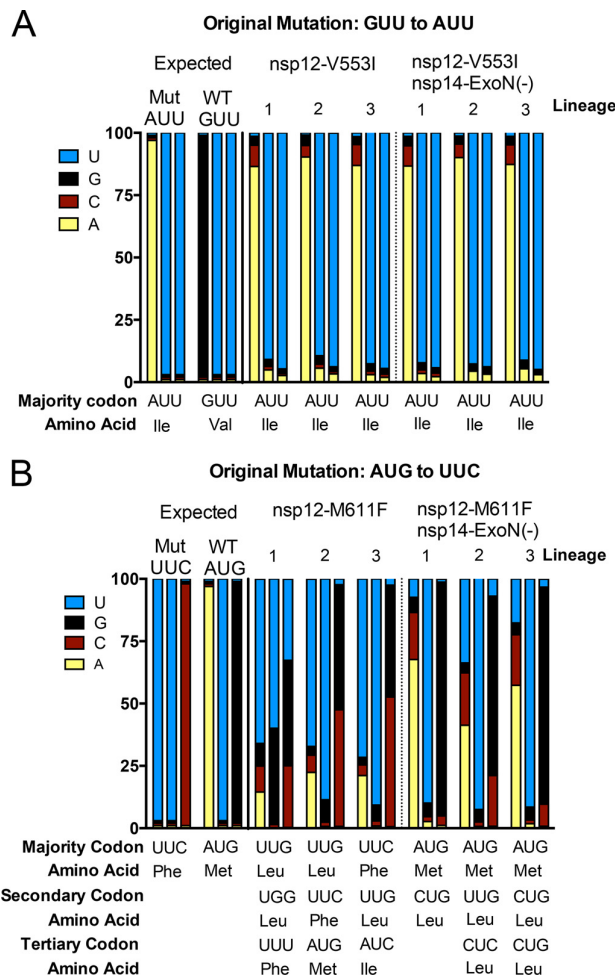


FIG 6 The nsp12-V553I mutation is stable across passages; however, nsp12-M611F is vulnerable to reversion. DBT cells were infected at an initial MOI of 0.01 and then blind passaged in triplicate for 5 passages. Total RNA was taken, and sequencing was performed across a 1.7-kb region of nsp12-RdRp that included both mutations. The percentage of each nucleotide present in each of the triplicate lineages after 5 passages is shown. Mutant viruses in the WT and nsp14-ExoN(-) backgrounds are shown. The original mutation for each of the viruses is shown above the graph, and the likely majority, secondary and tertiary codons present in the population are shown below the graph.

mutation that arose in the nsp14-ExoN(-) population was not present prior to the final low-MOI infection. It is possible that these mutations provided some benefit to the viruses, since they were fixed so rapidly in the population. We next determined whether either of the mutations resulted in a change in the types of mutations occurring during replication (Fig. 8H). Consistent with our previous studies, there were differences in the types of mutations incorporated when WT and nsp14-ExoN(-) backgrounds were compared (30). However, the addition of nsp12-V553I or nsp12-M611F did not alter these patterns in either the WT or nsp14-ExoN(-) background. Thus, nsp12-V553I results in an overall decrease in the accumulation of mutations over passages while nsp12-M611F seems to increase the number of mutations accumulated. These results confirm by sequence analysis the results from the 5-FU and 5-AZC resistance experiments for nsp12-

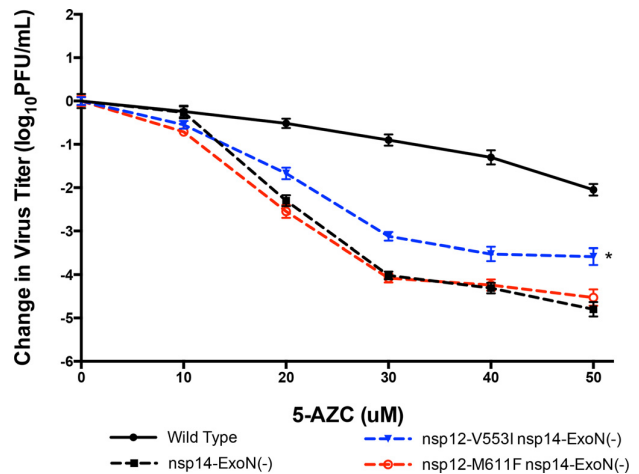
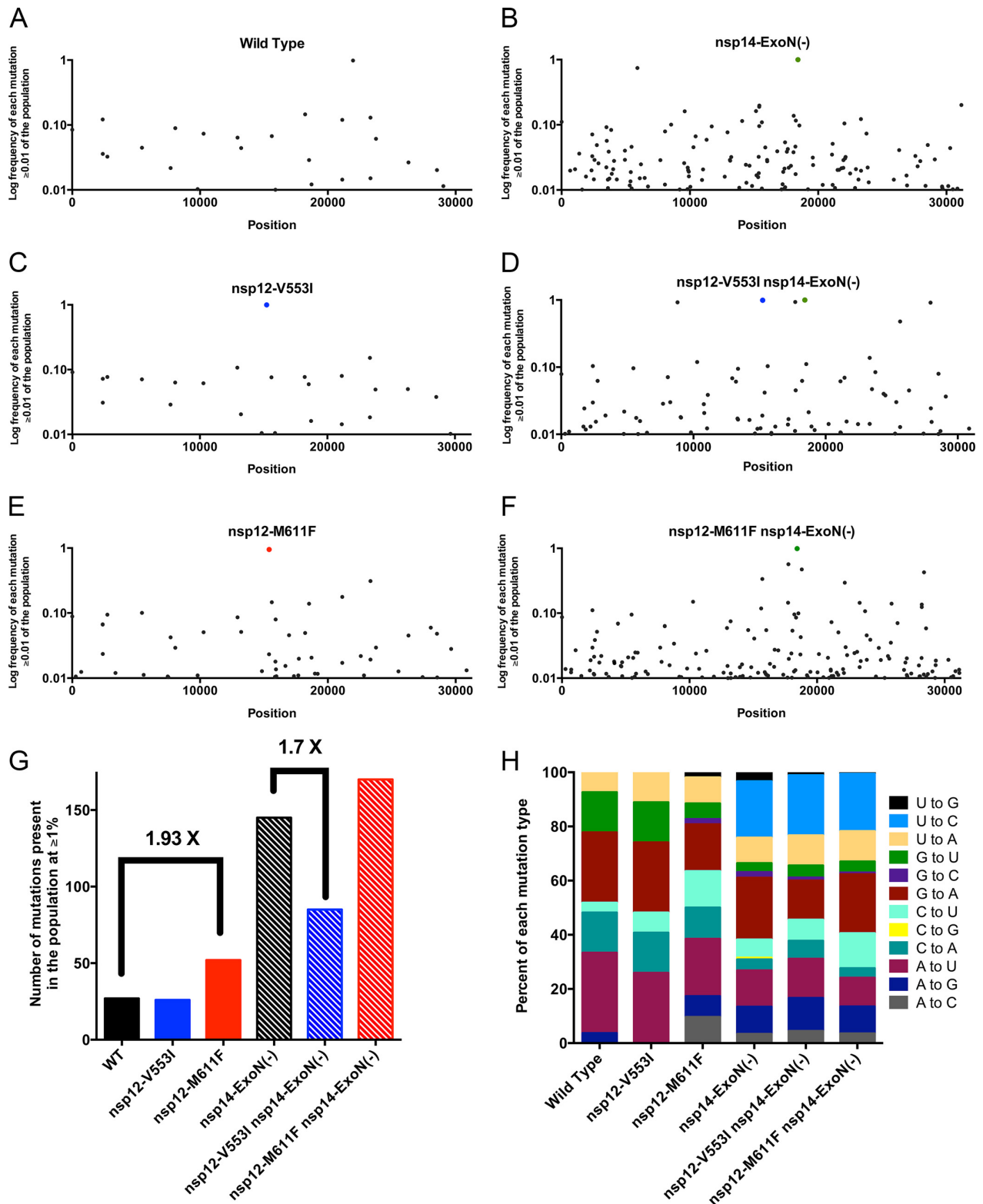


FIG 7 Resistance of MHV nsp12-RdRp V553I and M611F mutant viruses to 5-azacytidine in the nsp14-ExoN(-) background. The domain locations of mutations are indicated as follows: fingers, blues; palm, reds. DBT cells were pretreated with different concentrations of 5-AZC for 30 min. The treatment was removed, and the cells were infected with the indicated viruses at an MOI of 0.01. The medium containing 5-AZC was replaced 30 min p.i. Virus samples were taken at 32 h p.i., and the titer was determined by plaque assay. The data represent the results of 5 independent experiments, each with 2 replicates. The error bars represent SEM (\*,  $P < 0.05$  by ratio paired  $t$  test).

V553I, specifically, showing that the effects of nsp12-V553I are dependent on inactivation of nsp14-ExoN for their detection and that nsp12-V553I likely confers broad resistance to the incorporation of incorrect nucleotides. In contrast, these results further increase the complexity of the nsp12-M611F mutation in relation to incorporation of nucleotides and their analogs. We conclude that nsp12-M611F confers resistance to 5-FU but that this resistance is not likely to be due to broad resistance to the incorporation of alternate nucleotides.

## DISCUSSION

RdRp structures of divergent RNA viruses are structurally conserved and likely have common determinants of activity in the finger, palm, and thumb domains. Positive-strand RNA virus polymerases appear to utilize a common palm domain-based mechanism for active-site closure (23), and associated molecular determinants of fidelity in different RdRp domains have been proposed based on biochemical and mutagenesis studies (10). To date, there are no solved crystal structures of any CoV RdRp, and thus, direct comparison with other virus RdRp structures has not been possible. Further, regulation of CoV fidelity is likely dependent on multiple proteins, including the RdRp and proofreading ExoN. Thus, it was not clear that a CoV would phenotypically exhibit effects from mutating fidelity-determining residues located in the RdRp itself. In this study, we sought to determine whether we could use structure and mutagenesis data from distantly related RNA viruses to identify determinants of CoV nsp12-RdRp fidelity. Our results suggest that CoV RdRps do in fact participate in fidelity regulation at residues orthologous to those in the picornaviruses and likely in other RNA virus RdRps. The results also define, for the first time, a CoV RdRp determinant that increases resistance to multiple mutagens, decreases the accumulation of mutations over time, and so likely increases overall fidel-



**FIG 8** The nsp12-V553I mutation confers decreased accumulation of mutations in the nsp14-ExoN(-) background with no bias toward the exclusion of specific nucleotides. DBT cells were infected at an MOI of 0.01, and total RNA was collected. Deep sequencing was performed on the samples. The statistically significant mutations present at  $\geq 1\%$  of the total population are shown for the wild type, nsp14-ExoN, and nsp12-V553I or nsp12-M611F in both backgrounds. They are graphed according to their distribution across the genome (A to F), with intentionally introduced mutations shown with circles colored blue (nsp12-V553I), red (nsp12-M611F), or green [nsp14-ExoN(-)] (B to F); as the total number of mutations present in the population (G); and as the percentage of specific mutations present (H).

ity. Additionally, the data presented in this paper suggest that CoV RdRp-mediated increased fidelity is detectable only when nsp14-ExoN is inactive and only partially compensates for the loss of nsp14-ExoN high fidelity. Finally, both the nsp12-V553I and M611F mutations confer a replication cost in the WT background, but only nsp12-M611F confers a fitness disadvantage in the nsp14-ExoN(-) background. Together, the results suggest that nsp14-ExoN proofreading activity is epistatic to nsp12-RdRp fidelity but that in contrast, the replication defects of an inactive nsp14-ExoN are epistatic to replication defects in nsp12-RdRp.

**Determinants of nucleotide selectivity and fidelity in CoVs may be conserved with other RNA viruses.** Picornavirus functional RdRps contain only RdRp domains (58, 59). This is not the case for many viral RdRps, including CoVs (26, 39). In addition to the predicted RdRp core domain, all CoV nsp12 proteins contain a “CoV-specific” domain of over 350 amino acids at the N terminus of the protein. A nucleotidyltransferase activity was identified in this CoV-specific domain and has been shown to be important in SARS-CoV replication, but the specific function in replication remains to be determined (39). The N-terminal CoV-specific domain could not be modeled due to a lack of evolutionary homologues with known structures, but the predicted RdRp core domain could be modeled with high confidence using bioinformatic approaches. The structures that provided the best models for the CoV RdRp were from picornaviruses, including enterovirus 71, foot and mouth disease virus, and coxsackie virus B3. This allowed direct alignment and comparison of the known fidelity determinants in CVB3 with the MHV nsp12 core domain across both finger and palm domains. The recovery of mutations in these structurally conserved residues and their participation in CoV nucleotide selectivity support the hypothesis that there are determinants of base specificity conserved between CoVs and distantly related RNA viruses, specifically the picornaviruses. This supports the idea that all RdRps function similarly due to their structural conservation and despite the low level of sequence similarity and attached domains (54). It also suggests that the CoV RdRp domain folds in a manner similar to that of other RdRps, likely separate from the CoV-specific domain.

RdRp fidelity and nucleotide selectivity have been investigated extensively in picornaviruses, especially poliovirus (12, 15, 51, 53, 60, 61). However, even between picornaviruses, the impact on the fidelity of changes at identical or similar residues can vary dramatically (10). Mutations at the same residues in the RdRp of poliovirus and CVB3 affect fidelity differently; poliovirus mutations generally result in increased fidelity and CVB3 mutations in decreased fidelity (11). Specifically, the residue structurally orthologous to nsp12-M611 in poliovirus is F230, and in CVB3 it is I230; when F230 is mutated to Ile in poliovirus, it results in an increase in fidelity, but when I230 in CVB3 is mutated to Phe, Trp, or Val, it results in a decrease in fidelity (10). Similarly, the residue orthologous to nsp12-V553 in CVB3 is I176; when I176 is mutated to Val in CVB3, it results in a decrease in fidelity (10, 11). Although the specific substitutions differ, when nsp12-V553I and nsp12-M611F are introduced into the MHV genome, the resulting 5-FU resistance, and for nsp12-V553I a likely increase in fidelity, mimics results seen in poliovirus rather than CVB3. However, nsp12-M611F seemed to accumulate more mutations over time than controls, and many of the predicted mutations were recoverable only in the WT background; therefore, changes in nucleotide selectivity and fidelity were not tested, so it remains possible that the

mutations that were nonviable in the nsp14-ExoN(-) background had decreased fidelity and that nsp14-ExoN(-) treated with 5-FU defines the error threshold for CoVs.

**Coronavirus nsp12-RdRp and nsp14-ExoN cooperate to optimize both fidelity and replication kinetics.** In addition to nsp12-RdRp, CoVs encode nsp14-ExoN, which functions as a proofreading enzyme (29, 30, 33, 35). Beyond nsp14-ExoN, there are additional CoV-encoded nsp12s that potentially contribute to overall fidelity, such as the small-molecule modulator of nsp14-ExoN encoded in nsp10 (33, 62), nsp7, and nsp8, which together function as an elongation factor (63) and a primase (64, 65), and nsp13, which functions as a helicase (66). These proteins interact with each other and likely function as a multiprotein replication-fidelity complex (34, 38). The results of this study show that while determinants in nsp12-RdRp are likely capable of increasing fidelity, the changes are detectable only in the setting of loss of nsp14-ExoN proofreading and do not completely compensate for the impaired fidelity associated with nsp14-ExoN(-). If altering nucleotide selectivity determinants within nsp12-RdRp no longer significantly affects the overall nucleotide selectivity of the WT virus, then the evolutionary pressure on nsp12-RdRp may be more heavily weighted toward other aspects of replication, such as speed. Recent evidence suggests that RdRp fidelity and speed have an inverse relationship (10, 67). Therefore, if fidelity regulation by nsp14-ExoN occurs more rapidly than correct nucleotide selectivity by the RdRp, then CoV nsp12-RdRp may have been selected specifically for replication speed. Our data may support this hypothesis, as in the WT background, both nsp12-V553I and nsp12-M611F resulted in increased replication lag phases (Fig. 3A and E), though early RNA synthesis was not observably different in our system (Fig. 3C). However, this was not seen in the nsp14-ExoN(-) background, which could be explained by our data indicating that replication is already slowed when nsp14-ExoN is inactivated. One possible mechanism for this could be that the nsp14-ExoN(-) protein is trying to remove incorrect nucleotides and is stalling replication due to its inability to do so. In this case, nsp14-ExoN(-) would be epistatic to nsp12-RdRp in relation to speed, thereby obscuring decreases in replication speed caused by mutations in nsp12-RdRp itself.

**Conclusion.** Our results support the hypothesis that determinants of nucleotide selectivity are conserved across viral orders, identify the first likely increased fidelity determinant for CoV nsp12-RdRp, and demonstrate that nsp14-ExoN proofreading activity is epistatic to nsp12-RdRp nucleotide selectivity. The possibility that some fidelity determinants may be conserved across viral orders is an exciting discovery, as many fidelity determinants identified so far have resulted in attenuation (11, 15, 54, 68–70). It would be interesting to determine whether nsp12-V553I is also attenuated. We predict that it is, based on the fitness cost *in vitro*. However, there is also no clear increase in nucleotide selectivity or fidelity in the WT background. This may be due to a very minimal change that is not measurable even with the deep-sequencing technology used in this study. Alternatively, it could be that any change is simply overwhelmed by the high fidelity of intact ExoN proofreading. In any case, these mutations or other changes at these residues may allow selection of viruses that replicate with normal kinetics *in vitro* and *in vivo* yet confer attenuation in an animal setting. We know that the nsp14-ExoN(-) mutations confer genotypically and phenotypically stable attenuation *in vivo* (36). However, the concept of a high-level mutator as a mecha-

nism for attenuation in live viruses may be problematic. The identification of increased fidelity mutations in the RdRp that can partially, or potentially completely, compensate for the fidelity impairment of nsp14-ExoN(-) viruses, may allow the development of approaches that can benefit from the stability of the nsp14-ExoN(-) mutator phenotype while allowing more stability to the input genomes. Finally, these results, combined with those from previous work (33, 44), suggest that CoVs encode at least three proteins involved in fidelity (nsp12-RdRp, nsp14-ExoN, and nsp10), supporting the assembly of a multiprotein replicase-fidelity complex, as described previously (38). This increases the importance of establishing a biochemical model of the multiprotein complex to directly test the interactions of fidelity determinants, as well as potential inhibitors of each or all of these functions.

#### ACKNOWLEDGMENTS

We thank Xiaotao Lu for technical assistance. We also thank James Brett Case and Erica Andres for critical reviews of the manuscript.

National Institutes of Health (NIH) grant R01 AI108197 (M.R.D.) provided support for this work. N.R.S. was supported by NIH Virology Training Program grant T32 AI089554. E.C.S. was supported by NIH Childhood Infection Research Program grant T32 AI095202. This work was also supported by the Elizabeth B. Lamb Center for Pediatric Research.

#### FUNDING INFORMATION

This work, including the efforts of Nicole Sexton, was funded by HHS | NIH | National Institute of Allergy and Infectious Diseases (NIAID) (T32 AI089554). This work, including the efforts of Everett Clinton Smith, was funded by HHS | NIH | National Institute of Allergy and Infectious Diseases (NIAID) (T32 AI095202). This work, including the efforts of Nicole Sexton, Everett Clinton Smith, and Mark R. Denison, was funded by HHS | NIH | National Institute of Allergy and Infectious Diseases (NIAID) (R01 AI10897).

#### REFERENCES

- Sanjuán R, Nebot MR, Chirico N, Mansky LM, Belshaw R. 2010. Viral mutation rates. *J Virol* 84:9733–9748. <http://dx.doi.org/10.1128/JVI.00694-10>.
- Crotty S, Cameron CE, Andino R. 2001. RNA virus error catastrophe: direct molecular test by using ribavirin. *Proc Natl Acad Sci U S A* 98:6895–6900. <http://dx.doi.org/10.1073/pnas.111085598>.
- Castro C, Arnold JJ, Cameron CE. 2005. Incorporation fidelity of the viral RNA-dependent RNA polymerase: a kinetic, thermodynamic and structural perspective. *Virus Res* 107:141–149. <http://dx.doi.org/10.1016/j.virusres.2004.11.004>.
- Schaaper RM. 1993. Base selection, proofreading, and mismatch repair during DNA replication in *Escherichia coli*. *J Biol Chem* 268:23762–23765.
- Smith EC, Denison MR. 2012. Implications of altered replication fidelity on the evolution and pathogenesis of coronaviruses. *Curr Opin Virol* 2:519–524. <http://dx.doi.org/10.1016/j.coviro.2012.07.005>.
- Sanjuán R. 2012. From molecular genetics to phylodynamics: evolutionary relevance of mutation rates across viruses. *Plos Pathog* 8:e1002685. <http://dx.doi.org/10.1371/journal.ppat.1002685>.
- Domingo E. 2010. Mechanisms of viral emergence. *Vet Res* 41:38. <http://dx.doi.org/10.1051/vetres/2010010>.
- Eigen M. 1993. The origin of genetic information: viruses as models. *Gene* 135:37–47. [http://dx.doi.org/10.1016/0378-1119\(93\)90047-7](http://dx.doi.org/10.1016/0378-1119(93)90047-7).
- Arnold JJ, Vignuzzi M, Stone JK, Andino R, Cameron CE. 2005. Remote site control of an active site fidelity checkpoint in a viral RNA-dependent RNA polymerase. *J Biol Chem* 280:25706–25716. <http://dx.doi.org/10.1074/jbc.M503444200>.
- Campagnola G, McDonald S, Beaucourt S, Vignuzzi M, Peersen OB. 2015. Structure-function relationships underlying the replication fidelity of viral RNA-dependent RNA polymerases. *J Virol* 89:275–286. <http://dx.doi.org/10.1128/JVI.01574-14>.
- Gnädig NF, Beaucourt S, Campagnola G, Borderia AV, Sanz-Ramos M, Gong P, Blanc H, Peersen OB, Vignuzzi M. 2012. Cocksackievirus B3 mutator strains are attenuated in vivo. *Proc Natl Acad Sci U S A* 109:E2294–E2303. <http://dx.doi.org/10.1073/pnas.1204022109>.
- Vignuzzi M, Stone JK, Andino R. 2005. Ribavirin and lethal mutagenesis of poliovirus: molecular mechanisms, resistance and biological implications. *Virus Res* 107:173–181. <http://dx.doi.org/10.1016/j.virusres.2004.11.007>.
- Dapp MJ, Heineman RH, Mansky LM. 2013. Interrelationship between HIV-1 fitness and mutation rate. *J Mol Biol* 425:41–53. <http://dx.doi.org/10.1016/j.jmb.2012.10.009>.
- Smith EC, Sexton NR, Denison MR. 2014. Thinking outside the triangle: replication fidelity of the largest RNA viruses. *Annu Rev Virol* 1:111–132. <http://dx.doi.org/10.1146/annurev-virology-031413-085507>.
- Pfeiffer JK, Kirkegaard K. 2005. Increased fidelity reduces poliovirus fitness and virulence under selective pressure in mice. *Plos Pathog* 1:e11. <http://dx.doi.org/10.1371/journal.ppat.0010011>.
- Severson WE, Schmaljohn CS, Javadian A, Jonsson CB. 2003. Ribavirin causes error catastrophe during Hantaan virus replication. *J Virol* 77:481–488. <http://dx.doi.org/10.1128/JVI.77.1.481-488.2003>.
- Bruenn JA. 2003. A structural and primary sequence comparison of the viral RNA-dependent RNA polymerases. *Nucleic Acids Res* 31:1821–1829. <http://dx.doi.org/10.1093/nar/gkg277>.
- Lang DM, Zemla AT, Zhou CLE. 2013. Highly similar structural frames link the template tunnel and NTP entry tunnel to the exterior surface in RNA-dependent RNA polymerases. *Nucleic Acids Res* 41:1464–1482. <http://dx.doi.org/10.1093/nar/gks1251>.
- Butcher SJ, Grimes JM, Makeyev EV, Bamford DH, Stuart DI. 2001. A mechanism for initiating RNA-dependent RNA polymerization. *Nature* 410:235–240. <http://dx.doi.org/10.1038/35065653>.
- Ng KKS, Arnold JJ, Cameron CE. 2008. Structure-function relationships among RNA-dependent RNA polymerases. *Curr Top Microbiol Immunol* 320:137–156.
- Ferrer-Orta C, Arias A, Perez-Luque R, Escarmis C, Domingo E, Verdaguier N. 2007. Sequential structures provide insights into the fidelity of RNA replication. *Proc Natl Acad Sci U S A* 104:9463–9468. <http://dx.doi.org/10.1073/pnas.0700518104>.
- Ferrer-Orta C, Ferrero D, Verdaguier N. 2015. RNA-dependent RNA polymerases of picornaviruses: from the structure to regulatory mechanisms. *Viruses* 7:4438–4460. <http://dx.doi.org/10.3390/v7082829>.
- Gong P, Peersen OB. 2010. Structural basis for active site closure by the poliovirus RNA-dependent RNA polymerase. *Proc Natl Acad Sci U S A* 107:22505–22510. <http://dx.doi.org/10.1073/pnas.1007626107>.
- Brockway SM, Clay CT, Lu XT, Denison MR. 2003. Characterization of the expression, intracellular localization, and replication complex association of the putative mouse hepatitis virus RNA-dependent RNA polymerase. *J Virol* 77:10515–10527. <http://dx.doi.org/10.1128/JVI.77.19.10515-10527.2003>.
- Nagy PD, Barajas D, Pogany J. 2012. Host factors with regulatory roles in tombusvirus replication. *Curr Opin Virol* 2:691–698. <http://dx.doi.org/10.1016/j.coviro.2012.10.004>.
- te Velthuis AJW. 2014. Common and unique features of viral RNA-dependent polymerases. *Cell Mol Life Sci* 71:4403–4420. <http://dx.doi.org/10.1007/s00018-014-1695-z>.
- Peiris JSM, Lai ST, Poon LLM, Guan Y, Yam LYC, Lim W, Nicholls J, Yee WKS, Yan WW, Cheung MT, Cheng VCC, Chan KH, Tsang DNC, Yung RWH, Ng TK, Yuen KY, SARS Study Group. 2003. Coronavirus as a possible cause of severe acute respiratory syndrome. *Lancet* 361:1319–1325. [http://dx.doi.org/10.1016/S0140-6736\(03\)13077-2](http://dx.doi.org/10.1016/S0140-6736(03)13077-2).
- Zaki AM, van Boheemen S, Bestebroer TM, Osterhaus ADME, Fouchier RAM. 2012. Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. *N Engl J Med* 367:1814–1820. <http://dx.doi.org/10.1056/NEJMoa1211721>.
- Eckerle LD, Becker MM, Halpin RA, Li K, Venter E, Lu X, Scherbakova S, Graham RL, Baric RS, Stockwell TB, Spiro DJ, Denison MR. 2010. Infidelity of SARS-CoV Nsp14-exonuclease mutant virus replication is revealed by complete genome sequencing. *Plos Pathog* 6:e1000896. <http://dx.doi.org/10.1371/journal.ppat.1000896>.
- Smith EC, Blanc H, Vignuzzi M, Denison MR. 2013. Coronaviruses lacking exoribonuclease activity are susceptible to lethal mutagenesis: ev-



- idence for proofreading and potential therapeutics. *Plos Pathog* 9:e1003565. <http://dx.doi.org/10.1371/journal.ppat.1003565>.
31. Gorbalenya AE, Enjuanes L, Ziebuhr J, Snijder EJ. 2006. Nidovirales: evolving the largest RNA virus genome. *Virus Res* 117:17–37. <http://dx.doi.org/10.1016/j.virusres.2006.01.017>.
  32. Stenglein MD, Jacobson ER, Wozniak EJ, Wellehan JFX, Kincaid A, Gordon M, Porter BF, Baumgartner W, Stahl S, Kelley K, Towner JS, Derisi JL. 2014. Ball python nidovirus: a candidate etiologic agent for severe respiratory disease in Python regius. *mBio* 5:e01484–14. <http://dx.doi.org/10.1128/mBio.01484-14>.
  33. Bouvet M, Imbert I, Subissi L, Gluais L, Canard B, Decroly E. 2012. RNA 3'-end mismatch excision by the severe acute respiratory syndrome coronavirus nonstructural protein nsp10/nsp14 exoribonuclease complex. *Proc Natl Acad Sci U S A* 109:9372–9377. <http://dx.doi.org/10.1073/pnas.1201130109>.
  34. Smith EC, Denison MR. 2013. Coronaviruses as DNA wannabes: a new model for the regulation of RNA virus replication fidelity. *Plos Pathog* 9:e1003760. <http://dx.doi.org/10.1371/journal.ppat.1003760>.
  35. Eckerle LD, Lu X, Sperry SM, Choi L, Denison MR. 2007. High fidelity of murine hepatitis virus replication is decreased in nsp14 exoribonuclease mutants. *J Virol* 81:12135–12144. <http://dx.doi.org/10.1128/JVI.01296-07>.
  36. Graham RL, Becker MM, Eckerle LD, Bolles M, Denison MR, Baric RS. 2012. A live, impaired-fidelity coronavirus vaccine protects in an aged, immunocompromised mouse model of lethal disease. *Nat Med* 18:1820–1826. <http://dx.doi.org/10.1038/nm.2972>.
  37. Denison MR, Graham RL, Donaldson EF, Eckerle LD, Baric RS. 2011. Coronaviruses: an RNA proofreading machine regulates replication fidelity and diversity. *RNA Biol* 8:270–279. <http://dx.doi.org/10.4161/rna.8.2.15013>.
  38. Subissi L, Posthuma CC, Collet A, Zevenhoven-Dobbe JC, Gorbalenya AE, Decroly E, Snijder EJ, Canard B, Imbert I. 2014. One severe acute respiratory syndrome coronavirus protein complex integrates processive RNA polymerase and exonuclease activities. *Proc Natl Acad Sci U S A* 111:E3900–E3909. <http://dx.doi.org/10.1073/pnas.1323705111>.
  39. Lehmann KC, Gulyaeva A, Zevenhoven-Dobbe JC, Janssen GMC, Ruben M, Overkleef HS, van Veelen PA, Samborskiy DV, Kravchenko AA, Leontovich AM, Sidorov IA, Snijder EJ, Posthuma CC, Gorbalenya AE. 2015. Discovery of an essential nucleotidylating activity associated with a newly delineated conserved domain in the RNA polymerase-containing protein of all nidoviruses. *Nucleic Acids Res* 43:8416–8434. <http://dx.doi.org/10.1093/nar/gkv838>.
  40. Xu X, Liu Y, Weiss S, Arnold E, Sarafianos SG, Ding J. 2003. Molecular model of SARS coronavirus polymerase: implications for biochemical functions and drug design. *Nucleic Acids Res* 31:7117–7130. <http://dx.doi.org/10.1093/nar/gkg916>.
  41. Chen W, Baric RS. 1996. Molecular anatomy of mouse hepatitis virus persistence: coevolution of increased host cell resistance and virus virulence. *J Virol* 70:3947–3960.
  42. Yount B, Denison MR, Weiss SR, Baric RS. 2002. Systematic assembly of a full-length infectious cDNA of mouse hepatitis virus strain A59. *J Virol* 76:11065–11078. <http://dx.doi.org/10.1128/JVI.76.21.11065-11078.2002>.
  43. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE. 2015. The Phyre2 Web portal for protein modeling, prediction and analysis. *Nat Protoc* 10:845–858. <http://dx.doi.org/10.1038/nprot.2015.053>.
  44. Smith EC, Case JB, Blanc H, Isakov O, Shomron N, Vignuzzi M, Denison MR. 2015. Mutations in coronavirus nonstructural protein 10 decrease virus replication fidelity. *J Virol* 89:6418–6426. <http://dx.doi.org/10.1128/JVI.00110-15>.
  45. Isakov O, Bordería AV, Golan D, Hamenahem A, Celniker G, Yoffe L, Blanc H, Vignuzzi M, Shomron N. 2015. Deep sequencing analysis of viral infection and evolution allows rapid and detailed characterization of viral mutant spectrum. *Bioinformatics* 31:2141–2150. <http://dx.doi.org/10.1093/bioinformatics/btv101>.
  46. Li H, Durbin R. 2010. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* 26:589–595. <http://dx.doi.org/10.1093/bioinformatics/btp698>.
  47. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup. 2009. The sequence alignment/map format and SAMtools. *Bioinformatics* 25:2078–2079.
  48. Bordería AV, Isakov O, Moratorio G, Henningson R, Agüera-González S, Organtini L, Gnädig NF, Blanc H, Alcover A, Hafenstein S, Fontes M, Shomron N, Vignuzzi M. 2015. Group selection and contribution of minority variants during virus adaptation determines virus fitness and phenotype. *Plos Pathog* 11:e1004838. <http://dx.doi.org/10.1371/journal.ppat.1004838>.
  49. Pauly MD, Lauring AS. 2015. Effective lethal mutagenesis of influenza virus by three nucleoside analogs. *J Virol* 89:3584–3597. <http://dx.doi.org/10.1128/JVI.03483-14>.
  50. Van Slyke GA, Arnold JJ, Lugo AJ, Griesemer SB, Moustafa IM, Kramer LD, Cameron CE, Ciota AT. 2015. Sequence-specific fidelity alterations associated with West Nile virus attenuation in mosquitoes. *Plos Pathog* 11:e1005009. <http://dx.doi.org/10.1371/journal.ppat.1005009>.
  51. Verdaguer N, Ferrer-Orta C. 2012. Conformational changes in motif D of RdRPs as fidelity determinant. *Structure* 20:1448–1450. <http://dx.doi.org/10.1016/j.str.2012.08.014>.
  52. Levi LI, Gnädig NF, Beaucourt S, McPherson MJ, Baron B, Arnold JJ, Vignuzzi M. 2010. Fidelity variants of RNA dependent RNA polymerases uncover an indirect, mutagenic activity of amiloride compounds. *Plos Pathog* 6:e1001163. <http://dx.doi.org/10.1371/journal.ppat.1001163>.
  53. Pfeiffer JK, Kirkegaard K. 2003. A single mutation in poliovirus RNA-dependent RNA polymerase confers resistance to mutagenic nucleotide analogs via increased fidelity. *Proc Natl Acad Sci U S A* 100:7289. <http://dx.doi.org/10.1073/pnas.1232294100>.
  54. Yang X, Smidansky ED, Maksimchuk KR, Lum D, Welch JL, Arnold JJ, Cameron CE, Boehr DD. 2012. Motif D of viral RNA-dependent RNA polymerases determines efficiency and fidelity of nucleotide addition. *Structure* 20:1519–1527. <http://dx.doi.org/10.1016/j.str.2012.06.012>.
  55. Holland JJ, Domingo E, de la Torre JC, Steinhauer DA. 1990. Mutation frequencies at defined single codon sites in vesicular stomatitis virus and poliovirus can be increased only slightly by chemical mutagenesis. *J Virol* 64:3960–3962.
  56. Rozen-Gagnon K, Stapleford KA, Mongelli V, Blanc H, Failloux A-B, Saleh M-C, Vignuzzi M. 2014. Alphavirus mutator variants present host-specific defects and attenuation in mammalian and insect models. *Plos Pathog* 10:e1003877. <http://dx.doi.org/10.1371/journal.ppat.1003877>.
  57. Arias A, Arnold JJ, Sierra M, Smidansky ED, Domingo E, Cameron CE. 2008. Determinants of RNA-dependent RNA polymerase (in)fidelity revealed by kinetic analysis of the polymerase encoded by a foot-and-mouth disease virus mutant with reduced sensitivity to ribavirin. *J Virol* 82:12346–12355. <http://dx.doi.org/10.1128/JVI.01297-08>.
  58. Hansen JL, Long AM, Schultz SC. 1997. Structure of the RNA-dependent RNA polymerase of poliovirus. *Structure* 5:1109–1122. [http://dx.doi.org/10.1016/S0969-2126\(97\)00261-X](http://dx.doi.org/10.1016/S0969-2126(97)00261-X).
  59. Campagnola G, Weygant M, Scoggin K, Peersen O. 2008. Crystal structure of coxsackievirus B3 3Dpol highlights the functional importance of residue 5 in picornavirus polymerases. *J Virol* 82:9458–9464. <http://dx.doi.org/10.1128/JVI.00647-08>.
  60. Hobdley SE, Kempf BJ, Steil BP, Barton DJ, Peersen OB. 2010. Poliovirus polymerase residue 5 plays a critical role in elongation complex stability. *J Virol* 84:8072–8084. <http://dx.doi.org/10.1128/JVI.02147-09>.
  61. Korneeva VS, Cameron CE. 2007. Structure-function relationships of the viral RNA-dependent RNA polymerase: fidelity, replication speed, and initiation mechanism determined by a residue in the ribose-binding pocket. *J Biol Chem* 282:16135–16145. <http://dx.doi.org/10.1074/jbc.M610090200>.
  62. Chen Y, Su C, Ke M, Jin X, Xu L, Zhang Z, Wu A, Sun Y, Yang Z, Tien P, Ahola T, Liang Y, Liu X, Guo D. 2011. Biochemical and structural insights into the mechanisms of SARS coronavirus RNA ribose 2'-O-methylation by nsp16/nsp10 protein complex. *Plos Pathog* 7:e1002294. <http://dx.doi.org/10.1371/journal.ppat.1002294>.
  63. te Velthuis AJW, van den Worm SHE, Snijder EJ. 2012. The SARS-coronavirus nsp7+nsp8 complex is a unique multimeric RNA polymerase capable of both de novo initiation and primer extension. *Nucleic Acids Res* 40:1737–1747. <http://dx.doi.org/10.1093/nar/gkr893>.
  64. Imbert I, Guillemot J-C, Bourhis J-M, Bussetta C, Coutard B, Egloff M-P, Ferron F, Gorbalenya AE, Canard B. 2006. A second, non-canonical RNA-dependent RNA polymerase in SARS coronavirus. *EMBO J* 25:4933–4942. <http://dx.doi.org/10.1038/sj.emboj.7601368>.
  65. Xiao Y, Ma Q, Restle T, Shang W, Svergun DJ, Ponnusamy R, Sczakiel G, Hilgenfeld R. 2012. Nonstructural proteins 7 and 8 of feline coronavirus form a 2:1 heterotrimer that exhibits primer-independent RNA polymerase activity. *J Virol* 86:4444–4454. <http://dx.doi.org/10.1128/JVI.06635-11>.

66. Adedeji AO, Marchand B, de Velhuis AJW, Snijder EJ, Weiss S, Eoff RL, Singh K, Sarafianos SG. 2012. Mechanism of nucleic acid unwinding by SARS-CoV helicase. *PLoS One* 7:e36521. <http://dx.doi.org/10.1371/journal.pone.0036521>.
67. Regoes RR, Hamblin S, Tanaka MM. 2013. Viral mutation rates: modelling the roles of within-host viral dynamics and the trade-off between replication fidelity and speed. *Proc Biol Sci* 280:20122047. <http://dx.doi.org/10.1098/rspb.2012.2047>.
68. Vignuzzi M, Wendt E, Andino R. 2008. Engineering attenuated virus vaccines by controlling replication fidelity. *Nat Med* 14:154–161. <http://dx.doi.org/10.1038/nm1726>.
69. Lauring AS, Jones JO, Andino R. 2010. Rationalizing the development of live attenuated virus vaccines. *Nat Biotechnol* 28:573–579. <http://dx.doi.org/10.1038/nbt.1635>.
70. Coffey LL, Beeharry Y, Borderia AV, Blanc H, Vignuzzi M. 2011. Arbovirus high fidelity variant loses fitness in mosquitoes and mice. *Proc Natl Acad Sci U S A* 108:16038–16043. <http://dx.doi.org/10.1073/pnas.1111650108>.

**APPENDIX B:**

**THINKING OUTSIDE THE TRIANGLE REPLICATION FIDELITY OF THE  
LARGEST RNA VIRUSES**

# Thinking Outside the Triangle: Replication Fidelity of the Largest RNA Viruses

Everett Clinton Smith,<sup>1,3</sup> Nicole R. Sexton,<sup>2,3</sup>  
and Mark R. Denison<sup>1,2,3</sup>

<sup>1</sup>Department of Pediatrics, <sup>2</sup>Department of Pathology, Microbiology, and Immunology, and  
<sup>3</sup>Elizabeth B. Lamb Center for Pediatric Research, Vanderbilt University Medical Center,  
Nashville, Tennessee 37232; email: mark.denison@vanderbilt.edu

Annu. Rev. Virol. 2014. 1:111–32

First published online as a Review in Advance on  
June 2, 2014

The *Annual Review of Virology* is online at  
virology.annualreviews.org

This article's doi:  
10.1146/annurev-virology-031413-085507

Copyright © 2014 by Annual Reviews.  
All rights reserved

## Keywords

coronavirus, proofreading, exoribonuclease, virus evolution, quasispecies

## Abstract

When judged by ubiquity, adaptation, and emergence of new diseases, RNA viruses are arguably the most successful biological organisms. This success has been attributed to a defect of sorts: high mutation rates (low fidelity) resulting in mutant swarms that allow rapid selection for fitness in new environments. Studies of viruses with small RNA genomes have identified fidelity determinants in viral RNA-dependent RNA polymerases and have shown that RNA viruses likely replicate within a limited fidelity range to maintain fitness. In this review we compare the fidelity of small RNA viruses with that of the largest RNA viruses, the coronaviruses. Coronaviruses encode the first known viral RNA proofreading exoribonuclease, a function that likely allowed expansion of the coronavirus genome and that dramatically increases replication fidelity and the range of tolerated variation. We propose models for regulation of coronavirus fidelity and discuss the implications of altered fidelity for RNA virus replication, pathogenesis, and evolution.

**SARS-CoV:** severe acute respiratory syndrome coronavirus

**MERS-CoV:** Middle East respiratory syndrome coronavirus

**Replication fidelity:** the accuracy with which genetic information, either DNA or RNA, is copied

**CoV:** coronavirus

**Proofreading:** an error-correcting process involving the removal of a mismatched nucleotide during RNA or DNA synthesis

**Zoonotic:** describes an infectious agent capable of transmission between animals and humans

## INTRODUCTION

RNA viruses constitute some of the most ubiquitous and lethal human pathogens. The emergence of Ebola virus, new human influenza viruses, severe acute respiratory syndrome coronavirus (SARS-CoV), and Middle East respiratory syndrome coronavirus (MERS-CoV) underscores the capacity of RNA viruses to cause new human diseases. RNA viruses—notably HIV and influenza—also demonstrate emergence of resistance to antivirals and vaccines. Such adaptive potential results from several characteristics of RNA viruses including rapid replication cycles, enormous population size, and extensive genetic diversity. However, high adaptive potential is not without cost; low-fidelity replication imposes genome size constraints such that most RNA virus genomes do not exceed ~15 kb (1–3). In this review we highlight significant advances in understanding the determinants of RNA virus replication fidelity and the implications of altered fidelity for virus replication, fitness, and pathogenesis. We devote particular attention to the coronaviruses (CoVs), which encode genomes up to 32 kb long and thus are an exception to the constraint on RNA genome size and complexity. Additionally, we summarize recent experimental data and comparative genomics analyses demonstrating that CoVs have evolved a network of replicase proteins that increase the efficiency of RNA replication and regulate replication fidelity through a novel RNA proofreading activity. Finally, we discuss the possibility that CoV RNA replication involves a large multisubunit polymerase similar in organization to replicative DNA polymerase complexes.

## HUMAN CORONAVIRUSES: EMERGENCE, GENOME ORGANIZATION, AND REPLICATION

### Emergence and Human Disease

CoVs cause significant morbidity and mortality in humans (4), from mild common colds to lethal respiratory and systemic diseases. Five CoVs are known to circulate in humans (HCoVs): HCoV-NL63, HCoV-229E, HCoV-OC43, HCoV-HKU1, and MERS-CoV. Although SARS-CoV is not known to be present in human populations, a SARS-like CoV recently was discovered in Chinese horseshoe bats (*Rhinolophidae*) (5). This virus can use human ACE2 (angiotensin-converting enzyme 2) as a receptor and is >99% identical to SARS-CoV, indicating that the immediate SARS-CoV zoonotic precursor is likely present in bat populations. Molecular clock analyses suggest that all endemic HCoVs originated as zoonotic infections and emerged into humans over several hundred years: HCoV-NL63 between 1190 and 1449 CE (6), HCoV-229E between 1501 and 1883 CE (7, 8), and HCoV-OC43 between 1866 and 1918 CE (9). The timing of HCoV-HKU1 emergence remains unknown (10). SARS-CoV was identified in 2003 (11) and MERS-CoV in 2012 (12), both within a year of entering the human population, emphasizing the capacity of CoVs to cross species barriers and thrive in new host environments.

Though CoVs infect a large number of mammalian and avian species (13), the majority of the identified diversity within alpha- and betacoronaviruses exists in bats (14). CoVs are predicted to have coevolved with bats for millions of years, possibly since the evolutionary split between bats and birds (15, 16). All known HCoVs have proposed bat origins with the exception of HCoV-OC43, which emerged from a bovine CoV (6). However, movement of zoonotic CoVs into human populations may require passage or coemergence with other mammalian species, as exemplified by SARS-CoV emergence (17), which involved cyclical passage between civet cats and humans. The closest known relative of HCoV-229E recently was discovered in alpacas, suggesting interspecies transmission to humans might have been through alpacas (8) instead of directly from bats (7).

MERS-CoV is the most recently emerged HCoV, and its transmission history is still being defined. MERS-CoV originally was isolated in June 2012, and new cases continue to be identified, with more than 243 laboratory-confirmed cases as of April 2014 and more than 93 deaths. Emerging data suggest the involvement of an additional species in maintenance of MERS-CoV transmission to humans, and at present dromedary camels are the prime suspects (18, 19). There is clear evidence for limited human-human transmission (20, 21), and the continued evidence for new infections suggests the potential for further adaptation.

### Genome Organization and Replication Strategy

The order Nidovirales contains four families: the Arteriviridae, Coronaviridae, Mesoniviridae, and Roniviridae. CoVs comprise the Coronavirinae subfamily within the Coronaviridae family and demonstrate a remarkable diversity across many animal species. All CoVs have positive-sense single-stranded RNA [(+)ssRNA] genomes, which possess a 5' cap structure and a 3' poly(A) tail and are contained within a pleomorphic host membrane-derived envelope (reviewed in 22). The genome organization is similar for all CoVs, with the viral genome divided roughly into two major regions: the nonstructural protein genes and the structural and accessory protein genes (**Figure 1a**). Compared with other positive-strand RNA viruses, CoVs encode the most extensive ensemble of replicase and transcriptase proteins, which can include up to 16 nonstructural proteins (nsp1–16) (22). Proteins involved in viral replication and transcription compose the first two-thirds of the viral genome and are contained in two open reading frames, ORF1a and ORF1b. ORF1a is translated from genome RNA with every translational initiation, resulting in the expression of polyprotein 1a (pp1a), which is composed of nsp1–11. Translation of ORF1b requires a –1 ribosomal frameshift immediately following the nsp10 coding region, which occurs less than 40% of the time and results in the fusion polyprotein 1ab (pp1ab), which contains nsp1–16 (23, 24). Consequently, nsp1–11 are produced more abundantly than nsp12–16. Processing of both polyproteins by two or three viral proteases yields 16 mature nsps and several intermediate precursors, with the vast majority of these mature nsps known or predicted to function in virus replication complex formation and RNA synthesis (22, 25, 26).

CoV RNA synthesis can be divided into two general stages: genome replication and subgenomic mRNA synthesis (22). Genome replication is initiated following translation of the viral genome by host machinery and processing of replicase nsps by viral proteases. The (+)ssRNA genome is transcribed into a full-length (–)ssRNA intermediate, which is then used as the template for amplification of the (+)ssRNA genome (**Figure 1b**). Expression of the 3' structural and accessory proteins occurs from subgenomic mRNAs (sgmRNAs) initiated by a discontinuous transcription mechanism. Negative-strand subgenomic RNA [(–)sgRNA] templates are generated from the 3' end of the (+)ssRNA genome and are regulated by short transcriptional regulatory sequences (TRSs) located in the 5' untranslated region (UTR) of the (+)ssRNA genome, termed the leader TRS, and by those present immediately upstream of each 3' ORF (**Figure 1b**). Recognition of each 3' TRS by the viral RNA-dependent RNA polymerase (RdRp) can result in either read-through or dissociation of the (–)sgRNA template from the (+)ssRNA genome. The (–)sgRNA-RdRp complex then reassociates with the 5' leader TRS, thus generating a set of (–)sgRNAs that contain one or more ORFs as well as 3' and 5' termini identical to the negative-strand genome template. These (–)sgRNA templates are then used as the primary templates to generate sgmRNAs that possess 5' and 3' sequences identical to each other and to the genome. This nested set of mRNAs is the basis for the name of the order Nidovirales, of which CoVs are members. The sgmRNAs also are used as amplification templates for subsequent rounds of (–)sgRNA synthesis (27).

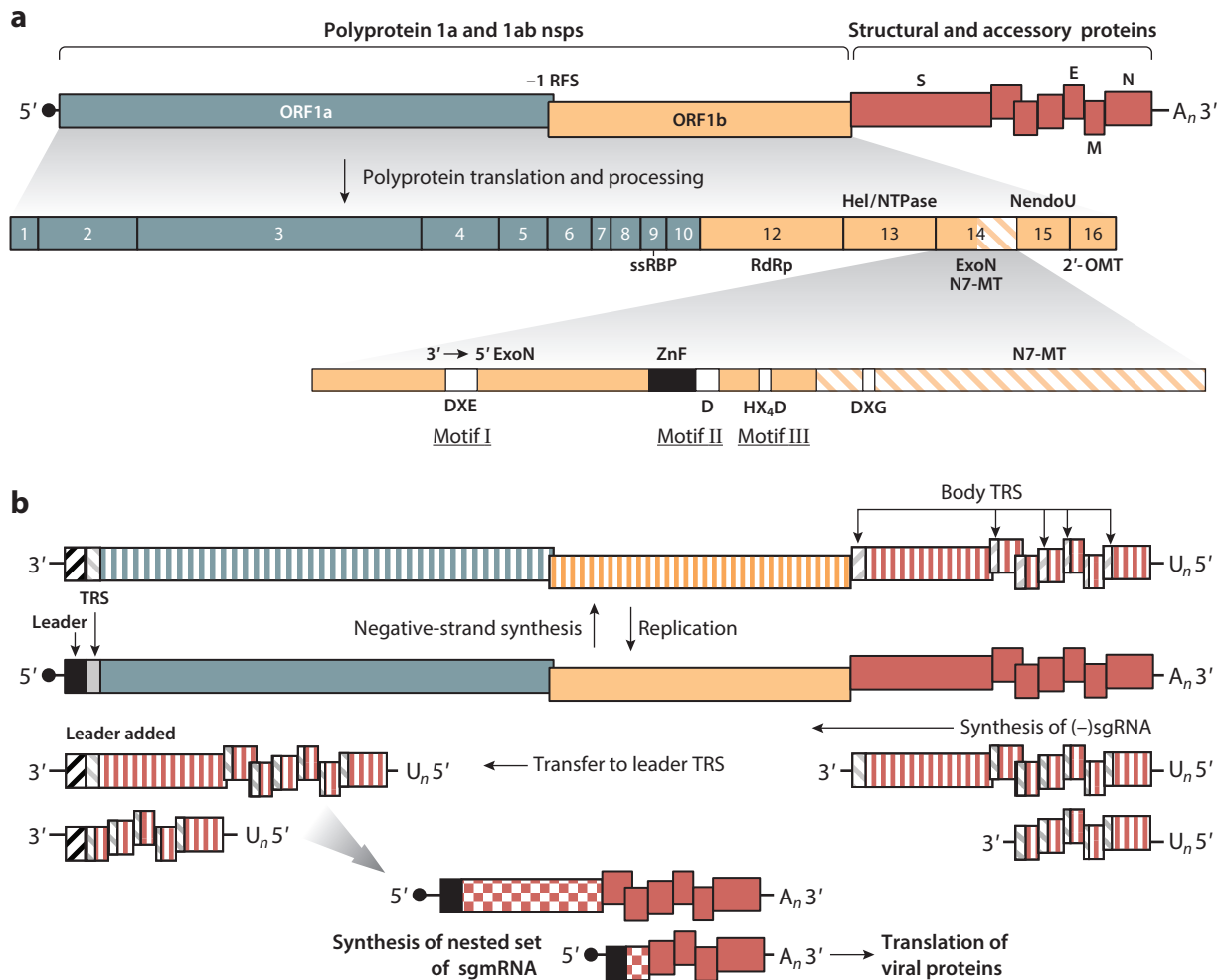
---

**(+)ssRNA:**  
positive-sense  
single-stranded RNA

**Nsp:** nonstructural  
protein

**RdRp:**  
RNA-dependent RNA  
polymerase

---



**Figure 1**

Coronavirus genome organization and replication strategy. (a) Open reading frame (ORF) 1ab encompasses roughly two-thirds of the genome and encodes the replicase nonstructural proteins (nsp1–16). Nsp11 is ~15 amino acids long with no known function and thus is not shown. The other one-third of the genome encodes structural and accessory proteins including the spike (S), envelope (E), matrix (M), and nucleocapsid (N) proteins. The structural and accessory proteins are shown for a generic CoV, as the accessory proteins vary both in number and in position between CoVs. The nonstructural proteins nsp1–11 are translated from ORF1a (*blue*), whereas translation of the ORF1b proteins (nsp12–16) only occurs following a –1 ribosomal frameshift (RFS). Polypeptides containing nsp1–11 (pp1a) or nsp1–16 (pp1ab) are cleaved by up to three viral proteases to generate individual nsps. The functions of several nsps are noted: nsp9, single-stranded RNA-binding protein (ssRBP); nsp12, RNA-dependent RNA polymerase (RdRp); nsp13, helicase (Hel) and NTPase; nsp14, 3'→5' exoribonuclease (ExoN) and N7-methyltransferase (N7-MT); nsp15, uridylylate-specific endonuclease (NendoU); nsp16, 2'-O-methyltransferase (2'-OMT). The active site residues of nsp14 involved in proofreading are distributed across three motifs (I–III). Additionally, a zinc finger domain (ZnF) is encoded, which is unique among the DEDD superfamily of exonucleases to which nsp14 ExoN belongs. (b) CoV replication (*top*) involves the production of full-length negative-strand intermediates (*striped*) that are used to generate new positive-strand genomes (*solid*). Transcription and translation (*bottom*) of the 3' ORFs requires the generation of negative-strand subgenomic RNAs [(-)sgRNAs] (*striped*) that are then used to make viral subgenomic messenger RNA (sgmRNA). Only the 5'-most ORF (*checkered*) is translated for each sgmRNA. Abbreviation: TRS, transcriptional regulatory sequence.



## Replicase Proteins and Novel RNA-Modifying Functions

Studies of the CoV RNA synthesis machinery have focused on the identification and functions of nsp1–16 (**Figure 1a**). Several CoV nsps have been shown to play prominent roles in RNA synthesis and modification by harboring one or more enzymatic activities: the nsp8 primase (28, 29), nsp12 RdRp (30–32), nsp13 helicase (Hel) and NTPase (33, 34), nsp14 N7-methyltransferase (N7-MT) (35) and exoribonuclease (ExoN) (36), nsp15 uridylyate-specific endoribonuclease (NendoU) (37), and nsp16 RNA 2'-O-methyltransferase (2'-OMT) (38). Other viral proteins without predicted or confirmed enzymatic activities also contribute to viral RNA synthesis and modification, including nsp7, which interacts with nsp8 to form a large hexadecameric ring-like structure (39), and nsp10, which is critical for viral RNA synthesis (40) and interacts with both nsp14 and nsp16 (41–44). The nsp12 RdRp is the catalytic core of the CoV RNA synthesis machinery and consists of at least two domains: an N-terminal domain unique to CoVs and a C-terminal catalytic domain that is predicted to adopt a structure similar to other viral RdRps (45). CoV nsp12 contains an SDD active site motif that is conserved in all members of the Nidovirales order (46). Work by de Velthuis et al. (30), using full-length nsp12 containing a native N terminus, demonstrated that nsp12 contains weak primer-dependent RdRp activity; another group (32) has reported that nsp12 RdRp can initiate de novo. Primase activity is likely provided by nsp8, which contains low-fidelity RdRp activity similar to known DNA-dependent DNA primases (28, 29). Even less is known about the unique N-terminal region, except that portions or all of the domain is required for in vitro RdRp activity (31) and that it appears to lack homology to any protein identified to date. The encoding of such a diverse repertoire of RNA-modifying functions suggests that CoVs have evolved a network of proteins that might be more analogous to cellular RNA and/or DNA synthesis machinery. This possibility, particularly relating to the mechanisms by which CoVs maintain the integrity of their large RNA genomes, is discussed further below.

---

**ExoN:**  
exoribonuclease

**Mismatch repair:** a postreplicative process distinct from proofreading that detects and repairs misincorporated or damaged bases

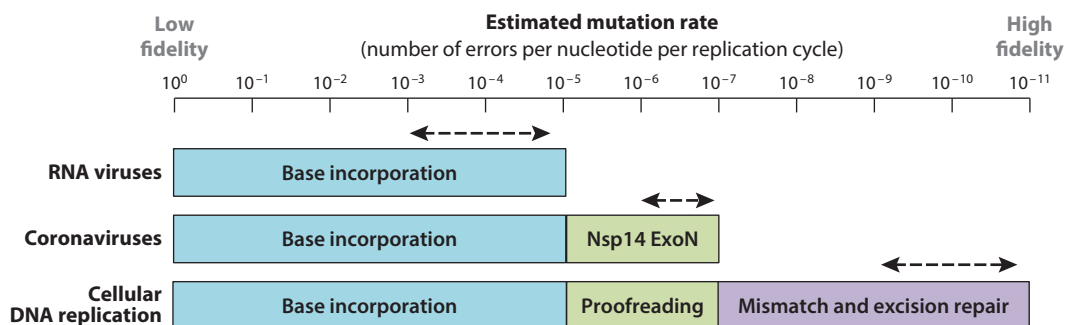
---

## RNA VIRUS REPLICATION FIDELITY: DETERMINANTS AND THE EFFECT ON VIRAL POPULATIONS

### A Comparison with Cellular Organisms

Both cellular life forms and all viruses must faithfully replicate their genomes to ensure the transmission of genetic information. For cellular organisms, faithful reproduction of genetic material is mediated by a network of proteins tasked with detecting and repairing errors during or subsequent to DNA synthesis. In its entirety, this process is referred to as replication fidelity, and it encompasses at least three critical steps: (a) selection and extension of the correct nucleotide onto the replicating DNA strand by a DNA polymerase, (b) removal of mismatched nucleotides by intrinsic 3'→5' ExoN activity within the DNA polymerase or by a closely associated 3'→5' ExoN-containing subunit, and (c) correction by the cellular mismatch repair system of errors that have escaped proofreading (**Figure 2**) (47–49). Each of these processes contributes to cellular DNA replication fidelity, but the relative contributions of these steps toward the estimated error rate of  $10^{-9}$  to  $10^{-11}$ , or one error per  $10^9$  to  $10^{11}$  nucleotides, vary. Correct nucleotide selection, recognition of a properly formed base pair, and extension by the DNA polymerase together provide a greater contribution to fidelity ( $10^{-5}$ ) than either proofreading ( $10^{-2}$ ) or the mismatch repair system ( $10^{-3}$ ), though the relative contributions are likely error type specific (47, 50, 51). The fidelity of DNA and RNA polymerases likely results more from polymerase dynamics—such as recognition of the shape of correct versus incorrect base pairs and conformational changes within the polymerase active site—than from the capacity of the polymerase to selectively discriminate between nucleotides (52–58). Beyond these mechanisms, DNA-based cellular organisms also have





**Figure 2**

Replication fidelity in RNA viruses and DNA-based organisms. Shown is the estimated range of mutation rates (*dashed arrows*) for RNA viruses, coronaviruses, and cellular DNA replication. The relative contributions of polymerase base incorporation (*cyan*), proofreading (*green*), and mismatch and excision repair (*purple*) toward the estimated mutation rates are shown. See the text for specific references regarding the estimated mutation rates. Abbreviation: nsp14 ExoN, nonstructural protein 14 exoribonuclease.

machinery dedicated to repairing damaged bases, removing UV-induced pyrimidine dimers, and rejoining double-strand breaks. Overall, cellular DNA replication is orchestrated by numerous accessory proteins and by an ensemble of DNA polymerases: 5 for *Escherichia coli*, 8 for *Saccharomyces cerevisiae*, and at least 16 for humans (recently reviewed in 47).

Compared with cellular organisms, RNA viruses are minimalists in terms of replication fidelity. The primary determinant of RNA virus genome replication, with the exception of the retroviruses, is the RdRp. Positive-strand RNA viruses with genomes <20 kb lack proofreading and other postreplicative repair mechanisms (59) and therefore rely on the viral RdRp for maintenance of genome sequence integrity. This reliance on the viral RdRp theoretically limits the fidelity of RNA viruses to an upper limit in the range of  $\sim 10^{-5}$ . In fact, most RNA viruses replicate with estimated error rates between  $10^{-3}$  and  $10^{-5}$ , which results in approximately one mutation per genome per round of replication for a typical  $\sim 10$ -kb genome (2, 60). This fidelity range is similar to that observed with exonuclease-deficient DNA polymerases, and in fact, error-prone Y family DNA polymerases polymerize with fidelities much lower than  $10^{-5}$  (49, 61–63). Thus, the low fidelity of RNA virus genome replication is likely not because RdRps polymerize RNA with significantly lower fidelity than their DNA polymerase counterparts; rather, low fidelity is primarily a result of the lack of proofreading activity (59, 64). This hypothesis is supported by the recent discovery that CoVs and other members of the Nidovirales order with genomes >20 kb encode a 3'→5' ExoN (36, 65) involved in maintaining CoV replication fidelity (66–69). The estimated mutation rates of murine hepatitis virus (MHV) and SARS-CoV are approximately  $2.5 \times 10^{-6}$  and  $9 \times 10^{-7}$  mutations per nucleotide per replication cycle, respectively (**Figure 2**). Genetic inactivation of CoV ExoN activity reduces these estimated mutation rates to the range observed in other RNA viruses and comparable to that of DNA polymerases lacking exonuclease activity (67, 68).

### Diversity of RNA Virus Populations

RNA viruses are among the most diverse replicative units in existence (70), and they demonstrate a remarkable capacity for adaptation due in part to high mutation rates during replication. RNA viruses replicate exponentially, resulting in populations that can theoretically include a substitution at every site in the genome (71). As a result of this replicative capacity, RNA viruses exist as populations of heterogeneous yet genetically related viruses, often referred to as mutant swarms,

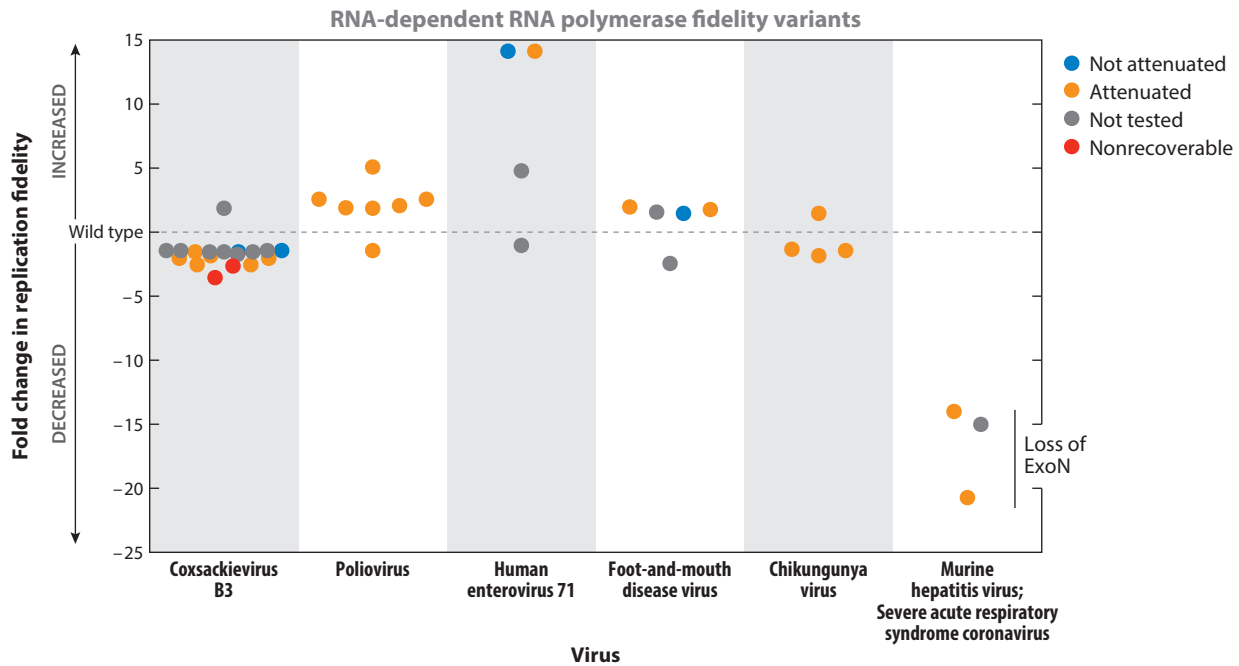
**MHV:** murine hepatitis virus

mutant clouds, or quasispecies. The defining principle of viral quasispecies evolution is that viral populations are composed of variants capable of interacting in a cooperative or antagonistic manner (see 71, 72 for recent reviews on viral quasispecies evolution). As a result, quasispecies evolution posits that it is then the viral population, and not an individual variant, that is the subject of natural selection. Mutant swarms have been shown to act as a collective during infection (73); one of the most definitive examples of this phenomenon occurs during poliovirus infection. Poliovirus is neurotropic in humans and mouse models. Poliovirus populations that have been genetically bottlenecked with reduced diversity lose their neurotropism, whereas expanding diversity of the poliovirus population by passage in the presence of RNA mutagen reestablishes neurotropism (74). This cooperativity might allow initially deleterious mutations or less fit variants to be maintained within a population, potentially facilitating more rapid adaptation under selective pressure or in new host environments (73, 75).

### Altered-Fidelity Variants

The replication fidelity of RNA viruses is likely evolutionarily constrained within a range that balances genome stability with the generation of sufficient genetic diversity. Therefore, the mechanisms by which RNA virus replication fidelity is maintained must be evolutionarily finely tuned to achieve these contrasting but important goals. A frequent misconception regarding viral mutation rates, particularly in response to a perceived danger of mutator variants (69, 76), is that increasing the mutation rate proportionally increases the rate at which viral populations adapt (77, 78). In contrast to this supposition, although the normally high mutation rates of RNA virus replication contribute to the adaptive capacity of RNA viruses, increasing the mutation rate beyond that of the wild-type virus often results in a decrease in fitness (69, 76, 79). This is primarily due to the probability that the majority of random mutations are deleterious (60, 71, 79). RNA viruses are thought to replicate at an error threshold; thus, even marginal increases in mutation rates could result in an excess of deleterious mutations within the population (see 72 for an excellent review). On the other hand, increasing fidelity also results in decreased pathogenesis and spread, either through a decrease in replication speed, an impaired ability to adapt, or both (71, 80–82).

Though many factors likely contribute to the high adaptive potential of RNA viruses, one of the key contributors is the viral RdRp. Much like their DNA polymerase counterparts, RdRps catalyze nucleotide polymerization and are the core machinery by which RNA viruses replicate their genomes. The structural and dynamic aspects of RdRp fidelity remain an active area of research, but early studies with the HIV reverse transcriptase and with *E. coli* and bacteriophage T4 DNA polymerases demonstrated that point mutations can increase or decrease polymerase fidelity, suggesting that changes to RdRp fidelity might be possible (83–85). The first example of an RdRp with altered fidelity was the poliovirus RdRp G64S mutation (74, 86). This mutation was identified independently by two groups following passage of poliovirus (PV) in the presence of the antiviral nucleoside analog ribavirin, and it resulted in a ~3-fold increase in replication fidelity. Later studies by both groups demonstrated that this single mutation resulted in viral attenuation *in vivo*, and that attenuation was likely the result of restricting the genetic diversity of the virus population (80, 81). These studies provided the first evidence that increased fidelity was attainable for an RNA virus. Subsequently, fidelity variants with mutations in viral RdRps have been isolated for other picornaviruses and arboviruses (**Figure 3** and **Table 1**) (67, 68, 74, 76, 80, 81, 86–99). Most reported altered-fidelity mutants deviate from the fidelity of the wild-type virus by a maximum of 4- to 5-fold and often are attenuated *in vivo* (**Figure 3**). A recent study using coxsackievirus B3 (CVB3) reported that mutations within the CVB3 RdRp that reduced fidelity by more than ~3-fold resulted in virus nonrecovery and suggested that this might represent the



**Figure 3**

Fidelity variants of RNA viruses. Published RNA virus fidelity variants (excluding retroviruses) are shown. For each variant, the fold change in replication fidelity was calculated using previously published data and compared with the wild type (*dotted horizontal line*). See the text and **Table 1** for specific references for each point. All points are colored according to whether or not the fidelity-altering mutation resulted in attenuation in vivo: blue, not attenuated; orange, attenuated; gray, attenuation not tested. The two red points denote mutations resulting in nonrecoverable viruses.

lower limit of RdRp fidelity for the picornaviruses (76). To our knowledge, no other group has reported RdRp mutants with lower fidelity. An upper limit for alterations in RdRp fidelity has yet to be described. These studies support the hypothesis that RNA viruses replicate within an evolutionarily selected range of fidelity, and that deviation outside of this range profoundly impacts virus fitness in vitro and in vivo.

### The CoV Exoribonuclease in Replication Fidelity and Pathogenesis

In contrast to the picornaviruses and arboviruses, CoVs encode a 3'→5' ExoN activity in nsp14 that is critical for replication fidelity. The larger members of the Nidovirales order (i.e., Coronaviridae and Roniviridae) with genomes of 26 to 32 kb encode ExoN, whereas it is absent in the smaller Arteriviridae family members with genomes <16 kb (25, 65). The recent identification of the Nam Dinh virus (NDiV), an insect nidovirus that encodes ExoN in its 20-kb genome, provides an important link in the transition from small to large nidoviruses (100, 101). CoV nsp14 is approximately 530 amino acids long and contains two enzymatic activities (**Figure 1a**): N-terminal ExoN activity and C-terminal N7-MT activity (35). The ExoN domain contains four acidic residues, DEDD, in three motifs, which are a defining characteristic of the DEDD superfamily of RNA and DNA exonucleases (102). In addition to these four invariant residues, CoV nsp14 ExoN contains a highly conserved histidine residue within motif III resulting in an HX<sub>4</sub>D arrangement (25, 36). Members of this DEDDh subgroup include proofreading enzymes such

**Table 1 Fidelity variants of RNA viruses**

Virus	Reference(s)	Mutation(s)	Protein	System <sup>a</sup>	Method <sup>b</sup>	Fold change <sup>c</sup>	Attenuated <sup>d</sup>	
CVB3	76	L241I	3Dpol	Virus	Fragment sequencing	-1.5	No	
		I230F, M145L, S299T	3Dpol	Virus	Fragment sequencing	-1.5	-	
		S299T, A372V	3Dpol	Virus	Fragment sequencing	-1.5	-	
		S164P	3Dpol	Virus	Fragment sequencing	-1.6	No	
		P48K	3Dpol	Virus	Fragment sequencing	-1.6	Yes	
		I230F, S299T	3Dpol	Virus	Fragment sequencing	-1.6	-	
		F232Y, S299T	3Dpol	Virus	Fragment sequencing	-1.6	-	
		F232Y, S299T, A372V	3Dpol	Virus	Fragment sequencing	-1.6	-	
		I230F, M145L	3Dpol	Virus	Fragment sequencing	-1.8	-	
		A239G	3Dpol	Virus	Fragment sequencing	-1.9	Yes	
		Y268W	3Dpol	Virus	Fragment sequencing	-2.1	Yes	
		Y268H	3Dpol	Virus	Fragment sequencing	-2.1	Yes	
		F232Y	3Dpol	Virus	Fragment sequencing	-2.6	Yes	
		I230F	3Dpol	Virus	Fragment sequencing	-2.6	Yes	
		F232V	3Dpol	Protein	Kinetic parameters	-2.7	-	
F232L	3Dpol	Protein	Kinetic parameters	-3.6	-			
PV	88, 98	S299T	3Dpol	Virus/protein	Multiple methods	-1.5	-	
		A372V	3Dpol	Virus/protein	Multiple methods	1.8	-	
		74, 80, 81, 86, 89	G64V	3Dpol	Virus	Fragment sequencing	1.8	Yes
			G64L	3Dpol	Virus	Fragment sequencing	1.8	Yes
			G64A	3Dpol	Virus	Fragment sequencing	2.0	Yes
G64S	3Dpol		Virus/protein	Multiple methods	2.0-4.0	Yes		
G64T	3Dpol		Virus	Fragment sequencing	2.5	Yes		
93	T362I	3Dpol	Virus/protein	Kinetic parameters	-1.5	Yes		
96	K359R	3Dpol	Virus/protein	Kinetic parameters	5.0	Yes		
HEV71	94, 95	G64N	3Dpol	Virus	Fragment sequencing	-1.1	-	
		G64T	3Dpol	Virus	Fragment sequencing	4.7	-	
		G64R	3Dpol	Virus	Fragment sequencing	14.0	-	
		S264L	3Dpol	Virus	Fragment sequencing	14.0	Yes	
FMDV	97	M296I	3Dpol	Virus/protein	Multiple methods	-2.5	-	
	91	R84H	3Dpol	Virus	Fragment sequencing	1.4	No	
	92	D5N	3Dpol	Virus	Fragment sequencing	1.5	No	
		A38V	3Dpol	Virus	Fragment sequencing	1.7	Yes	
		D5N, A38V, M194I, M296V	3Dpol	Virus	Fragment sequencing	1.9	Yes	
CHIKV	87	C483Y	Nsp4	Virus	Fragment sequencing	1.4	Yes	
	90	C483A	Nsp4	Virus	Fragment sequencing	-1.4	Yes	
		C483W	Nsp4	Virus	Fragment sequencing	-1.5	Yes	
		C483G	Nsp4	Virus	Fragment sequencing	-1.9	Yes	

(Continued)

**Table 1** (Continued)

Virus	Reference(s)	Mutation(s)	Protein	System <sup>a</sup>	Method <sup>b</sup>	Fold change <sup>c</sup>	Attenuated <sup>d</sup>
MHV	67	D89A, E91A	Nsp14	Virus	Full-genome sequencing	-15.0	-
SARS-CoV	68	D90A, E92A	Nsp14	Virus	Full-genome sequencing	-20.7	-
	69	D90A, E92A	Nsp14	Virus	Full-genome sequencing	-14.0	Yes

Abbreviations: CHIKV, chikungunya virus; CVB3, coxsackievirus B3; FMDV, foot-and-mouth disease virus; HEV71, human enterovirus 71; MHV, murine hepatitis virus; PV, poliovirus; SARS-CoV, severe acute respiratory syndrome coronavirus.

<sup>a</sup>“Virus” denotes the recovery of virus, whereas “protein” denotes recombinantly expressed protein.

<sup>b</sup>Sequencing was performed on either viral supernatants, viable plaques, or total intracellular viral RNA.

<sup>c</sup>Fold change was calculated using values reported in each reference. A range is reported if values were independently reported by more than one group.

<sup>d</sup>Dashes indicate that the attenuation phenotype is unknown.

as the  $\epsilon$  subunit of *E. coli* DNA polymerase III (Pol III). This similarity early on suggested a role for nsp14 ExoN in proofreading and/or other aspects of RNA processing (25). In contrast with its cellular counterparts, CoV ExoN also contains a unique zinc finger domain of unknown function between motifs I and II. Biochemical confirmation of ExoN activity using bacterially expressed nsp14 from HCoV-229E demonstrated that ExoN is capable of cleaving both ssRNA and dsRNA in a 3' → 5' direction (36). Recombinant HCoV-229E genomes containing mutations that inactivated ExoN activity did not allow recovery of replication-competent virus and exhibited profound defects in viral RNA synthesis. In contrast, MHV and SARS-CoV containing mutations in motifs I and III were replication competent, albeit with reductions in viral RNA synthesis (67, 68). Alanine substitution of the DEDDh residues does not impair N7-MT activity of purified nsp14 in vitro (35, 103); however, other mutations within nsp14 ExoN have been demonstrated to affect N7-MT activity in vitro, indicating that ExoN and N7-MT functions are evolutionarily linked or potentially serve a novel function in CoV RNA synthesis (103).

Recombinant MHV lacking ExoN activity (ExoN<sup>-</sup>) accumulated 15-fold more mutations compared with wild-type MHV with intact ExoN (ExoN<sup>+</sup>) (67). Recombinant SARS-CoV ExoN<sup>-</sup> also demonstrated an almost identical mutator phenotype in culture and during mouse infection using ExoN<sup>-</sup> variants of virulent mouse-adapted SARS-CoV (MA-SARS) (68, 69). In both cases, between 14- and 20-fold more mutations were present within the ExoN<sup>-</sup> viruses as compared with ExoN<sup>+</sup> viruses (Table 1 and Figure 3). ExoN inactivation profoundly attenuated the pathogenesis of MA-SARS in young, aged, and immunocompromised mice (67, 69). Both the genotype and phenotype of ExoN<sup>-</sup> MHV and SARS-CoV were stable over extended passage in culture and in mice. Furthermore, MA-SARS ExoN<sup>-</sup> did not revert to virulence even during persistent infection of SCID (severe combined immunodeficient) mice. Although the precise mechanism of fidelity regulation by ExoN remains to be defined, all available biochemical and virological evidence supports the conclusion that nsp14 ExoN provides a critical proofreading function during CoV replication (36, 42, 66–69, 104). Both MHV ExoN<sup>-</sup> and SARS-CoV ExoN<sup>-</sup> viruses demonstrated increased sensitivity to the RNA mutagen 5-fluorouracil (5-FU) as compared with ExoN<sup>+</sup> viruses (66). Next-generation sequencing of SARS-CoV RNA following treatment with 5-FU indicated that the ExoN<sup>-</sup> virus populations accumulated 40-fold more mutations compared with untreated ExoN<sup>-</sup> population, and 24-fold more mutations compared with 5-FU-treated wild-type ExoN<sup>+</sup> virus (66). These results together demonstrate a role for ExoN in maintaining

CoV replication fidelity, establish a link between CoV fidelity and pathogenesis, and provide the most direct evidence that ExoN is the first known proofreading enzyme encoded by an RNA virus.

It remains to be determined whether nsp14 ExoN mediates other functions in virus replication or host interactions. Arenaviruses such as Lassa fever virus (LASV) are the only other mammalian RNA viruses known to encode a 3'→5' ExoN (105). LASV nucleoprotein (NP) is a DEDDh ExoN that specifically degrades double-stranded RNA (dsRNA) and thus is a critical component of immune evasion (105). SARS-CoV nsp1 has been shown to induce endonucleolytic cleavage of host mRNAs (106); however, there is as yet no direct evidence that nsp14 ExoN degrades host mRNAs. CoV ExoN was shown to cleave both dsRNA and ssRNA *in vitro* (36), and ExoN-mediated cleavage of ssRNA resulted in larger cleavage products as compared with cleavage of dsRNA. These data suggest that nsp14 ExoN could potentially exert differential activity on diverse RNA species (36). More recent work demonstrated that the small nonenzymatic CoV protein nsp10 can stimulate ExoN activity by almost 35-fold and renders ExoN capable of cleaving 3' mismatched residues (42). Thus, the specificity and activity of ExoN on various viral or host RNA substrates conceivably could be determined through interactions with other viral proteins. Aside from nsp10, CoV nsp14 ExoN could also function in tandem with the viral endonuclease nsp15 NendoU to degrade RNA targets. In fact, CoV nsp15 NendoU has an uncharacterized interferon antagonist activity, and the arterivirus homolog of nsp15 was recently shown to inhibit interferon  $\beta$  induction (107, 108). The capacity of ExoN to cleave a variety of RNA substrates and the uncharacterized interferon antagonist activity identified for nsp15 NendoU suggest that these two proteins could be important for suppression of anti-CoV immune responses (108). If so, such a function for ExoN could contribute to the attenuation of MA-SARS ExoN<sup>-</sup> observed *in vivo*.

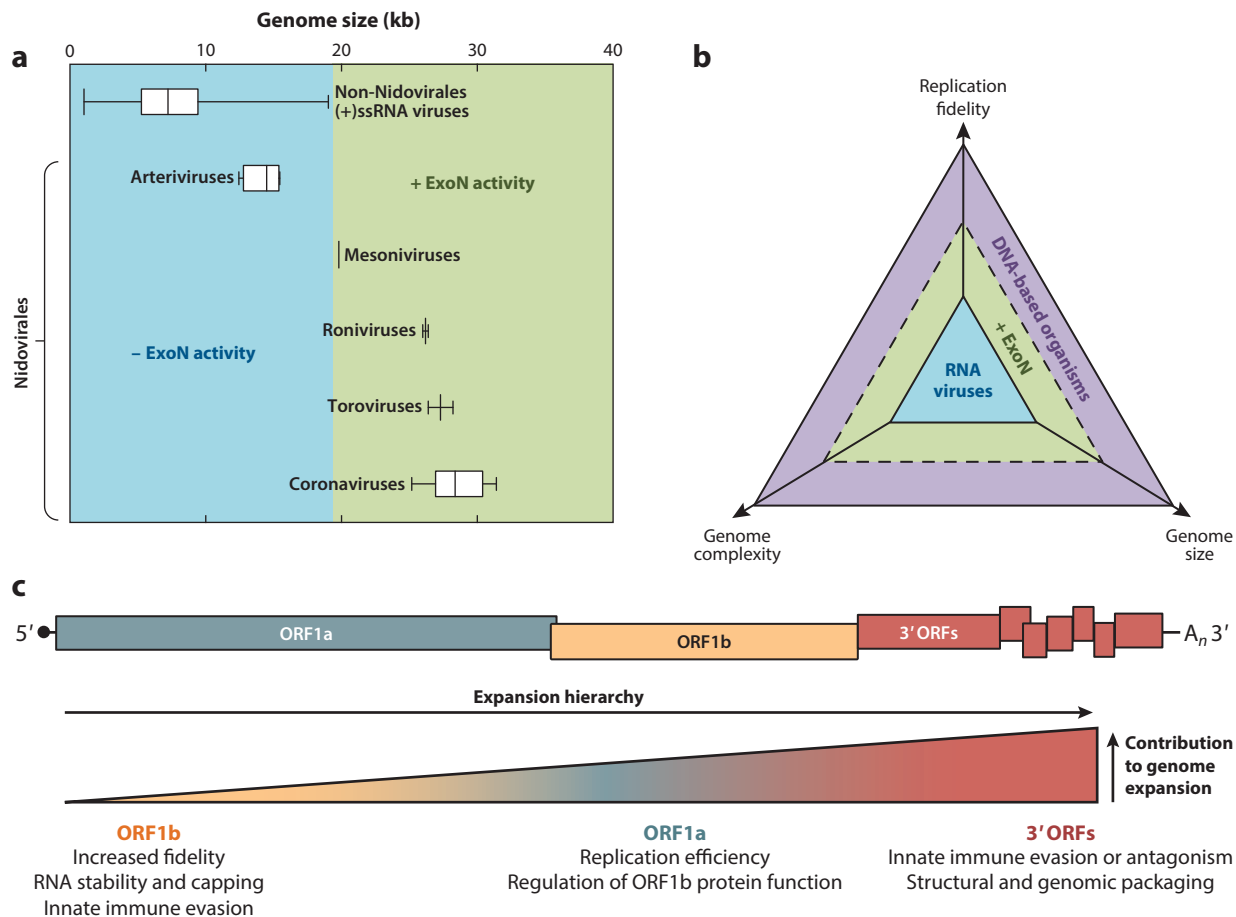
## CORONAVIRUS GENOME SIZE: SOLVING THE EIGEN PARADOX

### RNA Virus Genome Size and the Eigen Paradox

Unlike in DNA viruses and DNA-based organisms, the genome size distribution of ssRNA genomes from different virus families is quite narrow; the largest ssRNA virus genomes are ~32 kb in length (65). Excluding the nidoviruses, (+)ssRNA viruses range in size from ~2,300 to ~20,000 bases, with the majority of ssRNA virus genomes measuring ~10 kb (**Figure 4a**) (65). Assuming near-equivalent error rates, RNA viruses with larger genomes would be predicted to accumulate more mutations per genome during replication, and would in turn be predicted to accumulate more deleterious mutations, leading to virus extinction. This has led to a theoretical limitation to RNA virus genome size termed the Eigen paradox or Eigen trap, after Manfred Eigen's work describing self-replicating molecules (1, 109, 110). In an analysis of CoV evolution, Nga et al. (100) depicted the Eigen trap as a triangle on which genome size, replication fidelity, and genome complexity are located at the apices (**Figure 4b**). In the center of this triangle are the RNA viruses, which due to their low-fidelity replication are confined within the Eigen trap. Because they are unable to increase replication fidelity, RNA viruses are evolutionarily constrained to have genomes of relatively low complexity and small size.

However, a growing body of work suggests that CoVs have found a way to escape this trap. Recent comparative genomics studies of complete nidovirus genomes proposed that acquisition of ExoN allowed expansion of the smaller ancestor nidovirus genome (65, 100, 111). Once CoVs acquired mechanisms to increase their fidelity and limit the accumulation of deleterious mutations, increases in both genome size and complexity were possible. The acquisition of additional replicase proteins likely allowed both the continued expansion and the divergence of nidoviruses into present-day CoVs (**Figure 4c**) (111). Conceptually, expansion of ORF1a could have allowed





**Figure 4**

Size of positive-sense single-stranded RNA [(+)ssRNA] virus genomes and expansion of the coronavirus (CoV) genome. (a) Median genome size for (+)ssRNA viruses, excluding the nidoviruses, in comparison with each family within the order Nidovirales. Full-length genomes were obtained from the National Center for Biotechnology Information (NCBI) Viral Genomes Resource. Each bar depicts the median genome size and extends from the 25th to 75th percentile; the whiskers show the minimum and maximum values. Genomes containing (green) and lacking (cyan) 3'→5' exoribonuclease (ExoN) activity are shown. (b) The relationship between replication fidelity, genome complexity, and genome size. Fidelity, complexity, and size increase as the arrows move from the center to the edges of the triangle. RNA viruses have low replication fidelity, low genomic complexity, and small genomes and thus are constrained within the small Eigen triangle (or trap). Acquisition of ExoN likely helped CoVs (green) escape this trap by increasing replication fidelity. Panel b adapted with permission from Nga et al. (100). (c) A schematic of the CoV genome (top) depicts open reading frame (ORF) 1a in blue, ORF1b in orange, and the 3' ORFs in red. The order in which these regions are thought to have expanded—ORF1b→ORF1a→3' ORFs—is shown (bottom). The relative contribution to overall CoV genome size, as defined by the total number of bases added, is denoted by the height of the triangle. See Lauber et al. (111) for additional details. Some predicted activities and functions likely acquired as a result of genomic expansion are also shown.

for increased replication efficiency, possibly through the regulation of ORF1b proteins (111). This hypothesis is supported by studies showing the association of ORF1a products such as nsp7–10 with the ORF1b proteins (38, 41–44, 112–114). Furthermore, the lack of homologous nsp7–10 sequences outside of the coronaviruses and toroviruses suggests that acquisition of these proteins might have also facilitated genomic expansion of the “large” nidoviruses (65). Acquisition of

several 3' accessory ORFs, excluding essential structural proteins, was shown to have contributed the most to CoV genomic expansion, as defined by the total number of nucleotides added. Both the number and the function of these 3' ORFs vary tremendously among CoVs.

### A Model for a Putative Multisubunit Polymerase Complex

Whereas the simplest of RNA viruses encode only an RdRp, other RNA viruses encode a variety of RNA-modifying enzymes, such as RNA capping (or cap snatching) machinery and helicase activity. Due to their large genomes (65, 111), unique RNA-modifying functions, and putative proofreading capability (25, 65–69, 111), CoVs may have more complex RNA replication systems than those described for many other RNA viruses. Additionally, the limited polymerase activity observed for nsp12 RdRp (30–32) alone likely reflects the need for additional viral proteins in order to faithfully and rapidly replicate the CoV genome during infection. These observations suggest the possibility that CoVs employ a multisubunit polymerase complex for viral RNA synthesis. Such a complex has yet to be described experimentally, likely due to challenges of recapitulating in vitro a complex containing nsp12 RdRp and six or more additional replicase proteins. Although detailed biochemical and structural studies will be essential in understanding how the CoV replicase is assembled, a model can be proposed (104, 115) on the basis of known and predicted activities and by analogy to DNA polymerase complexes (**Figure 5**), specifically DNA Pol III.

One of the most extensively studied multisubunit polymerases is *E. coli* DNA Pol III, which is the major polymerase during chromosomal replication (recently reviewed in 116). *E. coli* DNA Pol III is a holoenzyme (DNA Pol III HE) that contains a catalytic core ( $\alpha\epsilon\theta$ ), a processivity factor ( $\beta_2$  sliding clamp), and a multisubunit clamp loader that loads  $\beta_2$  onto the DNA template (116). The catalytic core contains three subunits: the polymerase ( $\alpha$ ); the 3'→5' exonuclease ( $\epsilon$ ); and the small nonenzymatic  $\theta$  subunit, which stabilizes and stimulates  $\epsilon$  (116–118). Much like the Pol III core, CoV nsp12 RdRp likely interacts with nsp14 ExoN, as both proteins would need to be in close proximity for error removal and repair (**Figure 5**). The viral helicase nsp13 is likely upstream but closely associated with nsp12 RdRp (33) to ensure the availability of a single-stranded template. This model is consistent with previous studies describing that a large majority of CoV RNA is present as partially double-stranded RNA, suggesting that multiple RNA templates are being synthesized for each negative-strand template (119). Nsp15 and nsp16 could also be associated with this complex; however, given the undefined role of nsp15 during replication and the role of nsp16 in RNA capping, these proteins might form distinct complexes. Other proteins, particularly those from ORF1a, likely interact with the CoV polymerase core (112–114), which would be consistent with the hypothesis that some ORF1a proteins were acquired to regulate ORF1b proteins and/or to increase the efficiency of replicating an increasingly large CoV genome (111).

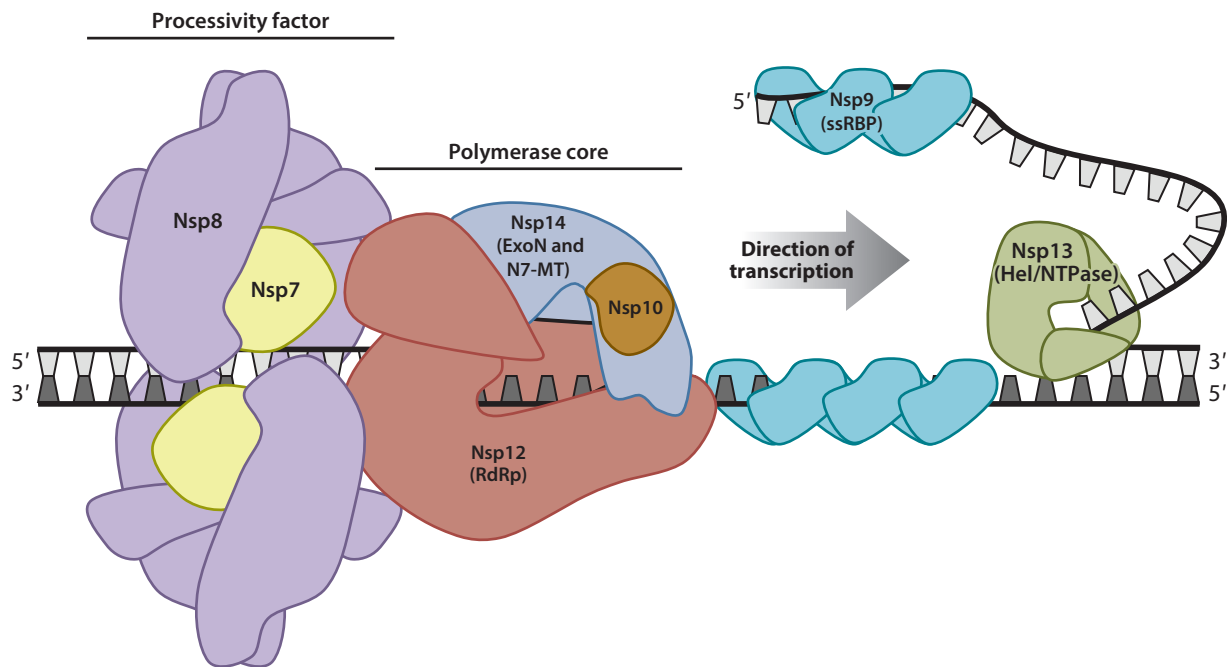
Though the functions of several of the ORF1a proteins are just beginning to be defined, several studies support the hypothesis that nsp7–10 are associated with this polymerase core. CoV nsp8 was shown to harbor RdRp and primase activity (28) and to interact with nsp7 to form a large toroidal hexadecamer structure (39). The ~30-Å central pore of the SARS-CoV nsp7-nsp8 hexadecamer is lined with positively charged amino acids, which allow the supercomplex to bind dsRNA. These data suggest that the nsp7-nsp8 hexadecamer could function as a processivity factor for the CoV polymerase during RNA replication, much like the  $\beta_2$  clamp within the *E. coli* Pol III HE (reviewed in 120) and other processivity factors. The small nonenzymatic nsp10 associates with both CoV methyltransferases: nsp14 N7-MT and nsp16 2'-OMT (41–44). Binding of nsp10 to nsp16 is required for 2'-OMT activity, an enzymatic activity that is critical in mitigating detection of CoV RNA by the innate immune system (121). Nsp10 also stimulates nsp14 ExoN activity in vitro (42), though the significance of an nsp10-nsp14 interaction during virus replication

---

**Processivity:** the average number of consecutive nucleotides a polymerase is capable of adding during a single template-binding event

---





**Figure 5**

Model of the putative coronavirus multisubunit polymerase. The viral genome is shown in light gray and negative-sense template RNA in dark gray. The viral helicase and NTPase (nsp13 Hel/NTPase; *green*) is shown removing a newly synthesized viral genome, and the viral single-stranded RNA-binding protein (nsp9 ssRBP; *cyan*) is shown protecting the single-stranded RNA. The remaining portions of the model are based on known structures (nsp7 and nsp8) and/or reported protein-protein interactions. The polymerase core is predicted to include the viral RNA-dependent RNA polymerase (nsp12 RdRp; *red*), the 3'→5' exoribonuclease and N7-methyltransferase (nsp14 ExoN and N7-MT; *blue*), and the nonenzymatic protein nsp10 (*orange*). This core is shown tightly associated with the putative processivity factor (nsp7 and nsp8; *yellow* and *purple*, respectively). Figure adapted with permission from Smith et al. (104).

has yet to be described. Nsp9 appears to function as an ssRNA-binding protein (122), suggesting that its function is to protect the single-stranded template as well as single-stranded newly synthesized RNA, a function analogous that of the SSB protein during DNA replication. Indeed, interactions between nsp9 and other replicase proteins have been observed (122). The continued elucidation of the functions of ORF1a proteins will be essential in establishing their importance during replication and their potential evolutionary linkage to specific ORF1b proteins.

## CONCLUDING REMARKS

RNA virus replication fidelity is emerging as a new field of study that incorporates polymerase biochemistry, in vitro evolution experiments, virus fitness studies, and bioinformatic analyses. The studies are yielding surprising insights that have important implications for how we think about virus replication, pathogenesis, and evolution, as well as translational approaches to virus inhibition and attenuation.

### Impact of Decreased Fidelity on Virus Fitness

Consider the following statement: An increased mutation rate favors the virus. This widely held view is based on two clear observations: (a) RNA viruses replicate with much lower fidelity than

DNA-based organisms, and (b) RNA viruses demonstrate rapid adaptation and emergence of resistance to antivirals and vaccines. The assumption built on these observations is that if decreased fidelity is beneficial for the virus, further decreases in fidelity are even better. Yet the data presented here for several virus families suggest that decreasing fidelity results in impaired replication, decreased competitive fitness, and attenuated virulence. Although only a few virus families have been investigated for the effect of decreased fidelity, it is certainly possible that examples will be identified where decreased fidelity favors the virus under specific circumstances, such as selection of resistance mutants to antivirals or revertants of conditional mutants.

### **New Models for RNA Synthesis and Regulation of Fidelity**

Data presented in this review support the hypothesis that CoVs use a multiprotein replication complex that incorporates a processivity factor, a proofreading exoribonuclease, an RNA-dependent RNA polymerase, and a helicase, as well as predicted stimulatory cofactors and capping activities. These data and models argue that CoV RNA synthesis, modification, and fidelity might be more akin to those of DNA-based organisms with multiprotein DNA replication complexes. This has significant implications for understanding CoV replication but also for investigating the interface between DNA- and RNA-based life. The fact that CoVs are the largest known replicating RNA-based organisms suggests the possibility that they may exist at the boundary of what is possible in an RNA virus that has to balance genome stability with the population diversity required for adaptation. Alternatively, the demonstrated capacity of CoVs for zoonotic infections and host-species movement might argue that CoVs use a much larger range of fidelity and genome complexity to explore a greatly expanded sequence and phenotype space.

### **Goldilocks? Maybe**

The Goldilocks metaphor is often used in cosmology and evolution to describe the fortuitous conditions that allowed the emergence and evolution of life. One is tempted to apply a similar metaphor to virus replication fidelity. Certainly the available data as presented in this review indicate that increasing or decreasing the fidelity of a number of divergent RNA viruses impairs fitness and is attenuating. However, the fact that fidelity can be moved off center—notably by at least 20-fold for CoVs—suggests that a range of fidelity is available to RNA viruses, potentially allowing for selection of more or less diverse populations. If Goldilocks is applicable, it would be to a range rather than a single optimal fidelity set point. CoVs represent the most obvious opportunity to explore this possibility, because they at minimum require the interaction of RdRp with a proofreading ExoN. This would represent at least two possible fidelity states, one high-fidelity state during RdRp-ExoN interaction and one low-fidelity state where RdRp interaction with ExoN is altered.

### **Fidelity as a Target for Inhibition and Attenuation**

Could altered-fidelity viruses be used as live attenuated vaccine candidates? The best data for the effect of altered fidelity are presented here and show that either increased or decreased fidelity is attenuating in a range of RNA virus families. Thus it is possible that fidelity regulation is so central to all aspects of virus replication and pathogenesis that altering fidelity may be a broadly applicable approach to stable attenuation. For CoVs it appears that genetic inactivation of ExoN results in both a genotype and a phenotype that are attenuated and are resistant to reversion *in vitro* and *in vivo* during passage or persistent infection. Further, inactivation of ExoN results in a profound increase in sensitivity to RNA mutagens for SARS-CoV. Thus, the process of replication

fidelity may be a new target for virus family-wide attenuation and inhibition of virus replication and pathogenesis.

#### SUMMARY POINTS

1. RNA viruses replicate with lower fidelity than DNA-based organisms due to the lack of mechanisms for error recognition and repair.
2. RNA viruses exist as populations of genetically related variants, also known as mutant swarms or quasispecies, that are the units of selection.
3. To date, small RNA viruses are thought to regulate fidelity principally through the RNA-dependent RNA polymerase.
4. The tolerated range of increased or decreased fidelity for RNA viruses lacking proofreading may be very narrow.
5. Incorporation of a proofreading exoribonuclease allowed expansion of nidovirus genome size and complexity, as observed in coronaviruses.
6. Coronaviruses tolerate a 20-fold decrease in fidelity; the limits of increased or decreased replication fidelity for CoVs have not been defined.
7. Coronaviruses may assemble a multiprotein complex containing RNA-dependent RNA polymerase and proofreading activities.
8. Fidelity determinants may represent highly conserved and nonredundant targets for viral inhibition and live-virus attenuation.

#### FUTURE ISSUES

1. For coronaviruses, identifying all of the components of a multiprotein replication complex by *in vitro* reconstitution and structural studies will allow prediction and testing of our model for assembly and function of the replication complex in RNA synthesis and fidelity.
2. It will be exciting to test whether RNA viruses explore the range of fidelity under different selective pressures, and even during the course of a single infectious cycle.
3. The contributions of host cell proteins to RNA virus replication fidelity should be explored to define whether different environments stimulate or impair virus replication fidelity.
4. It will be important to define the multiple contributing factors to replication fidelity for CoVs and other RNA viruses, particularly robustness to mutations and RNA-dependent RNA polymerase selectivity and speed.
5. Studies of other families of RNA viruses, including negative-strand segmented genomes such as that of influenza virus, are needed to better understand common and divergent mechanisms and ranges of tolerated replication fidelity.
6. The availability of increased-fidelity and decreased-fidelity strains of multiple RNA viruses will allow testing of the impact of fidelity on host-range expansion, adaptation, and experimental evolution.

7. Increasing availability and affordability of next-generation sequencing will allow in-depth analysis of the effects of fidelity and virus population diversity on virus replication, pathogenesis, and fitness.

## DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

## ACKNOWLEDGMENTS

This work was supported by NIH grants R01-AI108197 (M.R.D.), T32-AI095202 (M.R.D.), U19-AI109680 (M.R.D.), and F32-AI108102 (E.C.S.) from the National Institute of Allergy and Infectious Diseases.

## LITERATURE CITED

1. Eigen M. 1971. Selforganization of matter and the evolution of biological macromolecules. *Naturwissenschaften* 58:465–523
2. Sanjuán R, Nebot MR, Chirico N, Mansky LM, Belshaw R. 2010. Viral mutation rates. *J. Virol.* 84:9733–48
3. Bradwell K, Combe M, Domingo-Calap P, Sanjuán R. 2013. Correlation between mutation rate and genome size in riboviruses: mutation rate of bacteriophage Q $\beta$ . *Genetics* 195:243–51
4. Graham RL, Donaldson EF, Baric RS. 2013. A decade after SARS: strategies for controlling emerging coronaviruses. *Nat. Rev. Microbiol.* 11:836–48
5. Ge XY, Li JL, Yang XL, Chmura AA, Zhu G, et al. 2013. Isolation and characterization of a bat SARS-like coronavirus that uses the ACE2 receptor. *Nature* 503:535–38
6. Huynh J, Li S, Yount B, Smith A, Sturges L, et al. 2012. Evidence supporting a zoonotic origin of human coronavirus strain NL63. *J. Virol.* 86:12816–25
7. Pfefferle S, Oppong S, Drexler JF, Gloza-Rausch F, Ipsen A, et al. 2009. Distant relatives of severe acute respiratory syndrome coronavirus and close relatives of human coronavirus 229E in bats, Ghana. *Emerg. Infect. Dis.* 15:1377–84
8. Crossley BM, Mock RE, Callison SA, Hietala SK. 2012. Identification and characterization of a novel alpaca respiratory coronavirus most closely related to the human coronavirus 229E. *Viruses* 4:3689–700
9. Vijgen L, Keyaerts E, Moës E, Thoelen I, Wollants E, et al. 2005. Complete genomic sequence of human coronavirus OC43: molecular clock analysis suggests a relatively recent zoonotic coronavirus transmission event. *J. Virol.* 79:1595–604
10. Woo PCY, Lau SKP, Huang Y, Tsoi HW, Chan KH, Yuen KY. 2005. Phylogenetic and recombination analysis of coronavirus HKU1, a novel coronavirus from patients with pneumonia. *Arch. Virol.* 150:2299–311
11. Peiris JSM, Lai ST, Poon LLM, Guan Y, Yam LYC, et al. 2003. Coronavirus as a possible cause of severe acute respiratory syndrome. *Lancet* 361:1319–25
12. Zaki AM, van Boheemen S, Bestebroer TM, Osterhaus ADME, Fouchier RAM. 2012. Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. *N. Engl. J. Med.* 367:1814–20
13. Woo PCY, Lau SKP, Huang Y, Yuen KY. 2009. Coronavirus diversity, phylogeny and interspecies jumping. *Exp. Biol. Med.* 234:1117–27
14. Drexler JF, Corman VM, Drosten C. 2013. Ecology, evolution and classification of bat coronaviruses in the aftermath of SARS. *Antivir. Res.* 101:45–56
15. Wertheim JO, Chu DKW, Peiris JSM, Kosakovsky Pond SL, Poon LLM. 2013. A case for the ancient origin of coronaviruses. *J. Virol.* 87:7039–45

---

2. A recent compilation of experimentally determined viral mutation rates.

---

---

25. Identification of potential novel functions within the CoV genome, including nsp14 ExoN, using comparative genomics.

---

16. Calisher CH, Childs JE, Field HE, Holmes KV, Schountz T. 2006. Bats: important reservoir hosts of emerging viruses. *Clin. Microbiol. Rev.* 19:531–45
17. Hilgenfeld R, Peiris M. 2013. From SARS to MERS: 10 years of research on highly pathogenic human coronaviruses. *Antivir. Res.* 100:286–95
18. Haagmans BL, Al Dhahiry SHS, Reusken CBEM, Raj VS, Galiano M, et al. 2013. Middle East respiratory syndrome coronavirus in dromedary camels: an outbreak investigation. *Lancet Infect. Dis.* 14:140–45
19. Alagaili AN, Briese T, Mishra N, Kapoor V, Sameroff SC, et al. 2014. Middle East respiratory syndrome coronavirus infection in dromedary camels in Saudi Arabia. *mBio* 5:e00884–14
20. Cauchemez S, Fraser C, Van Kerkhove MD, Donnelly CA, Riley S, et al. 2014. Middle East respiratory syndrome coronavirus: quantification of the extent of the epidemic, surveillance biases, and transmissibility. *Lancet Infect. Dis.* 14:50–56
21. Breban R, Riou J, Fontanet A. 2013. Interhuman transmissibility of Middle East respiratory syndrome coronavirus: estimation of pandemic risk. *Lancet* 382:694–99
22. Perlman S, Netland J. 2009. Coronaviruses post-SARS: update on replication and pathogenesis. *Nat. Rev. Microbiol.* 7:439–50
23. Brierley I, Bournsnel ME, Binns MM, Bilimoria B, Blok VC, et al. 1987. An efficient ribosomal frame-shifting signal in the polymerase-encoding region of the coronavirus IBV. *EMBO J.* 6:3779–85
24. Plant EP, Pérez-Alvarado GC, Jacobs JL, Mukhopadhyay B, Hennig M, Dinman JD. 2005. A three-stemmed mRNA pseudoknot in the SARS coronavirus frameshift signal. *PLoS Biol.* 3:e172
25. **Snijder EJ, Bredendbeek PJ, Dobbe JC, Thiel V, Ziebuhr J, et al. 2003. Unique and conserved features of genome and proteome of SARS-coronavirus, an early split-off from the coronavirus group 2 lineage. *J. Mol. Biol.* 331:991–1004**
26. Ziebuhr J, Snijder EJ, Gorbalenya AE. 2000. Virus-encoded proteinases and proteolytic processing in the *Nidovirales*. *J. Gen. Virol.* 81:853–79
27. Wu HY, Brian DA. 2010. Subgenomic messenger RNA amplification in coronaviruses. *Proc. Natl. Acad. Sci. USA* 107:12257–62
28. Imbert I, Guillemot JC, Bourhis JM, Bussetta C, Coutard B, et al. 2006. A second, non-canonical RNA-dependent RNA polymerase in SARS coronavirus. *EMBO J.* 25:4933–42
29. te Velthuis AJW, van den Worm SHE, Snijder EJ. 2012. The SARS-coronavirus nsp7+nsp8 complex is a unique multimeric RNA polymerase capable of both de novo initiation and primer extension. *Nucleic Acids Res.* 40:1737–47
30. te Velthuis AJW, Arnold JJ, Cameron CE, van den Worm SHE, Snijder EJ. 2010. The RNA polymerase activity of SARS-coronavirus nsp12 is primer dependent. *Nucleic Acids Res.* 38:203–14
31. Cheng A, Zhang W, Xie Y, Jiang W, Arnold E, et al. 2005. Expression, purification, and characterization of SARS coronavirus RNA polymerase. *Virology* 335:165–76
32. Ahn DG, Choi JK, Taylor DR, Oh JW. 2012. Biochemical characterization of a recombinant SARS coronavirus nsp12 RNA-dependent RNA polymerase capable of copying viral RNA templates. *Arch. Virol.* 157:2095–104
33. Seybert A, Hegyi A, Siddell SG, Ziebuhr J. 2000. The human coronavirus 229E superfamily 1 helicase has RNA and DNA duplex-unwinding activities with 5'-to-3' polarity. *RNA* 6:1056–68
34. Ivanov KA, Ziebuhr J. 2004. Human coronavirus 229E nonstructural protein 13: characterization of duplex-unwinding, nucleoside triphosphatase, and RNA 5'-triphosphatase activities. *J. Virol.* 78:7833–38
35. Chen Y, Cai H, Pan J, Xiang N, Tien P, et al. 2009. Functional screen reveals SARS coronavirus nonstructural protein nsp14 as a novel cap N7 methyltransferase. *Proc. Natl. Acad. Sci. USA* 106:3484–89
36. Minskaia E, Hertzog T, Gorbalenya AE, Campanacci V, Cambillau C, et al. 2006. Discovery of an RNA virus 3'→5' exoribonuclease that is critically involved in coronavirus RNA synthesis. *Proc. Natl. Acad. Sci. USA* 103:5108–13
37. Ivanov KA, Hertzog T, Rozanov M, Bayer S, Thiel V, et al. 2004. Major genetic marker of nidoviruses encodes a replicative endoribonuclease. *Proc. Natl. Acad. Sci. USA* 101:12694–99
38. Decroly E, Imbert I, Coutard B, Bouvet M, Selisko B, et al. 2008. Coronavirus nonstructural protein 16 is a cap-0 binding enzyme possessing (nucleoside-2'*O*)-methyltransferase activity. *J. Virol.* 82:8071–84

39. Zhai Y, Sun F, Li X, Pang H, Xu X, et al. 2005. Insights into SARS-CoV transcription and replication from the structure of the nsp7-nsp8 hexadecamer. *Nat. Struct. Mol. Biol.* 12:980–86
40. Donaldson EF, Sims AC, Graham RL, Denison MR, Baric RS. 2007. Murine hepatitis virus replicase protein nsp10 is a critical regulator of viral RNA synthesis. *J. Virol.* 81:6356–68
41. Bouvet M, Debarnot C, Imbert I, Selisko B, Snijder EJ, et al. 2010. In vitro reconstitution of SARS-coronavirus mRNA cap methylation. *PLoS Pathog.* 6:e1000863
42. Bouvet M, Imbert I, Subissi L, Gluais L, Canard B, Decroly E. 2012. RNA 3'-end mismatch excision by the severe acute respiratory syndrome coronavirus nonstructural protein nsp10/nsp14 exoribonuclease complex. *Proc. Natl. Acad. Sci. USA* 109:9372–77
43. Decroly E, Debarnot C, Ferron F, Bouvet M, Coutard B, et al. 2011. Crystal structure and functional analysis of the SARS-coronavirus RNA cap 2'-O-methyltransferase nsp10/nsp16 complex. *PLoS Pathog.* 7:e1002059
44. Chen Y, Su C, Ke M, Jin X, Xu L, et al. 2011. Biochemical and structural insights into the mechanisms of SARS coronavirus RNA ribose 2'-O-methylation by nsp16/nsp10 protein complex. *PLoS Pathog.* 7:e1002294
45. Xu X, Liu Y, Weiss S, Arnold E, Sarafianos SG, Ding J. 2003. Molecular model of SARS coronavirus polymerase: implications for biochemical functions and drug design. *Nucleic Acids Res.* 31:7117–30
46. Gorbalenya AE, Pringle FM, Zeddam JL, Luke BT, Cameron CE, et al. 2002. The palm subdomain-based active site is internally permuted in viral RNA-dependent RNA polymerases of an ancient lineage. *J. Mol. Biol.* 324:47–62
47. Fijalkowska IJ, Schaaper RM, Jonczyk P. 2012. DNA replication fidelity in *Escherichia coli*: a multi-DNA polymerase affair. *FEMS Microbiol. Rev.* 36:1105–21
48. Kunkel TA, Bebenek K. 2000. DNA replication fidelity. *Annu. Rev. Biochem.* 69:497–529
49. Kunkel TA. 2004. DNA replication fidelity. *J. Biol. Chem.* 279:16895–98
50. Drake JW. 1991. A constant rate of spontaneous mutation in DNA-based microbes. *Proc. Natl. Acad. Sci. USA* 88:7160–64
51. Cuevas JM, Domingo-Calap P, Pereira-Gómez M, Sanjuán R. 2009. Experimental evolution and population genetics of RNA viruses. *Open Evol. J.* 3:9–16
52. Johnson SJ, Beese LS. 2004. Structures of mismatch replication errors observed in a DNA polymerase. *Cell* 116:803–16
53. Sintim HO, Kool ET. 2006. Remarkable sensitivity to DNA base shape in the DNA polymerase active site. *Angew. Chem. Int. Ed. Engl.* 45:1974–79
54. Franklin MC, Wang J, Steitz TA. 2001. Structure of the replicating complex of a pol  $\alpha$  family DNA polymerase. *Cell* 105:657–67
55. Castro C, Arnold JJ, Cameron CE. 2005. Incorporation fidelity of the viral RNA-dependent RNA polymerase: a kinetic, thermodynamic and structural perspective. *Virus Res.* 107:141–49
56. Cameron CE, Moustafa IM, Arnold JJ. 2009. Dynamics: the missing link between structure and function of the viral RNA-dependent RNA polymerase? *Curr. Opin. Struct. Biol.* 19:768–74
57. Moustafa IM, Shen H, Morton B, Colina CM, Cameron CE. 2011. Molecular dynamics simulations of viral RNA polymerases link conserved and correlated motions of functional elements to fidelity. *J. Mol. Biol.* 410:159–81
58. Yang X, Smidansky ED, Maksimchuk KR, Lum D, Welch JL, et al. 2012. Motif D of viral RNA-dependent RNA polymerases determines efficiency and fidelity of nucleotide addition. *Structure* 20:1519–27
59. Steinhauer DA, Domingo E, Holland JJ. 1992. Lack of evidence for proofreading mechanisms associated with an RNA virus polymerase. *Gene* 122:281–88
60. Drake JW, Holland JJ. 1999. Mutation rates among RNA viruses. *Proc. Natl. Acad. Sci. USA* 96:13910–13
61. Morrison A, Bell JB, Kunkel TA, Sugino A. 1991. Eukaryotic DNA polymerase amino acid sequence required for 3'→5' exonuclease activity. *Proc. Natl. Acad. Sci. USA* 88:9473–77
62. Schaaper RM, Radman M. 1989. The extreme mutator effect of *Escherichia coli* mutD5 results from saturation of mismatch repair by excessive DNA replication errors. *EMBO J.* 8:3511–16
63. McCulloch SD, Kunkel TA. 2008. The fidelity of DNA synthesis by eukaryotic replicative and translesion synthesis polymerases. *Cell Res.* 18:148–61

---

42. Demonstration of 3' mismatch excision by SARS-CoV nsp14 ExoN.

---



---

67. First experimental evidence that nsp14 ExoN is involved in CoV replication fidelity.

---

69. Demonstration that SARS-CoV lacking ExoN protects against a lethal wild-type SARS-CoV challenge in vivo.

---

72. A thorough review on viral quasispecies evolution and its clinical implications.

---

74. Demonstration that restricting viral population diversity can severely impact viral pathogenesis.

---

---

86. Description of the isolation of the first altered-fidelity variant of an RNA virus.

---

64. Arnold JJ, Cameron CE. 2004. Poliovirus RNA-dependent RNA polymerase (3Dpol): pre-steady-state kinetic analysis of ribonucleotide incorporation in the presence of  $Mg^{2+}$ . *Biochemistry* 43:5126–37
65. Gorbalenya AE, Enjuanes L, Ziebuhr J, Snijder EJ. 2006. *Nidovirales*: evolving the largest RNA virus genome. *Virus Res.* 117:17–37
66. Smith EC, Blanc H, Vignuzzi M, Denison MR. 2013. Coronaviruses lacking exoribonuclease activity are susceptible to lethal mutagenesis: evidence for proofreading and potential therapeutics. *PLoS Pathog.* 9:e1003565
67. Eckerle LD, Lu X, Sperry SM, Choi L, Denison MR. 2007. High fidelity of murine hepatitis virus replication is decreased in nsp14 exoribonuclease mutants. *J. Virol.* 81:12135–44
68. Eckerle LD, Becker MM, Halpin RA, Li K, Venter E, et al. 2010. Infidelity of SARS-CoV nsp14-exonuclease mutant virus replication is revealed by complete genome sequencing. *PLoS Pathog.* 6:e1000896
69. Graham RL, Becker MM, Eckerle LD, Bolles M, Denison MR, Baric RS. 2012. A live, impaired-fidelity coronavirus vaccine protects in an aged, immunocompromised mouse model of lethal disease. *Nat. Med.* 18:1820–26
70. Mokili JL, Rohwer F, Dutilh BE. 2012. Metagenomics and future perspectives in virus discovery. *Curr. Opin. Virol.* 2:63–77
71. Lauring AS, Andino R. 2010. Quasispecies theory and the behavior of RNA viruses. *PLoS Pathog.* 6:e1001005
72. Domingo E, Sheldon J, Perales C. 2012. Viral quasispecies evolution. *Microbiol. Mol. Biol. Rev.* 76:159–216
73. Ojosnegros S, Perales C, Mas A, Domingo E. 2011. Quasispecies as a matter of fact: viruses and beyond. *Virus Res.* 162:203–15
74. Vignuzzi M, Stone JK, Arnold JJ, Cameron CE, Andino R. 2006. Quasispecies diversity determines pathogenesis through cooperative interactions in a viral population. *Nature* 439:344–48
75. Lauring AS, Frydman J, Andino R. 2013. The role of mutational robustness in RNA virus evolution. *Nat. Rev. Microbiol.* 11:327–36
76. Gnädig NF, Beaucourt S, Campagnola G, Bordería AV, Sanz-Ramos M, et al. 2012. Coxsackievirus B3 mutator strains are attenuated in vivo. *Proc. Natl. Acad. Sci. USA* 109:E2294–303
77. Furió V, Moya A, Sanjuán R. 2005. The cost of replication fidelity in an RNA virus. *Proc. Natl. Acad. Sci. USA* 102:10233–37
78. Elena SF, Sanjuán R. 2005. Adaptive value of high mutation rates of RNA viruses: separating causes from consequences. *J. Virol.* 79:11555–58
79. Lee CH, Gilbertson DL, Novella IS, Huerta R, Domingo E, Holland JJ. 1997. Negative effects of chemical mutagenesis on the adaptive behavior of vesicular stomatitis virus. *J. Virol.* 71:3636–40
80. Pfeiffer JK, Kirkegaard K. 2005. Increased fidelity reduces poliovirus fitness and virulence under selective pressure in mice. *PLoS Pathog.* 1:e11
81. Vignuzzi M, Wendt E, Andino R. 2008. Engineering attenuated virus vaccines by controlling replication fidelity. *Nat. Med.* 14:154–61
82. Regoes RR, Hamblin S, Tanaka MM. 2013. Viral mutation rates: modelling the roles of within-host viral dynamics and the trade-off between replication fidelity and speed. *Proc. R. Soc. B* 280:20122047
83. Schaaper RM. 1998. Antimutator mutants in bacteriophage T4 and *Escherichia coli*. *Genetics* 148:1579–85
84. Muzyczka N, Poland RL, Bessman MJ. 1972. Studies on the biochemical basis of spontaneous mutation. I. A comparison of the deoxyribonucleic acid polymerases of mutator, antimutator, and wild type strains of bacteriophage T4. *J. Biol. Chem.* 247:7116–22
85. Gillin FD, Nossal NG. 1976. Control of mutation frequency by bacteriophage T4 DNA polymerase. I. The CB120 antimutator DNA polymerase is defective in strand displacement. *J. Biol. Chem.* 251:5219–24
86. Pfeiffer JK, Kirkegaard K. 2003. A single mutation in poliovirus RNA-dependent RNA polymerase confers resistance to mutagenic nucleotide analogs via increased fidelity. *Proc. Natl. Acad. Sci. USA* 100:7289–94
87. Coffey LL, Beechery Y, Bordería AV, Blanc H, Vignuzzi M. 2011. Arbovirus high fidelity variant loses fitness in mosquitoes and mice. *Proc. Natl. Acad. Sci. USA* 108:16038–43

88. Harrison DN, Gazina EV, Purcell DF, Anderson DA, Petrou S. 2008. Amiloride derivatives inhibit coxsackievirus B3 RNA replication. *J. Virol.* 82:1465–73
89. Arnold JJ, Vignuzzi M, Stone JK, Andino R, Cameron CE. 2005. Remote site control of an active site fidelity checkpoint in a viral RNA-dependent RNA polymerase. *J. Biol. Chem.* 280:25706–16
90. Rozen-Gagnon K, Stapleford KA, Mongelli V, Blanc H, Failloux AB, et al. 2014. Alphavirus mutator variants present host-specific defects and attenuation in mammalian and insect models. *PLoS Pathog.* 10:e1003877
91. Zeng J, Wang H, Xie X, Yang D, Zhou G, Yu L. 2013. An increased replication fidelity mutant of foot-and-mouth disease virus retains fitness in vitro and virulence in vivo. *Antivir. Res.* 100:1–7
92. Zeng J, Wang H, Xie X, Li C, Zhou G, et al. 2014. Ribavirin-resistant variants of foot-and-mouth disease virus: the effect of restricted quasispecies diversity on viral virulence. *J. Virol.* 88:4008–20
93. Liu X, Yang X, Lee CA, Moustafa IM, Smidansky ED, et al. 2013. Vaccine-derived mutation in motif D of poliovirus RNA-dependent RNA polymerase lowers nucleotide incorporation fidelity. *J. Biol. Chem.* 288:32753–65
94. Sadeghipour S, McMinn PC. 2013. A study of the virulence in mice of high copying fidelity variants of human enterovirus 71. *Virus Res.* 176:265–72
95. Sadeghipour S, Bek EJ, McMinn PC. 2013. Ribavirin-resistant mutants of human enterovirus 71 express a high replication fidelity phenotype during growth in cell culture. *J. Virol.* 87:1759–69
96. Weeks SA, Lee CA, Zhao Y, Smidansky ED, August A, et al. 2012. A polymerase mechanism-based strategy for viral attenuation and vaccine development. *J. Biol. Chem.* 287:31618–22
97. Arias A, Arnold JJ, Sierra M, Smidansky ED, Domingo E, Cameron CE. 2008. Determinants of RNA-dependent RNA polymerase (in)fidelity revealed by kinetic analysis of the polymerase encoded by a foot-and-mouth disease virus mutant with reduced sensitivity to ribavirin. *J. Virol.* 82:12346–55
98. Levi LI, Gnädig NF, Beaucourt S, McPherson MJ, Baron B, et al. 2010. Fidelity variants of RNA dependent RNA polymerases uncover an indirect, mutagenic activity of amiloride compounds. *PLoS Pathog.* 6:e1001163
99. Sierra M, Airaksinen A, González-López C, Agudo R, Arias A, Domingo E. 2007. Foot-and-mouth disease virus mutant with decreased sensitivity to ribavirin: implications for error catastrophe. *J. Virol.* 81:2012–24
100. Nga PT, Parquet MDC, Lauber C, Parida M, Nabeshima T, et al. 2011. Discovery of the first insect nidovirus, a missing evolutionary link in the emergence of the largest RNA virus genomes. *PLoS Pathog.* 7:e1002215
101. Zirkel F, Kurth A, Quan PL, Briese T, Ellerbrok H, et al. 2011. An insect nidovirus emerging from a primary tropical rainforest. *mBio* 2:e00077-11
102. Zuo Y, Deutscher MP. 2001. Exoribonuclease superfamilies: structural analysis and phylogenetic distribution. *Nucleic Acids Res.* 29:1017–26
103. Chen Y, Tao J, Sun Y, Wu A, Su C, et al. 2013. Structure-function analysis of severe acute respiratory syndrome coronavirus RNA cap guanine-N7-methyltransferase. *J. Virol.* 87:6296–305
104. Smith EC, Denison MR. 2013. Coronaviruses as DNA wannabes: a new model for the regulation of RNA virus replication fidelity. *PLoS Pathog.* 9:e1003760
105. Hastie KM, Kimberlin CR, Zandonatti MA, MacRae IJ, Saphire EO. 2011. Structure of the Lassa virus nucleoprotein reveals a dsRNA-specific 3' to 5' exonuclease activity essential for immune suppression. *Proc. Natl. Acad. Sci. USA* 108:2396–401
106. Huang C, Lokugamage KG, Rozovics JM, Narayanan K, Semler BL, Makino S. 2011. SARS coronavirus nsp1 protein induces template-dependent endonucleolytic cleavage of mRNAs: viral mRNAs are resistant to nsp1-induced RNA cleavage. *PLoS Pathog.* 7:e1002433
107. Shi X, Wang L, Li X, Zhang G, Guo J, et al. 2011. Endoribonuclease activities of porcine reproductive and respiratory syndrome virus nsp11 was essential for nsp11 to inhibit IFN- $\beta$  induction. *Mol. Immunol.* 48:1568–72
108. Frieman M, Ratia K, Johnston RE, Mesecar AD, Baric RS. 2009. Severe acute respiratory syndrome coronavirus papain-like protease ubiquitin-like domain and catalytic domain regulate antagonism of IRF3 and NF- $\kappa$ B signaling. *J. Virol.* 83:6689–705
100. A paper that helps close the evolutionary gap between the small and large nidoviruses.



---

**111. Description of the expansion hierarchy of the coronavirus genome using comparative genomics.**

---

109. Kun A, Santos M, Szathmáry E. 2005. Real ribozymes suggest a relaxed error threshold. *Nat. Genet.* 37:1008–11
110. Smith JM. 1979. Hypercycles and the origin of life. *Nature* 280:445–46
111. **Lauber C, Goeman JJ, Parquet MDC, Nga PT, Snijder EJ, et al. 2013. The footprint of genome architecture in the largest genome expansion in RNA viruses. *PLoS Pathog.* 9:e1003500**
112. von Brunn A, Teepe C, Simpson JC, Pepperkok R, Friedel CC, et al. 2007. Analysis of intraviral protein-protein interactions of the SARS coronavirus ORFome. *PLoS ONE* 2:e459
113. Pan J, Peng X, Gao Y, Li Z, Lu X, et al. 2008. Genome-wide analysis of protein-protein interactions and involvement of viral proteins in SARS-CoV replication. *PLoS ONE* 3:e3299
114. Imbert I, Snijder EJ, Dimitrova M, Guillemot JC, Lécine P, Canard B. 2008. The SARS-coronavirus PLnc domain of nsp3 as a replication/transcription scaffolding protein. *Virus Res.* 133:136–48
115. Subissi L, Imbert I, Ferron F, Collet A, Coutard B, et al. 2013. SARS-CoV ORF1b-encoded nonstructural proteins 12–16: replicative enzymes as antiviral targets. *Antivir. Res.* 101:122–30
116. Johansson E, Dixon N. 2013. Replicative DNA polymerases. *Cold Spring Harb. Perspect. Biol.* 5:a012799
117. Studwell-Vaughan PS, O'Donnell M. 1993. DNA polymerase III accessory proteins. V.  $\theta$  encoded by holE. *J. Biol. Chem.* 268:11785–91
118. Taft-Benz SA, Schaaper RM. 2004. The  $\theta$  subunit of *Escherichia coli* DNA polymerase III: a role in stabilizing the  $\epsilon$  proofreading subunit. *J. Bacteriol.* 186:2774–80
119. Sawicki D, Wang T, Sawicki S. 2001. The RNA structures engaged in replication and transcription of the A59 strain of mouse hepatitis virus. *J. Gen. Virol.* 82:385–96
120. Zhuang Z, Ai Y. 2010. Processivity factor of DNA polymerase and its expanding role in normal and translesion DNA synthesis. *Biochim. Biophys. Acta* 1804:1081–93
121. Züst R, Cervantes-Barragan L, Habjan M, Maier R, Neuman BW, et al. 2011. Ribose 2'-O-methylation provides a molecular signature for the distinction of self and non-self mRNA dependent on the RNA sensor Mda5. *Nat. Immunol.* 12:137–43
122. Egloff MP, Ferron F, Campanacci V, Longhi S, Rancurel C, et al. 2004. The severe acute respiratory syndrome–coronavirus replicative protein nsp9 is a single-stranded RNA-binding subunit unique in the RNA virus world. *Proc. Natl. Acad. Sci. USA* 101:3792–96

## REFERENCES

- Abudayyeh, O.O., Gootenberg, J.S., Konermann, S., Joung, J., Slaymaker, I.M., Cox, D.B.T., Shmakov, S., Makarova, K.S., Semenova, E., Minakhin, L., et al. (2016). C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. *Science* 353, aaf5573.
- Acevedo, A., and Andino, R. (2014). Library preparation for highly accurate population sequencing of RNA viruses. *Nat Protoc* 9, 1760–1769.
- Adedeji, A.O., Marchand, B., Velthuis, te, A.J.W., Snijder, E.J., Weiss, S., Eoff, R.L., Singh, K., and Sarafianos, S.G. (2012). Mechanism of nucleic acid unwinding by SARS-CoV helicase. *PLoS ONE* 7, e36521.
- Ahn, D.-G., Choi, J.-K., Taylor, D.R., and Oh, J.-W. (2012). Biochemical characterization of a recombinant SARS coronavirus nsp12 RNA-dependent RNA polymerase capable of copying viral RNA templates. *Arch Virol* 157, 2095–2104.
- Andino, R., Böddeker, N., Silvera, D., and Gamarnik, A.V. (1999). Intracellular determinants of picornavirus replication. *Trends Microbiol* 7, 76–82.
- Arias, A., Arnold, J.J., Sierra, M., Smidansky, E.D., Domingo, E., and Cameron, C.E. (2008). Determinants of RNA-Dependent RNA Polymerase (In)fidelity Revealed by Kinetic Analysis of the Polymerase Encoded by a Foot-and-Mouth Disease Virus Mutant with Reduced Sensitivity to Ribavirin. *J Virol* 82, 12346–12355.
- Arnold, J.J., Vignuzzi, M., Stone, J.K., Andino, R., and Cameron, C.E. (2005). Remote Site Control of an Active Site Fidelity Checkpoint in a Viral RNA-dependent RNA Polymerase. *J Biol Chem* 280, 25706–25716.
- Beaucourt, S., Bordería, A.V., Coffey, L.L., Gnädig, N.F., Sanz-Ramos, M., Beeharry, Y., and Vignuzzi, M. (2011). Isolation of fidelity variants of RNA viruses and characterization of virus mutation frequency. *J Vis Exp*.
- Becares, M., Pascual-Iglesias, A., Nogales, A., Sola, I., Enjuanes, L., and Zúñiga, S. (2016). Mutagenesis of Coronavirus nsp14 Reveals Its Potential Role in Modulation of the Innate Immune Response. *J Virol* 90, 5399–5414.
- Berkhout, B., and van Hemert, F. (2015). On the biased nucleotide composition of the human coronavirus RNA genome. *Virus Res* 202, 41–47.
- Bolles, M., Donaldson, E., and Baric, R. (2011). SARS-CoV and emergent coronaviruses: viral determinants of interspecies transmission. *Curr Opin Virol* 1, 624–634.
- Bordería, A.V., Isakov, O., Moratorio, G., Henningsson, R., Agüera-González, S., Organtini, L., Gnädig, N.F., Blanc, H., Alcover, A., Hafenstein, S., et al. (2015). Group Selection and Contribution of Minority Variants during Virus Adaptation Determines Virus Fitness and Phenotype. *Plos Pathog* 11, e1004838.

- Bouvet, M., Imbert, I., Subissi, L., Gluais, L., Canard, B., and Decroly, E. (2012). RNA 3'-end mismatch excision by the severe acute respiratory syndrome coronavirus nonstructural protein nsp10/nsp14 exoribonuclease complex. *Proc Natl Acad Sci USA* *109*, 9372–9377.
- Bouvet, M., Lugari, A., Posthuma, C.C., Zevenhoven, J.C., Bernard, S., Betzi, S., Imbert, I., Canard, B., Guillemot, J.-C., Lécine, P., et al. (2014). Coronavirus Nsp10, a Critical Co-factor for Activation of Multiple Replicative Enzymes. *J Biol Chem* *289*, 25783–25796.
- Brayton, P.R., Lai, M.M., Patton, C.D., and Stohlman, S.A. (1982). Characterization of two RNA polymerase activities induced by mouse hepatitis virus. *J Virol* *42*, 847–853.
- Breban, R., Riou, J., and Fontanet, A. (2013). Interhuman transmissibility of Middle East respiratory syndrome coronavirus: estimation of pandemic risk. *Lancet*.
- Bredenbeek, P.J., Pachuk, C.J., Noten, A.F.H., Charité, J., Luytjes, W., Weiss, S.R., and Spaan, W.J.M. (1990). The primary structure and expression of the second open reading frame of the polymerase gene of the coronavirus MHV-A59; a highly conserved polymerase is expressed by an efficient ribosomal frameshifting mechanism. *Nucleic Acids Res* *18*, 1825–1832.
- Briese, T., Mishra, N., Jain, K., Zalmout, I.S., Jabado, O.J., Karesh, W.B., Daszak, P., Mohammed, O.B., Alagaili, A.N., and Lipkin, W.I. (2014). Middle East respiratory syndrome coronavirus quasispecies that include homologues of human isolates revealed through whole-genome analysis and virus cultured from dromedary camels in Saudi Arabia. *MBio* *5*, e01146–14.
- Brockway, S.M., Clay, C.T., Lu, X.T., and Denison, M.R. (2003). Characterization of the expression, intracellular localization, and replication complex association of the putative mouse hepatitis virus RNA-dependent RNA polymerase. *J Virol* *77*, 10515–10527.
- Brunn, von, A., Teepe, C., Simpson, J.C., Pepperkok, R., Friedel, C.C., Zimmer, R., Roberts, R., Baric, R., and Haas, J. (2007). Analysis of intraviral protein-protein interactions of the SARS coronavirus ORFome. *PLoS ONE* *2*, e459.
- Calisher, C.H., Childs, J.E., Field, H.E., Holmes, K.V., and Schountz, T. (2006). Bats: important reservoir hosts of emerging viruses. *Clin Microbiol Rev* *19*, 531–545.
- Campagnola, G., McDonald, S., Beaucourt, S., Vignuzzi, M., and Peersen, O.B. (2015). Structure-function relationships underlying the replication fidelity of viral RNA-dependent RNA polymerases. *J Virol* *89*, 275–286.
- Campagnola, G., Weygandt, M., Scoggin, K., and Peersen, O. (2008). Crystal structure of coxsackievirus B3 3Dpol highlights the functional importance of residue 5 in picornavirus polymerases. *J Virol* *82*, 9458–9464.
- Carroll, S.S., Tomassini, J.E., Bosserman, M., Getty, K., Stahlhut, M.W., Eldrup, A.B., Bhat, B., Hall, D., Simcoe, A.L., LaFemina, R., et al. (2003). Inhibition of hepatitis C virus RNA replication by 2'-modified nucleoside analogs. *J Biol Chem* *278*, 11979–11984.

- Case, J.B., Ashbrook, A.W., Dermody, T.S., and Denison, M.R. (2016). Mutagenesis of S-Adenosyl-l-Methionine-Binding Residues in Coronavirus nsp14 N7-Methyltransferase Demonstrates Differing Requirements for Genome Translation and Resistance to Innate Immunity. *J Virol* *90*, 7248–7256.
- Castro, C., Arnold, J.J., and Cameron, C.E. (2005). Incorporation fidelity of the viral RNA-dependent RNA polymerase: a kinetic, thermodynamic and structural perspective. *Virus Res* *107*, 141–149.
- Cauchemez, S., Fraser, C., Van Kerkhove, M.D., Donnelly, C.A., Riley, S., Rambaut, A., Enouf, V., van der Werf, S., and Ferguson, N.M. (2014). Middle East respiratory syndrome coronavirus: quantification of the extent of the epidemic, surveillance biases, and transmissibility. *Lancet Infect Dis* *14*, 50–56.
- Chan, C.-M., Chu, H., Wang, Y., Wong, B.H.-Y., Zhao, X., Zhou, J., Yang, D., Leung, S.P., Chan, J.F.-W., Yeung, M.L., et al. (2016). Carcinoembryonic Antigen-Related Cell Adhesion Molecule 5 Is an Important Surface Attachment Factor That Facilitates Entry of Middle East Respiratory Syndrome Coronavirus. *J Virol* *90*, 9114–9127.
- Chen, P., and Shakhnovich, E.I. (2009). Lethal mutagenesis in viruses and bacteria. *Genetics* *183*, 639–650.
- Chen, P., Jiang, M., Hu, T., Liu, Q., Chen, X.S., and Guo, D. (2007). Biochemical characterization of exoribonuclease encoded by SARS coronavirus. *J Biochem Mol Biol* *40*, 649–655.
- Chen, W., and Baric, R.S. (1996). Molecular anatomy of mouse hepatitis virus persistence: coevolution of increased host cell resistance and virus virulence. *J Virol* *70*, 3947–3960.
- Chen, Y., Cai, H., Pan, J., Xiang, N., Tien, P., Ahola, T., and Guo, D. (2009). Functional screen reveals SARS coronavirus nonstructural protein nsp14 as a novel cap N7 methyltransferase. *Proc Natl Acad Sci USA* *106*, 3484–3489.
- Chen, Y., Su, C., Ke, M., Jin, X., Xu, L., Zhang, Z., Wu, A., Sun, Y., Yang, Z., Tien, P., et al. (2011). Biochemical and structural insights into the mechanisms of SARS coronavirus RNA ribose 2'-O-methylation by nsp16/nsp10 protein complex. *Plos Pathog* *7*, e1002294.
- Cheng, A., Zhang, W., Xie, Y., Jiang, W., Arnold, E., Sarafianos, S.G., and Ding, J. (2005). Expression, purification, and characterization of SARS coronavirus RNA polymerase. *Virology* *335*, 165–176.
- Chung, D.H., Sun, Y., Parker, W.B., Arterburn, J.B., Bartolucci, A., and Jonsson, C.B. (2007). Ribavirin Reveals a Lethal Threshold of Allowable Mutation Frequency for Hantaan Virus. *J Virol* *81*, 11722–11729.
- Coffey, L.L., Beeharry, Y., Bordería, A.V., Blanc, H., and Vignuzzi, M. (2011). Arbovirus high fidelity variant loses fitness in mosquitoes and mice. *Proc Natl Acad Sci USA* *108*, 16038–16043.

- Combe, M., Garijo, R., Geller, R., Cuevas, J.M., and Sanjuán, R. (2015). Single-Cell Analysis of RNA Virus Infection Identifies Multiple Genetically Diverse Viral Genomes within Single Infectious Units. *Cell Host Microbe* 18, 424–432.
- Crossley, B., Mock, R., Callison, S., and Hietala, S. (2012). Identification and Characterization of a Novel Alpaca Respiratory Coronavirus Most Closely Related to the Human Coronavirus 229E. *Viruses* 4, 3689–3700.
- Crotty, S., and Andino, R. (2002). Implications of high RNA virus mutation rates: lethal mutagenesis and the antiviral drug ribavirin. *Microbes Infect* 4, 1301–1307.
- Crotty, S., Cameron, C.E., and Andino, R. (2001). RNA virus error catastrophe: direct molecular test by using ribavirin. *Proc Natl Acad Sci USA* 98, 6895–6900.
- Crotty, S., Cameron, C., and Andino, R. (2002). Ribavirin's antiviral mechanism of action: lethal mutagenesis? *J. Mol. Med.* 80, 86–95.
- Crotty, S., Maag, D., Arnold, J.J., Zhong, W., Lau, J.Y.N., Hong, Z., Andino, R., and Cameron, C.E. (2000). The broad-spectrum antiviral ribonucleoside ribavirin is an RNA virus mutagen. *Nat Med* 6, 1375–1379.
- Curti, E., and Jaeger, J. (2013). Residues Arg283, Arg285, and Ile287 in the Nucleotide Binding Pocket of Bovine Viral Diarrhea Virus NS5B RNA Polymerase Affect Catalysis and Fidelity. *J Virol* 87, 199–207.
- Dapp, M.J., Heineman, R.H., and Mansky, L.M. (2013). Interrelationship between HIV-1 fitness and mutation rate. *J Mol Biol* 425, 41–53.
- Decroly, E., Debarnot, C., Ferron, F., Bouvet, M., Coutard, B., Imbert, I., Gluais, L., Papageorgiou, N., Sharff, A., Bricogne, G., et al. (2011). Crystal structure and functional analysis of the SARS-coronavirus RNA cap 2'-O-methyltransferase nsp10/nsp16 complex. *Plos Pathog* 7, e1002059.
- Denison, M.R., Graham, R.L., Donaldson, E.F., Eckerle, L.D., and Baric, R.S. (2011). Coronaviruses: An RNA proofreading machine regulates replication fidelity and diversity. *RNA Biol* 8, 270–279.
- Domingo, E., Escarmis, C., Sevilla, N., Moya, A., Elena, S.F., Quer, J., Novella, I.S., and Holland, J.J. (1996). Basic concepts in RNA virus evolution. *Faseb J* 10, 859–864.
- Domingo, E. (2010). Mechanisms of viral emergence. *Vet Res* 41, 38.
- Domingo, E., Sheldon, J., and Perales, C. (2012). Viral quasispecies evolution. *Microbiol Mol Biol Rev* 76, 159–216.
- Drake, J.W., and Holland, J.J. (1999). Mutation rates among RNA viruses. *Proc Natl Acad Sci USA* 96, 13910–13913.

- Drake, J.W., Charlesworth, B., Charlesworth, D., and Crow, J.F. (1998). Rates of spontaneous mutation. *Genetics* *148*, 1667–1686.
- Drexler, J.F., Corman, V.M., and Drosten, C. (2013). Ecology, evolution and classification of bat coronaviruses in the aftermath of SARS. *Antiviral Res* *101C*, 45–56.
- Dulin, D., Vilfan, I.D., Berghuis, B.A., Poranen, M.M., Depken, M., and Dekker, N.H. (2015). Backtracking behavior in viral RNA-dependent RNA polymerase provides the basis for a second initiation site. *Nucleic Acids Res* *43*, 10421–10429.
- Dveksler, G.S., Pensiero, M.N., Cardellichio, C.B., Williams, R.K., Jiang, G.S., Holmes, K.V., and Dieffenbach, C.W. (1991). Cloning of the mouse hepatitis virus (MHV) receptor: expression in human and hamster cell lines confers susceptibility to MHV. *J Virol* *65*, 6881–6891.
- East-Seletsky, A., O'Connell, M.R., Knight, S.C., Burstein, D., Cate, J.H.D., Tjian, R., and Doudna, J.A. (2016). Two distinct RNase activities of CRISPR-C2c2 enable guide-RNA processing and RNA detection. *Nature* *538*, 270–273.
- Eckerle, L.D., Becker, M.M., Halpin, R.A., Li, K., Venter, E., Lu, X., Scherbakova, S., Graham, R.L., Baric, R.S., Stockwell, T.B., et al. (2010). Infidelity of SARS-CoV Nsp14-exonuclease mutant virus replication is revealed by complete genome sequencing. *Plos Pathog* *6*, e1000896.
- Eckerle, L.D., Lu, X., Sperry, S.M., Choi, L., and Denison, M.R. (2007). High fidelity of murine hepatitis virus replication is decreased in nsp14 exonuclease mutants. *J Virol* *81*, 12135–12144.
- Eigen, M. (1993). The origin of genetic information: viruses as models. *Gene* *135*, 37–47.
- Elena, S.F., and Sanjuán, R. (2005). Adaptive value of high mutation rates of RNA viruses: separating causes from consequences. *J Virol* *79*, 11555–11558.
- Ferrer-Orta, C., Arias, A., Perez-Luque, R., Escarmis, C., Domingo, E., and Verdaguer, N. (2007). Sequential structures provide insights into the fidelity of RNA replication. *Proc Natl Acad Sci USA* *104*, 9463–9468.
- Ferrer-Orta, C., Ferrero, D., and Verdaguer, N. (2015). RNA-Dependent RNA Polymerases of Picornaviruses: From the Structure to Regulatory Mechanisms. *Viruses* *7*, 4438–4460.
- Forni, D., Cagliani, R., Mozzi, A., Pozzoli, U., Al-Daghri, N., Clerici, M., and Sironi, M. (2016). Extensive Positive Selection Drives the Evolution of Nonstructural Proteins in Lineage C Betacoronaviruses. *J Virol* *90*, 3627–3639.
- Freeman, M.C., Graham, R.L., Lu, X., Peek, C.T., and Denison, M.R. (2014). Coronavirus replicase-reporter fusions provide quantitative analysis of replication and replication complex formation. *J Virol* *88*, 5319–5327.
- Furio, V., Moya, A., and Sanjuán, R. (2005). The cost of replication fidelity in an RNA virus. *Proc Natl Acad Sci USA* *102*, 10233–10237.

- Ge, X.-Y., Li, J.-L., Yang, X.-L., Chmura, A.A., Zhu, G., Epstein, J.H., Mazet, J.K., Ben Hu, Zhang, W., Peng, C., et al. (2013). Isolation and characterization of a bat SARS-like coronavirus that uses the ACE2 receptor. *Nature* 1–16.
- Gnädig, N.F., Beaucourt, S., Campagnola, G., Bordería, A.V., Sanz-Ramos, M., Gong, P., Blanc, H., Peersen, O.B., and Vignuzzi, M. (2012). Coxsackievirus B3 mutator strains are attenuated in vivo. *Proc. Natl. Acad. Sci. U.S.a.* *109*, 13484–13485.
- Goebel, S.J., Miller, T.B., Bennett, C.J., Bernard, K.A., and Masters, P.S. (2007). A hypervariable region within the 3' cis-acting element of the murine coronavirus genome is nonessential for RNA synthesis but affects pathogenesis. *J Virol* *81*, 1274–1287.
- Goh, G.K.-M., Dunker, A.K., and Uversky, V. (2013). Prediction of Intrinsic Disorder in MERS-CoV/HCoV-EMC Supports a High Oral-Fecal Transmission. *PLoS Curr* *5*.
- Gohara, D.W., Arnold, J.J., and Cameron, C.E. (2004). Poliovirus RNA-dependent RNA polymerase (3Dpol): kinetic, thermodynamic, and structural analysis of ribonucleotide selection. *Biochemistry* *43*, 5149–5158.
- Gong, P., and Peersen, O.B. (2010). Structural basis for active site closure by the poliovirus RNA-dependent RNA polymerase. *Proc Natl Acad Sci USA* *107*, 22505–22510.
- Gong, P., Campagnola, G., and Peersen, O.B. (2009). A quantitative stopped-flow fluorescence assay for measuring polymerase elongation rates. *Anal Biochem* *391*, 45–55.
- Gorbalenya, A.E., Enjuanes, L., Ziebuhr, J., and Snijder, E.J. (2006). Nidovirales: Evolving the largest RNA virus genome. *Virus Res* *117*, 17–37.
- Graci, J.D., and Cameron, C.E. (2006). Mechanisms of action of ribavirin against distinct viruses. *Rev Med Virol* *16*, 37–48.
- Graci, J.D., Gnädig, N.F., Galarraga, J.E., Castro, C., Vignuzzi, M., and Cameron, C.E. (2012). Mutational Robustness of an RNA Virus Influences Sensitivity to Lethal Mutagenesis. *J Virol* *86*, 2869–2873.
- Graham, R.L., Becker, M.M., Eckerle, L.D., Bolles, M., Denison, M.R., and Baric, R.S. (2012). A live, impaired-fidelity coronavirus vaccine protects in an aged, immunocompromised mouse model of lethal disease. *Nat Med* *18*, 1820–1826.
- Graham, R.L., Donaldson, E.F., and Baric, R.S. (2013). A decade after SARS: strategies for controlling emerging coronaviruses. *Nat Rev Micro* *11*, 836–848.
- Gu, S., Li, W., Zhang, H., Fleming, J., Yang, W., Wang, S., Wei, W., Zhou, J., Zhu, G., Deng, J., et al. (2016). The  $\beta 2$  clamp in the Mycobacterium tuberculosis DNA polymerase III  $\alpha\beta 2\epsilon$  replicase promotes polymerization and reduces exonuclease activity. *Sci Rep* *6*, 18418.
- Hansen, J.L., Long, A.M., and Schultz, S.C. (1997). Structure of the RNA-dependent RNA polymerase of poliovirus. *Structure* *5*, 1109–1122.

Hemida, M., Perera, R., Wang, P., Alhammadi, M., Siu, L., Li, M., Poon, L., Saif, L., Alnaeem, A., and Peiris, M. (2013). Middle East Respiratory Syndrome (MERS) coronavirus seroprevalence in domestic livestock in Saudi Arabia, 2010 to 2013. *Euro Surveill* 18.

Hilgenfeld, R., and Peiris, M. (2013). From SARS to MERS: 10 years of research on highly pathogenic human coronaviruses. *Antiviral Res* 100, 286–295.

Hobdey, S.E., Kempf, B.J., Steil, B.P., Barton, D.J., and Peersen, O.B. (2010). Poliovirus polymerase residue 5 plays a critical role in elongation complex stability. *J Virol* 84, 8072–8084.

Holland, J.J., Domingo, E., la Torre, de, J.C., and Steinhauer, D.A. (1990). Mutation frequencies at defined single codon sites in vesicular stomatitis virus and poliovirus can be increased only slightly by chemical mutagenesis. *J Virol* 64, 3960–3962.

Hsue, B., and Masters, P.S. (1997). A bulged stem-loop structure in the 3'untranslated region of the genome of the coronavirus mouse hepatitis virus is essential for replication. *J Virol* 71, 7567–7578.

Huynh, J., Li, S., Yount, B., Smith, A., Sturges, L., Olsen, J.C., Nagel, J., Johnson, J.B., Agnihothram, S., Gates, J.E., et al. (2012). Evidence Supporting a Zoonotic Origin of Human Coronavirus Strain NL63. *J Virol*.

Imbert, I., Guillemot, J.-C., Bourhis, J.-M., Bussetta, C., Coutard, B., Egloff, M.-P., Ferron, F., Gorbalenya, A.E., and Canard, B. (2006). A second, non-canonical RNA-dependent RNA polymerase in SARS coronavirus. *Embo J* 25, 4933–4942.

Imbert, I., Snijder, E.J., Dimitrova, M., Guillemot, J.-C., Lécine, P., and Canard, B. (2008). The SARS-Coronavirus PLnc domain of nsp3 as a replication/transcription scaffolding protein. *Virus Res* 133, 136–148.

Isakov, O., Bordería, A.V., Golan, D., Hamenahem, A., Celniker, G., Yoffe, L., Blanc, H., Vignuzzi, M., and Shomron, N. (2015). Deep sequencing analysis of viral infection and evolution allows rapid and detailed characterization of viral mutant spectrum. *Bioinformatics* 31, 2141–2150.

Jacobson, I.M., Gordon, S.C., Kowdley, K.V., Yoshida, E.M., Rodriguez-Torres, M., Sulkowski, M.S., Shiffman, M.L., Lawitz, E., Everson, G., Bennett, M., et al. (2013). Sofosbuvir for hepatitis C genotype 2 or 3 in patients without treatment options. *N Engl J Med* 368, 1867–1877.

Jiang, C., Komazin-Meredith, G., Tian, W., Coen, D.M., and Hwang, C.B.C. (2009). Mutations that increase DNA binding by the processivity factor of herpes simplex virus affect virus production and DNA replication fidelity. *J Virol* 83, 7573–7580.

Kang, H., Feng, M., Schroeder, M.E., Giedroc, D.P., and Leibowitz, J.L. (2006). Putative cis-acting stem-loops in the 5' untranslated region of the severe acute respiratory syndrome coronavirus can substitute for their mouse hepatitis virus counterparts. *J Virol* 80, 10600–10614.

Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N., and Sternberg, M.J.E. (2015). The Phyre2



web portal for protein modeling, prediction and analysis. *Nat Protoc* 10, 845–858.

Kennedy, S.R., Schmitt, M.W., Fox, E.J., Kohn, B.F., Salk, J.J., Ahn, E.H., Prindle, M.J., Kuong, K.J., Shen, J.-C., Risques, R.-A., et al. (2014). Detecting ultralow-frequency mutations by Duplex Sequencing. *Nat Protoc* 9, 2586–2606.

Korneeva, V.S., and Cameron, C.E. (2007). Structure-function relationships of the viral RNA-dependent RNA polymerase: fidelity, replication speed, and initiation mechanism determined by a residue in the ribose-binding pocket. *J Biol Chem* 282, 16135–16145.

Kortus, M.G., Kempf, B.J., Haworth, K.G., Barton, D.J., and Peersen, O.B. (2012). A template RNA entry channel in the fingers domain of the poliovirus polymerase. *J Mol Biol* 417, 263–278.

Lauring, A.S., and Andino, R. (2010). Quasispecies Theory and the Behavior of RNA Viruses. *Plos Pathog* 6, e1001005.

Lauring, A.S., Acevedo, A., Cooper, S.B., and Andino, R. (2012). Codon usage determines the mutational robustness, evolutionary capacity, and virulence of an RNA virus. *Cell Host Microbe* 12, 623–632.

Lauring, A.S., Frydman, J., and Andino, R. (2013). The role of mutational robustness in RNA virus evolution. *Nat Rev Micro*.

Lauring, A.S., Jones, J.O., and Andino, R. (2010). Rationalizing the development of live attenuated virus vaccines. *Nat Biotechnol* 28, 573–579.

Le Cong, Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., et al. (2013). Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science* 339, 819–823.

Lee, C.H., Gilbertson, D.L., Novella, I.S., Huerta, R., Domingo, E., and Holland, J.J. (1997). Negative effects of chemical mutagenesis on the adaptive behavior of vesicular stomatitis virus. *J Virol* 71, 3636–3640.

Lee, J., Chowell, G., and Jung, E. (2016). A dynamic compartmental model for the Middle East respiratory syndrome outbreak in the Republic of Korea: A retrospective analysis on control interventions and superspreading events. *Journal of Theoretical Biology* 408, 118–126.

Lehmann, K.C., Gulyaeva, A., Zevenhoven-Dobbe, J.C., Janssen, G.M.C., Ruben, M., Overkleeft, H.S., van Veelen, P.A., Samborskiy, D.V., Kravchenko, A.A., Leontovich, A.M., et al. (2015). Discovery of an essential nucleotidylating activity associated with a newly delineated conserved domain in the RNA polymerase-containing protein of all nidoviruses. *Nucleic Acids Res* 43, 8416–8434.

Levi, L.I., Gnädig, N.F., Beaucourt, S., McPherson, M.J., Baron, B., Arnold, J.J., and Vignuzzi, M. (2010). Fidelity variants of RNA dependent RNA polymerases uncover an indirect, mutagenic activity of amiloride compounds. *Plos Pathog* 6, e1001163.

- Li, H., and Durbin, R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* 26, 589–595.
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078–2079.
- Longley, D.B., Harkin, D.P., and Johnston, P.G. (2003). 5-Fluorouracil: mechanisms of action and clinical strategies. *Nature Reviews Cancer* 3, 330–338.
- Lou, D.I., Hussmann, J.A., McBee, R.M., Acevedo, A., Andino, R., Press, W.H., and Sawyer, S.L. (2013). High-throughput DNA sequencing errors are reduced by orders of magnitude using circle sequencing. *Proc Natl Acad Sci USA* 110, 19872–19877.
- Löffler, M., Carrey, E.A., and Zameitat, E. (2016). Orotate (orotic acid): An essential and versatile molecule. *Nucleosides Nucleotides Nucleic Acids* 35, 566–577.
- Ma, Y., Wu, L., Shaw, N., Gao, Y., Wang, J., Sun, Y., Lou, Z., Yan, L., Zhang, R., and Rao, Z. (2015). Structural basis and functional analysis of the SARS coronavirus nsp14–nsp10 complex. *Proc Natl Acad Sci USA* 112, 9436–9441.
- Maki, H., and Kornberg, A. (1987). Proofreading by DNA polymerase III of *Escherichia coli* depends on cooperative interaction of the polymerase and exonuclease subunits. *Proc Natl Acad Sci USA* 84, 4389–4392.
- Makino, S., Keck, J.G., Stohlman, S.A., and Lai, M.M. (1986). High-frequency RNA recombination of murine coronaviruses. *J Virol* 57, 729–737.
- Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E., and Church, G.M. (2013). RNA-guided human genome engineering via Cas9. *Science* 339, 823–826.
- Masters, P.S. (2006). *The Molecular Biology of Coronaviruses* (Elsevier).
- Memish, Z.A., Mishra, N., Olival, K.J., Fagbo, S.F., Kapoor, V., Epstein, J.H., Alhakeem, R., Durosinsoun, A., Asmari, Al, M., Islam, A., et al. (2013). Middle East respiratory syndrome coronavirus in bats, Saudi Arabia. *Emerg Infect Diseases* 19, 1819–1823.
- Minskaia, E., Hertzog, T., Gorbalenya, A.E., Campanacci, V., Cambillau, C., Canard, B., and Ziebuhr, J. (2006). Discovery of an RNA virus 3'→5' exoribonuclease that is critically involved in coronavirus RNA synthesis. *Proc Natl Acad Sci USA* 103, 5108–5113.
- Mohanraju, P., Makarova, K.S., Zetsche, B., Zhang, F., Koonin, E.V., and van der Oost, J. (2016). Diverse evolutionary roots and mechanistic variations of the CRISPR-Cas systems. *Science* 353, aad5147–aad5147.
- Mokili, J.L., Rohwer, F., and Dutilh, B.E. (2012). Metagenomics and future perspectives in virus discovery. *Curr Opin Virol* 2, 63–77.

- Moustafa, I.M., Shen, H., Morton, B., Colina, C.M., and Cameron, C.E. (2011). Molecular dynamics simulations of viral RNA polymerases link conserved and correlated motions of functional elements to fidelity. *J Mol Biol* *410*, 159–181.
- Murphy, F.A. (2012). *The Foundations of Virology: Discoverers and Discoveries, Inventors and Inventions, Developers and Technologies*.
- Murray, K.E., and Nibert, M.L. (2007). Guanidine hydrochloride inhibits mammalian orthoreovirus growth by reversibly blocking the synthesis of double-stranded RNA. *J Virol* *81*, 4572–4584.
- Murray, K.O., Mertens, E., and Desprès, P. (2010). West Nile virus and its emergence in the United States of America. *Vet Res* *41*, 67.
- Ng, K.K.S., Arnold, J.J., and Cameron, C.E. (2008). Structure-function relationships among RNA-dependent RNA polymerases. *Curr Top Microbiol Immunol* *320*, 137–156.
- Ojosnegros, S., Perales, C., Mas, A., and Domingo, E. (2011). Quasispecies as a matter of fact: Viruses and beyond. *Virus Res* *162*, 203–215.
- Papadopoulou, C., Guilbaud, G., Schiavone, D., and Sale, J.E. (2015). Nucleotide Pool Depletion Induces G-Quadruplex-Dependent Perturbation of Gene Expression. *Cell Rep* *13*, 2491–2503.
- Parker, W.B. (2005). Metabolism and antiviral activity of ribavirin. *Virus Res* *107*, 165–171.
- Pauly, M.D., and Lauring, A.S. (2015). Effective lethal mutagenesis of influenza virus by three nucleoside analogs. *J Virol* *89*, 3584–3597.
- Peiris, J.S.M., Lai, S.T., Poon, L.L.M., Guan, Y., Yam, L.Y.C., Lim, W., Nicholls, J., Yee, W.K.S., Yan, W.W., Cheung, M.T., et al. (2003). Coronavirus as a possible cause of severe acute respiratory syndrome. *Lancet* *361*, 1319–1325.
- Penttinen, P.M., Kaasik-Aaslav, K., Friaux, A., Donachie, A., Sudre, B., Amato-Gauci, A.J., Memish, Z.A., and Coulombier, D. (2013). Taking stock of the first 133 MERS coronavirus cases globally--Is the epidemic changing? *Euro Surveill* *18*.
- Perlman, S., and Netland, J. (2009). Coronaviruses post-SARS: update on replication and pathogenesis. *Nat Rev Micro* *7*, 439–450.
- Pfefferle, S., Oppong, S., Drexler, J.F., Gloza-Rausch, F., Ipsen, A., Seebens, A., Müller, M.A., Annan, A., Vallo, P., Adu-Sarkodie, Y., et al. (2009). Distant relatives of severe acute respiratory syndrome coronavirus and close relatives of human coronavirus 229E in bats, Ghana. *Emerg Infect Diseases* *15*, 1377–1384.
- Pfeiffer, J.K., and Kirkegaard, K. (2003). A single mutation in poliovirus RNA-dependent RNA polymerase confers resistance to mutagenic nucleotide analogs via increased fidelity. *Proc Natl Acad Sci USA* *100*, 7289–7294.

- Pfeiffer, J.K., and Kirkegaard, K. (2005). Increased fidelity reduces poliovirus fitness and virulence under selective pressure in mice. *Plos Pathog* *1*, e11.
- Price, A.A., Sampson, T.R., Ratner, H.K., Grakoui, A., and Weiss, D.S. (2015). Cas9-mediated targeting of viral RNA in eukaryotic cells. *Proc Natl Acad Sci USA* *112*, 6164–6169.
- Raaben, M., Whitley, P., Bouwmeester, D., Setterquist, R.A., Rottier, P.J.M., and de Haan, C.A.M. (2008). Improved microarray gene expression profiling of virus-infected cells after removal of viral RNA. *BMC Genomics* *9*, 221.
- Raj, V.S., Mou, H., Smits, S.L., Dekkers, D.H.W., Müller, M.A., Dijkman, R., Muth, D., Demmers, J.A.A., Zaki, A., Fouchier, R.A.M., et al. (2013). Dipeptidyl peptidase 4 is a functional receptor for the emerging human coronavirus-EMC. *Nature* *495*, 251–254.
- Regoes, R.R., Hamblin, S., and Tanaka, M.M. (2012). Viral mutation rates: modelling the roles of within-host viral dynamics and the trade-off between replication fidelity and speed. *Proc R Soc London B* *280*, 20122047.
- Ren, Z. (2016). Molecular events during translocation and proofreading extracted from 200 static structures of DNA polymerase. *Nucleic Acids Res* *44*, 7457–7474.
- Reusken, C.B., Raj, V.S., Koopmans, M.P., and Haagmans, B.L. (2016). Cross host transmission in the emergence of MERS coronavirus. *Curr Opin Virol* *16*, 55–62.
- Rozen-Gagnon, K., Stapleford, K.A., Mongelli, V., Blanc, H., Failloux, A.-B., Saleh, M.-C., and Vignuzzi, M. (2014). Alphavirus mutator variants present host-specific defects and attenuation in mammalian and insect models. *Plos Pathog* *10*, e1003877.
- Sampson, T.R., Saroj, S.D., Llewellyn, A.C., Tzeng, Y.-L., and Weiss, D.S. (2013). A CRISPR/Cas system mediates bacterial innate immune evasion and virulence. *Nature* *497*, 254–257.
- Sanjana, N.E., Shalem, O., and Zhang, F. (2014). Improved vectors and genome-wide libraries for CRISPR screening. *Nat Meth* *11*, 783–784.
- Sanjuán, R. (2012). From molecular genetics to phylodynamics: evolutionary relevance of mutation rates across viruses. *Plos Pathog* *8*, e1002685.
- Sanjuán, R., Nebot, M.R., Chirico, N., Mansky, L.M., and Belshaw, R. (2010). Viral mutation rates. *J Virol* *84*, 9733–9748.
- Schaaper, R.M. (1993). Base selection, proofreading, and mismatch repair during DNA replication in *Escherichia coli*. *J Biol Chem* *268*, 23762–23765.
- Scheuermann, R.H., and Echols, H. (1984). A separate editing exonuclease for DNA replication: the epsilon subunit of *Escherichia coli* DNA polymerase III holoenzyme. *Proceedings of the National ...*

- Schmitt, M.W., Kennedy, S.R., Salk, J.J., Fox, E.J., Hiatt, J.B., and Loeb, L.A. (2012). Detection of ultra-rare mutations by next-generation sequencing. *Proc Natl Acad Sci USA* *109*, 14508–14513.
- Sevajol, M., Subissi, L., Decroly, E., Canard, B., and Imbert, I. (2014). Insights into RNA synthesis, capping, and proofreading mechanisms of SARS-coronavirus. *Virus Res* *194*, 90–99.
- Severson, W.E., Schmaljohn, C.S., Javadian, A., and Jonsson, C.B. (2003). Ribavirin Causes Error Catastrophe during Hantaan Virus Replication. *J Virol* *77*, 481–488.
- Sexton, N.R., Smith, E.C., Blanc, H., Vignuzzi, M., Peersen, O.B., and Denison, M.R. (2016). Homology-Based Identification of a Mutation in the Coronavirus RNA-Dependent RNA Polymerase That Confers Resistance to Multiple Mutagens. *J Virol* *90*, 7415–7428.
- Seybert, A., Posthuma, C.C., van Dinten, L.C., Snijder, E.J., Gorbalenya, A.E., and Ziebuhr, J. (2005). A complex zinc finger controls the enzymatic activities of nidovirus helicases. *J Virol* *79*, 696–704.
- Shi, S.T., and Lai, M.M.C. (2005). Viral and cellular proteins involved in coronavirus replication. *Curr Top Microbiol Immunol* *287*, 95–131.
- Shirato, K., Maeda, K., Tsuda, S., Suzuki, K., Watanabe, S., Shimoda, H., Ueda, N., Iha, K., Taniguchi, S., Kyuwa, S., et al. (2012). Detection of bat coronaviruses from *Miniopterus fuliginosus* in Japan. *Virus Genes* *44*, 40–44.
- Shuaib, W., Stanazai, H., Abazid, A.G., and Mattar, A.A. (2016). Re-Emergence of Zika Virus: A Review on Pathogenesis, Clinical Manifestations, Diagnosis, Treatment, and Prevention. *Am J Med* *129*, 879.e7–879.e12.
- Smith, E.C., and Denison, M.R. (2012). Implications of altered replication fidelity on the evolution and pathogenesis of coronaviruses. *Curr Opin Virol* *2*, 519–524.
- Smith, E.C., and Denison, M.R. (2013). Coronaviruses as DNA wannabes: a new model for the regulation of RNA virus replication fidelity. *Plos Pathog* *9*, e1003760.
- Smith, E.C., Blanc, H., Vignuzzi, M., and Denison, M.R. (2013). Coronaviruses lacking exoribonuclease activity are susceptible to lethal mutagenesis: evidence for proofreading and potential therapeutics. *Plos Pathog* *9*, e1003565.
- Smith, E.C., Case, J.B., Blanc, H., Isakov, O., Shomron, N., Vignuzzi, M., and Denison, M.R. (2015). Mutations in Coronavirus Nonstructural Protein 10 Decrease Virus Replication Fidelity. *J Virol* *89*, 6418–6426.
- Smith, E.C., Sexton, N.R., and Denison, M.R. (2014). Thinking Outside the Triangle: Replication Fidelity of the Largest RNA Viruses. *Annu Rev Virol* *1*, 111–132.
- Snijder, E.J., Decroly, E., and Ziebuhr, J. (2016). The Nonstructural Proteins Directing Coronavirus RNA Synthesis and Processing. *Adv Virus Res* *96*, 59–126.

Stapleford, K.A., Rozen-Gagnon, K., Das, P.K., Saul, S., Poirier, E.Z., Blanc, H., Vidalain, P.-O., Merits, A., and Vignuzzi, M. (2015). Viral Polymerase-Helicase Complexes Regulate Replication Fidelity To Overcome Intracellular Nucleotide Depletion. *J Virol* 89, 11233–11244.

Stenglein, M.D., Jacobson, E.R., Wozniak, E.J., Wellehan, J.F.X., Kincaid, A., Gordon, M., Porter, B.F., Baumgartner, W., Stahl, S., Kelley, K., et al. (2014). Ball python nidovirus: a candidate etiologic agent for severe respiratory disease in *Python regius*. *MBio* 5, e01484–14.

Stockman, L.J., Bellamy, R., and Garner, P. (2006). SARS: systematic review of treatment effects. *PLoS Med* 3, e343.

Subissi, L., Posthuma, C.C., Collet, A., Zevenhoven-Dobbe, J.C., Gorbalenya, A.E., Decroly, E., Snijder, E.J., Canard, B., and Imbert, I. (2014). One severe acute respiratory syndrome coronavirus protein complex integrates processive RNA polymerase and exonuclease activities. *Proc Natl Acad Sci USA* 111, E3900–E3909.

Thi, E.P., Mire, C.E., Lee, A.C.H., Geisbert, J.B., Zhou, J.Z., Agans, K.N., Snead, N.M., Deer, D.J., Barnard, T.R., Fenton, K.A., et al. (2015). Lipid nanoparticle siRNA treatment of Ebola-virus-Makona-infected nonhuman primates. *Nature* 521, 362–365.

Thorp, E.B., and Gallagher, T.M. (2004). Requirements for CEACAMs and cholesterol during murine coronavirus cell entry. *J Virol* 78, 2682–2692.

Tsetsarkin, K.A., Vanlandingham, D.L., McGee, C.E., and Higgs, S. (2007). A single mutation in chikungunya virus affects vector specificity and epidemic potential. *Plos Pathog* 3, e201.

Ulrich, A., Andersen, K.R., and Schwartz, T.U. (2012). Exponential megapriming PCR (EMP) cloning--seamless DNA insertion into any target plasmid without sequence constraints. *PLoS ONE* 7, e53360.

van Boheemen, S., de Graaf, M., Lauber, C., Bestebroer, T.M., Raj, V.S., Zaki, A.M., Osterhaus, A.D.M.E., Haagmans, B.L., Gorbalenya, A.E., Snijder, E.J., et al. (2012). Genomic characterization of a newly discovered coronavirus associated with acute respiratory distress syndrome in humans. *MBio* 3.

van Doremalen, N., Bushmaker, T., and Munster, V.J. (2013). Stability of Middle East respiratory syndrome coronavirus (MERS-CoV) under different environmental conditions. *Euro Surveill* 18.

van Hemert, F., and Berkhout, B. (2016). Nucleotide composition of the Zika virus RNA genome and its codon usage. *Virol J* 13, 95.

van Hemert, F., van der Kuyl, A.C., and Berkhout, B. (2016). Impact of the biased nucleotide composition of viral RNA genomes on RNA structure and codon usage. *J Gen Virol* 97, 2608–2619.

Van Slyke, G.A., Arnold, J.J., Lugo, A.J., Griesemer, S.B., Moustafa, I.M., Kramer, L.D., Cameron, C.E., and Ciota, A.T. (2015). Sequence-Specific Fidelity Alterations Associated with

West Nile Virus Attenuation in Mosquitoes. *Plos Pathog* *11*, e1005009.

Velthuis, te, A.J.W., Arnold, J.J., Cameron, C.E., van den Worm, S.H.E., and Snijder, E.J. (2009). The RNA polymerase activity of SARS-coronavirus nsp12 is primer dependent. *Nucleic Acids Res* *38*, 203–214.

Velthuis, te, A.J.W., van den Worm, S.H.E., and Snijder, E.J. (2012). The SARS-coronavirus nsp7+nsp8 complex is a unique multimeric RNA polymerase capable of both de novo initiation and primer extension. *Nucleic Acids Res* *40*, 1737–1747.

Velthuis, te, A.J.W. (2014). Common and unique features of viral RNA-dependent polymerases. *Cell Mol Life Sci* *71*, 4403–4420.

Verdaguer, N., and Ferrer-Orta, C. (2012). Conformational changes in motif D of RdRPs as fidelity determinant. *Structure* *20*, 1448–1450.

Vignuzzi, M., Stone, J.K., and Andino, R. (2005a). Ribavirin and lethal mutagenesis of poliovirus: molecular mechanisms, resistance and biological implications. *Virus Res* *107*, 173–181.

Vignuzzi, M., Stone, J.K., Arnold, J.J., Cameron, C.E., and Andino, R. (2005b). Quasispecies diversity determines pathogenesis through cooperative interactions in a viral population. *Nature* *439*, 344–348.

Vignuzzi, M., Wendt, E., and Andino, R. (2008). Engineering attenuated virus vaccines by controlling replication fidelity. *Nat Med* *14*, 154–161.

Vijgen, L., Keyaerts, E., Moës, E., Thoelen, I., Wollants, E., Lemey, P., Vandamme, A.-M., and Van Ranst, M. (2005). Complete genomic sequence of human coronavirus OC43: molecular clock analysis suggests a relatively recent zoonotic coronavirus transmission event. *J Virol* *79*, 1595–1604.

Walsh, D., Mathews, M.B., and Mohr, I. (2013). Tinkering with translation: protein synthesis in virus-infected cells. *Cold Spring Harb Perspect Biol* *5*, a012351–a012351.

Wang, J., Dong, H., Chionh, Y.H., McBee, M.E., Sirirungruang, S., Cunningham, R.P., Shi, P.-Y., and Dedon, P.C. (2016). The role of sequence context, nucleotide pool balance and stress in 2'-deoxynucleotide misincorporation in viral, bacterial and mammalian RNA. *Nucleic Acids Res* gkw572.

Weiss, S.R., and Navas-Martin, S. (2005). Coronavirus pathogenesis and the emerging pathogen severe acute respiratory syndrome coronavirus. *Microbiol Mol Biol Rev* *69*, 635–664.

Wertheim, J.O., Chu, D.K.W., Peiris, J.S.M., Kosakovsky Pond, S.L., and Poon, L.L.M. (2013). A case for the ancient origin of coronaviruses. *J Virol*.

Woo, P.C.Y., Lau, S.K.P., Huang, Y., Tsoi, H.-W., Chan, K.H., and Yuen, K.Y. (2005). Phylogenetic and recombination analysis of coronavirus HKU1, a novel coronavirus from

patients with pneumonia. *Arch Virol* 150, 2299–2311.

Woo, P.C.Y., Lau, S.K.P., Huang, Y., and Yuen, K.-Y. (2009). Coronavirus diversity, phylogeny and interspecies jumping. *Exp Biol Med* 234, 1117–1127.

Xiao, Y., Ma, Q., Restle, T., Shang, W., Svergun, D.I., Ponnusamy, R., Sczakiel, G., and Hilgenfeld, R. (2012). Nonstructural proteins 7 and 8 of feline coronavirus form a 2:1 heterotrimer that exhibits primer-independent RNA polymerase activity. *J Virol* 86, 4444–4454.

Xiao, Y., Rouzine, I.M., Bianco, S., Acevedo, A., Goldstein, E.F., Farkov, M., Brodsky, L., and Andino, R. (2016). RNA Recombination Enhances Adaptability and Is Required for Virus Spread and Virulence. *Cell Host Microbe* 19, 493–503.

Xu, X., Liu, Y., Weiss, S., Arnold, E., Sarafianos, S.G., and Ding, J. (2003). Molecular model of SARS coronavirus polymerase: implications for biochemical functions and drug design. *Nucleic Acids Res* 31, 7117–7130.

Yang, X., Smidansky, E.D., Maksimchuk, K.R., Lum, D., Welch, J.L., Arnold, J.J., Cameron, C.E., and Boehr, D.D. (2012). Motif D of viral RNA-dependent RNA polymerases determines efficiency and fidelity of nucleotide addition. *Structure* 20, 1519–1527.

Yount, B., Denison, M.R., Weiss, S.R., and Baric, R.S. (2002). Systematic assembly of a full-length infectious cDNA of mouse hepatitis virus strain A59. *J Virol* 76, 11065–11078.

Zaki, A.M., van Boheemen, S., Bestebroer, T.M., Osterhaus, A.D.M.E., and Fouchier, R.A.M. (2012). Isolation of a Novel Coronavirus from a Man with Pneumonia in Saudi Arabia. *N Engl J Med* 367, 1814–1820.

Zeng, J., Wang, H., Xie, X., Li, C., Zhou, G., Yang, D., and Yu, L. (2014). Ribavirin-Resistant Variants of Foot-and-Mouth Disease Virus: The Effect of Restricted Quasispecies Diversity on Viral Virulence. *J Virol*.